

A Complex Systems Perspective on Economic Modelling: From Macroeconomic Agent-Based Models to Labor Market Bottlenecks

Max Knicker

▶ **To cite this version:**

Max Knicker. A Complex Systems Perspective on Economic Modelling : From Macroeconomic Agent-Based Models to Labor Market Bottlenecks. Economics and Finance. Institut Polytechnique de Paris, 2025. English. NNT : 2025IPPAX010 . tel-05117230

HAL Id: tel-05117230

https://theses.hal.science/tel-05117230v1

Submitted on 17 Jun 2025

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

INSTITUT POLYTECHNIQUE DE PARIS

NNT : 2025IPPAX010

Thèse de doctorat

ÉCOLE POLYTECHNIQUE

IP PARIS

A Complex Systems Perspective on Economic Modelling

From Macroeconomic Agent-Based Models to Labor Market Bottlenecks

Thèse de doctorat de l'Institut Polytechnique de Paris
préparée à l'École polytechnique

École doctorale n°626 École doctorale de l'Institut Polytechnique de Paris (EDIPP)
Spécialité de doctorat: Physique

Thèse présentée et soutenue à Palaiseau, le 2 Juin 2025, par

Max Sina Knicker

Composition du Jury :

Isabelle L. Salle University of Ottawa	Présidente
Jean-Pierre Nadal École Normale Supérieure, EHESS	Rapporteur
Andrea Roventini Scuola Superiore Sant'Anna	Rapporteur
Dany Lang Université de Paris XIII	Examinateur
Michael Benzaquen CNRS & École Polytechnique	Directeur de thèse
Jean-Philippe Bouchaud Capital Fund Management	Co-directeur de thèse

Here's to Love.

Acknowledgements

In 1675, Isaac Newton famously wrote to Robert Hooke, *If I have seen further, it is by standing on the shoulders of giants.* This sentiment resonates deeply with me, as the completion of this thesis would not have been possible without the support, guidance, and inspiration of many people. I am profoundly grateful to those who have shaped this journey, providing both intellectual and personal encouragement along the way.

First and foremost, I extend my deepest thanks to Michael Benzaquen, whose unwavering support and encouragement allowed me to pursue my ideas without hesitation. He always listened with patience, never dismissing a thought as too ambitious or unfeasible, but instead pushing me to explore it further. His ability to provide the right words at the right moment and his constant focus on finding solutions have been invaluable to me. I am particularly grateful for his guidance in helping me structure my thoughts before speaking, that is an ability I continue to refine.

I am equally grateful to Jean-Philippe Bouchaud, who always found time to discuss my ideas, no matter how rudimentary they might have been. His boundless curiosity and pure passion for uncovering new phenomena have been an endless source of inspiration. He taught me to embrace intellectual exploration, even without a clear destination in mind.

I want to thank Karl Naumann-Woleske, who introduced me to the world of heterodox economics at the beginning of my PhD. I fondly remember the countless hours spent walking through Paris, Pisa, and Vienna, debating different schools of thought in economics. Those discussions shaped my enthusiasm for contributing to the field.

A special thanks to Elia Moretti, my brother-in-arms throughout this PhD journey. Sharing both the struggles and joys of this process with him has been invaluable. His ability to listen and provide fresh perspectives has helped shape me not just as a researcher, but also as a person. Our discussions in Ancona and Quebec City pushed the boundaries of my thinking, and for that, I am grateful.

I also want to thank Max Topel, who became a close friend during his time visiting the lab. His unwavering dedication to sports was a reminder that every challenge, academic or physical, is there to be mastered. Similarly, I am thankful to Guillaume Maitrier, who brought an infectious positivity to our office every single day. I hope he never loses that gift.

Beyond that, I am grateful to the entire Econophysics Lab -Jérôme, Samy, Rudy, Cécilia, Nirbhay, Salma, Antoine, Ruben, Fabian, Victor, Thomas, Jutta, Ortensia and Natascha- for creating an intellectually stimulating and supportive environment. Our countless

discussions over lunch, in group meetings, and at lab retreats ranged from science to philosophy, sports, and politics, making every day a learning experience.

I also want to thank Francesco Zamponi with whom I collaborated on my first paper, as well as Florian, Andrea, and Simone I met at various conferences, including our shared experience at GeoInno in Manchester. Their insights and camaraderie welcomed me into the broader network and complexity science community.

My gratitude extends to the CFM Chair of Econophysics & Complex Systems which provided an ideal space to explore new ideas without the usual constraints of academia. Likewise, I am thankful to Capital Fund Management which not only hosted me as a researcher but also offered an intellectually enriching environment, with discussions and seminars that broadened my perspectives.

None of this would have been possible without the unconditional support of my family. My mother, who instilled in me from an early age the belief that anything is possible, and whose love knows no bounds. My father, whose infinite patience and curiosity about everything in the world influenced me profoundly. My brother, who has always been there to shield me from challenges, fighting battles so that I wouldn't have to.

Finally, my deepest gratitude to Sophia Heger, who stood by my side without hesitation. Through struggles and moments of joy alike. Her unwavering optimism and problem-solving mindset continue to inspire me every day.

To my friends, too many to name, from Paris, Munich, and Boston: even across distances, your presence has been a source of energy and renewal whenever I needed it most.

I am incredibly grateful to all of you. In every aspect of my life, I have been able to count on your belief in me, your open ears, and your helping hands.

Foreword

In 2016, shortly after starting my bachelor's degree in physics, I found myself staring at a blank page, trying to write a motivational letter for a scholarship. Having never written one before, I asked myself the fundamental question: *Why do I do what I do?* As I reflected, the words started accumulating on the page. The opening paragraph reads:

> *Since I was in the 7th grade, I wanted to study physics. I was fascinated by the world of formulas which are feasible for nearly every nature-related topic, in technological areas or in daily life. This fascination is conserved till now and as far as I see, it will never decay. Physics means for me learning every day something new, improving my analytical and problem-solving skills and the ability to understand complex issues. My vision is to develop theories which are on the one hand applicable for a broad range of people and on the other hand describe the nature of life.*

Looking back, I'm struck by how precisely these words still capture my motivation for studying physics. I was also quietly amused by the subtle insider joke I had unknowingly included. Despite this heartfelt statement, I didn't receive the scholarship, but it did not change anything about my motivation. Yet, in the years that followed, almost every application I wrote included some version of these sentences. Not because I was lazy, but because my perspective at the age of 20 hasn't changed. I am still fascinated. I am still curious. I am still eager to learn. And I have always aspired to use my skills to develop models that are both applicable and useful to a broad range of people. I hope this PhD thesis reflects that ambition.

Like many scientists, I have always regarded the Nobel Prize as a north star, a recognition of those who, after decades of mastering their field, leave a lasting impact on generations to come. The Physics and Chemistry Nobel medal depicts the goddess Isis, holding a cornucopia, her eyes veiled as the Genius of Science gently lifts the covering. Along the medal's edge, an inscription reads: *Inventas vitam iuvat excoluisse per artes—It is beneficial to have improved human life through discovered arts.* I realized then that my perception of science had long been guided by this very idea.

Although my younger self was uncertain about which concrete path to follow, everything became clear in 2018 when I stumbled upon the textbook *Statistical Physics of Financial Markets* [1]. I did not fully understand every detail back then, but I knew I wanted to work in a field where I could merge my fascination for physics with real-world applications. A field that could contribute to improving people's lives. I am incredibly grateful that, whether because of or despite the first lines of my motiva-

tional letter, I was given the opportunity to pursue this passion. Michael Benzaquen and Jean-Philippe Bouchaud took me in, first as an intern and later as a PhD student, for which I owe them my deepest gratitude. They provided me with the resources, patience, and mentorship that allowed me to explore the questions that had fascinated me for nearly a decade. They encouraged me to follow my curiosity while always offering guidance whenever I felt lost.

The work presented in this thesis is the result of that journey—an exploration of complex systems, mathematical modeling, and socio-economic interactions, shaped by years of curiosity, fascination, and the unwavering willingness to learn new things.

Publications

[2] Karl Naumann-Woleske, **Max Sina Knicker**, Michael Benzaquen, Jean-Philippe Bouchaud. *Exploration of the Parameter Space in Macroeconomic Models.* Routledge International Handbook of Complexity Economics. Routledge, 2024. 190-207.
[3] **Max Sina Knicker**, Karl Naumann-Woleske, Jean-Philippe Bouchaud, and Francesco Zamponi. *Post-COVID inflation and the monetary policy dilemma: An agent-based scenario analysis.* Journal of Economic Interaction and Coordination (2024): 1-55.
[4] **Max Sina Knicker**, Karl Naumann-Woleske, Michael Benzaquen. *The Structure of Occupational Mobility in France.* Journal of Statistical Mechanics: Theory and Experiment, 2025(5), 053403.

Résumé en Français

L'alliance des méthodes « à la physicienne » issue des systèmes complexes avec la modélisation économique offre un cadre puissant pour explorer les interdépendances complexes au sein des systèmes économiques. Grâce aux outils de modélisation basés sur les agents et de science des réseaux, ce travail apporte de nouvelles perspectives concernant les défis économiques contemporains. Cette approche, inspirée par la physique, met en lumière la nature dynamique et interconnectée des comportements économiques, proposant des cadres novateurs pour analyser et traiter les complexités des systèmes économiques.

Les modèles basés sur les agents (ABM) proposent une approche ascendante pour simuler les dynamiques macroéconomiques, allant du microscopique au macroscopique. Ces modèles semblent particulièrement efficaces pour comprendre les phénomènes économiques majeurs tels que les cycles, les tendances inflationnistes et les crises. En étendant le modèle Mark-0, cette recherche intègre un mécanisme d'ancrage des attentes pour étudier les moteurs de l'inflation post-COVID et l'efficacité des politiques visant à la contenir. L'analyse révèle que le succès de la politique monétaire repose davantage sur l'ancrage des attentes que sur les effets économiques directs des ajustements des taux d'intérêt. Des dynamiques clés, telles que la spirale salaire-prix, apparaissent comme des mécanismes critiques influençant les résultats économiques, tandis que l'analyse met en évidence des arbitrages entre le contrôle de l'inflation et le chômage. Les résultats soulignent également qu'en l'absence de mesures fiscales solides, la reprise économique pourrait stagner, voire s'aggraver en une récession plus profonde.

Pour améliorer la robustesse et l'interprétabilité des ABM, une approche novatrice inspirée par la physique statistique est introduite afin d'explorer systématiquement les espaces de paramètres à haute dimension. En s'appuyant sur le cadre de la *sloppiness*, cette étude identifie un ensemble restreint de combinaisons de paramètres clés — les « directions rigides » — qui influencent de manière significative le comportement du modèle, tout en dévalorisant les paramètres les moins impactants. Cette démarche conduit au développement d'un algorithme efficace qui capture l'ensemble des dynamiques possibles du modèle avec un coût computationnel minimal. Appliquée au modèle Mark-0, la méthode restitue le spectre complet des dynamiques de chômage, offrant une approche structurée pour évaluer la sensibilité des paramètres et améliorer l'utilité des ABM.

La seconde partie de la recherche se concentre sur les marchés du travail, considérés comme des systèmes adaptatifs complexes. Interpréter le marché du travail comme

un réseau d'occupations interconnectées met en évidence des déséquilibres structurels, avec des transitions et des pourcentages d'occupation d'emplois distribués en loi de puissance. Cette structure hiérarchique révèle que certaines professions agissent comme des hubs de mobilité, tandis que d'autres restent isolées. Pour quantifier ces dynamiques, de nouvelles métriques — accessibilité et transférabilité — sont développées, offrant une compréhension nuancée de la mobilité des travailleurs. Ces métriques identifient des rôles critiques dans le marché du travail, tels que les professions *condensatrices*, qui servent de goulets d'étranglement en raison de leur forte accessibilité mais faible transférabilité, entravant ainsi la mobilité interprofessionnelle. Une analyse de l'état stationnaire révèle des inefficacités à long terme, les distributions observées des travailleurs s'écartant significativement des états d'équilibre. L'intégration de données sur la similarité des compétences enrichit l'analyse, montrant que les professions à haute transférabilité facilitent les transitions intercommunautaires, tandis que les rôles à haute accessibilité agissent comme des portes d'entrée vers de nouveaux secteurs. Ces perspectives trouvent des applications pratiques, allant de la résolution des inadéquations de compétences à la conception de programmes de formation ciblés, en passant par l'amélioration de la fluidité du marché du travail face aux évolutions technologiques et économiques.

Cette thèse propose des outils et méthodologies prenant origine dans la physique statistique et la théorie des réseaux pour démêler les complexités économiques, identifier les moteurs clés et traiter les inefficacités structurelles. Elle pourra ainsi servir de base pour des interventions politiques innovantes dans les systèmes économiques interconnectés.

Acknowledgements iii

Foreword v

Publications vii

Résumé en Français viii

Introduction 1

1. Introduction 2
1.1. Macroeconomic Modelling . . . 3
1.2. Modelling Labor Markets . . . 7

Part I: Exploring Macroeconomic Agent-Based Models & Inflation 11

2. Agent-Based Models and Parameter Exploration 12
2.1. Introduction to Agent-Based Models . . . 13
2.2. Systematic Parameter Exploration . . . 19

3. Post-COVID Inflation & the Monetary Policy Dilemma: An Agent-Based Scenario Analysis 23
3.1. The Mark-0 Model . . . 27
3.2. Policy Channels in Mark-0 . . . 33
3.3. Stationary Dynamics: the Role of the Central Bank . . . 37
3.4. The COVID Shock and its Aftermath in the Absence of Monetary Policy . . . 38
3.5. Monetary Policy Response to Inflationary Shocks . . . 43
The Beveridge Curve and Labor Market Adjustments Post-Shock . . . 48
3.6. Fiscal Stimulus . . . 49
3.7. Model Sensitivity: The Dangers of a Wage-Price Spiral . . . 51
3.8. Conclusions . . . 53

4. Exploration of the Parameter Space in Macroeconomic Models 57
4.1. Macroeconomic Models Considered . . . 60
4.2. Are Macroeconomic Models Sloppy? . . . 61
4.3. A New Method for Systematic Phase Exploration . . . 66
4.4. An Application to the Mark-0 Agent-Based Model . . . 68
4.5. Extension with KL-Divergence . . . 71
4.6. Conclusions . . . 73

Part II: Occupational Mobility & the Structure of Worker Transitions 75

5. Labor Markets as Complex Systems: A Quantitative Framework 76
5.1. French Labor Market Data . . . 77
5.2. Power Law Distributions in the Labor Market . . . 79
5.3. Occupational Flow Ratios in the Labor Market . . . 82
5.4. Occupational Transition Matrix . . . 84
5.5. Markov Chain . . . 85
5.6. Markovian Assumption and Empirical Data . . . 86
5.7. Steady State of the Labor Market . . . 89
5.8. Detailed Balance of the Labor Market . . . 91
5.9. Dynamical Properties of the Labor Market . . . 92
5.10. On Measurement Noise of Transition Matrices . . . 95

5.11. Network Properties of the Labor Market . . . 97

6. The Structure of Occupational Mobility in France 105
6.1. Literature Review . . . 107
6.2. Occupational Transition Network of France . . . 108
6.3. Identifying Bottlenecks: A Data-driven Taxonomy . . . 112
6.4. Informing Policy using Accessibility and Transferability . . . 119
6.5. Conclusions . . . 122

7. Skill Network of the Labor Market 124
7.1. French Skill Data . . . 124
7.2. Network Properties of Skill Network . . . 126
7.3. Relationship of Skills and Transition Rates . . . 127
7.4. Relationship of Skills and Accessibility and Transferability . . . 128
7.5. Skill Effects on Accessibility and Transferability . . . 129
7.6. Expert-Assessed Skill Network . . . 130
7.7. Routine, Non-Routine, and Cognitive-Manual Classification . . . 131
7.8. Accessibility-Transferability Ratios and Routine Occupations . . . 133

PART III: THE INTERPLAY OF SKILL, LUCK & COMPETITION IN SUCCESS 135

8. Beyond Luck: Measuring Skill and Competition in Venture Capital and Analyst Forecasts 136
8.1. Introduction . . . 136
8.2. Literature Review . . . 137
8.3. A Framework for Analyzing Competitive Performance . . . 139
8.4. Skill and Randomness in Venture Capital Investments . . . 141
Measuring of True Skill . . . 146
8.5. Skill and Randomness in Analyst Forecasts . . . 147
8.6. Skill & Competition: A Comparative Analysis . . . 153
8.7. Conclusions . . . 154

CONCLUSIONS 157

9. Perspectives & Conclusions 158

Bibliography 162

APPENDIX 180

A. Part I 181
A.1. Hessian of Schelling Model within the boundaries . . . 181
A.2. Extension KL divergence at the boundaries . . . 184
A.3. Table of Parameters . . . 185
A.4. Counterfactuals . . . 186
A.5. Monetary Policy . . . 186
A.6. Importance of Expectation Anchoring . . . 187
A.7. Monetary Policy with Stronger Central Bank Reaction . . . 188
A.8. Sensitivity of Monetary Policy to α_c and α_Γ . . . 188
A.9. Helicopter Money . . . 189

A.10. Windfall Tax 190
A.11. Sloppiness 192
A.12. Sensitivity to the Indexation Parameters g_p and g_w 194

B. Part II **195**
B.1. Cluster Methods & Labor Data 195
B.2. Temporal Stability 196
B.3. Entry and Exit from the Occupation Network 197
B.4. BRIM Communities 198
B.5. Taxonomy Clusters 212
B.6. Inter- vs. Intra-Community Transitions 227
B.7. Age Effect on Accessibility and Transferability 227
B.8. Policy Evaluation with Matrix Perturbation 229
B.9. Task-based Cluster 232

List of Figures

2.1. **1-dimensional Cellular Automata model:** (a,c) Schematic overview of Rule 30 (126) for determining the next state in a 1D cellular automaton, defined as $f_{30}(x_{i-1}^t, x_i^t, x_{i+1}^t) = x_i^{t+1}$ (Eq. 2.1), with a neighborhood size of 2. There are $2^3 = 8$ possible neighborhood states, and the rule f_{30} specifies the next state for each of these configurations. Empty cells represent dead cells ($x_i = 0$), while blue cells indicate living cells ($x_i = 1$). (b,d) Simulation of the cellular automaton governed by Rule 30 (126), starting with an initial condition of a single living cell positioned at the center of the line at $t = 0$ and evaluated for 30 time steps. 13

2.2. **Neighborhood configuration in Schelling model:** Configurations of a city in the Schelling model as extended by Grauwin et al. (2009) [63], consisting of $Q = 36$ blocks, each containing $H = 100$ cells, with an initial density of $\rho_0 = 0.5$. (a) Stationary state of the model with parameters $m = 0.5$, $\alpha = 1$, and $T \to 0$. (b) Stationary state of the model with parameters $m = 0.5$, $\alpha = 0$, and $T \to 0$. 14

2.3. **Talent vs. Capital distribution:** The relationship between talent and final capital in the Talent vs. Luck model. Each point represents an individual's talent and accumulated capital after 40 time steps (years). The system is initialized with talent distributed normally ($\mu_T = 0.6$, $\sigma_T = 0.1$) across $N = 500$ individuals, each starting with an equal initial capital of $C(0) = 10$. The parameters are choosen as in Pluchino et al. (2018) [66]. The simulation is run 5,000 times, and all individuals from all runs are plotted. The color intensity of the points indicates the number of lucky events encountered by each individual. The purple distribution in the top panel shows the talent for the wealthiest individual after each run, with a mean talent of $\langle T \rangle_{max(C_i)} = 0.67$. 15

2.4. **Visualization of the loss function in parameter space:** Contour plot of the loss function for a two-dimensional model parameterized by Φ_1 and Φ_2. Movement along the stiff axis results in significant changes to the loss function, indicating strong sensitivity to parameter variations in this direction. In contrast, movement along the sloppy axis leads to minimal changes in the model's dynamics, demonstrating insensitivity to parameter variations in this direction. Figure adapted from Gutenkunst et al. (2007) [93]. 19

3.1. **Economic dashboard for the three Central Bank scenarios described in sections 3.3, 3.3 in absence of any policy and shock:** Blue lines: Inactive Central Bank scenario ($\pi^\star = \phi_\pi = 0$, $\rho^\star = 1.2\%$ p.a.). Orange lines: Reactive Central Bank with Anchored Expectation scenario ($\pi^\star = 2.4\%$ p.a., $\phi_\pi = 1$, $\tau^T = 0.95$). Green lines: Reactive Central Bank with Floating Expectation scenario. The complete parameter set can be seen in table A.1. The two insets correspond to the demand/output ratio and the bankruptcy rate. 37

3.2. **Empirical shocks**: Fitting of shock scenarios to macro data from the US up to end of 2022, when the present analysis was conducted. The empirical series (dashed lines) are indexed to January 2020. (Left Panel) Empirical series for the Personal Consumption Expenditure (orange, dashed line) and Industrial Production Index (blue, dashed line). Solid lines represent our model for the shocks to consumption propensity c (orange) and productivity ζ (blue). Dark grey area corresponds to the COVID Shock and the grey area to the Supply Chain Shock in the shock scenarios that we apply to Mark-0. (Right Panel) WTI Crude oil spot price during the periods of shock (green, dashed line). Modelling of the artificial price series (green, solid) in the Mark-0 model for a period of two and a half years (light grey area). Data retrieved from FRED (2023) [134], FRED (2023) [135], FRED (2023) [136]. 39

3.3. **Economic dashboard for the three shocks in the Inactive Central Bank scenario and in the absence of any policies:** The dynamics for the three shock scenarios, COVID only (blue), COVID and Supply Chain shock (orange) and all shocks (green) when the Central Bank is inactive and no policy is applied. The areas shaded in grey indicate the duration of the three shocks: the COVID shock lasting until the end of the dark grey area, the supply chain shock until the end of the grey area, and the energy price shock until the end of the light grey area. The dashed black line represents pre-shock averages. In the first two cases, the economy is able to recover on its own (although with significant fluctuations that last for a few years), but significantly gets worse when the energy shock is turned on. 40
3.4. **Counterfactual COVID shock in the Inactive Central Bank scenario and in the absence of any policies:** Unemployment (left) and inflation (right) in the Inactive Central Bank scenario without any policy, with an amplified COVID shock by a factor κ and Supply Chain shock. $\kappa = 1$ (golden line) corresponds to the COVID shock considered in the previous section, with a very low level of unemployment. With $\kappa = 1.1$ (pink line), the unemployment and inflation dynamics change only very slightly, as the shock is mild. As soon as $\kappa \geq 1.2$, the economy is unable to recover spontaneously, and output collapses as demonstrated in the $\kappa = 1.2$ case (blue line). 42
3.5. **Economic dashboard for a stronger COVID shock in the Inactive Central Bank scenario, with and without Easy-Credit policy:** The dynamics of the Inactive CB scenario in case (ii) (COVID and supply chain shock but no energy price shock), for a COVID shock strength of $\kappa = 1.2$ without Easy-Credit policy (blue) and $\kappa = 1.2$ with Easy-Credit policy, $\mu = 1.3$ (orange). We see that policy is effective in preventing the economic collapse in the case of the stronger COVID shock. 43
3.6. **Unemployment and Inflation with all shocks, Easy-Credit policy, and three distinct monetary policy scenarios**. (Blue lines) Inactive Central Bank scenario. (Orange lines) Reactive Central Bank with Anchored Expectation, with anchor parameter $\tau^T = 0.95$. A Taylor rule policy successfully decreases peak inflation but increases peak unemployment. (Green lines) Reactive Central Bank with Floating Expectation. Here, trust is eroded during the high inflation period. Monetary policy then fails at reducing inflation, with the risk of hyper-inflation lurking (see section 3.7); unemployment remains lower because of the higher inflation. 46
3.7. **The Beveridge Curve and Labor Market Adjustments Across Policy Scenarios:** Relationship between unemployment and the required workforce Eq.3.29 across different monetary policy regimes and all shocks. The three panels correspond to: (a) Inactive CB, (b) Reactive CB with anchored trust, and (c) Reactive CB with floating trust. The color gradient represents time progression. Across all scenarios, the model generates a counterclockwise rotation of the Beveridge Curve. 48
3.8. **Helicopter Money in the Reactive Central Bank with Floating Expectation scenario, Easy-Credit policy, and all shocks.** Unemployment (left) and inflation (right) for a helicopter drop of size κ_H times savings one month after the price shock. Already with $\kappa_H \geq 0.2$ unemployment is reduced almost to zero. A further increase of Helicopter Money only increases the duration of the high-inflation period, without further reducing unemployment. Note the small unemployment and inflation "ripples" persisting several years after the initial shock. 49
3.9. **Windfall Tax in the Reactive Central Bank with Floating Expectation scenario, Easy-Credit policy, and all shocks.** Unemployment (left) and inflation (right) for a Windfall Tax of $\delta_e + \Delta\delta_e$ one year before the end of the price shock with a duration of two years. With $\Delta\delta_e \approx 4\%$ unemployment is reduced strongly. A further increase of tax does only increase unemployment again. For even larger $\Delta\delta_e$, inflation increases because of increased demand due to increased savings. 50

3.10. **Hyperinflation tipping points.** Left: Stable inflation vs. Hyperinflation in the plane ($g_p = g_w$, τ^T) in the Anchored Expectation scenario and in absence of exogenous shocks. Note that inflation can be stable even when $g_p = g_w > 1$ for strong enough anchoring. Middle: Stable inflation vs. Hyperinflation in the plane ($g_p = g_w$, α_I) in the Floating Expectation case and in absence of exogenous shocks. Right: inflation rate (top) and unemployment (bottom) as a function of time, in the Floating Expectation case, $\alpha_I = 0.3$ and different values of $g_p = g_w$, after the three COVID shocks. When indexation is too strong, an hyperinflation regime sets in at the end of the energy price shock. 52

4.1. (LHS) Eigenvalue Spectra relative to the largest eigenvalue (λ/λ_1), and (RHS) absolute cosine similarities between each eigenvector and its closest bare parameter axis. Note that the ranking of the cosine similarities does not necessarily correspond to that of the eigenvalues, see Table 4.2. The models include in red the Mark-0 ABM [152] with all variables evaluated in the full employment phase (a, $\Theta = 2.5$, $\rho^\star = 1\%$), with the remaining parameters from Gualdi et al. (2015) [152]. In blue, the Dynamic Stochastic General Equilibrium models of (b) Galí (2015) [20] chpt. 2, (c) chpt. 3, and (d) chpt. 8, as well as the estimated models of (e) Smets and Wouters (2007) [19], (f) Del Negro et al. (2015) [175], and (g) Carlstrom et al. (2017) [176] evaluated at the authors' reported parameters. 62

4.2. (LHS) Eigenvalue Spectra relative to the largest eigenvalue (λ/λ_1), (Center) the absolute cosine similarities between each eigenvector and its most aligned bare parameter axis, and (Right) the unemployment rate for different phase, for the Mark-0 ABM [152]. At variance with Figure 4.1 we restrict the estimation to only the unemployment rate as a variable, with parameters in: the full employment phase (blue, with FE, $\Theta = 2.5$, $\rho^\star = 1\%$, and FE II, $\Theta = 5.0$, $\rho^\star = 1\%$), the residual unemployment phase (red, with RU, $\Theta = 2.5$, $\rho^\star = 5\%$, and RU II, $\Theta = 5.0$, $\rho^\star = 5\%$), and the full unemployment phase (black, FU, $\Theta = 2.5$, $\rho^\star = 15\%$, and FU II, $\Theta = 5.0$, $\rho^\star = 12.5\%$). The EC phase (green) is evaluated at $\Theta = 1.3$, $\rho^\star = 1\%$. All remaining parameters are from Gualdi et al. (2015) [152]. 63

4.3. Heatmaps of the components of the stiffest 5 eigenvectors of the Mark-0 ABM focused on the unemployment rate with all parameters. The sign of the eigenvector is such that the angle with the direction c_0 is positive. All components with a magnitude between -0.1 and 0.1 have been set to zero for exposition purposes. The directions were evaluated in the full employment phase (FE, $\Theta = 2.5$, $\rho^\star = 1\%$, and FE II, $\Theta = 5.0$, $\rho^\star = 1\%$), the residual unemployment phase (RU, $\Theta = 2.5$, $\rho^\star = 5\%$, and RU II, $\Theta = 5.0$, $\rho^\star = 5\%$), and the full unemployment phase (FU, $\Theta = 2.5$, $\rho^\star = 15\%$, and FU II, $\Theta = 5.0$, $\rho^\star = 12.5\%$). All remaining parameters are from Gualdi et al. (2015) [152]. 64

4.4. Unemployment dynamics in different phases in their original form (black), with a unit log step in the stiffest (red) and sloppiest (green) directions identified in Figure 4.3, i.e. $\log \Phi' = \log \Phi + v_i$. The panels are: (left) the full employment phase (FE, $\Theta = 2.5$, $\rho^\star = 1\%$), (center) the residual unemployment phase (RU, $\Theta = 2.5$, $\rho^\star = 5\%$), and (right) the full unemployment phase (FU, $\Theta = 2.5$, $\rho^\star = 15\%$). All remaining parameters are from Gualdi et al. (2015) [152]. . 65

4.5. (LHS) Phase diagram of the Mark-0 model in the $(\log \Theta, \log \rho^\star)$ plane with projected eigenvectors. The eigenvectors, v_1 (red) and v_2 (blue), are based on the 2-dimensional Hessian with only Θ and $\rho^\star$ using unemployment as output variable. Dashed black lines indicate the locations of phase transitions. (RHS) logarithm of the ratio of the first to the second eigenvalue at different points in the phase space. The black dots indicate $\ln(\lambda_2/\lambda_1) < -10$. Simulation parameters: $T = 30,000$, $S = 20$, $T_{\text{eq}} = 10,000$. The parameters of the model, other than Θ and $\rho^\star$, are from Gualdi et al. (2015) [152] 65

4.6. (LHS) Phase diagram of the Mark-0 model without central bank in the $(\log \Theta, \log \rho^\star)$ plane with two likely algorithm paths for three different starting values. Solid lines show the most likely path following only v_1, while dotted lines indicate an alternate path with mixed steps. Dashed black lines indicate the locations of phase transitions. (b) Dynamics of the unemployment rate for different steps along the v_1 path for each starting point. Simulation parameters: $T = 30,000$, $S = 20$, $T_{\text{eq}} = 10,000$. The parameters of the model, other than Θ and $\rho^\star$, are those of Gualdi et al. (2015) [152]. Algorithm parameters are $\varepsilon_{\min} = 0.3$, $\varepsilon = 0.1$, $\varepsilon_{\max} = 1.0$. 69
4.7. Dynamics of the unemployment rate for different steps in the 3D case (Θ, $\rho^\star$, α_Γ) and the multi-dimensional ($P = 14$) case. We consider again the three most likely paths with 5 steps. The initial parameters of the algorithm in the (Θ, $\rho^\star$)-space are shown in the top right-hand corner of each panel and correspond to those of Figure 4.6). Simulation parameters: $T = 30,000$, $S = 20$, $T_{\text{eq}} = 10,000$. All other parameters are those of Gualdi et al. (2015) [152]. 70

5.1. **Occupational Transition Network of France**: Graph visualization of the weighted and directed labor market network derived from the transition probability matrix $\mathcal{P}_{i,j}$ (see Eq. 5.8), computed from data spanning the years 2012 to 2020. Each node symbolizes an occupation, with links that illustrate transitions between them. Node sizes correspond to the occupation's workforce size, while node color indicates its BRIM community [187] (see Section 3). Line widths are proportional to the transition probability $\mathcal{P}_{i,j}$. The layout is based on the OpenOrd algorithm [188], and the graph was generated using Gephi [189]. 77
5.2. **Power-law distributions in the French Labor Market**: (a) Log-log plot of the empirical probability density function for occupation sizes. The inset displays the same data with linear-linear axes. (b) Log-log plot of the empirical probability density function for transition sizes between occupations. The inset shows the data with linear-linear axes. Black dotted lines represent fitted power-law trends (Eq. 5.2) as a guidance for the eye. Different colors correspond to distributions for various years. 79
5.3. **Cumulative distributions of the French Labor Market**: (a) Log-log plot of the empirical cumulative density function for occupation sizes, with a dotted line representing the fit to Equation 5.4 and an exponent $\mu = 1.15$. The inset shows the corresponding Lorenz curve, with an average Gini coefficient of $\langle g \rangle = 0.6$ across years. (b) Log-log plot of the empirical cumulative density function for transition sizes between occupations, with a dotted line representing the fit to Equation 5.4 and an exponent $\mu = 1.38$. The inset displays the Lorenz curve, with an average Gini coefficient of $\langle g \rangle = 0.85$ across years. 80
5.4. **Distribution of occupational flow ratios**: (a) Empirical probability density function of the occupational flow ratio r (Eq. 5.5). The dotted lines represent the fit using a Student's t-distribution (Eq. 5.6), with the parameter ν of the fit shown in the inset. (b) Joint distribution of the back-and-forth transitions N_{ij} and N_{ji} as a scatter plot. The inset depicts the symmetry index for the distribution (Eq. 5.7). Different colors correspond to different years. 82
5.5. **Distribution of the Occupational Transition Matrix $\mathcal{P}^t_{ij}$**: (a) Empirical cumulative distribution function of transition probabilities p_{ij}. The dotted lines represent power-law fits (Eq. 5.4) for two distinct regimes, with exponents $\mu_1 = 1.0$ and $\mu_2 = 3.1$. Inset shows a violin plot of the fraction of movers ($\sum_{i \neq j} N_{ij}$) for each year. (b) Scatter plot of the joint distribution for the detailed balance condition (Eq. 5.24). The inset displays the symmetry index of the spectral decomposition V_t (see eqs. 5.7 and 5.25). Different colors correspond to different years. 84
5.6. **Comparison of next-state predictions with $\mathcal{P}^t$ and $\hat{\mathcal{P}}$**: Plot of the relative error Eq. 5.14 for different forecasting horizons k across various years. Each point corresponds to the relative error of the prediction $\vec{x}'_{t+k} = \left(\Pi_{\tau=t}^{t+k-1} \mathcal{P}^\tau\right) \vec{x}_t$, plotted at the year $t + k$. The dotted line represents the line of equality, where predictions using $\mathcal{P}^t$ are equal to those using $\hat{\mathcal{P}}$. 86

5.7. **Pairwise distance matrix**: (a) Pairwise Wasserstein distance matrix between the steady states $W(\pi_t, \pi_{t'})$ (Eq. 5.23). The dotted line provides a visual delimitation between the steady states of $\mathcal{P}^t$ and the steady state of the average matrix $\hat{\mathcal{P}}$. (b) Pairwise Wasserstein distance matrix between the actual states $W(x_t, x_{t'})$ (Eq. 5.23). The dotted line provides a visual delimitation between the actual states of $\mathcal{P}^t$ and the steady state of the average matrix $\hat{\mathcal{P}}$ 88
5.8. **Eigenvalue spectrum of Transition Matrices**: Scatter plot of the eigenvalues of the transition matrices $\mathcal{P}_t$ and $\hat{\mathcal{P}}$ in the complex plane, with the real and imaginary components shown. Different colors correspond to different years. The inset depicts the relaxation time $T_{relax} \propto \frac{1}{1-\lambda_2}$ in units of years of the system for each year and the average matrix. 93
5.9. **Generalized cosine similarity**: Generalized cosine similarity between the subspaces spanned by the leading k eigenvectors of V_t and $\hat{V}$ across different years. The black solid line represents the generalized cosine similarity for two random matrices [206]. The inset provides a zoomed view of the plot for the first $k \geq 30$ eigenvectors. 94
5.10. **Reversed generalized cosine similarity**: Generalized cosine similarity between the subspaces spanned by the k eigenvectors starting from the smallest eigenvalue of V_t and $\hat{V}$ across different years. The black solid line indicates the generalized cosine similarity for two random matrices [206]. The inset provides a detailed view of the plot for the first $k \geq 30$ reversed eigenvectors. 95
5.11. **Effect of measurement noise**: (Top row) Eigenvalue spectrum of the sampled transition matrices (for all random seeds), where different sample sizes N_s are shown in different colors. The black points represent the true eigenvalue spectrum of $\hat{\mathcal{P}}$. The bottom row presents a box plot of the generalized cosine similarity between the sampled matrices (for all random seeds) and $\hat{\mathcal{P}}$. Different colors represent different sample sizes N_s. 96
5.12. **Indegree distribution**: Distribution of weighted in-degree centralities (Eq. 5.36) for each node in the labor market network $\mathcal{G}$. 97
5.13. **Betweenness centrality distribution**: Distribution of betweenness centralities (Eq. 5.37) of each node in the labor market network $\mathcal{G}$. 98
5.14. **Closeness centrality distribution:** Distribution of closeness centralities (Eq. 5.38) for each node in the labor market network $\mathcal{G}$. 98
5.15. **Clustering coefficient distribution:** Distribution of local clustering coefficients (Eq. 5.39) for each node in the labor market network $\mathcal{G}$. 99
5.16. **Communities by 1-digit PCS codes**: Visualization of the occupational transition matrix $\hat{\mathcal{P}}$, where color intensity represents transition probabilities. Solid lines indicate the clustering given by the 1-digit PCS codes. 99
5.17. **Community detection with Louvain**: Visualization of the occupational transition matrix $\hat{\mathcal{P}}$, where color intensity represents transition probabilities. Solid lines indicate the clustering results with the Louvain Clustering algorithms. The clustering reflects the best partition based on modularity from 500 random seeds. Color bars adjacent to the communities serve as visual guides. 100
5.18. **Community detection with BRIM**: Visualization of the occupational transition matrix $\hat{\mathcal{P}}$, where color intensity represents transition probabilities. Solid lines indicate the clustering results with the BRIM Clustering algorithms. The clustering reflects the best partition based on modularity from 500 random seeds. Color bars adjacent to the communities serve as visual guides. 101
5.19. **Community detection with Infomap**: Visualization of the occupational transition matrix $\hat{\mathcal{P}}$, where color intensity represents transition probabilities. Solid lines indicate the clustering results with the Infomap Clustering algorithms. The clustering reflects the best partition based on modularity from 500 random seeds. Color bars adjacent to the communities serve as visual guides. 101

5.20. **Community detection with Spectral Clustering**: Visualization of the occupational transition matrix $\hat{\mathcal{P}}$, where color intensity represents transition probabilities. Solid lines indicate the clustering results with the Spectral Clustering algorithms. The clustering reflects the best partition based on modularity from 500 random seeds. Color bars adjacent to the communities serve as visual guides. 102
5.21. **Comparison between community detection algorithms**: (a) Box plot illustrating the modularity (Eq. B.4) achieved by each clustering algorithm as a function of the number of clusters. (b) Box plot depicting the Gini coefficient (see Section 5.2) for each clustering algorithm as a function of the number of clusters. Each box plot represents results from 500 random seeds. . 103

6.1. **Occupational Transition Matrix and steady state distribution** (a) Temporal stability analysis from 2012-2020 where each box corresponds to the distribution of standardized coefficient $\beta_{i,t}$ (see Appendix B.2) estimated by the regression in Eq. B.6. The grey line in each box corresponds to the median of the distribution. (b) Bar plot of percentage difference between the actual share of workers in each occupation x and the share of workers implied by the steady state π. (c) Transition probability matrix $\mathcal{P}_{i,j}$, where each entry represents the fraction of individuals in $(T-1)$ (columns) who transitioned to occupation in (T) (rows). Self-loops are excluded. Red lines and colored bars indicate the BRIM communities. 108
6.2. **Inflow and outflow dynamics of the Labor Market**: Each node in the network represents an occupational BRIM community: (1) Blue-collar workers, (2) Public, Health & Education, (3) Engineers & Executives, (4) Sales & Retail (5) Transport & Logistic (6) Agriculture, (7) Artist and Creative Work (8) Medical Service and (U) Out of the occupation network. Flows indicate transitions to and from outside the labor market normalized by the total number of transitions from and to the community. 111
6.3. **Schematic visualization of the degree of Transferability and Accessibility.** An illustrative representation of a transitional occupation matrix $\tilde{\mathcal{P}}_{i,j}$, where a grey square signifies a connection between two occupations ($\tilde{\mathcal{P}}_{i,j} > 10^{-2}$), while a white square denotes no connection. (a) Transferability of an occupation is linked to the number of occupations a worker can transition to from a given occupation in time T-1. (b) Accessibility of an occupation is associated with the number of occupations from which a worker could have transitioned to a given occupation at time T. 113
6.4. **Sorted, Transition Probability Matrix.** X-axis corresponds to occupations in (T-1) and y-axis to occupations in (T). (a) Diversification and Ubiquity ordering (b) Accessibility and Transferability ordering (c) In-Block Accessibility and Transferability ordering. Each Block, detected by the BRIM algorithm (red lines), has a higher NODF than the raw matrix. The intensity of the coloring corresponds to higher probability flow from (T-1) to (T). 114
6.5. **Accessibility and Transferability Space**. (a) Each point represents a single occupation and is located in the space by its corresponding accessibility (y-axis) and transferability (x-axis) metric. The dashed lines denote the mean-shift clustering boundaries at $\Theta^A = 0.0031$ and $\Theta^T = 49.9$. The histograms show the distribution along the accessibility (horizontal) and transferability (vertical) dimensions. (b) Proposal of a new taxonomy based on the degree of accessibility and transferability of an occupation into four categories. The arrows schematically represent the potential number of incoming and outgoing transitions for the different categories. 115
6.6. **Occupation size effect on Accessibility and Transferability.** Scatter plot depicts the relationship between the accessibility (a), and transferability (b) and the number of employees in a given occupation. A dashed line serves as a visual guide and represents the linear fit of the points. There is a significant and positive rank correlation of 0.83 between accessibility and number of employees and a significant and negative rank correlation of -0.41 observed between number of employees and transferability. 117

6.7. **Correlation of Accessibility and betweenness centrality**. (a) Heatmap of the occupational transition matrix, with rows colored according to the accessibility of each respective occupation. (b) Scatter plot depicting the relationship between the accessibility of an occupation and its betweenness centrality, a measure of bottlenecks, showing a significant positive rank correlation of 0.57. Only condenser occupations are shown. A dashed line serves as a visual guide and represents the linear fit of the points. 117
6.8. **Correlation of Transferability and closeness centrality**. (a) Heatmap of the occupational transition matrix, with columns colored according to the transferability of each respective occupation. (b) Scatter plot depicting the relationship between the transferability of an occupation and its closeness centrality, a measure of a nodes ability to spread information in a graph, revealing a significant positive rank correlation of 0.27. Only condenser occupations are shown. A dashed line serves as a visual guide and represents the linear fit of the points. 118
6.9. **Gender effect on Accessibility and Transferability.** Scatter plot depicts the relationship between Female-to-Male ratio (a) the degree of transferability and (b) the degree of accessibility for each occupation. Color coding corresponds to the introduced occupation taxonomy. A dashed line serves as a visual guide and represents the linear fit of the points. There is a significant and negative rank correlation of -0.29 observed between Female-to-Male ratio and transferability and a significant and positive rank correlation of 0.19 between Female-to-Male ratio and accessibility (legend of (a,b)). 119
6.10. **Reducing constraints through effective policy making** (a) Scatter plot showing the eigenvector value difference between the destination and source occupation $(\vec{v}_2)_{l^+} - (\vec{v}_2)_{l^-}$ against the eigenvector value of the target occupation $(\vec{v}_2)_k$ (see Eq.6.4). The color coding indicates the change in the second eigenvalue, $\Delta\lambda_2$. (b) Scatter plot showing the relation between log-difference of source and target occupation in transferability $\log(T_{l^+}) - \log(T_{l^-})$ and accessibility $\log(A_{l^+}) - \log(A_{l^-})$. The color coding indicates the change in the second eigenvalue, $\Delta\lambda_2$. . . 121

7.1. **Network properties of the Skill Network**: Distribution of network properties for the skill networks constructed using cosine similarity (green) and θ-similarity (blue), including (a) degree, (b) betweenness centrality, (c) closeness centrality, and (d) clustering coefficient. 126
7.2. **Correlation between skill similarity and differences in Accessibility and Transferability**: Scatter plots showing (a) the accessibility difference $|\log(A_i) - \log(A_j)|$ and (b) the transferability difference $|\log(T_i) - \log(T_j)|$ as functions of skill similarity (measured using cosine similarity and θ-similarity) between each pair of occupations i and j. Lines represent linear fits, with the corresponding correlation values displayed in the legend. 128
7.3. **Skill effect on Accessibility and Transferability.** Scatter plot depicts the relationship between (a) d^p the product of inter- and intra-community skill similarity and the degree of accessibility and (b) d^r the ratio of inter- and intra-community skill similarity and the degree of transferability for each occupation. A dashed line serves as a visual guide and represents the linear fit of the points. There is a significant and positive rank correlation of 0.27 observed between the product of inter- and intra-community of skill similarity and accessibility and a significant and positive rank correlation of 0.19 between the ratio of inter- and intra-community skill similarity and the degree of transferability for each occupation. 129
7.4. **Expert-Assessed Skill Network**: Visualization of the expert-assessed skill network, where each node represents an occupation in the ROME classification. Node size corresponds to its betweenness centrality, and node color indicates its 1-digit ROME cluster. Edge weights are determined by the relevance index α: $w_{ij} = \frac{1}{2}$ for $\alpha = 1$, $w_{ij} = \frac{1}{3}$ for $\alpha \in [2, 5]$, and $w_{ij} = 0$ otherwise. The top row displays the distributions of network metrics, including degree, betweenness (Eq. 5.37), closeness (Eq. 5.38) centralities, and the clustering coefficient (Eq. 5.39). 130

7.5. **Expert-Assessed Skill Network with task-based classification**: Expert-assessed skill network constructed as in Figure 7.4, with nodes color-coded according to the task-based classification. The top row displays the clustering with the highest modularity achieved by BRIM [212], Infomap [213], and Spectral Clustering [214] algorithms. The box plot illustrates the Adjusted Mutual Information (AMI) scores (Eq. 7.7) for all random seeds, comparing the clustering results with the task-based classification. 132
7.6. **Accessibility and Transferability with task-based classification**: (a) Accessibility-Transferability space, as defined in Section 6.3, with each point color-coded according to the task-based classification. (b) Box plot of the bottleneck ratio $\log_{10} \frac{A_i}{T_i}$ for occupation i within each task-based cluster. 133

8.1. **Distributions of Venture Capital data:** Distribution of startup multiples (grey) and VC multiples (colored) across funding stages. The distributions are shown for (a) Seed to Series A, (b) Series A to B, and (c) Series B to C. The inset displays the distribution of the number of startups invested per VC. All distributions are shown on a log-log scale. 142
8.2. **Comparison of VC multiples to random sampling:** Distribution of actual VC multiples (colored) compared to the random sampling of portfolios (black) across funding stages. The distributions are shown for (a) Seed to Series A, (b) Series A to B, and (c) Series B to C. The random sampling preserves the number of startups invested per VC but assigns startups randomly to portfolios. All distributions are normalized for comparability. 142
8.3. **Luck and skill impact factor plane:** Estimated parameters for luck (α) and skill impact factor (β) for each sector and funding round. Sectors are distinguished by color, while marker styles indicate funding rounds: circles for Seed to Series A, squares for Series A to B, and triangles for Series B to C. Arrows connect consecutive funding rounds within the same sector, pointing in the direction of higher rounds. 143
8.4. **Skill vs. multiple at the VC level:** The skill values of individual VCs (x-axis) are plotted against their portfolio multiples (y-axis). Different funding rounds are distinguished by color and marker style: blue circles for Seed to Series A, green squares for Series A to B, and red triangles for Series B to C. The color intensity, as indicated by the color bar, reflects the number of companies in each VC's portfolio. The dotted horizontal line represents a portfolio multiple of 1, and the vertical dashed lines indicate the mean benchmark positions across all vintages and funding rounds, providing visual guidance. 145
8.5. **Skill vs. multiple with randomly sampled portfolios:** The scatter plots show the relationship between skill values and portfolio multiples for both real and randomly sampled VC portfolios. Black points represent real VCs, while colored points correspond to randomly assembled portfolios, with each color indicating a different random seed. The solid lines depict the linear fits for each dataset. The inset displays the distribution of slopes from the linear fits of the random portfolios, with the solid black line indicating the slope of the real system for comparison. . . 146
8.6. **Forecast bias across horizons:** The ex-post forecast bias $b_{\alpha,t-\theta}$ is shown for different years (colored lines) across forecasting horizons θ, with the black line representing the average bias across all years. Positive values indicate overoptimistic forecasts, also found by Guedj and Bouchaud (2005) [264] for the period of 1987-2004. The inset illustrates average absolute bias across horizons. 147
8.7. **Distribution of forecast biases:** The distribution of individual analyst forecast biases $b^i_{\alpha,t,\theta}$ is shown for horizons $\theta \in \{0, 5, 11\}$ months before the earnings announcement date. Inset shows the distribution of the number of companies rated by analysts N_i. 148
8.8. **Random shuffling of analyst predictions:** The distribution of the percentage difference of predictions ($\Delta^i_{t-\theta}$) is shown for actual analyst performance (colored) compared to a random predictor baseline (black) for three forecasting horizons: (a) $\theta = 0$, (b) $\theta = 5$, and (c) $\theta = 11$ months before the earnings announcement. 149

8.9. **Luck vs. skill impact factor for public markets:** The estimated parameters of luck (α) and skill impact factor (β) are plotted for different forecasting horizons (θ), where each color represents a different horizon. The arrows indicate the progression across horizons, moving closer to the earnings announcement date. 151
8.10. **Skill vs. accuracy for analysts:** Scatter plots show the relationship between skill (x-axis) and accuracy (y-axis) for individual analysts, with each panel representing different forecasting horizons: (a) $\theta = 0$ (blue circles), (b) $\theta = 5$ (green squares), and (c) $\theta = 11$ (red triangles). Dashed vertical lines represent benchmarks corresponding to the 10th, 25th, and 50th percentiles of the accuracy distribution, linking skill values to fixed levels of accuracy for comparison. . . . 151
8.11. **Skill vs. multiple with randomly sampled portfolios:** The scatter plots show the relationship between skill values and prediction accuracy. Black points represent real analysts, while colored points correspond to randomly assembled portfolios. The solid lines depict the linear fits for each dataset. The inset displays the distribution of slopes from the linear fits of the random portfolios, with the solid black line indicating the slope of the real system for comparison. . . 152
8.12. **Luck vs. skill impact factor VCs vs. Analysts:** Density plot of estimated parameters of luck (α) and skill impact factor (β) are plotted for each funding round and sector for VCs in green and for each horizon for Analysts in blue. 153

A.1. **Economic dashboard for the Inactive Central Bank scenario without any policy and with a stronger COVID shock:** The dynamics for a COVID shock strength of $\kappa = 1.0$ (blue), $\kappa = 1.1$ (orange) and $\kappa = 1.2$ (green) in the Inactive Central Bank scenario without any policy. Once the COVID shock gets too strong, the economy collapses in the long run. 186
A.2. **Dynamics for the three shocks for the Inactive Central Bank scenario:** The dynamics for the three shocks, COVID only (blue), COVID and Supply Chain shock (orange) and all shocks (green) for the Inactive Central Bank scenario with Easy-Credit policy. The areas shaded in grey indicate the duration of the three shocks: the COVID shock lasting until the end of the dark grey area, the supply chain shock until the end of the grey area, and the price shock until the end of the light grey area. 186
A.3. **Dynamics for the three shocks for the Reactive Central Bank with Anchored Expectation scenario:** The dynamics for the three shocks, COVID only (blue), COVID and Supply Chain shock (orange) and all shocks (green) to the scenario with reactive Central Bank and Anchored Expectation of economic agents with an Easy-Credit policy. The areas shaded in grey indicate the duration of the three shocks: the COVID shock lasting until the end of the dark grey area, the supply chain shock until the end of the grey area, and the price shock until the end of the light grey area. 186
A.4. **Dynamics for the three shocks for the Reactive Central Bank with Floating Expectation scenario:** The dynamics for the three shocks, COVID only (blue), COVID and Supply Chain shock (orange) and all shocks (green) to the scenario with reactive Central Bank and Floating Expectation of economic agents with an Easy-Credit policy. The areas shaded in grey indicate the duration of the three shocks: the COVID shock lasting until the end of the dark grey area, the supply chain shock until the end of the grey area, and the price shock until the end of the light grey area. 187
A.5. **Dynamics for all scenarios and all shocks:** The full dashboard for Figure 3.6 All dynamics are with Easy-Credit policy in the Inactive Central Bank scenario (blue), Reactive Central Bank with Anchored Expectation scenario (orange) and Reactive Central Bank with Floating Expectation scenario (green). The areas shaded in grey indicate the duration of the three shocks: the COVID shock lasting until the end of the dark grey area, the supply chain shock until the end of the grey area, and the price shock until the end of the light grey area. 187

A.6. **Dynamics with de-Anchored Expectation** ($\tau^T = 0$): All dynamics are with Easy-Credit policy and all shocks, including the Inactive Central Bank scenario (blue), Reactive Central Bank with Anchored Expectation scenario (orange dashed) and Reactive Central Bank with no trust scenario $\tau^T = 0$ (red dash-dotted). The areas shaded in grey indicate the duration of the three shocks: the COVID shock lasting until the end of the dark grey area, the supply chain shock until the end of the grey area, and the price shock until the end of the light grey area. Note that without expectation anchoring, the inflation path is very similar with (red dashed-dotted) and without (blue) monetary policy. 187
A.7. **Economic dashboard for the stronger Reactive Central Bank with Anchored Expectation scenario and all shocks.** The dynamics for a Central Bank strength of $\phi_\pi = 1.0$ (blue) and $\phi_\pi = 2.0$ (orange). A stronger central bank is not able to decrease inflation due to an external price shock, but with stronger policies, the consumption of households reduces, which leads to an increase in unemployment. 188
A.8. **Economic dashboard for the stronger Reactive Central Bank with Floating Expectation scenario and all shocks.** The dynamics for a Central Bank strength of $\phi_\pi = 1.0$ (blue) and $\phi_\pi = 2.0$ (orange) with Floating Expectation of economic agents. A stronger Central bank can reduce excess inflation at the cost of increased unemployment. The central bank must maintain a balance between falling inflation and rising unemployment. 188
A.9. **Sensitivity of α_c for the Reactive Central Bank with Anchored Expectation scenario and all shocks.** As already described in [115], larger α_c lead to a greater magnification of price trends (cet. par.). This amplification, in turn, causes the fluctuations in unemployment and inflation. Furthermore, larger α_c increases consumption which reduces unemployment due to higher demand. 188
A.10 **Sensitivity of α_Γ for the Reactive Central Bank with Anchored Expectation scenario and all shocks.** Larger values of α_Γ increases the influence of financial fragility on the hiring/firing policy of firms (cet. par.). This in turn leads to a larger downward adjustment of the workforce which increases unemployment. 189
A.11 **Helicopter Money in the Inactive Central Bank scenario and all shocks.** Unemployment (left) and inflation (right) for a helicopter drop of size $\kappa_H S$ one month after the price shock. Already with $\kappa_H \geq 0.2$ unemployment is reduced almost to zero. A further increase of Helicopter Money only increases inflation without reducing unemployment more. 189
A.12 **Helicopter Money in the Reactive Central Bank with Anchored Expectation scenario.** Unemployment (left) and inflation (right) for a helicopter drop of size $\kappa_H S$ one month after the price shock. With $\kappa_H \geq 0.3$ unemployment is reduced almost to zero. A further increase of Helicopter Money only increases inflation without reducing unemployment further. However, the Central Bank policy manages to keep inflation under control. 189
A.13 **Economic dashboard for Helicopter Money in the Inactive Central Bank scenario and all shocks.** Full dynamics for scenario without Helicopter Money (blue), for Helicopter Money with $\kappa_H = 0.2$ (orange) and $\kappa_H = 0.6$ (green) one month after the price shock and with Easy-Credit policy. 189
A.14 **Economic dashboard for Helicopter Money in the Reactive Central Bank with Anchored Expectation scenario and all shocks.** Full dynamics for scenario without Helicopter Money (blue), for Helicopter Money with $\kappa_H = 0.3$ (orange) and $\kappa_H = 0.6$ (green) one month after the price shock and with Easy-Credit policy. 190
A.15 **Economic dashboard for Helicopter Money in the Reactive Central Bank with Floating Expectation and all shocks.** Full dynamics for scenario without Helicopter Money (blue), for Helicopter Money with $\kappa_H = 0.2$ (orange) and $\kappa_H = 0.6$ (green) one month after the price shock and with Easy-Credit policy. 190

A.16 **Windfall Tax in the Inactive Central Bank scenario.** Unemployment (left) and inflation (right) for Windfall Tax of $\delta_e + \Delta\delta_e$ one year before the end of the price shock with a duration of two years. With $\Delta\delta_e \approx 4\%$ unemployment is reduced strongly. A further increase of tax does only increase unemployment again. For larger dividends, inflation increases because increased demand due to increased savings. . . . 190
A.17 **Windfall Tax in the Reactive CB with Anchored Expectation scenario.** Unemployment (left) and inflation (right) for Windfall Tax of $\Delta\delta_e$ one year before the end of the price shock with a duration of two years. With $\Delta\delta_e \approx 6\%$ unemployment is reduced strongly. A further increase of $\Delta\delta_e$ leads to a significant resurgence of unemployment. . . . 191
A.18 **Economic dashboard for Windfall Tax in the Inactive Central Bank scenario, all shocks.** Full dynamics for the Inactive Central Bank scenario without Windfall Tax (blue), for tax of $\Delta\delta_e = 6\%$ (orange) and $\Delta\delta_e = 12\%$ (green) one year before the end of the price shock with a duration of two years. . . . 191
A.19 **Economic dashboard for Windfall Tax in the Reactive Central Bank with Anchored Expectation scenario and all shocks.** Full dynamics for scenario without Windfall Tax (blue), for tax of $\Delta\delta_e = 6\%$ (orange) and $\Delta\delta_e = 12\%$ (green) one year before the end of the price shock with a duration of two years. . . . 191
A.20 **Economic dashboard for Windfall Tax in the Reactive Central Bank with Floating Expectation scenario and all shocks.** Full dynamics for scenario without Windfall Tax (blue), for tax of $\Delta\delta_e = 6\%$ (orange) and $\Delta\delta_e = 12\%$ (green) one year before the end of the price shock with a duration of two years. . . . 192
A.21 Eigenvalue Spectrum for three scenarios and all shocks. Sloppiness parameter $\epsilon = 0.01$, $n_{seeds} = 50$ target variable inflation (red) unemployment (blue), (a,b) Inactive Central Bank, (c,d) Reactive Central Bank with Anchored Expectation, (e,f) Reactive Central Bank with Floating Expectation. . . . 193
A.22 **Eigenvectors of the Hessian matrix for the Inactive Central Bank Scenario and all shocks.** The length of the eigenvectors is normed to 1 and the colors indicate the target variables, unemployment (blue) and inflation (red). . . . 193
A.23 **Eigenvectors of the Hessian matrix for the Reactive Central Bank with Anchored Expectation scenario and all shocks.** The length of the eigenvectors is normed to 1 and the colors indicate the target variables, unemployment (blue) and inflation (red). . . . 193
A.24 **Eigenvectors of the Hessian matrix for the Reactive Central Bank with Floating Expectation scenario and all shocks.** The length of the eigenvectors is normed to 1 and the colors indicate the target variables, unemployment (blue) and inflation (red). . . . 193
A.25 **Sensitivity of $g_p = g_w = const$ for the Reactive Central Bank with Anchored Expectation scenario and all shocks.** Increasing $g_p = g_w$ increases slightly inflation (cet. par.), causing slightly higher interest rates which leads to a reduction of consumption budget that reduces demand and therefore increases unemployment. . . . 194
A.26 **Sensitivity of $\frac{g_p+g_w}{2} = const$ for the reactive Central Bank with Anchored Expectation scenario and all shocks.** (blue) When $g_w > g_w$, there is a higher bargaining power of worker which increases wages (cet. par.), therefore there is a higher consumption budget and higher demand which decreases unemployment. (yellow) For larger market power ($g_p > g_w$), there is the reverse effect of decreased wages causing a decrease in consumption budget and demand which therefore increases unemployment. . . . 194

B.1. **Regression Parameters of Linear Regression in eq. B.6.** (a) Box plot of the coefficient $\beta_{i,t}$ for occupation i and year t. (b) Box plot of the intersection coefficient $\alpha_{i,t}$ for occupation i and year t. 197

B.2. **Entry and Exit Dynamics in the Labor Market**: Distribution of inflows (green) and outflows (red) for each occupation (on the x-axis) over the sample period for (a) all transitions, (b) workers below 30 years old, and (c) workers above 60 years old. The y-axis represents the proportion of transitions relative to the total number of workers in each occupation. Throughout the sample period, there is a slightly higher influx of workers into the labor market, amounting to 2% more inflows. Among the age group below 30, the influx is 25% higher, while for the age group below 60, outflows exceed inflows by a factor of two. 197
B.3. **Correlation of Accessibility and Transferability Metrics with Inter- vs. Intra-Community Transitions**. (a) Scatter plot depicting the relationship between the accessibility of an occupation and weighted inter- vs. intra-community transitions, showing a significant positive rank correlation of 0.27. (b) Scatter plot depicting the relationship between the transferability of an occupation and weighted inter- vs. intra-community transitions, revealing a significant positive rank correlation of 0.13. The color of the points corresponds to the respective accessibility and transferability taxonomy described in Section 6.3. A dashed line serves as a visual guide and represents the linear fit of the points. 227
B.4. **Age Effect on Accessibility and Transferability.** Scatter plot depicts the relationship between age median and (a) the degree of transferability and (b) the degree of accessibility for each occupation. Color coding corresponds to the introduced occupation taxonomy. A dashed line serves as a visual guide and represents the linear fit of the points. There is no significant rank correlation observed between age and transferability nor between age and accessibility (legend of (a,b)). 227
B.5. **Confusion Matrix for Gender and Age.** Panel (a) shows the confusion matrix for taxonomy clustering in the male and female networks, excluding diagonal elements (occupations assigned to the same cluster in both networks). A significant number of occupations in the male Condenser cluster are spread across Hubs, Diffusers, and Channels in the female network, indicating a broader dispersion in accessibility and transferability for females. Panel (b) displays the confusion matrix for clustering in age groups below 30 and above 60, excluding diagonal elements. Many occupations in the Condenser cluster for the 60+ age group are distributed across Hubs, Diffusers, and Channels for the under-30 group, reflecting a wider dispersion in accessibility and transferability for younger individuals. 228
B.6. **Matrix Perturbation of the French Labor Market.** (a) Scatter plot showing eq. B.20, demonstrating a strong positive correlation that supports the validity of the assumptions made for first-order matrix perturbation. The color coding of points corresponds to the sign cases of $\Delta\lambda$ (see eq. B.21). (b) Scatter plot showing the relationship between the eigenvalue difference and the value of the eigenvector for the target occupation, $(\vec{v}_2)_k$. (c) Scatter plot showing the eigenvalue difference against the difference in eigenvector values between the destination and source occupations, $((\vec{v}_2)_{l^+} - (\vec{v}_2)_{l^-})$. 230
B.7. **Correlation of Accessibility and Transferability Metrics with Changes in Eigenvalue** (a) Scatter plot showing the log difference in accessibility metrics between destination and source occupations against the change in the second eigenvalue. (b) Scatter plot showing the log difference in transferability metrics between destination and source occupations against the change in the second eigenvalue. (c) Scatter plot showing the log difference in accessibility metrics between destination and source occupations against the difference in the corresponding values of the second eigenvector. (d) Scatter plot showing the log difference in transferability metrics between destination and source occupations against the difference in the corresponding values of the second eigenvector. 231

Introduction

1. Introduction

In September 1987, the Santa Fe Institute hosted a landmark workshop titled *The Economy as an Evolving Complex System* [5] bringing together natural scientists and economists to bridge their fields. Pedagogical lectures introduced economists to tools for analyzing complex dynamical systems and natural scientists to economic analysis. This interdisciplinary exchange revealed innovative pathways to tackle complex challenges like global capital flows, economic cycles, and systemic instabilities. Discussions explored economic systems as complex adaptive systems, focusing on themes like innovation, coevolution, feedback loops, and the dynamics of networks such as trade flows and business cycles. Foundational concepts, including lock-in, path dependence, and phase transitions, were identified as key to understanding economic behavior.

The workshop not only shaped research agendas for decades but also demonstrated the profound value of interdisciplinary collaboration, overcoming barriers between disciplines and fostering a shared vocabulary. In the years since, the intersection of natural scientists and economics has yielded transformative insights. Statistical physics has illuminated emergent behaviors and collective dynamics in economic systems, offering new perspectives on wealth inequality, firm size distributions, and market volatility. Ideas like phase transitions and critical points have enriched our understanding of stability and fragility in financial systems. While economics contributes centuries of knowledge in studying economic systems, institutional depth and behavioral aspects of agents, physics can add innovative modeling techniques and approximation methods to make models more tractable. Together, these fields have forged a fertile ground for progress, enhancing theoretical frameworks and advancing practical solutions to real-world challenges.

Building on these foundations, this thesis delves deeper into the application of complex systems theory to economics, revisiting and expanding upon ideas first broached at Santa Fe and beyond. The first part focuses on a macroeconomic agent-based models to address the question: *How do fiscal and monetary policies influence inflationary dynamics in a complex macroeconomic environment, particularly in the context of post-COVID economic shocks?* By adopting an approach inspired by statistical physics, I explore how macroeconomic agent-based models can serve as meaningful scenario generators, helping to uncover root causes of inflation and guiding the development of targeted, effective policies.

In the second part, I turn to a detailed study of labor markets, presenting a novel perspective on how to model these systems. Specifically, I answer the question: *How can structural bottlenecks in the French labor market be quantified and identified?* Using a complex systems approach, I develop meaningful metrics to identify bottleneck occupations, offering insights into improving labor mobility and economic efficiency.

In the third part, I focus on the dynamics of skill and luck in public and private markets, presenting a new framework to disentangle these factors. Specifically, I address the question: *What is the interplay between skill and luck in driving the success in public and private markets?* Using the skill-luck-depth model, I analyze venture capital investments and public market analyst forecasts to assess the interplay of skill, randomness, and competition. This approach not only offers a nuanced understanding of performance drivers but also provides actionable metrics for benchmarking and improving decision-making in these highly competitive environments.

The outcomes of this research underscore the essential role of interdisciplinary approaches in addressing complex economic challenges. By integrating tools from fields like physics and computational modeling, this thesis demonstrates how economic systems can be better understood as dynamic, interconnected systems. Moving beyond traditional economic models, techniques such as network analysis and agent-based simulations enable policymakers and researchers to uncover actionable insights and craft solutions that are both innovative and effective for real-world problems.

1.1. Macroeconomic Modelling

Neoclassical Models The foundations of modern macroeconomic modeling lie in the neoclassical tradition, which emerged in the late 19th and early 20th centuries. This school of thought brought mathematical rigor to economics, influenced by Newtonian mechanics [6]. Pioneers such as Léon Walras and Alfred Marshall developed and applied the concept of equilibrium to markets, positing that supply and demand interact dynamically through prices to clear markets [7, 8]. Neoclassical economics highlights microfoundations, focusing on rational agents optimizing utility or profit, and explore how markets allocate resources efficiently.

A key application of this framework was to explain employment, which neoclassical economists viewed through the lens of labor market equilibrium. They argued that wages adjust to balance the supply and demand for labor, implying that unemployment, if it occurred, was temporary. Falling wages were expected to increase labor demand, eventually restoring full employment [9]. While this elegant model captured the mechanics of market adjustment, it overlooked critical complexities, such as wage rigidities and systemic shocks.

The limitations of the neoclassical approach became particularly evident during the Great Depression of the 1930s, when economies experienced sustained unemployment despite falling wages [10]. Neoclassical theory failed to account for such prolonged disequilibrium, exposing a fundamental gap in its framework.

Keynesian Economics This gap set the stage for the revolutionary insights of Keynesian economics. In his seminal work, *The General Theory of Employment*, Keynes (1937) [11] challenged the neoclassical assumption of equilibrium, introducing a paradigm shift in macroeco-

nomic thinking. Keynes argued that aggregate demand—comprising consumption, investment, and government spending—played a central role in determining output and employment. He emphasized that insufficient demand could result in persistent high unemployment, a scenario neoclassical models were unable to capture [12].

Keynesian economics introduced concepts such as the consumption function, the multiplier effect, and liquidity preference, providing a framework to understand the economy as a dynamic, interdependent system where short-term fluctuations could have profound and lasting effects [9]. This perspective addressed the shortcomings of neoclassical models and laid the groundwork for modern macroeconomic policy, advocating for active government intervention to stabilize business cycles and mitigate economic downturns.

By the mid-20th century, Keynesian economics had become the dominant paradigm for macroeconomic policy-making, particularly in the post-war era. The development of large-scale macroeconometric models further transformed Keynesian theory, translating its qualitative insights into empirical tools. These models used historical data to estimate relationships between macroeconomic variables such as consumption, investment, and employment, incorporating hundreds of equations to capture the interconnectedness of economic sectors. This provided a detailed empirical framework for understanding macroeconomic dynamics and deriving policy recommendations.

However, this reliance on historical relationships revealed a critical weakness. The limitation became starkly evident during the stagflation of the 1970s—a period characterized by high inflation and unemployment—which the models were unable to predict or explain [13].

Lucas Critique and DSGE models The fragility of traditional macroeconometric models' assumptions led to significant criticism, most notably in the form of the Lucas critique. Lucas (1976) [14] argued that these models failed because they ignored how agents' behavior adapts to policy changes. Lucas identified a fundamental flaw: macroeconometric models treated relationships between variables as fixed and stable over time. However, these relationships are contingent on the expectations and behavior of economic agents, which evolve in response to changes in policy or economic conditions. For instance, consistent use of expansionary monetary policy to reduce unemployment might lead individuals to anticipate higher inflation, prompting behavioral adjustments that ultimately negate the policy's intended effects. This insight rendered traditional models unreliable for analyzing the impact of new policies and demanded macroeconomic modeling based on *microfoundations*.

The response to Lucas's critique culminated in the development of Dynamic Stochastic General Equilibrium (DSGE) models [15]. DSGE models sought to address these limitations by grounding macroeconomic relationships in microfoundations, explicitly modeling the behavior of individual agents, such as households and firms, optimizing utility or profit over time. A critical innovation was the incorporation of rational expectations, ensuring that agents' expectations about the

future aligned with the structure of the model itself rather than being extrapolated solely from past data [16].

DSGE models typically assume a representative agent, simplifying the behavior of all households or firms into that of a single, *typical* entity. While this abstraction facilitates analytical tractability, it has been subject to criticism for ignoring heterogeneity[17][1]. These models integrate elements of Real Business Cycle (RBC) theory with some aspects of New Keynesian economics, wherein economic agents maximize utility subject to intertemporal budget constraints. [2]

1: Although recent advancements in DSGE models have introduced heterogeneity [18]

2: We refer to the work of Smets and Wouters (2007) [19] and Galí (2015) [20] for an in-depth introduction to the details of DSGE models.

The theoretical rigor and policy orientation of DSGE models made them the dominant tool for central banks and policymakers [21]. By simulating the effects of monetary or fiscal policy under various scenarios, these models provided a structured framework for policy evaluation. Central banks, such as the Federal Reserve, European Central Bank, and Bank of England, adopted DSGE models to guide decisions on interest rates, inflation targeting, and economic stabilization.

Failure of DSGE However, the 2008 financial crisis exposed critical limitations in DSGE models, particularly their inability to account for large-scale disruptions. These models inherently assume that economic crises arise from exogenous shocks, which propagate through an economy fundamentally in equilibrium. This framework proved inadequate for capturing the endogenous vulnerabilities and systemic risks that characterized the financial crisis.

Several limitations became apparent, such as the reliance on the representative agent assumption, which overlooks the critical heterogeneity among households, firms, and financial institutions [17]. Moreover, the assumption of rational expectations, where agents are presumed to perfectly anticipate future conditions, severely constrained these models' ability to account for the uncertainty, feedback loops, and cascading failures that characterize real-world crises. For a comprehensive discussion of the shortcomings of DSGE models, see Dosi and Roventini (2019) [22], who argue that these models are fundamentally inadequate to predict or effectively analyze crises.

Recent advancements aim to address their shortcomings [23, 24]. The incorporation of bounded rationality, where agents rely on heuristics or adaptively form expectations, allows models to better capture the unpredictability and volatility observed during crises [25]. Another significant development is the rise of Heterogeneous Agent New Keynesian (HANK) models, which replace the representative agent with diverse households and firms. HANK models enable the analysis of distributional effects, liquidity constraints, and differential impacts of shocks, addressing key critiques while retaining the microfoundations and equilibrium structure that make DSGE models valuable for policy evaluation [26].

Complexity Economics Complexity economics offers a transformative framework for understanding the economy, conceptualizing it as a complex adaptive system composed of interacting, heterogeneous agents such as households, firms, banks, and governments [27, 28].

Rooted in the foundational ideas of the Santa Fe Institute, this approach was formally articulated in the workshop *The Economy as an Evolving Complex System* where six defining features were identified: (i) dispersed interaction among agents, (ii) the absence of a global controller, (iii) cross-cutting hierarchical organization, (iv) continual adaptation, and (v) perpetual novelty and (vi) out-of-equilibrium dynamics [5].

In this framework, agents are not assumed to be perfectly rational or homogeneous. Instead, they rely on adaptive and heuristic decision-making rules that evolve over time in response to their environment and interactions. This stands in stark contrast to traditional models such as Dynamic Stochastic General Equilibrium (DSGE) frameworks, which simplify the economy by aggregating behavior into a representative agent and assuming a tendency toward equilibrium. Complexity economics, by contrast, emphasizes non-equilibrium dynamics, where macroeconomic phenomena emerge organically from the localized interactions and adaptive behaviors of agents, reflecting the economy's constantly evolving nature.

A key strength of complexity economics is its natural alignment with the Lucas critique. By grounding macroeconomic behavior in microfoundations, it inherently addresses Lucas's critique to model agents whose decisions adapt dynamically to changes in policy and economic conditions. Complexity models do not rely on fixed, aggregate relationships; instead, they allow for agents' behaviors to evolve in response to both their environments and the actions of others. This flexibility enables complexity economics to capture the evolving nature of real-world economic systems, going beyond the limitations of DSGE models that often treat systemic shifts as exogenous shocks[29, 30].

Central to the complexity economics framework is Agent-Based Modeling (ABM), a powerful tool for simulating economies from the bottom up. ABMs explicitly model the actions and interactions of diverse agents within dynamic, non-equilibrium environments, enabling the study of emergent phenomena such as financial crises, inequality, or systemic risks. By incorporating heterogeneity and network interactions, ABMs reveal how aggregate outcomes arise from micro-level behaviors and interdependencies, offering insights that traditional equilibrium-based models cannot provide[22, 31].

Despite their potential, building non-equilibrium models presents significant challenges. The inherent heterogeneity of agents and their individualized decision-making processes make these models computationally intensive. Calibration often requires either substantial assumptions about parameters or vast amounts of data, and even when data is available, achieving accurate calibration can be difficult. These challenges underscore the need for innovative approaches to enhance ABM applicability and robustness.

This Thesis In this context, my research contributes to addressing some of the central challenges in complexity economics. Chapter 3 builds an agent-based model that incorporates adaptive expectations, a mechanism crucial for capturing real-world dynamics. This model was applied to study post-COVID macroeconomic dynamics, specifi-

cally inflationary pressures. The findings highlight that the effectiveness of monetary policy depends significantly on the anchoring of expectations. For instance, while strong inflation anchoring by central banks stabilizes prices, it can suppress consumption and employment, demonstrating the complex trade-offs in policy design. This contribution aligns with complexity economics' emphasis on modeling systems as adaptive and evolving while providing actionable insights for policymakers dealing with non-linear economic environments.

Chapter 4 addresses the persistent issue of parameter uncertainty in ABMs, a limitation often cited in the complexity economics literature. Drawing inspiration from systems biology, we developed a novel metric that systematically explores the high-dimensional parameter space of ABMs. This approach identifies critical parameter combinations that significantly influence macroeconomic outcomes, enabling a more efficient exploration of possible system dynamics, such as phase transitions in unemployment or inflation. By reducing the reliance on exhaustive calibration and improving the interpretability of model results, this methodology enhances the robustness and applicability of ABMs.

1.2. Modelling Labor Markets

General Equilibrium The study of labor markets has traditionally been grounded in general equilibrium models, which assume that wage adjustments would ensure market clearing. However, the persistent coexistence of high unemployment and job vacancies posed a challenge to this framework. The seminal Diamond-Mortensen-Pissarides (DMP) model addressed this gap by introducing a partial equilibrium search and matching framework, in which job seekers and firms incur costs and delays while searching for suitable matches[32].

Diamond (1982) [33] demonstrated how externalities in the matching process could lead to inefficiencies, showing that unemployment rates are not uniquely determined in markets with frictions. Building on this, Mortensen (1982) [34] incorporated wage bargaining under imperfect information, linking wage dynamics to job creation and destruction. Pissarides (1985) [35] added a dynamic perspective, explaining how unemployment, vacancies, and wages evolve following economic shocks. Their introduction of labor-market tightness—the ratio of vacancies to unemployment—was pivotal in explaining the Beveridge curve, a well-known macroeconomic relationship illustrating the inverse relationship between unemployment and vacancies. The DMP model had a big influence on labor economics by explaining the interplay of job creation, destruction, and matching. Even though Diamond initially admitted that *one cannot draw policy conclusions directly from such a model* [33], its successors models became instrumental in guiding labor policies, bridging theoretical models with real-world labor complexities.

Skill-biased Technological Change This framework, however, did not account for the transformative effects of technological advancements, particularly automation and digitization, which play pivotal

roles in shaping economic growth and business cycles[36]. As Simon Kuznets noted in his Nobel Prize lecture, *rapid changes in production structure are inevitable*[37], emphasizing the profound impact of technological progress. The concept of a bias in production factors between capital and labor, first introduced by Hicks (1932) [38], later evolved into the study of Skill-Biased Technological Change (SBTC)[39]. SBTC explains how technology complements non-routine, cognitive tasks, typically performed by high-skill workers, while substituting for routine tasks, often carried out by middle-skill workers[40]. This dynamic has driven job polarization, where high-skill, high-wage and low-skill, low-wage jobs have grown, while middle-skill occupations have steadily declined.

From a Schumpeterian perspective, this structural change reflects creative destruction[41], where technological progress enables some firms and workers to become more productive while forcing others—unable to adapt or acquire new skills—into obsolescence or unemployment. David (2013) [42] demonstrated how automation displaces routine-intensive roles, while Frey and Osborne (2017) [43] quantified the broader automation risk, estimating that nearly half of U.S. jobs are vulnerable to computerization. Acemoglu and Autor (2011) [40] further highlighted that it is not entire jobs but specific tasks that evolve dynamically with technology, reshaping employment patterns and wage structures. Krüger (2008) [36] aligns with this view by analyzing how technological progress historically reallocated labor from agriculture to industry and, more recently, to services, accelerating structural changes in labor markets.

These shifts have widened wage gaps and created significant challenges for displaced workers, underscoring the urgent need for policies focused on retraining, skill development, and labor market adaptability. SBTC not only illustrates the differential impacts of technology on labor markets but also highlights the growing importance of incorporating skill data into labor market models to better address these ongoing transformations.

Network Models Parallel to these developments, network approaches have redefined the way labor markets are modeled, offering granular insights into job mobility, skill transferability, and structural constraints. Unlike aggregate models, network frameworks analyze the relationships and interdependencies between agents—workers, firms, and industries—revealing patterns and bottlenecks that shape labor market dynamics. Labor flow networks, as shown by Guerrero and Axtell (2013) [44], highlight the outsized role of central firms as hubs of innovation and opportunity, while Neffke and Henning (2013) [45] demonstrated that most job transitions occur between skill-related industries, underscoring the barriers posed by skill mismatch. Schmutte (2014) [46] further refined this perspective by identifying segmented mobility networks, where workers transition freely within clusters of jobs but face significant barriers moving between clusters, exacerbating skill mismatches during downturns. Bocquet (2022) [47] further improves the modeling by combining network approach with a search and matching framework studying the adjustment of labor markets due to technological shocks. Rio-Chanona et al. (2021) [48]

builds a data driven network agent based model to study the effect of automation on the labor market. By modeling labor markets as networks, researchers gain a granular understanding of job mobility, bottlenecks, and resilience, offering insights for workforce development to enhance model driven policy making for more resilient labor markets.

Stochastic Markov Models Simultaneously to search theory and network approaches, Markov models have served as a complimentary method for studying labor markets through detailed labor flow data, offering a stochastic framework to model transitions between discrete labor market states, such as employment, unemployment, and inactivity. These models focus on the probabilities of moving between these states over time, capturing the dynamic nature of labor market flows in a mathematically tractable manner [49].

Traditional Markov models, already discussed in 1985[50], assumed first-order transitions, where the probability of moving between states depends only on the current state. This straightforward approach became instrumental in early studies of aggregate labor market dynamics, providing a baseline for understanding key phenomena such as unemployment duration or inter-generational social mobility [49]. Over time, Markov models were extended to account for multi-spell and multi-state transitions, incorporating factors like duration dependence and worker heterogeneity. For instance, longer durations of unemployment have been shown to significantly reduce the probability of transitioning to employment [51]. These extensions also revealed how individual characteristics, such as education, age, or gender, influence labor market outcomes[52].

The development of Hidden Markov Models (HMMs) marked a significant advancement in this field. By incorporating unobservable states, such as worker types or hidden market segments, HMMs have provided better empirical fit to labor market dynamics, particularly in capturing long-term consequences and path-dependencies of unemployment[53]. These models have been used to uncover latent structures, such as the division between primary and secondary labor markets that face disproportionately impact of shocks for certain worker groups, further enriching the analysis of labor market transitions.

Markov models have also provided insights into the effects of labor market policies. By simulating how changes in unemployment insurance, retraining programs, or hiring subsidies influence transition probabilities, these models have informed the design of interventions aimed at reducing long-term unemployment and improving job stability[54]. Recent advances in the integration of Markov models with richer data sources, such as administrative records and longitudinal surveys, as well as the use of machine learning techniques, have further enhanced their predictive power and applicability[55].

Despite their strengths, Markov models have certain limitations. Their reliance on aggregate probabilities can obscure important interaction effects and feedback mechanisms within labor markets. Additionally, the ambiguity in identifying hidden Markov states can lead to varying interpretations of results, which may result in conflicting pol-

icy recommendations. Nevertheless, Markov models provide a valuable complement to search theory and network approaches, offering a probabilistic framework for understanding labor state transitions.

Econometric Models Finally, econometric models have long been used for analyzing labor market dynamics, providing statistically rigorous methods to estimate the effects of time-varying variables in a data-driven fashion, such as unemployment insurance or training programs, on the likelihood of state transitions[56]. These models estimate how factors like education, unemployment insurance, and training programs influence labor market transitions. Advances in data availability, particularly linked microdata from labor force surveys, have enabled econometric approaches to reveal critical insights, such as how employment exit rates correlate with economic development and how information frictions create barriers in developing labor markets[57].

This diversity of approaches—from search and matching theory to network frameworks, stochastic models, and econometrics—underscores the complexity of labor markets. Together, they provide a rich foundation for understanding labor market dynamics and developing policies to address ongoing structural challenges.

This Thesis In this context, my research enhances the understanding of labor markets by applying network science and statistical methods to tackle challenges related to mobility, structural imbalances, and skill mismatches, building on the extensive foundation of existing labor market research.

Chapter 5 examines the hierarchical structure of the French labor market through occupational transition networks, uncovering power-law distributions in occupation sizes and flows. These findings align with network-based studies by Neffke and Henning (2013) [45], Bocquet (2022) [47], which highlight the role of hubs and bottlenecks in workforce mobility. Additionally, the analysis reveals that the transition matrix adheres to detailed balance, enabling a comparison of spectral and dynamical properties over time.

Chapter 6 introduces accessibility and transferability metrics to evaluate labor mobility, identifying bottleneck occupations as those with high accessibility but low transferability. These metrics offer practical tools for addressing structural imbalances and inform the design of targeted interventions to mitigate mobility barriers within the labor market.

Chapter 7 integrates skill similarity with mobility analysis, demonstrating how skills influence occupational flows and emphasizing the importance of combining skill and transition data. The findings show that high-accessibility occupations function as entry points, while high-transferability occupations bridge occupational communities, driven by their skill compositions.

Part I: Exploring Macroeconomic Agent-Based Models & Inflation

Agent-Based Models and Parameter Exploration

2.

Section 2.2 is adapted from Naumann-Woleske et al. (2024) [2] *Exploration of the Parameter Space in Macroeconomic Agent-Based Models*, with minor modifications to the text to align with the context of this thesis. All other sections are entirely new. The original publication was a collaborative effort with *K. Naumann-Woleske*, *M. Benzaquen* and *J.-P. Bouchaud* and was later extended by *K. Naumann-Woleske* in his thesis [58].

Agent-Based Models are transformative tools for studying economic systems, capturing the dynamics of heterogeneous agents and the emergent macroeconomic phenomena arising from their interactions. Section 2.1 traces the origins of ABMs to foundational ideas in complexity science, including von Neumann's cellular automata and Conway's Game of Life, which illustrated how simple, localized rules could generate complex, system-level behaviors. These early concepts laid the groundwork for more realistic macroeconomic ABMs, which now tackle diverse economic challenges.

While ABMs have demonstrated significant explanatory power, challenges persist, including high data requirements, designing realistic heuristic rules, and navigating the vast parameter spaces inherent in these models. To address these challenges, Section 2.2 introduces the concept of a novel framework for parameter exploration. This approach identifies the most influential parameters driving model dynamics, enabling systematic analysis and enhancing the interpretability and applicability of ABMs in understanding complex economic systems.

2.1. Introduction to Agent-Based Models

Cellular Automata and the Game of Life

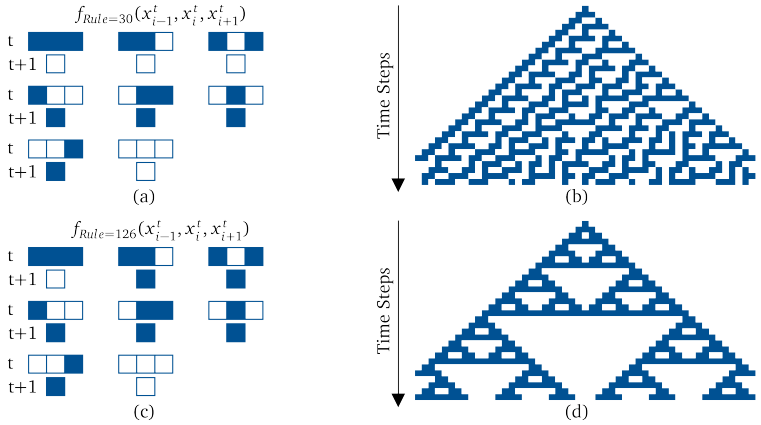


Figure 2.1.: 1-dimensional Cellular Automata model: (a,c) Schematic overview of Rule 30 (126) for determining the next state in a 1D cellular automaton, defined as $f_{30}(x_{i-1}^t, x_i^t, x_{i+1}^t) = x_i^{t+1}$ (Eq. 2.1), with a neighborhood size of 2. There are $2^3 = 8$ possible neighborhood states, and the rule f_{30} specifies the next state for each of these configurations. Empty cells represent dead cells ($x_i = 0$), while blue cells indicate living cells ($x_i = 1$). (b,d) Simulation of the cellular automaton governed by Rule 30 (126), starting with an initial condition of a single living cell positioned at the center of the line at $t = 0$ and evaluated for 30 time steps.

The origins of modelling complex systems through simple rules can be traced back to the work of John von Neumann, who, in the 1940s, introduced the concept of cellular automata [59]. Von Neumann envisioned a mathematical framework where simple, discrete units (or *cells*) interact according to predefined rules, leading to complex behaviors over time. Although von Neumann's work was primarily theoretical, it laid the groundwork for studying how local interactions among individual components could produce emergent patterns.

Building on the foundation of complexity science, John Conway introduced the Game of Life in 1970, a cellular automaton that became an iconic demonstration of how complexity arises from simplicity [60]. The Game of Life operates on a grid of cells, each of which can be in one of two states: *alive* or *dead*. Over discrete time steps, the state of each cell evolves based on a simple set of rules determined by the states of its neighboring cells. For example, a cell *lives* or *dies* depending on how many of its neighbors are alive.

Formally, the binary state of a cell, $x_i^t \in \{0, 1\}$, at time t can be described by a function f, which depends on the states of its neighboring cells j :

$$f(x_{i-j}^t, \ldots, x_{i-1}^t, x_i^t, x_{i+1}^t, \ldots, x_{i+j}^t) = x_i^{t+1}. \tag{2.1}$$

This function determines the state of cell i at the next time step, $t + 1$. In its simplest form, cells are arranged on a 1D line, where each cell interacts with its two immediate neighbors. For such a configuration, a rule must specify the outcome of $2^3 = 8$ possible neighborhood states, resulting in $2^{2^3} = 256$ possible rule sets or interaction configurations (see the $2^3 = 8$ possible configurations in Figure 2.1(a)).[1] Despite the apparent simplicity of these rules—such as *if my neighbor cells are alive, I will continue living too*—the system produces astonishingly complex behaviors, including stable structures, oscillations, and even patterns capable of performing computational logic. Figure 2.1(a-b) illustrates the evolution of the cellular automaton under Rule 30, showcasing its chaotic behavior,[2] while Figure 2.1(c-b) shows the evolution under Rule 126, which reveals identifiable patterns.

1: The rules for the cellular automaton are determined by converting the rule number into its binary representation. This binary sequence specifies the outcomes for all possible neighborhood configurations, which are arranged in a specific order, from bottom to top, as shown in Figure 2.1(a),(c). For instance, rule 30 in decimal (30_{10}) is represented as 00011110_2 in binary, while rule 126 (126_{10}) is represented as 01111110_2. Each binary digit corresponds to the outcome of a specific neighborhood configuration. For example, in rule 30, the rightmost digit of the binary sequence is 0. This means that for the first configuration (000, where the cell and both its neighbors are dead), the result for the middle cell in the next state is also 0 (dead). Similarly, the second-rightmost digit is 1, indicating that for the second configuration (001, where only the right neighbor is alive), the result for the middle cell in the next state is 1 (alive). This pattern continues for all configurations.

2: According to Knudsen (1994) [61], Rule 30 satisfies the formal definitions of chaos.

The Game of Life illustrates a core principle of complexity science and agent-based modelling (ABM): complex, unpredictable phenomena can emerge from deterministic, localized interactions. The richness of these emergent behaviors underscores how simple rules can generate a vast array of outcomes, from static equilibrium-like patterns to dynamic, chaotic configurations. Fascinatingly, certain rules—such as Rule 30—are observed in nature, appearing on the shell patterns of the conus textile snail, showcasing the profound connection between mathematical models and the natural world [62].

Schelling's Model of Segregation

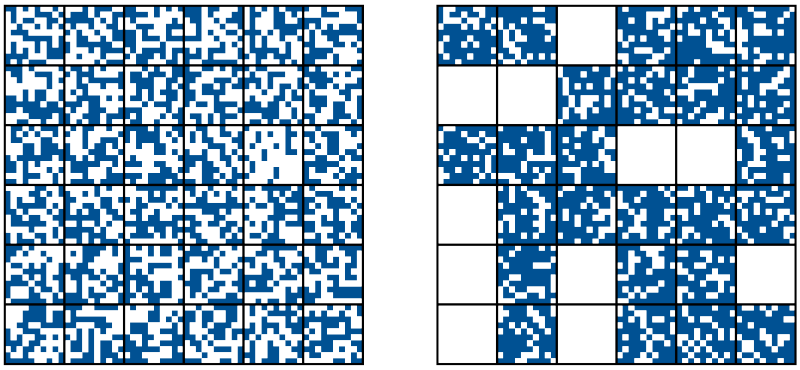


Figure 2.2.: Neighborhood configuration in Schelling model: Configurations of a city in the Schelling model as extended by Grauwin et al. (2009) [63], consisting of $Q = 36$ blocks, each containing $H = 100$ cells, with an initial density of $\rho_0 = 0.5$. (a) Stationary state of the model with parameters $m = 0.5$, $\alpha = 1$, and $T \to 0$. (b) Stationary state of the model with parameters $m = 0.5$, $\alpha = 0$, and $T \to 0$.

The perspective of simple models was extended into social science and economics by models like Thomas Schelling's segregation model, a pioneering agent-based model that illustrated how micro-level preferences can lead to unintended macro-level outcomes [64]. Schelling's model demonstrated that even agents with a slight preference for diversity might, through local interactions, collectively produce segregated neighborhoods. Building on this foundation, the version of the model introduced by Grauwin et al. (2009) [63] extends these ideas to explore the role of cooperation and altruism in shaping urban dynamics.

In this extended version, a city is represented as a grid of Q blocks, each containing H cells or flats. Each cell can house a single agent, and the density of agents in block q is defined as $\rho_q = n_q/H$, where n_q is the number of agents in the block. Agents derive satisfaction—or utility—from their block's density, captured by a utility function $u(\rho_q)$. The collective utility of the city is then given by:

$$U(x) = H \sum_q \rho_q u(\rho_q), \tag{2.2}$$

where x represents the city's overall configuration.[3]

Agents move between blocks based on a probabilistic decision rule that balances their personal utility and the collective welfare. At each step, an agent and a vacant cell are selected, and the agent moves with probability:

$$P_{xy} = \frac{1}{1 + e^{-G/T}}, \tag{2.4}$$

3: The individual utility is defined as

$$u(\rho) = \begin{cases} 2\rho, & \text{if } \rho \leq \frac{1}{2} \\ m + 2(1-m)(1-\rho), & \text{if } \rho > \frac{1}{2} \end{cases} \tag{2.3}$$

with the asymmetry parameter m.

where T introduces randomness in decision-making (reflecting social or contextual factors), and G is the gain from moving, calculated as:

$$G = \Delta u + \alpha(\Delta U - \Delta u). \tag{2.5}$$

Here, Δu is the change in the agent's own utility, ΔU is the change in collective utility, and $\alpha \in [0, 1]$ is the altruism parameter, which determines the weight an agent places on collective welfare. When $\alpha = 0$, agents act purely selfishly, mirroring Schelling's original model. When $\alpha = 1$, agents prioritize the collective welfare.

Figure 2.2 shows the neighborhood configurations of the Schelling model under two different parameter settings. Panel (a) represents the collective case with $\alpha = 1$, where in the stationary state agents are distributed homogeneously. In contrast, panel (b) shows the selfish case with $\alpha = 0$, where agents act individually, leading to a segregated configuration with some empty blocks and other blocks with high density. Interestingly, even though agents in the selfish scenario aim to maximize their utility, the resulting segregation leaves them with a lower overall utility compared to the collective, homogeneous configuration.

Moreover, this typ of Schelling models, retain a degree of analytical tractability (see [63, 65]), a quality traditionally associated with DSGE models. This demonstrates that agent-based models, while often implemented as computational simulations, can in limited cases be understood analytically, bridging the gap between simulation-based and theoretical approaches.

The Schelling model is particularly insightful for economics as it illustrates how heterogeneity and local interactions among agents can generate emergent phenomena. It reveals that systemic outcomes—such as housing segregation, or the clustering of economic activity—can arise unintentionally, diverging from the anticipated goals of individual agents. These outcomes are driven by dynamic feedback loops between agents and their environment, producing emergent behaviors that are central to complexity economics and agent-based modelling.

Talent vs. Luck Model

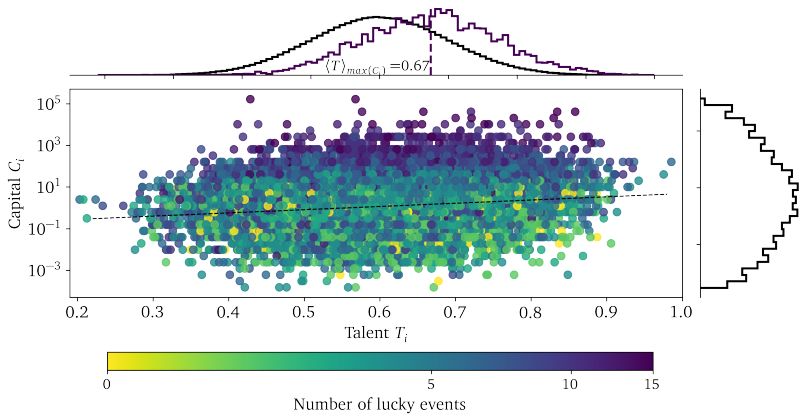


Figure 2.3.: Talent vs. Capital distribution: The relationship between talent and final capital in the Talent vs. Luck model. Each point represents an individual's talent and accumulated capital after 40 time steps (years). The system is initialized with talent distributed normally ($\mu_T = 0.6$, $\sigma_T = 0.1$) across $N = 500$ individuals, each starting with an equal initial capital of $C(0) = 10$. The parameters are choosen as in Pluchino et al. (2018) [66]. The simulation is run 5,000 times, and all individuals from all runs are plotted. The color intensity of the points indicates the number of lucky events encountered by each individual. The purple distribution in the top panel shows the talent for the wealthiest individual after each run, with a mean talent of $\langle T \rangle_{max(C_i)} = 0.67$.

The interplay of randomness and individual traits has long been a topic of interest in understanding socio-economic outcomes. The Talent vs. Luck model, introduced by Pluchino et al. (2018) [66], exemplifies how agent-based modelling can challenge conventional views of meritocracy by quantifying the roles of skill and chance in determining success. This model extends the analytical reach of agent-based approaches into questions of inequality and fairness, illustrating how seemingly meritocratic systems can yield outcomes heavily influenced by randomness.

The model simulates a population of N agents ($N = 500$), each endowed with an inherent level of talent T ($\mathcal{N}(\mu = 0.6, \sigma = 0.1)$) and an initial amount of capital C ($C(0) = 10\$$). Agents interact with random lucky or unlucky events over a simulated career spanning 40 years. Lucky events, which double an agent's capital, occur with a probability proportional to their talent (if rand$[0, 1] < T_i$, then $C_i(t) = 2C_i(t-1)$). Unlucky events halve their capital regardless of talent ($C_i(t) = \frac{C_i(t-1)}{2}$), simulating the unpredictable adversities of real life. Through repeated interactions, agents accumulate capital as a proxy for success, shaped by both their intrinsic talent and the stochastic effects of external events.

After simulating the model, success follows a highly unequal power-law distribution, despite talent being normally distributed across the population (see Figure 2.3). The agents with the greatest success are not necessarily the most talented but moderately talented individuals who benefited from a disproportionate share of lucky events. In 5000 random simulations with 500 individuals and a Gaussian distributed talent with $\mathcal{N}(\mu = 0.6, \sigma = 0.1)$, the mean talent of the wealthiest individual after each run is $\langle T \rangle_{max(C_i)} = 0.67$ (see purple distribution in top panel of Figure 2.3). Conversely, highly talented agents who encounter few lucky events may remain relatively unsuccessful, highlighting a significant misalignment between talent and success (see color of scatter plot in Figure 2.3). This challenges the meritocratic assumption that talent alone determines outcomes.

The Talent vs. Luck model holds relevance for economic modelling by demonstrating how agent-based methods can bridge micro-level individual traits with macro-level systemic outcomes. Much like Schelling's segregation model, it reveals that emergent patterns—such as inequality or market dominance—can arise unintentionally from dynamic feedback loops between agents and their environment. These findings highlight the limitations of deterministic models and underscore the importance of incorporating stochastic and non-linear dynamics into socio-economic analysis.

Building on a similar conceptual view, our work on skill vs. luck in Chapter 8 extends the ideas of this framework to the domains of venture capital and public markets, where skill and chance similarly interplay to shape *success* outcomes. Instead of talent, we focus on the measurable skills of VCs and analysts—the ability of VCs to identify and invest in high-growth startups, and the ability of analysts in forecasting financial metrics. Lucky events in this context correspond to external opportunities, such as a VC coming across a promising startup or an analyst benefiting from favorable market conditions.

Macroeconomic Agent-Based Models

Agent-based macroeconomic models have evolved from simplified models like the Schelling or the TvL model, into sophisticated tools capable of capturing the intricate dynamics of modern economies. For instance, frameworks like the CATS model [67] delve into financial networks, examining how shocks propagate through interconnected systems and amplify systemic risks. Meanwhile, EURACE and Eurace@Unibi [68] incorporate stock-flow consistency and multi-regional dynamics, enabling detailed analysis of trade flows, labor mobility, and fiscal policies. In contrast, the Keynes meets Schumpeter (KS) framework links innovation-driven growth with demand-side fluctuations [69], offering insights into the interplay between long-term technological advancements and short-term economic cycles. For more an in-depth review of macroeconomic agent-based models I refer to Dawid and Delli Gatti (2018) [70].

Beyond these general frameworks, individual models exemplify the practical utility of ABMs in addressing specific macroeconomic challenges. One such example is the open economy model developed by Poledna et al. (2023) [71], which uses national accounts and sector-specific data to forecast key economic variables like GDP, inflation, and unemployment. This model not only matches but often outperforms traditional approaches such as DSGE and VAR models in handling shocks and producing accurate out-of-sample forecasts, demonstrating the growing potential of ABMs for policy evaluation and economic forecasting (see for a similar approach [72]). Beyond their quantitative capabilities, ABMs increasingly incorporate insights from behavioral economic research, making them more and more realistic. For example, the model developed by Salle et al. (2013) [73] integrates adaptive expectations and bounded rationality to explore how central bank communication and credibility shape macroeconomic stability, providing critical insights into the role of expectations in policy effectiveness.

ABMs have also been used to address pressing interdisciplinary challenges. The Dystopian Schumpeter Meeting Keynes model [74], developed by Lamperti et al. (2018) [75], bridges economic and climate systems, showing how economic activities drive climate change and how climate shocks feedback into economic dynamics. By illustrating the insufficiency of carbon taxes alone and emphasizing the need for coordinated green innovation policies, this model highlights the ability of ABMs to tackle multi-system interactions and provide actionable insights into sustainability and growth.

Labor market dynamics are another area where ABMs offer insights. Rio-Chanona et al. (2021) [48] use a data-driven occupational mobility network to analyze the effects of automation on labor markets, revealing how network frictions exacerbate unemployment and highlighting the importance of targeted retraining policies. Similarly, Fair and Guerrero (2023) [76] develop a labor flow network model that captures job-to-job transitions influenced by structural shocks, providing insights into how shifts in job distributions across industries and regions impact unemployment and worker mobility.

The model developed by Bardoscia et al. (2025) [77] further exempli-

fies the policy relevance of ABMs. Developed in collaboration with the Bank of England, this model evaluates the effects of prudential policies on the UK housing and credit markets. It shows how measures like capital requirements and loan-to-income caps influence lending, housing transactions, and broader economic outcomes, highlighting the complex interplay between regulatory policies and macroeconomic stability (see Hommes et al. (2024) [78] for an ABM co-developed with the Bank of Canada and Mérő et al. (2023) [79] with the Bank of Hungary).

These examples illustrate the adaptability of ABMs in modelling diverse economic phenomena, from financial networks to climate change and labor markets. As these models become increasingly sophisticated, they are not only matching but often surpassing traditional frameworks in accuracy and explanatory power. This growing capability positions ABMs as essential tools for policymakers, offering robust, data-driven understanding to address the complexities of modern economies.

Challenges of Agent-based Models

While Agent-Based Models offer insights into economic systems, their practical application comes with significant challenges. These hurdles—ranging from data requirements to the design of heuristic rules and parameter calibration (see Pangallo and Rio-Chanona (2024) [80] for a review). However, ongoing advancements are steadily addressing these issues, enhancing the robustness and relevance of ABMs.

One major challenge lies in the reliance on detailed, micro-level data to define agent behaviors, interaction networks, and initial conditions. Collecting such granular data has traditionally been a significant hurdle. Researchers are now leveraging big data sources, such as transaction-level datasets, social media activity, and real-time financial information, to empirically ground their models. Machine learning techniques further facilitate this process by extracting patterns and calibrating complex models from large datasets [81, 82]. Open data initiatives and collaborations between academic and governmental institutions are also expanding access to the necessary data, fostering a more supportive ecosystem for ABM development [76, 78, 83].

Designing heuristic rules for agent behavior presents another challenge. While these rules capture bounded rationality, they often rely on oversimplified assumptions that may lack empirical or theoretical foundations. Recent efforts integrate insights from behavioral economics [84], psychology [85, 86], and experimental data from laboratory settings [87, 88] to create more realistic decision-making processes.

Computational demands remain a significant barrier, particularly for large-scale simulations involving thousands or even millions of agents. Advances in parallel computing and the development of more efficient algorithms are enabling faster simulations without compromising complexity. These technological improvements are not only making ABMs more accessible but also paving the way for large-scale ap-

plications, such as creating digital twins of economies, capable of simulating real-world dynamics at a 1-to-1 scale [71, 72].

One of the most persistent challenges remains the systematic exploration of vast and complex parameter spaces. The sheer number of parameters and their interactions often makes it difficult to identify which configurations drive meaningful macroeconomic dynamics. Researchers are addressing this issue using sensitivity analysis [89], advanced sampling techniques [90], and surrogate modelling [91]. Optimization algorithms, including evolutionary strategies, are also being employed to efficiently navigate large parameter spaces [92]. However, those techniques often require substantial computational resources and can struggle to uncover the full range of a model's behavioral *phases*. Addressing this challenge is essential to fully harness the potential of ABMs for policy analysis and theoretical insights. In this context, we introduce a novel sloppiness-based framework for parameter exploration, offering an efficient and systematic approach to navigate high-dimensional parameter spaces. By leveraging this framework, we aim to identify critical parameters and uncover key phase transitions, enhancing both the interpretability and robustness of ABMs.

2.2. Systematic Parameter Exploration

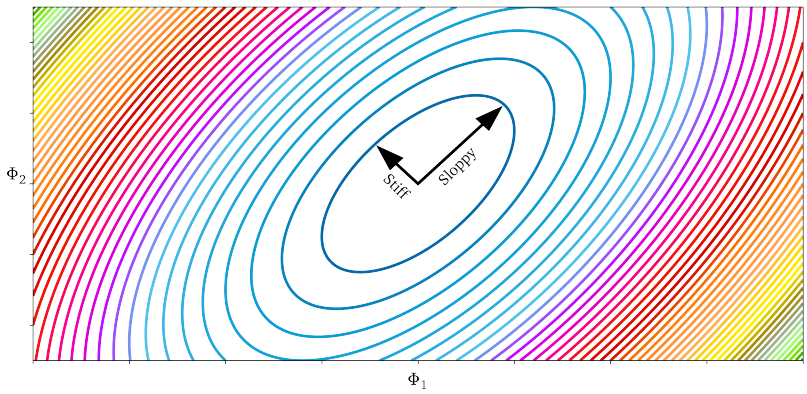


Figure 2.4.: Visualization of the loss function in parameter space: Contour plot of the loss function for a two-dimensional model parameterized by Φ_1 and Φ_2. Movement along the stiff axis results in significant changes to the loss function, indicating strong sensitivity to parameter variations in this direction. In contrast, movement along the sloppy axis leads to minimal changes in the model's dynamics, demonstrating insensitivity to parameter variations in this direction. Figure adapted from Gutenkunst et al. (2007) [93].

Systematically exploring complex parameter spaces in agent-based models (ABMs), raises the question on *how does one determine which parameters are most important for driving model dynamics*? To address this, it is helpful to step back from ABMs and examine the broader context of models used to fit data.

Suppose we have data points $y(t_i)$ that we aim to fit using a model $y(\Phi, t)$, parameterized by $\Phi = \{\Phi_1, \Phi_2, \ldots, \Phi_n\}$. A simple and common way to evaluate the model's quality is to define a loss function $\mathcal{L}(\Phi)$, such as the squared error, which quantifies the discrepancy between the model predictions and the data:

$$\mathcal{L}(\Phi) = \sum_i \big(y(\Phi, t_i) - y_i\big)^2 \tag{2.6}$$

At the best-fit parameters, the loss landscape can be visualized as forming ellipsoidal contour surfaces in the parameter space (see Figure 2.4).

The geometry of these ellipsoids provides insights into the sensitivity of the loss function to variations in the parameters. Along the major axis of the ellipsoid (the *sloppy* direction), the loss function changes minimally across a wide range of parameter values, indicating that these parameters have little influence on the model's behavior. Conversely, along the minor axis (the *stiff* direction), the loss function changes rapidly, identifying parameter combinations that are critical for maintaining the quality of the fit. The terms *stiff* and *sloppy* were introduced by Brown and Sethna (2003) [94], who first examined this concept in the context of systems biology models. It is important to note that *sloppy* does not imply the model is of poor quality; rather, it reflects a hierarchy of parameter importance, with some directions in parameter space being less influential than others.

The stiff and sloppy directions in parameter space can be mathematically characterized by the eigenvalues and eigenvectors of the Hessian matrix of the loss function:

$$H^{\mathscr{L}}_{i,j}(\Phi) = \frac{\partial^2 \mathscr{L}(\Phi)}{\partial \Phi_i \partial \Phi_j} \tag{2.7}$$

The eigenvectors of $H^{\mathscr{L}}_{i,j}(\theta)$ indicate the directions of the ellipsoid's axes, while the eigenvalues λ_n determine the lengths of the axes, that are $\propto 1/\sqrt{\lambda_n}$ [93]. Small eigenvalues correspond to sloppy directions, where the loss function remains nearly unchanged, whereas large eigenvalues correspond to stiff directions, where small parameter changes cause significant deviations in the loss.

While this description provides a simplified view in two dimensions, in higher-dimensional parameter spaces, the loss landscape often resembles a ribbon-like structure [95]. In nonlinear models, the loss function may also be non-smooth, adding further complexity. Nonetheless, numerous studies across diverse fields have shown that many models exhibit *sloppiness*, characterized by eigenvalue spectra spanning several orders of magnitude [96]. This means that some directions in parameter space are vastly more influential than others, with many parameters contributing negligibly to the loss.

Importantly, the specific choice of the loss function (e.g., squared error in Eq. 2.6) is not unique—other metrics can also be considered, depending on the modelling goals [97, 98].

Defining Sloppiness for Macroeconomic Agent-based Models

We adapt the sloppiness framework to agent-based models, using a mean-squared loss function to identify stiff and sloppy directions. This approach provides a systematic way to analyze the sensitivity of ABMs to their parameters and to understand how parameter importance influences model dynamics.

In a first step, we define stiff and sloppy directions in parameter space formally. The eigenvector corresponding to the largest eigenvalue represents the stiffest linear combination of parameters, while the eigen-

vector corresponding to the smallest eigenvalue is the sloppiest direction.

More precisely, let $\Phi = \{\Phi_1, \ldots, \Phi_P\}$ represent a point in P-dimensional parameter space. We quantify the change in observable outcomes due to a shift δ in parameters from Φ to $\Phi + \delta$ by the mean-squared loss function:

$$\mathscr{L}(\Phi, \delta) = \frac{1}{2SKT} \sum_s \sum_k \sum_t \left(\frac{y_{s,k,t}(\Phi + \delta) - y_{s,k,t}(\Phi)}{\left\| y_{s,k}(\Phi) \right\|} \right)^2, \tag{2.8}$$

where $y_{s,k,t}(\Phi)$ is the realisation of output variable $k \in \{1, \ldots, K\}$ at time $t \in \{1, \ldots, T\}$ for random realisation $s \in \{1, \ldots, S\}$. The term $\left\| y_{s,k}(\Phi) \right\|$ is a normalisation to re-scale outputs to the same unit, and is taken to be the mean realisation. To avoid numerical issues, the output variables should all be of a similar order of magnitude.

Unlike Eq. 2.6, this loss function compares simulated outputs, $y_{s,k,t}(\Phi)$, to the changes caused by shifting the parameters to $\Phi + \delta$.

The Hessian matrix for the mean-squared loss at point Φ is defined slightly different to Eq. 2.7 as

$$H^{\mathscr{L}}_{i,j}(\Phi) := \left. \frac{d^2 \mathscr{L}(\Phi, \delta)}{d \log \Phi_i \, d \log \Phi_j} \right|_{\delta = 0}, \tag{2.9}$$

and we consider parameters $\Phi_i, \Phi_j \in \Phi = \{\Phi_1, \ldots, \Phi_P\}$, where P is the number of parameters. We take the derivative with respect to the $\log$ parameters to focus on relative parameter changes, as most parameters have different units. Evaluating the Hessian at a predefined target point Φ yields

$$H^{\mathscr{L}}_{i,j}(\Phi) = \frac{1}{SKT} \sum_s \sum_k \sum_t \frac{1}{\left\| y_{s,k,t}(\Phi) \right\|^2} \frac{d y_{s,k,t}(\Phi)}{d \log \Phi_i} \frac{d y_{s,k,t}(\Phi)}{d \log \Phi_j}, \tag{2.10}$$

where the second order term vanishes as we evaluate $H^{\mathscr{L}}_{i,j}$ at $\delta = 0$. This reduces the computational burden to $\mathcal{O}(PS)$ function calls of the underlying model. The target point Φ is the parameter combination that the modeller wants to investigate. Furthermore, since Eq. (A.52) is a sum of projectors, it is immediate to see that the numerically determined matrix H is by construction positive definite.

Due to the simulation based nature of ABMs, we calculate the first-order derivatives numerically using a central difference approach.[4] While numerical derivatives can be troublesome in sloppy models, we find that the eigenvalues of the Hessian matrix converge to a set of steady ratios as the simulation time T increases, which can be aided by introducing an equilibrating time T_{eq} before which all observations are dropped. We have generally observed eigenvalues to converge sequentially, with the largest eigenvalue converging first. Consequently, the values of the smallest eigenvalues are likely to be more noisy and closer together. In contrast to T and T_{eq}, the number of seeds S has little effect on convergence beyond a certain minimum threshold.

To assess the parameter directions of importance, we consider the eigenvectors $\{v_1, \ldots, v_P\}$ of the numerically approximated Hessian

4: Central Differences imply $\frac{d y_{s,k,t}(\Phi^\star)}{d \log \Phi_i} \approx [y_{s,k,t}(\log \Phi^\star + \delta e_i) - y_{s,k,t}(\log \Phi^\star - \delta e_i)]/2\delta$, where e_i is a vector of zeros such that only element i equals to one, and δ is the step-size taken in log space. The choice of δ is arbitrary and potentially model dependent. In our case, we have chosen $\delta = 0.1$ for the agent-based models to ensure a sufficiently large step to avoid noise, while remaining small enough to avoid numerical instability. For the DSGE models we found $\delta = 0.1$ to be unstable, and so reduced to $\delta = 0.001$.

matrix (Eq. (A.52)). The eigenvectors correspond to the sorted eigenvalues $\{\lambda_1, \ldots, \lambda_P | \lambda_i \geq \lambda_j \; \forall \; i < j\}$, where λ_1 is the largest eigenvalue. The first eigenvector, v_1, represents the linear combination of parameters corresponding to the stiffest direction in parameter space. Before applying the sloppiness approach to studying ABMs, we first extend the Mark0 agent-based model in Chapter 3 to explore post-COVID inflationary dynamics. Using the sloppiness framework, we identify critical mechanisms, pinpointing the Wage-Price spiral as a key driver for crises. In Chapter 4, we expand the analysis to additional models and develop an algorithm to systematically explore the phase spaces of agent-based models, offering deeper insights into the dynamics they are able to generate.

Key Takeaways

1. **Navigating Complexity in Economic Models:** Agent-based models offer a powerful way to explore macroeconomic dynamics, but their complexity makes it challenging to extract clear insights. Understanding which parameters truly shape economic outcomes is crucial for making these models more practical and interpretable..
2. **Exploration of Key Drivers of Economic Behavior:** By applying a statistical physics-inspired approach, we shows that only a few key parameters significantly influence macroeconomic trends. Identifying these *stiff* parameter directions simplifies model exploration and provides a more efficient way to analyze policy interventions.

Post-COVID Inflation & the Monetary Policy Dilemma: An Agent-Based Scenario Analysis

3.

This chapter is taken from Knicker et al. (2024) [3] *Post-COVID inflation and the monetary policy dilemma: an agent-based scenario analysis*. To align with the objectives of this thesis, I modified several parts to provide additional context. The section on the Beveridge curve is entirely new. The original work was a collaborative effort with *K. Naumann-Woleske* and *J.-P. Bouchaud* and *F. Zamponi*.

In Chapter 2, we introduced agent-based models and their diverse applications formally. In this chapter, we extend an existing agent-based model, the Mark0 model developed by Gualdi et al. (2015) [99], by incorporating a trust mechanism that captures economic agents' confidence in the Central Bank. Using this enhanced framework, we examine the post-COVID inflationary dynamics.

Inflation has captured global attention since the onset of the COVID-19 pandemic (from now on, simply "COVID") in 2020. In the United States annual inflation reached 4.8% in 2021 and peaked at 9.1% in June 2022,[1] while Europe experienced highs of 11.5% in October 2022.[2] Competing narratives have emerged to explain the mechanisms driving this inflationary surge, which has persisted longer than expected. Policymakers have been blindsided by inadequate models, with the Bank of England admitting it had "big lessons to learn" from failure to forecast inflation using existing models.[3] This chapter aims to explore the influence of fiscal and monetary policies on prevailing inflationary dynamics within a complex macroeconomic environment, modeled using the Mark-0 Agent-based Model that is based on alternative modelling foundations.

During the period of high inflation in 2021 and 2022, several theories emerged to explain the underlying causes of persisting inflationary trends. These interpretations led to differing views on appropriate policy responses, ranging from monetary policy interventions such as interest rate hikes to more targeted fiscal policies and price controls. Our analysis contributes to the debate on the appropriate policy responses to post-COVID inflation by providing a flexible framework to assess different policy options in the context of various inflation drivers, including demand-pull, cost-push, and profit-driven inflation. Our framework can accommodate varying behavioral foundations within a complex economy, including agents' trust in the Central Bank's clout and the anchoring of inflation expectations. Our main conclusions are that (i) the economic recovery after the shocks, especially in absence of mitigating fiscal policies, can be much more sluggish than expected – or even fall into deep recessions beyond dangerous tipping points or *dark corners*[100]; (ii) the policy response (both

1: see e.g. U.S. Bureau of Labor Statistics, Consumer Price Index for All Urban Consumers: All Items in U.S. City Average [CPIAUCSL], retrieved from FRED, Federal Reserve Bank of St. Louis; May 24, 2023.

2: see e.g. Eurostat EuroIndicators, report n. 31/2023, https://ec.europa.eu/eurostat/documents/2995521/16310161/2-17032023-AP-EN.pdf

3: see for example the Financial Times, https://www.ft.com/content/b972f5e3-4f03-4986-890d-5443878424ac

from the government, the Central Bank, and other public authorities) has to navigate a narrow path, facing the trade-off between high inflation (with the risk of a runaway scenario) and high unemployment (with the risk of an economic collapse). In particular too weak a fiscal stimulus is ineffective, and too large a stimulus fuels high inflation; (iii) the success of monetary policy in curbing inflation is primarily due to expectation anchoring, rather than to direct impact of interest rate hikes; (iv) the two most sensitive model parameters (in terms of the inflation outcome) are those describing wage and price indexation, or in other words the bargaining power of workers and the market power of firms.

An initially dominant view of the post-COVID inflation was based on the *too much money chasing too few goods* theory of inflation [see e.g. 101, 102], also known as demand-pull inflation. This narrative focuses on the large amount of fiscal stimulus, such as the Coronavirus Aid, Relief and Economy Security Act (the "CARES Act") or the American Rescue Plan in the U.S., that have led to excess demand due to increases in disposable income while supply has not adjusted, thus pushing up prices. Some, such as Ferguson and Storm (2023) [103] agree with the demand-pull inflation analysis but contend that "the final cause of the inflationary surge in the U.S. [...] was in large measure the unequal (wealth) effects of ultra-loose monetary policy during 2020-2021".

In contrast to the excess demand interpretation, scholars such as Stiglitz and Regmi (2022) [104] conclude that "today's inflation is largely driven by supply shocks and sectoral demand shifts, not by excess aggregate demand". This claim is supported by the study of Cavallo and Kryvtsov (2023) [105], who present evidence spanning the years 2020 to 2022 illustrating the significant inflationary consequences of unexpected disruptions in product availability and stockouts across various sectors. Taking this one step further, Guerrieri et al. (2022) [106] and Ferrante et al. (2023) [107] consider the case wherein a negative supply shock on one (or several) sectors can actually lead to a larger drop in demand than the original shock, and consequently a recession. In the New Keynesian model of Ferrante et al. (2023) [107] demand reallocation explains a large share of the post-COVID U.S. inflation. The energy and food price shocks following COVID and the Russian invasion of Ukraine in 2022 were major causes of inflation, but this time from a cost-push inflation perspective wherein firms pass on increases in costs to consumers through prices. The strength of the inflationary surge was in large part due to the systemically important nature of the sectors where inflation occurred: energy and food [108]. These observations suggest an alternative cost-push scenario that is outside the realm of Central Bank's policy toolkit. This is also the conclusion of Bernanke and Blanchard (2023) [109], where a granular decomposition of inflation over different factors is proposed.[4]

Finally, a recent debate has emerged around profit-driven inflation, wherein sellers have increased prices beyond the increase in costs they face, thus expanding their profit margins. This is based on the observation that profits have increased sharply in the current inflationary climate, as compared to the observed amount of cost increases, even when including concerns of a wage-price spiral, and the excess de-

4: Similarly to our own approach, Bernanke and Blanchard (2023) [109] include wage bargaining, labor tightness and trust anchoring in their framework. Among the most important differences, however, are the absence of (i) supply-demand imbalances in the price setting mechanism and (ii) possible price gouging effects. Furthermore, the multiple equilibria and corresponding *tipping points* found within our model do not exist in their simplified specification.

mand cited by a monetarist perspective [104, 110]. Weber and Wasner (2023) [111] analysed firms' earnings calls and posit an interpretation based on imperfect competition and market power. Specifically, in an environment of cost increases and supply bottlenecks, firms' market power temporarily increases. This may be enhanced by the fact that when inflation is high, uncertainty about prices increases, such that consumers may be prone to accepting unreasonably high prices (an effect sometimes called "consumer discombobulation"), thus allowing companies to enlarge their margins.

All of these different interpretations lead to different guidelines for how policy should respond, and whether it should at all. From a standard monetarist perspective, the Central Bank should raise interest rates in order to push down demand, thus solving the excess demand scenario. On the other hand, in a cost-push scenario, the Central Bank's rate has no effect on the external cost increases and may actually harm the situation if firms choose to pass on the increased costs of debt to consumers. On this, Stiglitz and Regmi (2022) [104] suggest "monetary policy, then, is too blunt an instrument because it will greatly reduce inflation only at the cost of unnecessarily high unemployment, with severe adverse distributive consequences", as in their view inflation is not due to excess demand. In reality, the U.S. Federal Reserve has hiked rates at the fastest rate since Paul Volcker was chairman, in line with a monetarist view. Simultaneously, in 2023 inflation has begun easing, which begs the question of whether the reduction in inflation throughout 2022 is due to monetary policy or to external factors such as the easing of energy prices as with the oil crises of the 1970s [112, 113], or else to a spontaneous tendency of inflation to self-heal due to economic forces. But since the Central Bank's mandate to keep inflation low and stable, they are expected to act, generally through interest rate mechanisms and communication.[5] The question remains, to what effect? Are dips in inflation due to exogenous or endogenous factors or to Central Bank policy? What are the consequences of raising rates in this environment? Are the consequences of not acting on inflation with monetary policy greater than those when one does? After all, Bruno and Easterly (1998) [114] found that "countries can manage to live with relatively high – around 15-30 percent – inflation for long periods". It is these questions upon which we aim to shed some light by considering various scenarios generated through an Agent-based Model, that allows us to run various counterfactual scenarios.

5: There is an ongoing impression that monetary policy is the *only game in town* when it comes to inflation response. See e.g. ECB economist Isabel Schnabel's recent speech about the "Risk of Stubborn Inflation", https://www.ecb.europa.eu/press/key/date/2023/html/ecb.sp230619$__$1$\sim$2c0bdf2422.en.html

To conduct our analysis, we use the Mark-0 Agent-based Model (ABM) originally proposed by Gualdi et al. (2015) [99], and extended in Gualdi et al. (2017) [115], Bouchaud et al. (2018) [116] to study monetary policy and inflation, with an early application to the effects of the COVID pandemic by Sharma et al. (2020) [117] and a model-exploration study by Naumann-Woleske et al. (2024) [2]. Agent-based Models are computational scenario-generators in which a large number of individual agents, in our case firms, interact based on a set of behavioral rules [see 16, 70, 118, for more detailed introductions]. These models allow us to study (possibly) out-of-equilibrium systems by generating emergent macroeconomic dynamics, many of which have been successfully replicated [see 69, for instance]. Agent-Based Mod-

els also allow one, with minimal cost, to study the impact of several effects, such as the (de-)anchoring of trust in the Central Bank policies, which could be highly relevant in practice but are very hard to take into account within standard rational equilibrium models.

The Mark-0 ABM is a simplified model of a closed macroeconomy that nonetheless generates a wide variety of phenomena, from stable low unemployment and inflation, to endogenous crises that may oscillate regularly or punctuate long periods of recovery, or even to runaway inflation. The philosophy behind the model is to generate qualitatively plausible scenarios. In this context, Gualdi et al. (2017) [115] studied the efficacy of monetary policy in maintaining low unemployment and inflation, and "find that provided the economy is far from phase boundaries (or 'dark corners' [100]) such policies can be successful, whereas too aggressive policies may in fact, unwillingly, drive the economy to an unstable state, where large swings of inflation and unemployment occur." An analysis using the Mark-0 model in fact predicted, as early as June 2020, that the post-COVID recovery could be more sluggish than expected and lead to a period of sustained inflation [117].

In this chapter we present a more detailed study of the post-COVID recovery by considering three distinct shocks occurring in the Mark-0 model: (1) a COVID-shock that negatively impacts firms productivity and consumer demand for the period of lockdowns, (2) a supply-chain shock on firms' productivity that mimics the after-effects of COVID on global value chains, and (3) an energy price shock that is exacerbated by the Russian invasion of Ukraine. Each of these shocks is calibrated to macroeconomic time-series for the United States, such that the magnitude and duration match the observed data. We then study the effects of these shocks under different assumptions about the activity and efficacy of monetary policy, including the strength of monetary policy reactions to inflation, the ability of the Central Bank to influence firms' expectations, and the structure of firm decision-making on prices and wages.

Our first results confirm that in the presence of properly calibrated shocks, the economic recovery in absence of any mitigating policy is extremely sluggish, taking at least several years (if not much more) for the economy to return to the pre-crisis equilibrium. We note that this happens with shocks that are calibrated on the observed post-COVID macroeconomic time series, which already incorporate the effect of actually implemented policies; we expect that in absence of any policies the impact would have been even greater. This result thus confirms the necessity of some sort of fiscal policy to avoid the economy spiraling into a crisis or a very sluggish recovery.

A second set of results concerns the impact of such policies. We find that the model properly accounts for the inflation-unemployment trade-off, which leads to a narrow window in which policy can be efficient. For example, a disproportionate response of the Central Bank to inflation leads to unnecessary high unemployment, and an oversized injection of Helicopter Money leads to an unnecessary high inflation. Finally, we find that the anchoring (or de-anchoring) of economic agents' trust in the Central Bank strongly influences the dynam-

ics of inflation, much more than the direct influence of interest rates on economic activity. This is due to the fact that the decision of households to consume, and of firms to increase wages and economic activity is based on the gap between interest rates and expected inflation, which is either close to, or very different from the target inflation, depending on expectation anchoring. This confirms the necessity for the Central Bank to manage expectations. However, keeping expectations strongly anchored to the Central Bank's target may come at a cost: if households believe that inflation is under control while rates are going up, consumption might be hobbled. Similarly, if firms believe that inflation is under control when it is not, high interest rates will hurt employment while wages will lag behind, again leading to a drop in consumption. These effects, perhaps paradoxically, lead to increased unemployment compared to the case where expectations are floating. We conclude that each policy must be carefully tuned to achieve the desired result, and we believe that our modelling tool can guide policy makers in this respect.

Our primary objective is to provide possible scenarios and counterfactuals. As discussed at length in the previous papers [99, 116, 117], our ambition is not to provide precise predictions based on a fully calibrated model, but rather a tool for decision makers to help them apprehend different possible outcomes and anticipate unintended consequences and potential counter-intuitive impacts of their policies. We hope that Mark-0 can be usefully added to the policymakers toolbox and help them navigate a radically complex world (see e.g. King and Kay (2020) [119], Bouchaud (2021) [120]).

The remainder of the chapter is structured as follows: Section 3.1 gives an overview of the Mark-0 model and the adaptations made for this chapter. Pursuing this, Section 3.2 then outlines the policy channels that the model contains: interest rate, expectation management, access to credit, Helicopter Money and Windfall Tax. Section 3.3 discusses the model's dynamics in absence of exogenous shocks. In Section 3.4 we calibrate our three shocks, and show how they affect the macroeconomic dynamics with or without an easy credit policy, but without any monetary policy interventions by the Central Bank. In Section 3.5 we introduce and discuss monetary policy through interest rates and expectations management, and highlight its effects. In Section 3.6 we add fiscal policy by examining two distinct kind of stimuli, Helicopter Money and a Windfall Tax. Section 3.7 presents a discussion of the robustness of our model to variations of its parameters, and highlights the risk of hyperinflation. Finally, Section 6.5 summarizes our results and lays out some perspectives and ideas for future work.

3.1. The Mark-0 Model

The Mark-0 Model is a stylized macroeconomic agent-based model that is qualitatively very rich, as many plausible – and sometimes un-intuitive – economic phenomena can emerge at the macro scale. The model was created as a minimalist reduction of the Mark family of Agent-Based Models originally developed by Gaffeo et al. (2008) [121] and Gatti et al. (2011) [122]. Mark-0 was proposed by Gualdi et al.

(2015) [99], and expanded to monetary policy questions by Gualdi et al. (2017) [115] and Bouchaud et al. (2018) [116]. More recently, an analysis of possible economic responses to the COVID-pandemic were studied using Mark-0 by Sharma et al. (2020) [117]. Unless stated otherwise, we use the model description and formalism of Sharma et al. (2020) [117]. In this section, we outline the relevant equations for the model and explain the economic intuition behind all of them.

Model Overview

The model comprises a large set of firms producing a homogeneous consumption good that is purchased by a representative household sector. The household sector goes to the market with a consumption budget C_B computed as a fraction of savings, $S(t)$, wages $W(t)$, payouts from the energy sector $\delta_e \mathscr{E}_e(t)$, and interest on deposits $\rho_d(t)S(t)$,

$$C_B(t) = c(t)\left[S(t) + W(t) + \delta_e \mathscr{E}_e(t) + \rho_d(t)S(t)\right], \tag{3.1}$$

where $c(t) \in (0,1)$ is the consumption propensity out of wealth and income, defined as

$$c(t) = c_0\left[1 + \alpha_c\left(\hat{\pi}(t) - \rho_d(t)\right)\right], \qquad \alpha_c \geq 0, \tag{3.2}$$

where $\hat{\pi}(t)$, specified by Eq. (3.23) below, is the expected future inflation. Eq. (3.2) mimics the behaviour of the classical Euler equation, as the household consumes more when real interest rates, $\rho_d(t) - \hat{\pi}(t)$, are lower. The savings of households over time can then be written as:

$$S(t+1) = S(t) + W(t) + \rho_d(t)S(t) - C(t) + \Delta(t) + \delta_e \mathscr{E}_e(t) \tag{3.3}$$

with dividends received from firms profits $\Delta(t)$ and actual consumption $C(t) \leq C_B(t)$ that is given by the matching of demand and production.

Households choose to split their consumption budget between N_F different firms (labeled by an index $i = 1, 2, \cdots, N_F$) based on an intensity of choice model

$$D_i(t) = \frac{C_B(t)}{p_i(t)} \frac{\exp(-\beta p_i(t))}{\sum_j \exp(-\beta p_j(t))}, \quad \text{s.t.} \quad \sum_i p_i(t) D_i(t) \equiv C_B(t), \tag{3.4}$$

with a price sensitivity β.[6]

6: Good differentiation could easily be included at this stage by replacing p_i by p_i/ψ_i, where ψ_i is the preference for good i.

Because the model considers out-of-equilibrium situations, demand D_i for good i and production Y_i may not match, leading to a realized consumption c_i^R given by $\min(D_i, Y_i)$. Meanwhile, firms produce consumption goods based on a linear production function,

$$Y_i(t) = \zeta(t) N_i(t), \tag{3.5}$$

dependent only on the firm's employed labour force N_i and time-dependent labour-productivity ζ. The unemployment rate $u(t)$ is then given by $u(t) = 1 - \sum_i N_i(t)/N$, where N is the number of workers. At each time step, corresponding to one month throughout this chapter, firms can adapt to the current economic environment by choosing three firm-specific variables: the target production $Y_i(t)$, price $p_i(t)$,

and wage offered to their employees, $w_i(t)$. We describe these in turn, but note that the time step of one month and the different update parameters γ, g are assumed to be small such that production, prices and wages evolve slowly between t and $t+1$, barring the role of external shocks of the COVID type, as discussed in section 3.4 below.

1. *Production update.* Production adapts to the observed gap between supply and demand at the previous time step as follows:

$$\begin{aligned} &\text{If } Y_i(t) < D_i(t) \quad \Rightarrow \quad Y_i(t+1) = Y_i(t) + \min\{\eta_i^+(D_i(t) - Y_i(t)), \zeta u_i^\star(t)\} \\ &\text{If } Y_i(t) > D_i(t) \quad \Rightarrow \quad Y_i(t+1) = Y_i(t) - \eta_i^-[Y_i(t) - D_i(t)] \end{aligned} \tag{3.6}$$

where $u_i^\star(t)$ is the maximum number of unemployed workers that are available to the firm i at time t, which increases as the firm's wage $W_i(t)$ relative to the production-weighted average wage $\overline{w}(t)$ increases[7]

7: The production weighted average wage is defined as $\overline{w}(t) = \frac{\sum_i W_i(t) Y_i(t)}{\sum_i Y_i(t)}$.

$$u_i^\star(t) = \frac{N e^{\beta W_i(t)/\overline{w}(t)}}{\sum_i e^{\beta W_i(t)/\overline{w}(t)}} u(t). \tag{3.7}$$

In other words, the maximum available workforce to firm i is the total number of unemployed agents, times a wage dependent factor that favours high-paying firms.
The speed at which firms hire and fire workers depends on their level of financial fragility Φ_i, defined as the debt-to-sales ratio, where debt is negative cash balance:[8]

8: Note that this definition of fragility Φ slightly differs from that used in our previous publication, in particular the normalisation by the bankruptcy threshold Θ.

$$\Phi_i(t) = -\frac{1}{\Theta} \frac{\min\{\mathscr{C}_i(t), 0\}}{\min\left(p_i(t)D_i(t), p_i(t)Y_i(t)\right)}. \tag{3.8}$$

Non-indebted firms have zero fragility and Θ is the maximum debt-to-sales ratio allowed by firms creditors, beyond which firms are declared bankrupt.[9] Firms that are close to bankruptcy (i.e. $\Phi \approx 1$) are arguably faster to fire and slower to hire, and vice-versa for healthy firms. The coefficients of the hiring and firing rates $\eta_i^\pm$ for firm i (belonging to $[0,1]$) are given by:

9: For the detailed bankruptcy settlement, see Gualdi et al. (2017) [115], Bouchaud et al. (2018) [116]

$$\eta_i^\pm = [\![\eta_0^\pm \left(1 \mp \Gamma(t)\Phi_i(t)\right)]\!], \tag{3.9}$$

where $\eta_0^\pm$ are fixed coefficients, identical for all firms, and $[\![x]\!] = x$ when $x \in (0,1)$, $[\![x]\!] = 1$ for $x \geq 1$ and $[\![x]\!] = 0$ when $x \leq 0$. The factor $\Gamma > 0$ measures how the financial fragility of firms influences their hiring/firing policy, since a larger value of Φ_i then leads to a faster downward adjustment of the workforce when the firm is over-producing, and a slower (more cautious) upward adjustment when the firm is under-producing. Γ itself depends on the inflation-adjusted interest rate and takes the following form:

$$\Gamma(t) = \max\left[\alpha_\Gamma\left(\rho_\ell(t) - \hat{\pi}(t)\right), \Gamma_0\right], \qquad \Gamma_0, \alpha_\Gamma \geq 0, \tag{3.10}$$

where $\rho_\ell(t)$ is the rate at which firms can borrow, Γ_0 and α_Γ are free, non negative parameters, the latter being similar to α_c that captures the influence of the real interest rate on the hiring/fir-

ing policy of firms.

2. *Price update*. Compared to the previous versions of the Mark-0 model, in this chapter the firms' price update, $\Delta p_i(t+1) = (p_i(t+1) - p_i(t))/p_i(t)$, explicitly takes into account the demand-output gap, and a contribution for the change in exogenous (e.g. energy) price $\Delta p_{e,\text{ema}}$ (in %), to with:

$$\Delta p_i(t+1) = \gamma \xi_i(t) \cdot \frac{D_i(t)}{Y_i(t)} + g_p \hat{\pi}(t) + g_e \Delta p_{e,\text{ema}}(t) \qquad \text{if} \quad \begin{cases} D_i(t) > Y_i(t) \\ p_i(t) < \overline{p}(t) \end{cases} \tag{3.11}$$

$$\Delta p_i(t+1) = -\gamma \xi_i(t) \cdot \frac{Y_i(t)}{D_i(t)} + g_p \hat{\pi}(t) + g_e \Delta p_{e,\text{ema}}(t) \qquad \text{if} \quad \begin{cases} D_i(t) < Y_i(t) \\ p_i(t) > \overline{p}(t) \end{cases} \tag{3.12}$$

$$\Delta p_i(t+1) = g_p \hat{\pi}(t) + g_e \Delta p_{e,\text{ema}}(t) \qquad \text{otherwise,} \tag{3.13}$$

where

- γ is a parameter and $\xi_i(t)$ are independent uniform $U[0,1]$ random variables scaled up by the actual demand-output ratio $D_i(t)/Y_i(t)$ (or its inverse), in order to mimic an increased pressure on prices in case of supply (or demand) gluts. When $D_i(t) \approx Y_i(t)$, this reduces to the rule used in previous papers, see Gualdi et al. (2015) [99];
- $\hat{\pi}(t)$ is the expected long-term inflation in % per month, that firms take into account in their price setting mechanism. This term also accounts for changes in price due to expected changes in wages and input costs, which are part of the expected inflation. A factor $g_p > 1$ models an inflation gouging behaviour, while $g_p < 1$ means that only part of inflationary costs are passed to the customers.[10]
- $\Delta p_{e,\text{ema}}(t)$ is the exponentially weighted moving average (with parameter ω, as in Eq. (3.22) below) of the exogenous price variations that are partially or fully passed on to final customers. The weighted average reflects that changes in energy price are not instantaneously transmitted to customers but rather distributed gradually over several months.

10: Note that in principle, g_p (and g_w) might depend on the firm or on the sector. We have not considered this possibility here.

The magnitudes of the corresponding price adjustments are determined by parameters: γ for supply-demand imbalance, g_p for inflation expectations, and g_e for changes in the exogenous price. In this regard, g_e can be thought of as the effective energy-share of production that firms want to pass on to customers, whereas $g_p > 1$ would correspond to the much discussed concept of "greedflation".

Note that $\overline{p}$[11] is the consumer price index, defined as

$$\overline{p} = \frac{\sum_i c_i^R p_i}{\sum_i c_i^R} , \qquad c_i^R = \min(D_i, Y_i) , \tag{3.14}$$

with c_i^R the realized consumption of product i.

11: Here again, our present definition of $\overline{p}$ slightly differs from that used in the previous papers, where we used a production weighted index instead of a consumption based index. We conform in this chapter with the more standard definition of inflation based on the Consumer Price Index (CPI).

3. *Wage update.* In similar fashion, firms update wages as

$$\Delta w_i(t+1) = \gamma(1-u(t))\,(1-\Gamma(t)\Phi_i(t))\;\xi_i'(t) + g_w\hat{\pi}(t) \qquad \text{if} \quad \begin{cases} D_i(t) > Y_i(t) \\ \mathscr{P}_i(t) > 0 \end{cases} \tag{3.15}$$

$$\Delta w_i(t+1) = -\gamma u(t)\,(1+\Gamma(t)\Phi_i(t))\;\xi_i'(t) + g_w\hat{\pi}(t) \qquad \text{if} \quad \begin{cases} D_i(t) < Y_i(t) \\ \mathscr{P}_i(t) < 0 \end{cases} \tag{3.16}$$

$$\Delta w_i(t+1) = g_w\hat{\pi}(t) \qquad \text{otherwise} \tag{3.17}$$

with $\Delta w_i(t+1) = (w_i(t+1) - w_i(t))/w_i(t)$, and where $u(t)$ represents the unemployment rate, $\xi_i'(t)$ is a $U[0,1]$ random realisation, and g_w is the wage sensitivity to expected inflation, which could be seen as worker's bargaining power. Firm profits $\mathscr{P}_i$ include the revenue earned on their cash, the interest paid on their loans and the cost of energy $g_e p_e Y_i$:

$$\mathscr{P}_i = p_i(t)\min\{Y_i(t), D_i(t)\} - w_i(t)Y_i(t) + \rho_d \max\{\mathscr{C}_i(t), 0\} - \rho_\ell \min\{\mathscr{C}_i(t), 0\} - g_e p_e(t) Y_i(t)\,. \tag{3.18}$$

The conditions of wage change therefore depend on the demand-supply imbalance, current firm profit $\mathscr{P}_i(t)$ and firm financial health, and also on the tension on the labour market, since lower unemployment leads to higher wage increase. This allows the model to reproduce Phillips curve effects. When firms experience positive profits and maintain a positive cash balance, they distribute dividends to households as a fraction δ of their cash balance, as in Eq. (3.3):

$$\Delta(t) = \delta \sum_i \mathscr{C}_i(t)\,\mathbb{I}\Big(\mathscr{P}_i(t) > 0\Big)\mathbb{I}\Big(\mathscr{C}_i(t) > 0\Big)\,, \tag{3.19}$$

where $\mathbb{I}(E)$ is the indicator function of event E.

4. *Energy Sector.* In this chapter, we consider an exogenously varying energy price, $p_e(t)$. Before the explicit shock introduced in Section 3.4, we consider the real energy price to be constant and equal to the average price of goods, i.e. $p_e(t) = \overline{p}(t)$. In each period, firms pay a total amount $g_e p_e(t) \sum_i Y_i(t)$ to the energy sector. Subsequently, the energy sector pays a fraction of its accumulated profits to households at a rate δ_e. Therefore the cash balance $\mathscr{C}_e$ of the energy sector writes as

$$\mathscr{C}_e(t+1) = \mathscr{C}_e(t) + g_e p_e(t) \sum_i Y_i(t) - \delta_e \mathscr{C}_e(t) \tag{3.20}$$

with the fraction δ_e of the energy profits that is paid out to households as income. Here, δ_e can be understood as dividends, share sales, or other channels through which energy sector profits circulate back into the economy. The way we introduce the energy sector in this chapter as an accounting identity with exogenous prices is arguably simplistic, and can certainly be improved in a future version of the model.[12]

12: Note that in the model there is a resource constraint such that the total money M (that is created by the Central Bank) is kept fixed during the simulation The balance sheet of the banking sector can then be written as $M = S(t) + \mathscr{C}^+ - \mathscr{C}^- + \mathscr{C}_e$, with the household savings S and the cash balance of firms $\mathscr{C}^{+/-}$.

Inflation Expectations and Monetary Policy

The equations above depend on inflation rate expectations $\hat{\pi}(t)$, that we now define. We consider that the measure of realized inflation is given by the change in consumption-weighted average price,

$$\pi(t) = \frac{\overline{p}(t) - \overline{p}(t-1)}{\overline{p}(t-1)}, \tag{3.21}$$

where $\overline{p}$ is defined in Eq. (3.14).

In the model, agents form expectations of future inflation partly on the basis of past realisations and partly on the basis of their trust in the Central Bank ability to enforce its inflation target. More precisely, they use an exponentially weighted moving average over realised inflation,

$$\pi_{\text{ema}}(t) = (1-\omega)\pi_{\text{ema}}(t-1) + \omega\pi(t), \tag{3.22}$$

where ω sets a memory time over which agents perceive realized inflation, equal to $-[\log(1-\omega)]^{-1} \approx \omega^{-1}$ for small ω. Together with the Central Bank's communicated inflation target, $\pi^\star$, all agents form the same inflation expectation, given by a weighted average of $\pi_{\text{ema}}(t)$ and $\pi^\star$:

$$\hat{\pi}(t) = (1-\tau^T(t))\pi_{\text{ema}}(t) + \tau^T(t)\pi^\star, \tag{3.23}$$

where $\tau^T(t)$ is the degree to which expectations are anchored around the Central Bank's target.[13] Consistently with the definition of π_{ema} and the long-term nature of the inflation target, we interpret $\hat{\pi}(t)$ as the expectation of *long-term* inflation.[14]

13: For empirical work on the question of trust, see e.g. Christelis et al. (2020) [123].

14: This is in contrast to the common formulations in DSGE models that operate with one-period-ahead expectations ($t \to t+1$).

As far as monetary policy is concerned, we assume that the Central Bank sets the baseline interest rate, ρ_0, via a classical Taylor rule based on observed (realized) inflation π_{ema}:

$$\rho_0 = \rho^\star + \phi_\pi \left(\pi_{\text{ema}}(t) - \pi^\star\right), \tag{3.24}$$

with reaction strength ϕ_π. The baseline rate ρ_0 is then translated into a time-dependent rate on loans, $\rho_\ell(t)$, and deposits, $\rho_d(t)$, adjusting for the cost of bankruptcies. These interest rates are defined as

$$\rho_\ell(t) = \rho_0(t) + f\frac{\mathscr{D}(t)}{\mathscr{E}^-(t)}, \tag{3.25}$$

$$\rho_d(t) = \frac{\rho_0(t)\mathscr{E}^-(t) - (1-f)\mathscr{D}(t)}{S + \mathscr{E}^+(t)} \tag{3.26}$$

with the positive cash balance $\mathscr{E}^+(t) = \sum_i \max(\mathscr{E}_i, 0)$, firms total debt $\mathscr{E}^-(t) = -\sum_i \min(\mathscr{E}_i, 0)$ and the total costs accrued due to firms defaulting $\mathscr{D}(t)$. The parameter f determines how the consequences of defaults are allocated to lenders and depositors, interpolating between the costs borne entirely by borrowers ($f = 1$) and those shouldered entirely by depositors ($f = 0$). In this chapter, we do not consider a double mandate for the Central Bank as in Gualdi et al. (2017) [115], instead focusing only on inflation. We will comment about this below, as strict inflation control may turn out to be highly detrimental to unemployment.

Finally, we allow inflation expectation anchoring to evolve dynami-

cally via[15]

$$\tau^T(t+1) = (1-\omega)\tau^T(t) + \omega \exp\left[-\alpha_I \frac{|\pi(t) - \pi^\star|}{\pi^\star \phi_\pi}\right]. \tag{3.27}$$

This equation aims at capturing the fact that the degree of expectation anchoring depends on how closely the realised inflation actually matches the Central Bank target. This is factored into an exponentially weighted moving average with memory time $\approx \omega^{-1}$: realized and target inflation must differ significantly, and for sufficiently long times, for agents to lose trust in the Central Bank. A larger α_I means economic agents lose trust in the Central Bank more abruptly as the gap between realized inflation and inflation target becomes significant. We have included the factor ϕ_π in the denominator to emphasize the fact that stronger commitment of the Central Bank should decrease the sensitivity of anchoring on realized inflation, but of course this extra factor can be reabsorbed into α_I.

The calibration of the parameters we choose for the model is based on the results and studies done in the previous work around Mark0 (Gualdi et al. (2015) [99], Gualdi et al. (2017) [115], Bouchaud et al. (2018) [116], Sharma et al. (2020) [117]). Here, we discuss only the parameters that are relevant to the current work, in particular those identified by the sensitivity analysis in Section 3.7. The parameters used in this study can be found in the Table A.1.

15: Note that in the previous paper (Bouchaud et al. (2018) [116]), the anchoring parameter τ^T was assumed to be time independent, a case we will call "Anchored Expectation" henceforth.

3.2. Policy Channels in Mark-0

In the augmented Mark-0 model, the central authorities can influence macroeconomic dynamics through (i) the manipulation of interest rates, (ii) anchoring inflation expectations, and (iii) regulating the amount of debt firms can accumulate. Additionally, there is the possibility of fiscal policy in the form of direct injection of cash in the economy ("Helicopter Money") and Windfall Tax on the energy sector. These manipulations take effect through firms' financial fragility, which directly determines the probability of bankruptcies and indirectly changes the propensity of firms to hire and fire workers, as well as the wage setting behavior. Interest rates and cash injection also have effects on the willingness of consumers to spend. In addition, there is an expectations channel whereby firms and households may or may not trust the Central Bank's communicated inflation target depending on how close actual inflation is to the target. We now consider these mechanics in turn, as these will turn out to be important in response to the COVID related shocks that we will introduce in the next section.

The Interest Rate Channel

The Central Bank's baseline interest rate ρ_0 additively influences both deposit (ρ_d) and lending (ρ_ℓ) rates. Through this channel, the interest rate affects firms' wages and hiring strategies, as well as households' spending.

The impact on households is straightforward and parallels that of the standard Euler equation of intertemporal substitution: higher interest

rates, all else being equal, reduce the propensity to consume out of income and wealth through parameter α_c (Eq. (3.2)), effectively decreasing current demand in favor of later consumption. This decrease leads firms to reduce or maintain prices due to excess supply (see Eq. (3.12)), thereby lowering inflation.

The impact of interest rates on firm behavior occurs through the coefficient α_Γ, defined in Eq. (3.10), which influences firms' production target and wage adjustment. In all cases α_Γ affects firm behavior in combination with the firm's fragility Φ (where $\Phi = 1$ implies bankruptcy, see Eq. (3.8)). An increase in the baseline rate will increase Γ, which implies that firms have stronger reactions to excess demand or excess supply. Unhealthy firms ($\Phi > 0$) will therefore more cautiously expand production when needed (η^+), but more abruptly fire staff in the case of excess supply (η^-). Hence, an increase of the baseline rate ρ_0 will tend to decrease production and increase unemployment as fragile firms feel the brunt of the cost of debt. For similar reasons, an increase of ρ_0 implies a stronger downward pressure on wages.

Therefore, in Mark-0 an increase of the baseline rate ceteris paribus induces a reduction of consumption and an increase of unemployment. The success of a rate hike in curbing inflation depends on the strength of the response of households and of firms to interest rates, and hence on the value of the parameters α_c and α_Γ. If, as some authors argue [124], the reaction of households and of firms to interest rates are subdued, large hikes would be necessary for monetary policy to have a significant effect on inflation, in part by bringing about a recession [104].

The Expectation Channel

With rising global inflation, the debate about the importance of inflation expectations for macroeconomic dynamics has been rekindled. Some, such as Rudd (2022) [125], argue that short-term inflation expectations play no role in macroeconomic dynamics, while others, like Reis (2021) [126], maintain their importance but question whether they are *de-anchored* from Central Bank's targets. In the present version of the Mark-0, long-term inflation expectations $\hat{\pi}$ play a role in price and wage settings: formally, the factor g_p in Eqs. (3.11) and (3.12) determines the fraction of expected inflation reflected in firms' next-period prices. In a similar style, the factor g_w in Eqs. (3.15) and (3.16) controls the rate of wage adaptation to inflation in response to firms' expected inflation, and is thus adjacent to the idea of labour-bargaining power.

The long term inflation expectation, $\hat{\pi}$, is determined as a mixture of past realized inflation and the Central Bank target, with the Central Bank's effectiveness in reaching its target dictating the level of "anchoring" τ^T (see Eq. (3.23) and (3.27)). With this formulation, economic agents lose confidence in Central Bank actions when inflation significantly and persistently deviates from the Central Bank's inflation target. The parameter α_I represents the sensitivity of trust to the Central Bank's ability to control inflation.

De-anchored beliefs imply firms' inflation expectation is based solely on past observation, running a risk of a wage-price spiral or a hyper-

inflation runaway, depending on the imbalance between the parameters g_p and g_w. The simplest case is equality in bargaining and market power, $g_w = g_p$, which implies that on average both prices and wages are raised in the same proportion to expected inflation, thus having no real effects on the economy. One then faces either stable inflation when $g_w = g_p < 1$ or possible hyper-inflation when $g_w = g_p > 1$, see section 3.7.

A more interesting case is $g_p \neq g_w$, implying a differentiation in bargaining vs. pricing power. For $g_p > g_w$, labour bargaining power is lower than firms pricing power. In this case, firms raise prices by more than wages, eroding the purchasing power of workers. This results in lower demand that makes price hikes less likely until firms' excess demand becomes excess supply, reversing the cycle. For $g_w > g_p$, on the other hand, an unstable wage spiral may set in, with higher wages driving demand up, leading to further increases of prices as firms face excess demand. The situation $g_p > 1$, at least for firms with a healthy balance sheet, would correspond to "greedflation", i.e. firms trying to use inflation to hide increases in their profits.

Credit Regulatory Policy

The last channel for the monetary authority to affect the economy within the framework of Mark-0 is by means of the default threshold, Θ. Earlier work on the Mark-0 model found that the default threshold is a key parameter causing a phase transition between a regime of high unemployment and/or endogenous crises (low Θ), and of stable full employment (higher Θ conditional on low interest rates) [see 99]. This is because higher thresholds give firms more time to adjust their production and price strategy to return to a profitable state, assuming interest payments are low. Empirically, both in the CARES act and the CAA there were relaxations of the bankruptcy laws in the United States.[16] Direct state aid to firms may also prevent bankruptcies from occurring. In the realm of our model these policies are mimicked by increasing the bankruptcy threshold Θ of firms during the shock and reducing it again after the shock. The "Easy-Credit policy" proposed in Sharma et al. (2020) [117] amounts to set the value of Θ such that it remains commensurate to the current average firm fragility, to wit

16: CARES refers to the Coronavirus Aid, Relief, and Economic Security Act, and the CAA refers to the Consolidated Appropriations Act of 2021.

$$\Theta(t) = \max(\mu\langle\phi\rangle(t), \Theta_0), \tag{3.28}$$

where Θ_0 is the bankruptcy threshold before the shock, μ a multiplier, and $\langle\phi\rangle(t)$ the firm-wide average debt-to-sales ratio at time t, meaning that only the firms whose debt-to-sales ratio exceeds $\mu\langle\phi\rangle$ must file for bankruptcy. Throughout the following we will use $\mu = 1.3$ and $\Theta_0 = 3.2$, which represents a compromise between "too robust" economies when Θ_0 is large, and "too fragile" economies when Θ_0 is small.

This implies that with easier credit access, firms can accumulate debt during the shock without the fear of bankruptcy. This, in turn, affects the firing and hiring policies of firms as they can continue with their business as usual, leading to a hiring and firing as if they were not fragile (see Eq. (3.8)). Sharma et al. (2020) [117] have shown that such policy is effective in speeding up the recovery to pre-shock levels after

severe COVID-like shocks, without such policy long-lasting recessions would otherwise ensue.

Helicopter Money

When interest rates are low and access to credit is loose, but shocks nonetheless persist, central authorities may turn to less conventional policies such as "Helicopter Money" to address economic downturn. Introduced by Friedman (1969) [127], this implies an injection of money directly into the hands of economic agents to increase spending and thus economic output.[17] While some view Helicopter Money as a potential cause of hyperinflation and currency devaluation, it can be an effective response to financial crises and pandemics by stimulating real economic activity [129]. In this chapter we consider Helicopter Money as introduced by Friedman: a distribution of newly printed money by the Central Bank directly to households, resulting in a decrease in the Central Bank's net worth and an increase in the net worth of households. This injection of money is modeled as a one-time increase of savings of households by a factor $\kappa_H > 1$, such that $S \rightarrow \kappa_H S$ instantaneously after all shocks. The increased savings of households boost demand through an increased consumption budget C_B (see Eq. (3.4)) and finally end up on the balance sheet of firms where it is then redistributed to households as wages and dividends.

17: By contrast, quantitative easing implies that money created by the Central Bank is used to purchase government bonds and distributed through the government, whereas Helicopter Money is directly distributed in the economy [128].

Windfall Tax

Whereas Helicopter Money stimulates consumption by injecting external funds into households savings, a tax on "windfall" profits, which refer to large unforeseen profits [130], achieves a similar outcome by redistributing money through taxation. Following the Russian invasion of Ukraine, there has been a significant surge in fossil fuel prices, leading to unexpectedly high returns for utilities and fossil fuel producing companies [131], and profit-driven price markups for other firms [111], leading to strong debate amongst policymakers about implementing a Windfall Tax. The specific design of the Windfall Tax as a rent-sharing fiscal policy may vary depending on the particular circumstances [132].

In this chapter, we implement a "Windfall Tax" as a temporary increase of the fraction of energy sector cash balance that is re-injected to the savings of households, i.e. $\delta_e \rightarrow \delta_e + \Delta\delta_e$. This tax has a duration of two years, commencing one year prior to the end of the price shock. The primary objective of increasing δ_e is to boost consumption by enhancing households' savings through the redistribution of the increased cash balance in the energy sector that follow the energy price increases.

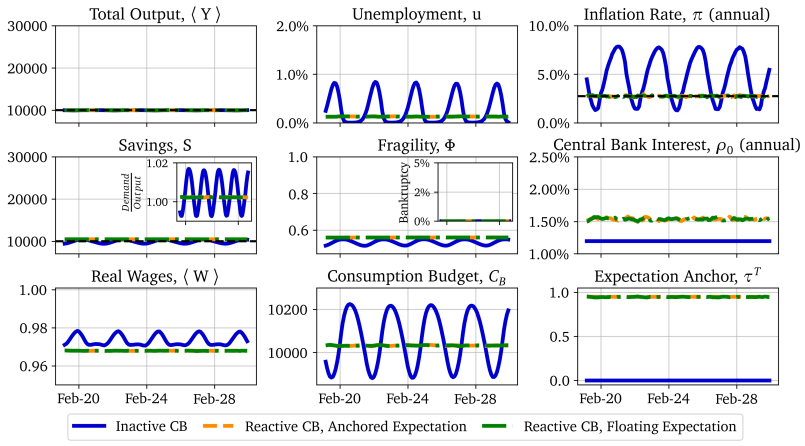


Figure 3.1.: Economic dashboard for the three Central Bank scenarios described in sections 3.3, 3.3 in absence of any policy and shock: Blue lines: Inactive Central Bank scenario ($\pi^\star = \phi_\pi = 0$, $\rho^\star = 1.2\%$ p.a.). Orange lines: Reactive Central Bank with Anchored Expectation scenario ($\pi^\star = 2.4\%$ p.a., $\phi_\pi = 1$, $\tau^T = 0.95$). Green lines: Reactive Central Bank with Floating Expectation scenario. The complete parameter set can be seen in table A.1. The two insets correspond to the demand/output ratio and the bankruptcy rate.

3.3. Stationary Dynamics: the Role of the Central Bank

The Economic Equilibrium Without Monetary Policy

To begin our analysis, we consider the economy without shocks. Our first choice of parameters corresponds to a Central Bank that does not react to inflation (i.e. $\pi^\star = \phi_\pi = 0$) while maintaining low baseline interest rates ($\rho^\star = 1.2\%$ p.a.). By consequence, agents form inflation expectations based only on their own observations ($\tau^T = 0$, see Eq. (3.23)). In this scenario, called "Inactive CB", and for the choice of parameters detailed in Table A.1, the economy spontaneously evolves into a steady state of full capacity, as evidenced by the low, oscillating unemployment rate around 0.4% (see Figure 3.1). There are small endogenous output and unemployment fluctuations of roughly two years in duration. Simultaneously, there are stronger oscillations in the demand and supply for goods, and in real wages. These fluctuations cause firms to adjust both their production and prices, thus generating a small "business cycle" with periodic imbalance between demand and output due to price adjustments. But in the absence of an active Central Bank target, the amplitude of the resulting inflation oscillations is found to be substantial, ranging between 1.5% and 8% p.a., suggesting that monetary policy indeed iron out business cycles. As already found by Gualdi et al. (2015) [99], these cycles emerge endogenously from the oscillating feedback loop of prices, demand and savings. In this scenario, called "Inactive CB", and for the choice of parameters detailed in Table A.1, the economy spontaneously evolves into a steady state of full capacity, as evidenced by the low, oscillating unemployment rate around 0.4% (see Figure 3.1). There are small endogenous output and unemployment fluctuations of roughly two years in duration. Simultaneously, there are stronger oscillations in the demand and supply for goods, and in real wages. These fluctuations cause firms to adjust both their production and prices, thus generating a small "business cycle" with periodic imbalance between demand and output due to price adjustments. But in the absence of an active Central Bank target, the amplitude of the resulting inflation oscillations is found to be substantial, ranging between 1.5% and 8% p.a., suggesting that monetary policy may iron out business cycles, as indeed found below. As already discussed in Gualdi et al. (2015) [99],

these cycles emerge endogenously from the oscillating feedback loop of prices, demand and savings.

Note however that depending on the choice of parameters, the economy may settle into a much less favourable state. In particular, when the hire-to-fire ratio is too small or when the baseline interest rate is too high, the economy collapses, see Gualdi et al. (2015) [99], Bouchaud et al. (2018) [116]. Similarly, when the bankruptcy threshold Θ_0 is too low, unemployment remains at a relatively high level.

Stabilizing Inflation through Monetary Policy

At this point, we introduce an active Central Bank and study its impact on inflation and unemployment. Specifically, we set the Central Bank's inflation target to $\pi^\star = 2.4\%$ p.a. and its reaction to inflation variation to $\phi_\pi = 1$. In this configuration, we assume that the Central Bank works with a well-anchored inflation expectation fixed at $\tau^T = 0.95$. In the pre-COVID period core inflation was quite stable as a result of an arguably successful monetary policy [133]. Simultaneously, the Central Banks had been acting in a transparent and measurable way despite the historically low interest rates [126], suggesting a high degree of trust. We call this constant τ^T scenario "Reactive CB, Anchored Expectation" in the following. In this case, Figure 3.1 shows that inflation is reduced to an average of $\langle \pi_{AT} \rangle \approx 2.7\%$, close to the 2.4% target, while nearly completely suppressing the business cycle with a very low stable unemployment. On the other hand, there is a slight reduction in real wages and higher interest rates.

Finally, we introduce a third scenario, referred to as "Reactive CB, Floating Expectation", where we allow agents' trust in the Central Bank to vary depending on the perceived success of the CB to bring inflation to roost, see $\tau^T(t)$ defined in Eq. (3.27). We choose as parameter values $\omega = 0.2$ (corresponding to a memory time of 5 months) and a sensitivity to off-target inflation of $\alpha_I = 0.4$. Again, all other parameters are taken from Table A.1. In the absence of external shocks, this configuration leads to a similarly high degree of trust in the Central Bank, and is otherwise identical to the case where trust is anchored – simply because inflation is on target.

3.4. The COVID Shock and its Aftermath in the Absence of Monetary Policy

Modelling the Three COVID Shocks

In order to assess the effects of recovery and monetary policies in response to the events of 2020-2022, we introduce in Mark-0 three shocks calibrated on U.S. data: (1) a COVID shock, (2) a supply-chain shock, and (3) an energy price shock. The data was retrieved from the Federal Reserve Bank of St. Louis, and are shown in Figure 3.2 where we have indexed values to January 2020 and calibrated the model until end of 2022.[18] We now discuss these shocks in turn:

1. **COVID**: The first major COVID outbreak emerged in the US in

18: This date corresponds to the moment we started our study. We have not tried to re-run the model including more recent data.

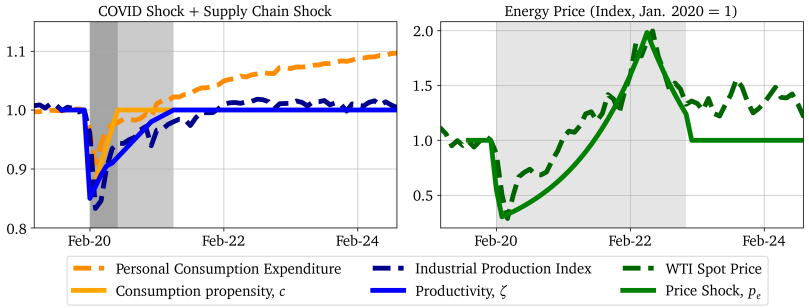


Figure 3.2.: Empirical shocks: Fitting of shock scenarios to macro data from the US up to end of 2022, when the present analysis was conducted. The empirical series (dashed lines) are indexed to January 2020. (Left Panel) Empirical series for the Personal Consumption Expenditure (orange, dashed line) and Industrial Production Index (blue, dashed line). Solid lines represent our model for the shocks to consumption propensity c (orange) and productivity ζ (blue). Dark grey area corresponds to the COVID Shock and the grey area to the Supply Chain Shock in the shock scenarios that we apply to Mark-0. (Right Panel) WTI Crude oil spot price during the periods of shock (green, dashed line). Modelling of the artificial price series (green, solid) in the Mark-0 model for a period of two and a half years (light grey area). Data retrieved from FRED (2023) [134], FRED (2023) [135], FRED (2023) [136].

February 2020, prompting the US government to declare a public health emergency, followed by the implementation of stay-at-home orders in March of the same year. The COVID outbreak led to a significant reduction in personal consumption expenditure by households, falling by around 15% in February 2020 (Figure 3.2, dashed orange line), which was of the same order of magnitude as the decline in the Industrial Production Index during the same month (dashed blue line). It took about 5 months for aggregate Personal Consumption Expenditure to recover to its pre-pandemic level. In accordance with this, we consider a shock to the consumption propensity $c(t)$ of households by 15%, which appears in Eq. (3.2), recovering linearly over 5 periods (solid orange line in Figure 3.2).

2. **Supply Chains**: During the initial COVID outbreak, firms in the US laid off a large number of workers, reducing their production to a similar degree as the personal consumption expenditure. However, while personal consumption recovered within 5 months, it took an additional 10-15 months for industrial production to return to its pre-pandemic state, as can be seen in the industrial production index (dashed blue line, Figure 3.2). This was due to a plethora of idiosyncratic supply chain disruptions, such as logistics and transportation difficulties, semiconductor shortages, pandemic-related restrictions on economic activity, and labor shortages that led to the slower recovery of production in the industrial sector [137]. In the context of this chapter, we model this by a shock to firm productivity $\zeta(t)$ (defined in Eq. (3.5)) of an initial magnitude of 15%, with a recovery of 15 months, see solid blue line in Figure 3.2.
3. **Energy Prices**: Finally, we consider the energy price shock. The reduction of demand and production throughout the pandemic led to a supply glut in energy markets, which led to a steep decrease in energy prices, such as oil, by up to 70% for immediate delivery of West Texas Intermediate crude oil (Figure 3.2, dashed green line). As the recovery period began, external factors such as extreme weather conditions in various parts of the

world and maintenance work that had been postponed during the pandemic caused a surge in demand, and the energy prices thus rebounded quickly [138]. Unfortunately, with the Russian invasion of Ukraine in February 2022 this rebound was further exacerbated to a global energy crisis, due to Russia's position as a major global exporter of natural gas and oil. By June 2022, the WTI crude oil spot price had peaked, rising nearly 100% compared to pre-pandemic levels. Following this, the global recovery and adjustment of energy markets has led to a sharp easing of energy prices. For Mark-0, these processes are external to the model, such that we introduce here an exogenous price shock to firms' price update Eqs. (3.11)-(3.13). Specifically, firms' prices change by an additional exogenous factor $g_e \Delta p_{e,\text{ema}}(t)$, where g_e is a constant factor akin to the energy-share in production, and $\Delta p_{e,\text{ema}}(t)$ is an exponentially weighted moving average of the time-dependent monthly percentage change in energy prices. Thus, the transmission of the change in energy prices to firms' product prices is smoothed, as expected to be the case in reality.

The form of the moving average is the same as for trust in Central Bank Eq. (3.27) and inflation expectations Eq. (3.22), with the very same memory time parameter ω. Our artificial energy price series is shown in Figure 3.2 as the solid green line, and is based on the WTI Spot Price.[19] Alternative calibrations with the energy component of the U.S. Consumer Price Index or the Henry Hub Natural Gas price lead to similar shapes and magnitudes. In the following we will choose $g_e = 3.25\%$, i.e. half of the energy share of the GDP in the US, where the factor 2 accounts for inventories and partial substitutability. We also posit that the profits of the energy sector are transferred back to households at an effective rate of $\delta_e = 4\%$ per time step, unless stated otherwise.

19: Prior to the shock, we set $p_e(t) = \overline{p}(t)$.

The Effect of Shocks Without Stabilisation Policies

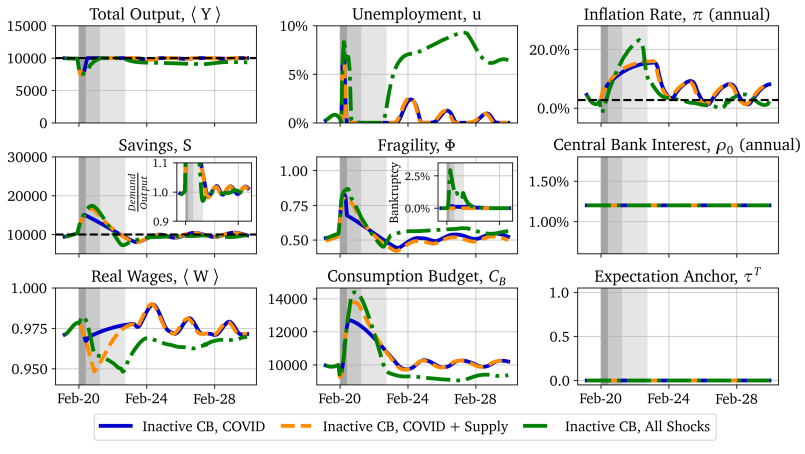


Figure 3.3.: Economic dashboard for the three shocks in the Inactive Central Bank scenario and in the absence of any policies: The dynamics for the three shock scenarios, COVID only (blue), COVID and Supply Chain shock (orange) and all shocks (green) when the Central Bank is inactive and no policy is applied. The areas shaded in grey indicate the duration of the three shocks: the COVID shock lasting until the end of the dark grey area, the supply chain shock until the end of the grey area, and the energy price shock until the end of the light grey area. The dashed black line represents pre-shock averages. In the first two cases, the economy is able to recover on its own (although with significant fluctuations that last for a few years), but significantly gets worse when the energy shock is turned on.

To develop some intuition, we now build three counterfactual cases where (i) only the COVID shock, (ii) COVID + supply chain shocks and

(iii) COVID + supply chain + energy price shocks hit the Mark-0 economy, without any stabilisation policy put in place. It should be noted that this exercise is somewhat optimistic, as our shocks have been calibrated on an economy where emergency policies were actually implemented. This is discussed further below. Moreover, in the Inactive Central Bank scenario discussed in this section, economic agents form their inflation expectations solely on the basis of realized inflation, which means that there is no anchoring of expectations ($\tau^T = 0$, see Eq. (3.23)). All other parameters of this scenario can be found in Table A.1.

1. **Case (i):** After the impact of only the COVID shock, the economy remains operational with full capacity but experiences a persistent and high inflation that peaks at 15.5% and only starts to recede by the end of 2023, with significant fluctuations that persist for a few years (Figure 3.3, blue). All other observables return to their pre-shock levels by the end of 2023, with small oscillations around the steady state in accordance with the Inactive Central Bank scenario without shocks.
2. **Case (ii):** When adding the supply chain shock (Figure 3.3, orange), the dynamics remain very similar: the economy returns to its steady state without fiscal or monetary policy. Remarkably, inflationary dynamics are almost identical, albeit with a slightly higher peak value (16.0%) reached slightly earlier in 2023. The biggest difference is the prolongation of the recovery period, which leads to a steeper real wage dip.
3. **Case (iii):** Taking now the energy price shock into consideration, the situation abruptly changes. Initially, we observe a short deflation period due to the steep drop in energy prices, increasing real wages, supply-demand imbalance and the fragility of firms due to the decreased prices.[20] Immediately following this, we observe an explosion in inflation peaking at 23.2% mid-2023. Firms fragility increases due to the price shock which causes bankruptcies (peaking at 3 % rate), unemployment and a decrease of wages. In the long run, this disruption is so strong that wages recover only very slow. This starts a feedback cycle of low demand that leads to a decrease of output, which causes a reduction in savings and therefore demand drops. This feedback cycle consistently increases unemployment, and decreases savings and wages, which takes the economy significantly longer than 10 years to recover fully with a high and persistent unemployment between 6%-8%.

20: We consider here a direct price-shock transmission. In practice, energy is purchased in various forms and often with inventories and financial contracts to insure against price volatility. In this respect, it might make more sense to define the effective price of energy $p_e(t)$ in the production process as an exponential moving average of the WTI spot price. This would smooth out the initial dip and lead to more realistic inflation time series. We leave this for a later study.

In the realm of the Mark-0 model, the COVID and supply chain shocks would thus have had a minor long term impact on output, but would have led to substantial medium term inflation due to excess savings, as predicted using the Mark-0 model as early as June 2020 in Sharma et al. (2020) [117]. The energy price shock, however, is the last straw on the camel's back and, in the absence of monetary and fiscal policies, has a strongly detrimental long-term impact on the economy. This is a consequence of the existence of discontinuous transition boundaries (a.k.a. *tipping points*) in parameter space, as emphasized in Gualdi et al. (2015) [99] and, within the specific context of the COVID shock, in Sharma et al. (2020) [117]. As a scenario-generating tool for policymak-

ers, our model demonstrates that in the absence of mitigating policies, the full sequence of COVID, supply chain and energy shocks can trigger a negative feedback cycle, manifested as a downward wage spiral, that results in a collapse of demand and a full blown crisis. However, we see that the first two shocks, as modelled above, are of sufficiently mild amplitude not to trigger a complete collapse of the economy. In the following section, we study counterfactuals with larger "bare" (unmitigated) shock amplitudes, where mitigating policy measures are indeed needed to prevent a catastrophic collapse of output beyond a certain tipping amplitude of the initial COVID shock alone.

Sensitivity to Shock Magnitude and the Role of Easy-Credit Policy

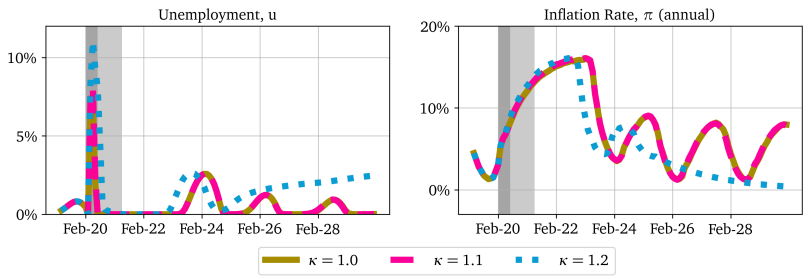


Figure 3.4.: Counterfactual COVID shock in the Inactive Central Bank scenario and in the absence of any policies: Unemployment (left) and inflation (right) in the Inactive Central Bank scenario without any policy, with an amplified COVID shock by a factor κ and Supply Chain shock. $\kappa = 1$ (golden line) corresponds to the COVID shock considered in the previous section, with a very low level of unemployment. With $\kappa = 1.1$ (pink line), the unemployment and inflation dynamics change only very slightly, as the shock is mild. As soon as $\kappa \geq 1.2$, the economy is unable to recover spontaneously, and output collapses as demonstrated in the $\kappa = 1.2$ case (blue line).

Our model is based on data that already incorporates policy measures, which were indeed in place in 2020-21 to mitigate shocks. As such, it is possible that the unmitigated shocks were actually more severe than what we observe retrospectively in data. Here, we investigate counterfactual (ii)-scenarios with different shock magnitudes, where we shut down the final energy shock but scale the amplitude of the observed COVID shock by a factor κ (see Figure 3.2, left panel dark grey area). This stronger shock could have taken place if for example the CARES act and the CAA where not authorized.[21] As κ increases, firms go progressively bankrupt due to their high financial fragility Φ, leading to a higher unemployment rate during the shock. Households save money during the shock, resulting in higher demand, leading to upward pressure on prices and somewhat higher inflation. Therefore in the long run, the scenarios with intensified shocks $\kappa \lesssim 1.2$ very slightly increases inflation and long-term unemployment. However, for larger κ, a tipping point is reached, as in Sharma et al. (2020) [117]: more initial bankruptcies and firms' hesitation to increase wages after the shock leads to a collapse of output and deflation, see Figure 3.4 (for a full dashboard see Figure A.1 in Appendix A.4). Eventually, the shock is too strong for the economy to recover, and the negative feedback loop of a downward wage spiral causes the economy to collapse.

21: The supply chain shock is left unchanged in the present exercise, but we observe almost the same phenomenon without such additional shock.

The default threshold of firms Θ, holds a pivotal role in the economic downturn; if chosen too low an excessive number of firm bankruptcies ensues, which can be detrimental to the overall economy. In this

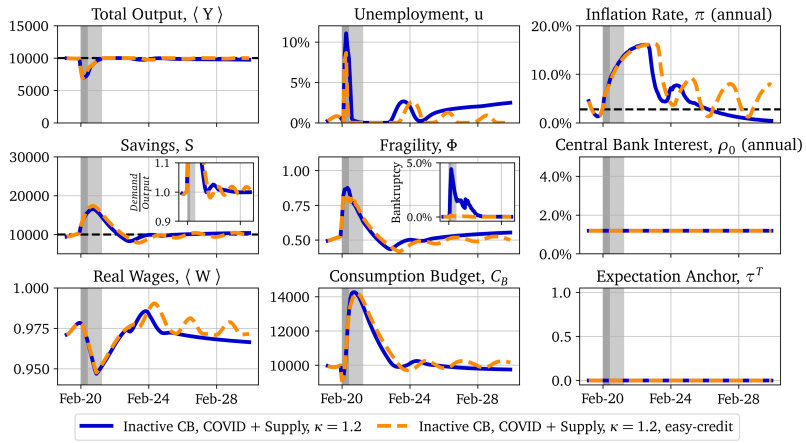


Figure 3.5.: Economic dashboard for a stronger COVID shock in the Inactive Central Bank scenario, with and without Easy-Credit policy: The dynamics of the Inactive CB scenario in case (ii) (COVID and supply chain shock but no energy price shock), for a COVID shock strength of $\kappa = 1.2$ without Easy-Credit policy (blue) and $\kappa = 1.2$ with Easy-Credit policy, $\mu = 1.3$ (orange). We see that policy is effective in preventing the economic collapse in the case of the stronger COVID shock.

case, allowing firms to accumulate more debt by easing bankruptcy rules is highly effective to keep the economy on an even keel. In fact, as shown in Figure 3.5, the "Easy-Credit" policy defined by Eq. (3.28) with $\mu = 1.3$ for the bankruptcy threshold Θ manages to substantially reduce the impact of strong COVID-related shocks, even with an increased initial COVID shock ($\kappa = 1.2$) that would collapse the economy on its own. With Easy-Credit policy, firms can maintain wages and do not need to fire employees to remain solvent, allowing the economy to recover. However, unemployment still peaks above 8% (with a second hump mid 2024) and inflation reaches 16.3% at the end of 2022. As we will discuss in section 3.7 below, the main issue with such a period of high inflation is the risk of de-anchoring inflation expectations, opening the path to possible hyper-inflation.

In order to control inflation, one needs to consider the effect of monetary policy. We thus now turn to the study of the same sequence of three non-amplified shocks, with Easy-Credit policy and with a fully active Central Bank, distinguishing between the case of Anchored Expectation and the case of Floating Expectation in its ability to curb inflation.

3.5. Monetary Policy Response to Inflationary Shocks

In the previous section, we showed that the loosening of regulatory bankruptcy policy during severe shocks prevents the Mark-0 economy from collapsing, but does not address the issue of high inflation rates (Section 18). To study inflation mitigation policies, we now introduce a Taylor-rule based monetary policy into the mix, i.e. we combine the situation explored in Section 3.4 (all shocks and Easy-Credit policy without Central Bank) with the different monetary policy scenarios discussed in Section 3.3. All numerical experiments henceforth are run with the Easy-Credit policy described in Eq. (3.28) with $\mu = 1.3$, as in Section 3.4. As the economy returns to its steady state, the Easy-Credit policy becomes equivalent to the fixed default threshold, such that its duration depends on the recovery period.[22] Note that in what follows, we consider the three shocks together, keeping the amplitude of the COVID shock to the one calibrated on the observed data (i.e. $\kappa = 1$

22: Eq. (3.28) indeed implies $\Theta(t) = \Theta_0$ in the steady state when the average fragility is low, which is the case for the choice of parameters made in this chapter.

in the language of Section 3.4). As already discussed, such calibration may underestimate the severity of the unmitigated shock, since the US government immediately implemented Easy-Credit and Helicopter Money measures to alleviate the COVID shock. Note also that unemployment is not an explicit target of our Central Bank, in the sense that the Taylor-rule is only responsive to realized inflation and not to unemployment.

An Inactive Central Bank

We begin our study by considering again the situation with an Inactive Central Bank as in Section 3.3, but now with all shocks together and with the Easy-Credit policy being implemented (Figure 3.6, solid blue line; see Figure A.5 in Appendix A.5 for all macroeconomic time-series). In this case, over the first few months, the combination of shocks causes a spike in unemployment that peaks around 8% during a deflationary phase, but the economy quickly recovers to full employment, yet at the price of a sustained inflationary period that peaks above 20% before slowly reverting to equilibrium, consistently with the discussion of Section 3.4: the Easy-Credit policy can avoid full collapse, but only at the price of high inflation.

Monetary Policy without Anchoring

Before introducing trust dynamics, we consider the case where agents form their inflation expectations without considering the Central Bank's target inflation, but the Central Bank is nonetheless actively pursuing monetary policy. We thus set $\tau^T = 0$ such that inflation expectations $\hat{\pi}(t)$ are simply an exponentially weighted moving average over realized inflation. In this case, we find that the Central Bank initially decreases rates before rapidly increases interest rates as inflation increases (see Appendix A.6 for the corresponding dashboard). However, there is very little resulting effect on inflation, which displays roughly the same dynamics as when the Central Bank is inactive. In other words, the direct impact of higher interest rates on inflation is weak, and expectation anchoring will be dominant in this respect.

However, the de-Anchored Expectations lead to a doubling in magnitude of the initial spike in the unemployment rate, together with a higher (and oscillating) peak following the end of the shocks, as compared to the case where monetary policy is inactive. This is due to the spike in debt induced by bankruptcies accompanying the initial wave of shocks.

We can thus draw a first conclusion that in the absence of any sort of expectation anchoring the main impact of monetary policy is detrimental to the unemployment rate, which is in fact not within the Central Bank's mandate in our model setup.

Monetary Policy with Anchored Expectation

We next consider a monetary policy experiment with a responsive Central Bank in the "Anchored Expectation" scenario. In the Mark-0 model, as in real life, the Central Bank's aim is to keep inflation on

target. Here, following Section 3.3, the inflation target is chosen to be $\pi^\star = 2.4\%$ p.a. and the Taylor rule strength is $\phi_\pi = 1$. Moreover, we optimistically assume that the Central Bank has successfully communicated to all economic actors that it has inflation under control, thereby convincing them to believe that long-term inflation will be close to the Central Bank's inflation target, with a fixed anchor weight $\tau^T = 0.95$.

In the case of Anchored Expectation and variable interest rates with all three shocks (Figure 3.6, dashed orange line), the Central Bank is successful in taming inflation: peak inflation is lower than in the case without an active central bank (peak at 8.8% vs 23.2% in the Inactive Central Bank scenario) and quickly tapers off. Although the inflation rate is not in line with the Central Bank's target throughout the crisis, it appears that monetary policy is able to reign in inflation. The reduction in inflation is not primarily due to the impact of interest rate policy but rather to the strong anchoring of expectations, which significantly dampens expected (and thus realized) inflation, as upward pressure on both prices and wages are reduced. Such anchoring moderates the price increases after the initial COVID shock, but does not dampen the deflationary dynamics towards the end of the energy price shock, as this is driven by a strongly increasing unemployment that puts downward pressure on wages (see Figure A.5 in Appendix A.5 for all macroeconomic timeseries).

The initial price shock in the first few months causes bankruptcies that are less numerous with a reactive Central Bank, hence also reducing the initial unemployment spike slightly (peak at 8.1% in the Inactive Central Bank scenario and 7.4% with Active Central Bank with Anchored Expectation). However, at the end of the shock the Central Bank effort to control inflation results, within an Anchored Expectation scenario, in a peak unemployment of 12.8 % in Feb 2024, 5% higher than in the Inactive Central Bank scenario. Despite realized inflation surging during shocks, firms inflation expectations stay anchored to the central bank's target. This results in a substantial decline in real wages during the shock and a sluggish recovery thereafter. Furthermore, firms' wariness of the cost of debt is increased (through parameter Γ, Eq. (3.10)), leading to an increase of the firing rate. Similarly, households expect a controlled level of inflation and hence reduce consumption in the face of high interest rates that favour savings.

All these effects lead to higher unemployment, impeding the overall post-shock recovery. We thus see that in the Mark-0 model, the Central Bank is stuck between a rock and a hard place with its inflation-unemployment trade-off, meaning that in the context of these shocks, monetary policy is not a panacea: some form of Keynesian stimulus or "Helicopter Money" is needed in conjunction with monetary policy to restore the economy to its pre-crisis steady-state – see section 3.6 below. In this regard, a Central Bank with an explicit dual-mandate would potentially navigate the trade-off better.

Monetary Policy with Floating Expectation

The scenario of high and Anchored Expectation in monetary policy for long-term inflation expectations might be a good approximation

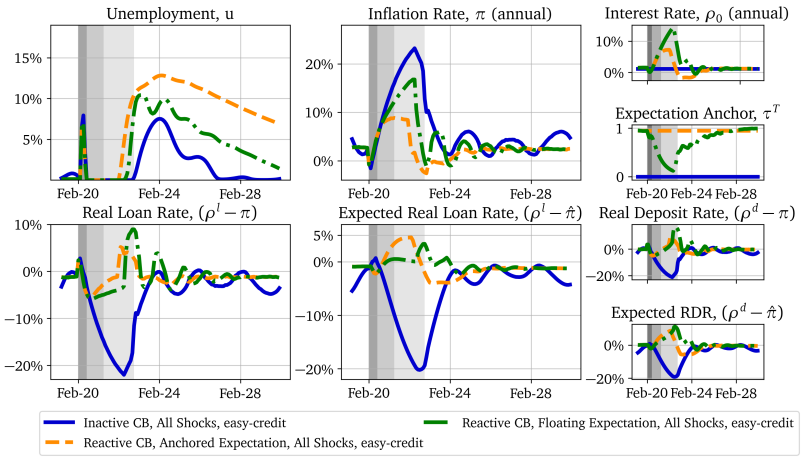


Figure 3.6.: Unemployment and Inflation with all shocks, Easy-Credit policy, and three distinct monetary policy scenarios. (Blue lines) Inactive Central Bank scenario. (Orange lines) Reactive Central Bank with Anchored Expectation, with anchor parameter $\tau^T = 0.95$. A Taylor rule policy successfully decreases peak inflation but increases peak unemployment. (Green lines) Reactive Central Bank with Floating Expectation. Here, trust is eroded during the high inflation period. Monetary policy then fails at reducing inflation, with the risk of hyper-inflation lurking (see section 3.7); unemployment remains lower because of the higher inflation.

during long periods of stability, but in periods of shocks, the anchoring τ^T may decrease when inflation deviates strongly from the Central Bank's target [126]. We model such a Floating Expectation effect through Eq. (3.27), where the anchoring parameter decreases when the spread between observed and target inflation increases. We set $\alpha_I = 0.4$ and a memory time of ~ 5 months ($\omega = 0.2$). These values imply that when inflation reaches 4 times the Central Bank's target $\pi^\star$, trust in the Central Bank falls to approximately a third of its initial value after approximately one year. Larger values of α_I would lead to an even steeper de-anchoring of inflation expectations. Such a loss of faith in monetary authorities further increases realized inflation as economic agents expect higher inflation in the future and take this into account when setting prices and wages. This may lead to a self-fulfilling feedback loop between expected inflation and real inflation through a wage-price spiral (see section 3.7).

Compared to the Anchored Expectation case, the surge of inflation after all COVID shocks leads to a loss in trust that almost vanishes (i.e. $\tau^T \to 0$) during the energy price shock. The consequence of this is a higher realized inflation rate, that reaches a peak value only slightly below the Inactive Central Bank scenario (Figure 3.6, right).[23] Unemployment, on the other hand, turns out to be *higher* than with an Inactive Central Bank (see Figure 3.6, left), but *lower* than when trust is well anchored. As noted above, low expected inflation is detrimental to economic activity when interest rates are high. Therefore, perhaps paradoxically, loss of trust improves the situation in terms of unemployment, but is of course detrimental to inflation, with the lurking possibility of runaway situations. This again illustrates the quandary faced by Central Banks, and the importance of endowing them with a double mandate.

A strong factor of the loss in trust is the exogenous nature of the energy price movements. Despite the strong hike up to a real interest rate of 3.4% per year (16.8% nominal target rate), inflation remains far above target. Once these shocks have passed, the Central Bank is arguably quite successful in keeping inflation closer to its steady state value. However, similarly to the Anchored Expectation case, the economy now contends with a lower, but persistent unemployment problem. The unemployment here is driven by the energy price shock that

23: For the full economic dashboard in the case of dynamic expectations, see Appendix A.5, Fig. A.4, and compare with Fig. A.3.

causes a strong surge in inflation, which erodes household savings and reduces real wages by 4% compared to its steady state value, thus resulting in a drop in demand (see Figure A.4 in Appendix A.5). In contrast to the Anchored Expectation scenario, unemployment in the Floating Expectation scenario recovers to full employment faster. As noted above, this paradoxical result is driven by the comparatively higher real wages and consumption propensity, which allows for a faster recovery of the consumption budget and consequently output.

As with both the Inactive Central Bank and Anchored Expectation scenario, strong exogenous inflation outside of the Central Bank's sphere of influence leads to a severe contraction of the economy, which recovers only slowly. The convergence to pre-crisis equilibrium takes more than 18 years in the Anchored Expectation scenario and 11 years in the Floating Expectation scenario. This suggests that in the face of exogenous price shocks, fiscal policy shoring up consumer's budget or reducing energy price effects on firms and households can be a more effective tool than monetary policy.

Sensitivity to the Strength of the Monetary Policy

In circumstances of crises, as with our shocks, Central Banks should possibly react more strongly to inflation and keep expectations anchored. We thus consider a larger Central Bank Taylor reactivity parameter $\phi_\pi = 2.0$ (see Appendix A.7, Figure A.7 and A.8 for details).

In the case of Anchored Expectation, an increase in ϕ_π hardly changes inflation, but stronger interest rate hikes result in significantly higher unemployment rates. Hence, stronger monetary policy is clearly detrimental in this case. By contrast, in the Floating Expectation scenario with dynamic trust, an increase in ϕ_π signals a stronger commitment of the Central Bank to control inflation. Therefore, all else being the same, inflation expectations are lower which, in a self-fulfilling manner, cuts realized inflation from 16.8% to 11.2% and keeps trust better anchored. However, as in the Anchored Expectation case, the sharper interest rate hikes by the Central Bank lead to a significant increase in the unemployment rate (from 12.8% to 20.5%) and a much longer recovery time for the economy as a whole. Again, our model confirms and quantifies the observation made by Stiglitz and Regmi (2022) [104], quoted in the introduction: monetary policy interest tools tend to be too blunt, curbing inflation at the cost of unnecessarily high unemployment.

Sensitivity to Transmission Channels

Within our modelling framework, the impact of monetary policy relies on the efficiency of three transmission channels: (a) expectation anchoring, discussed above; (b) sensitivity of consumption on real interest rate, through parameter α_c; (c) sensitivity of firms' hiring and wage policies on real interest rate, through parameter α_Γ. We have run some simulations to check the dependence of the state of the economy on these last two parameters in the Anchored Expectation case, see Appendix A.8. The conclusion of this exercise is that increasing

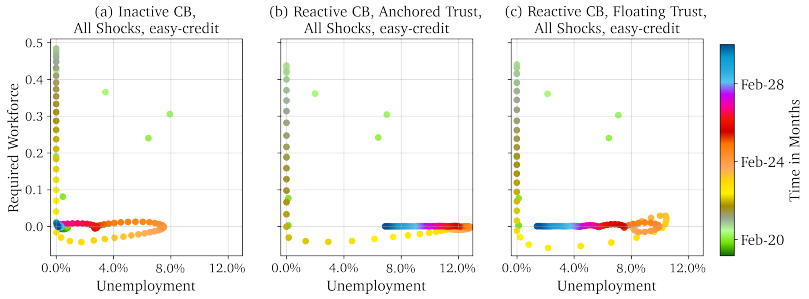


Figure 3.7.: The Beveridge Curve and Labor Market Adjustments Across Policy Scenarios: Relationship between unemployment and the required workforce Eq.3.29 across different monetary policy regimes and all shocks. The three panels correspond to: (a) Inactive CB, (b) Reactive CB with anchored trust, and (c) Reactive CB with floating trust. The color gradient represents time progression. Across all scenarios, the model generates a counterclockwise rotation of the Beveridge Curve.

α_c and α_Γ has a minor direct effect on inflation but a significant effect on unemployment, depending on the sign of the real interest rate $\rho - \hat{\pi}$. As expected, a raise of interest rates degrades economic activity, all the more so when the sensitivity of firms and households to the real rate is higher. As already stated, the main transmission channel through which monetary policy impacts inflation is expectation anchoring, rather than directly through cost of loans or income on savings.

The Beveridge Curve and Labor Market Adjustments Post-Shock

The Beveridge Curve describes the inverse relationship between job vacancies and unemployment [139]. Typically, in an expansionary phase, high job vacancies correspond to low unemployment, while in downturns, vacancies decline as unemployment rises. A key empirical feature of labor market adjustments following economic shocks is the counterclockwise rotation of the Beveridge Curve: unemployment rises sharply at the onset of a crisis, while vacancies remain low. As the economy recovers, vacancies increase before unemployment begins to decline, creating a distinct lag in the adjustment process [140].

Despite not explicitly modelling vacancy rates, our Mark0 model reproduces the Beveridge Curve and its counterclockwise rotation during the shocks. The model does not include vacancy rate directly, therefore we approximate the vacancy rate as required workforce using the demand-output imbalance, defined as:

$$V_t = \frac{\sum_i (D_i(t) - Y_i(t))}{Y_{\max}} \tag{3.29}$$

where $\sum_i D_i(t)$ denotes aggregate demand, $\sum_i Y_i(t)$ represents production and $Y_{\max}$ corresponds to the economy's output under full employment. Since production in our model is directly linked to employment (see Eq. 3.5), this serves as a reasonable proxy for labor market tightness. Figure 3.7 shows the Beveridge curve for all shocks and the three scenarios of (a) Inactive CB, (b) Reactive CB, anchored trust and (c) Reactive CB, floating trust. Across all shocks and all scenarios analyzed, the model produces the expected negative relationship between vacancies and unemployment and moreover the counterclockwise rotation of the Beveridge Curve emerges endogenously. At the onset of 2020, the COVID shock causes a sharp increase in required workforce while unemployment remains stable. This is followed by a fast recov-

ery, but subsequently, unemployment begins to rise. As the shocks subside, unemployment gradually declines again, accompanied by a slight increase in workforce demand.

This finding reinforces the robustness of our model in capturing key macroeconomic mechanisms. The counterclockwise rotation highlights a delayed labor market recovery, consistent with real-world data [141], where firms hiring is slow even after demand conditions improve. This contributes to persistent unemployment and sluggish post-shock recoveries, further validating the policy trade-offs. Our results suggest that monetary and fiscal policy must account for these labor market frictions to avoid prolonged economic downturns.

3.6. Fiscal Stimulus

Section 3.5 showed that in the presence of exogenous inflation drivers, the Central Bank can reduce inflation only at the cost of high unemployment. The mechanisms responsible for unemployment are the increased burden of debt that weighs on wages and increases the speed at which firms lay off workers. This in turn lowers demand, an effect amplified by the erosion of household's purchasing power by the lingering inflation. One way to tackle this issue is to ensure that demand does not drop as severely by increasing it through fiscal stimulus (see Section 16). We test this policy device here by considering different kinds of fiscal stimuli. Unless otherwise indicated, the fiscal stimuli are applied on the scenario of Section 22, i.e. with all shocks, Easy-Credit policy, and an Active Central Bank with Floating Expectation, which we believe to be the scenario closest to reality.

The Effects of Helicopter Money

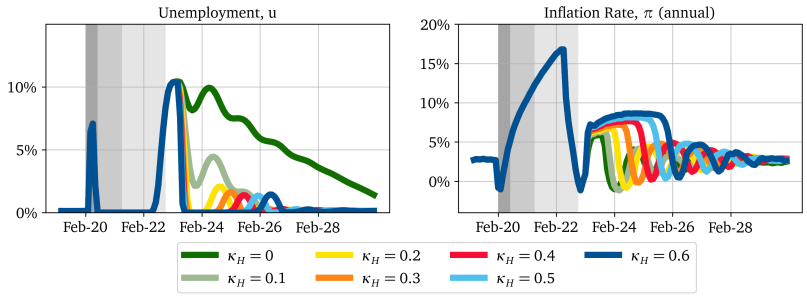


Figure 3.8.: Helicopter Money in the Reactive Central Bank with Floating Expectation scenario, Easy-Credit policy, and all shocks. Unemployment (left) and inflation (right) for a helicopter drop of size κ_H times savings one month after the price shock. Already with $\kappa_H \geq 0.2$ unemployment is reduced almost to zero. A further increase of Helicopter Money only increases the duration of the high-inflation period, without further reducing unemployment. Note the small unemployment and inflation "ripples" persisting several years after the initial shock.

Following Sharma et al. (2020) [117], we begin by considering a one-time increase of the households' budget by a factor $\kappa_H \in [0\%, 60\%]$ one months after the end of the energy price shock (July 2023). The goal is to find an optimal multiplier κ_H that reduces unemployment while keeping excess inflation to a minimum in amplitude and duration. Here we find that this minimum is reached for $\kappa_H \approx 20\%$ (see Figure 3.8 for the Reactive Central Bank with Floating Expectation case).[24] This choice of κ_H matches the magnitude of the "Emergency Money for the People Act", which entailed the US government pro-

24: See Appendix A.9, Figures A.11, A.12 for the other two monetary policy scenarios, and Figures A.13, A.14 and A.15 for the complete economic dashboards of all scenarios.

viding direct payments of 2000 USD per month for a maximum of 12 months to support individuals during the COVID pandemic.

In all considered scenarios, the fiscal stimulus package leads to a quick recovery of production to its pre-shock levels and unemployment at near 0%, thus successfully eliminating the steep recession and long recovery following the energy price shock, as shown in Figure 3.8. However, policymakers face an inflation-unemployment trade-off, as the stimulus generates a spurt of inflation. In the absence of monetary authority, inflation due to the injection of money rises to 15.8% (see Appendix A.9, Figure A.11). However, this is an endogenous inflation, which means that in the Floating Expectation case, the Central Bank can raise rates to curb consumption propensity and keep inflation below 10% p.a. (Figure 3.8). Increasing the stimulus leads to a higher inflation peak and a longer duration of high inflation.

The inflationary period after the stimulus is in part due to a short-term de-anchoring of expectations as the demand stimulus leads to higher prices in the context of a tight labour market preventing production increases. Yet, here one can take the position that the economy is in good shape from a macroeconomic perspective, with near-zero unemployment despite an above-target inflation.[25] In this case, the Anchored Expectation case would be much more favorable as the high unemployment problem is resolved with only a minimal rise in inflation, around 7% annualized over 8 months, as soon as $\kappa_H \gtrsim 0.2$.

25: Because Mark-0 treats households at the aggregate level, we do not analyze the distributional consequences of inflation here, though inflation is always and everywhere differential in its effects.

The Effects of a Windfall Tax

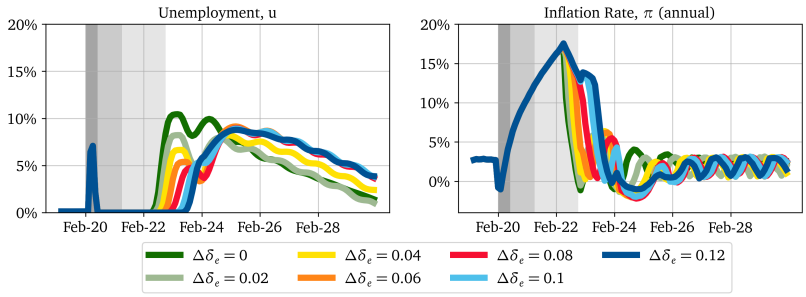


Figure 3.9.: Windfall Tax in the Reactive Central Bank with Floating Expectation scenario, Easy-Credit policy, and all shocks. Unemployment (left) and inflation (right) for a Windfall Tax of $\delta_e + \Delta\delta_e$ one year before the end of the price shock with a duration of two years. With $\Delta\delta_e \approx 4\%$ unemployment is reduced strongly. A further increase of tax does only increase unemployment again. For even larger $\Delta\delta_e$, inflation increases because of increased demand due to increased savings.

Another stimulus that central authorities can use to alleviate the impact of the price shock is a Windfall Tax. This policy redistributes the excess profits generated by the energy sector as a result of rising energy prices to household savings with the aim to stimulate consumption and consequently reduce unemployment. We test the effectiveness of the policy, which in our modelling consists in an increase of the energy sector payout rate $\delta_e \to \delta_e + \Delta\delta_e$ one year before the end of the price shock with a duration of two years, with $\Delta\delta_e \in [0\%, 12\%]$.[26]

Our aim is to design the fiscal stimulus in a manner that minimizes both unemployment and inflation. We find that for $\Delta\delta_e = 2\% - 4\%$ the excess profits of the energy sector are redistributed such that unemployment is significantly reduced at all times while inflation remains under control. Further increasing $\Delta\delta_e$ reduces the unemployment in 2023, but at the cost of increasing unemployment in 2024 after the

26: We remind here that δ_e is the fraction of the energy sector cash balance redistributed to households at each time step, see Eq. (3.20).

Windfall Tax has ended (see Figure 3.9).[27] As long as the tax is active, the increase in δ_e results in higher savings, leading to an increased consumption budget and therefore higher demand. This has the effect of reducing unemployment as firms expand production, as well as reducing firm fragility due to increased profitability. However, as the Windfall Tax increases in the Floating Trust scenario, inflation also increases and expectations remain de-anchored for longer, compelling the Central Bank to increase interest rates. As the tax ends, the household's consumption budget has been reduced to below the case without a Windfall Tax, and the real interest rate remains high. This leads to a negative spiral of lower demand and output, higher unemployment, leading to lower income, until the system equilibrates again.

27: For more detail, see Appendix A.10, Figures A.16, A.17.

The inflationary effects of the increase in savings due to the Windfall Tax remain minor for small values of $\Delta\delta_e$. However, once the Windfall Tax exceeds $\Delta\delta_e \sim 10\%$, a longer period of high inflation after the shock is observed (see Figure 3.9 and Appendix A.10, Figure A.20) as inflation expectations remain de-anchored for a longer period.

The Anchored Expectation scenario presents a different challenge. Again, up to a threshold, the Windfall Tax reduces the initial and long-term unemployment spike, with $\Delta\delta_e = 4\%$ still remaining a robust choice (see Appendix A.10, Figures A.17, A.19). However, this time at the cost of inflation (or less deflation) on the way back to the pre-crisis steady state. In this regard, policymakers might opt instead to minimize the volatility of inflation to let it smoothly return to the Central Bank's target rate, which actually suggests a higher Windfall Tax of $\Delta\delta_e = 8\%$ for instance.

We conclude that when implementing a Windfall Tax as a fiscal stimulus, excessive tax levels can have the unintended consequence of exacerbating long-term unemployment, while a measured approach can both reduce unemployment and smooth inflation dynamics.

3.7. Model Sensitivity: The Dangers of a Wage-Price Spiral

"Sloppiness" Analysis

To develop a deeper understanding of the reaction to shocks of our model economy, we explore the parameter sensitivity of this configuration of the Mark-0 model using the "sloppy model" methodology put forth by Sethna and collaborators (see Transtrum et al. (2015) [142] for an introduction) and applied to macroeconomic models by Naumann-Woleske et al. (2024) [2]. This method identifies the *stiffest* parameter combinations, along which one observes the largest change in outcomes for small perturbations. In particular, we consider the parameter sensitivity of the dynamics of the inflation rate and of the unemployment rate (for methodological details, see Appendix A.11 and Naumann-Woleske et al. (2024) [2]). Overall, this approach provides a more comprehensive understanding of parameter sensitivity by considering the curvature of the model manifold, as compared to one-at-a-time parameter sensitivity analysis.

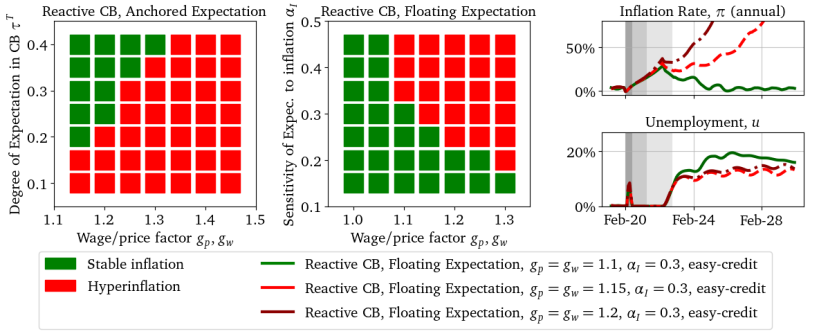


Figure 3.10.: Hyperinflation tipping points. Left: Stable inflation vs. Hyperinflation in the plane ($g_p = g_w$, τ^T) in the Anchored Expectation scenario and in absence of exogenous shocks. Note that inflation can be stable even when $g_p = g_w > 1$ for strong enough anchoring. Middle: Stable inflation vs. Hyperinflation in the plane ($g_p = g_w$, α_I) in the Floating Expectation case and in absence of exogenous shocks. Right: inflation rate (top) and unemployment (bottom) as a function of time, in the Floating Expectation case, $\alpha_I = 0.3$ and different values of $g_p = g_w$, after the three COVID shocks. When indexation is too strong, an hyperinflation regime sets in at the end of the energy price shock.

With respect to both inflation and unemployment, the balance of wage indexation (bargaining power) g_w to price indexation (market power) g_p, has by far the strongest influence on the outcome dynamics for the inactive CB and Floating Expectation scenarios (see detailed results in Appendix A.11 and A.12). Beyond inflation itself, eigenvectors are combinations of various pricing parameters including γ, g_p and g_e for both unemployment and inflation. Central Bank parameters are present but not significantly. This aligns with our findings that the Mark-0 economy always maintains a tight connection between unemployment and inflation, indicating that it may not be possible to minimize one without affecting the other (see Appendix A.11, Figures A.23, A.24).

Gualdi et al. (2015) [99] have demonstrated the significance of the default threshold of firms Θ in determining the economic phase of the Mark-0 model. This parameter exhibits a non-linear behavior, where minimal alterations have negligible effects on the overall dynamics until a critical "tipping point" is reached. Beyond this threshold, the dynamics undergo substantial transformations. It is important to note that with the "sloppiness" analysis, we can only identify local sensitivity in parameter combinations, where slight deviations from these combinations result in noticeable changes in observed dynamics. As already found in Naumann-Woleske et al. (2024) [2], multiple phase transitions occur along the parameter axis of Θ, although minor perturbations do not alter the curvature of the loss function, therefore we do not find Θ to be crucial in our results. However it is crucial to distinguish between local sensitivity, which can be determined by the sloppiness approach, and global sensitivity to parameters, which can trigger tipping points in the model Gualdi et al. (2015) [99].

The Risk of a Hyperinflation Spiral

We can interpret the sensitivity to g_w and g_p in light of the dangers of a hyperinflation episode resulting from a wage-price spiral. Alvarez et al. (2022) [143] cite a concern that hyperinflation may occur if firms increase wages in response to higher inflation, leading to an increase in purchasing power and ultimately feeding into a wage-price spiral in the current macroeconomic environment. This feedback loop is influenced by the indexation of prices and wages to firms' inflation expectations [144], i.e. by the value of parameters g_p and g_w in our model. In

the simplest case where bargaining power and market power are equal ($g_p = g_w$), the economy reaches a stable inflationary state, marked by cyclical fluctuations due to mismatches in supply and demand (see Figure 3.1 in Section 3.1, for which $g_w = g_p = 0.8$). When indexation is weak ($g_p = g_w < 1$), neither wages nor prices fully incorporate inflation expectations, and fluctuations in the inflation rate due to mismatches in demand and supply are dampened. Conversely, for strong indexation ($g_p = g_w > 1$), the economy may enter a state of hyperinflation, though as wages and prices increase equally fast, there are no real effects (omitting, of course, the cost of inflation itself due to "menu costs"). Only when $g_p \neq g_w$ does the economy collapse.

Introducing monetary policy implies that hyperinflation can be staved off due to the anchoring of inflation expectations. In the case of Anchored Expectation, inflation remains stable until a critical point $g_p = g_w < g^\star$ where $g^\star > 1$ depends on the commitment of the Central Bank as well as the strength of expectation anchoring to the its target (see Figure 3.10, left panel). The same holds true with Floating Expectation (middle panel). However, in this case an interesting range of parameters where inflation remains stable for $g^\star > g_w = g_p > 1$, until a strong enough shock occurs, which triggers a loss of trust in the Central Bank, tipping the economy into a hyperinflation phase (as illustrated in Figure 3.10, right panel). This is precisely the scenario that Central Bankers want to avoid by doing "whatever it takes".

If bargaining and market power differ ($g_p \neq g_w$), the effect on unemployment is significant (see Appendix A.12, Figure A.26). In the Anchored Expectation case, when market power is greater than bargaining power – a.k.a. "greedflation" – ($g_p > g_w$), unemployment increases after the shock and only slowly decreases as wages adjust at a slower pace to pre-crisis levels, leading to sluggish demand recovery. Consequently, firms do not see the need to increase their production faster. In contrast, if bargaining power exceeds market power ($g_w > g_p$), wages recover more quickly, resulting in an earlier recovery of pre-crisis purchasing power, which helps to reduce unemployment. This conclusion is in line with the result of section 3.6 on the impact of Helicopter Money or, more generally, of Keynesian stimulus: increasing the consumption budget of households is quite efficient at reviving an ailing economy, but only if hyperinflation can be avoided.

3.8. Conclusions

In this chapter, we have expanded the Mark-0 Agent-based Model to assess the inflationary dynamics following the COVID pandemic and energy crisis of 2020-2023. Our results highlight the narrow path of monetary policy to balance unemployment and inflation, as well as the benefits of joint monetary and fiscal policy packages.

We extended the Mark-0 model used in Sharma et al. (2020) [117] in two directions (Section 3.1): First, a dynamically evolving trust in the Central Bank's ability to control inflation, such that expected inflation remains anchored to the Central Bank's target if trust is high, but persistent off-target inflation realisations lead to a loss of trust in the Central Bank. Second, a simple "exogenous" energy sector from which

firms buy energy that allows us to introduce an energy price, and slowly re-inject the energy profits into the economy as dividend payouts. This slightly expanded Mark-0 model allows us to investigate several policy channels (Section 3.2): (a) monetary policy via the management of interest rates and expectations, (b) an Easy-Credit regulatory policy, and fiscal policies such as (c) Helicopter Money and (d) a Windfall Tax.

The Mark-0 steady state economy is then perturbed by three calibrated shocks (Section 3.4): a consumption propensity drop (representing lockdowns), a productivity drop (representing supply chains disruptions), and an energy price shock (first a drop, representing reduced demand during lockdowns, followed by a rise due do demand recovery exacerbated by the Russian invasion of Ukraine). We show in Section 3.4 that the Easy-Credit policy, which was one of the first emergency responses to the COVID pandemic, is able to alleviate the impact of the shocks, but at the price of high and sustained inflation, as observed in the data and predicted as early as June 2020 by Sharma et al. (2020) [117] within the Mark-0 framework.

We investigated whether monetary policy alone can control a surge of inflation, and at what cost to unemployment (Section 3.5). Our results show that if agents' expectations are fully de-anchored, i.e. expected inflation coincides with past realized inflation and not to the Central Bank target, monetary policy has very little impact on inflation. If on the contrary expectations remain fully anchored, inflation remains closer to the Central Bank target, but at the price of a strong recession leading to a wave of unemployment. Meanwhile, if agents' expectations evolve dynamically, inflation rises well above target during the shock leading to de-Anchored Expectations. Monetary policy is then again essentially ineffective in controlling inflation while causing unemployment to increase in comparison to a baseline case with no monetary policy at all.

This conclusion is robust to variations in parameters related to the channels of monetary policy. Our framework thus supports the view that the efficacy of monetary policy comes at the cost of unnecessarily high unemployment [104], in particular, quite paradoxically, if inflation expectations are strongly anchored (see the discussion in Section 3.5).

Within Mark-0, trust anchoring is thus a crucial determinant of the success of the Central Bank inflation mitigation policy, far more important than the direct economic impact of higher interest rates (noting that we do not model in detail a financial sector). This resonates with Bernanke's statement: "Expectations matter so much that a Central Bank may be able to help make policy more effective by working to shape those expectations",[28] and points to the importance of *narratives* in shaping expectations [145]. In turn, fine-tuned fiscal policy combined with monetary policy can be successful in controlling inflation while keeping unemployment around acceptable levels (Section 3.6). However, this requires a high degree of precision to be effective: too weak a fiscal stimulus is ineffective, and too large a stimulus leads to further high inflation. Finally, in Section 3.7 we show that the pricing power of firms and the bargaining power of workers play a cru-

28: see Bernanke, B. (2013), *Communication and Monetary Policy*, speech at *National Economists Club Annual Dinner*. (https://www.federalreserve.gov/newsevents/speech/bernanke20131119a.pdf)

cial role. Depending on their relation, the economy can experience a runaway hyperinflation wage-price spiral. This points to an additional difficulty in properly calibrating monetary and fiscal policy due to a variety of possible tipping points or "dark corners" lurking around.

Overall, this study of the Mark-0 economy shows that (i) the economic recovery can be very sluggish due to self-fulfilling expectations and other non-linear feedback loops, (ii) there is a tension between inflation and unemployment that is robustly captured by the model, (iii) in the Mark-0 economy, monetary policy is effective at controlling inflation provided trust is anchored (but not because of real economic effects of higher interest rates), and (iv) fiscal policy can alleviate some of the negative unemployment effects of inflation-focused monetary policy. While our study did not exhaust all possible cases and parameters, we believe that it illustrates both the realism and the richness of the Mark-0 economy. It offers a versatile tool with which policy makers can easily play in order to forge their intuition about what may happen if they turn this knob or add that policy measure. Indeed, almost all possible narratives that emerged during the recent debate about post-COVID inflation can be captured and reproduced by our model with proper choices of parameters and shock specifications. Furthermore, the procedure proposed in Section 3.7 can be used to rigorously establish which predictions of the model are robust to the choice of parameters, most of which are difficult to properly calibrate on data.

Our model is obviously incomplete and improvable on many counts. In particular, it could be extended to include (i) a dual mandate central bank as in Gualdi et al. (2017) [115], Bouchaud et al. (2018) [116] to address negative unemployment side-effects, (ii) a disaggregated household sector (e.g. distinguishing wage and rent earners, and by accumulated wealth) to assess distributional implications of policy, (iii) production networks and energy supply chains with differentiated products and commodities [as in 146, for instance], which introduces firm distribution effects and systemic risk, (iv) a financial sector to more closely model interest rate pass-through, lending choices and financial systemic risk, and (v) more detailed labour market structures to account for the change in occupational structure post-COVID and demographic trends. However, we do believe, as we already argued in Gualdi et al. (2015) [99], that the Mark-0 model should already be part of the toolkit of Central Banks, if only as an inspiring *scenario generator*, or "telescope for the mind", especially in times of great modelling uncertainty during which it is crucial to be at least "roughly right" and avoid being blindsided by spurious Black Swans.[29]

29: See for example Mark Buchanan, "This Economy Does Not Compute", New York Times, October 2008, Gualdi et al. (2015) [99], King and Kay (2020) [119], Bouchaud (2021) [120], or the recent Financial Times piece cited in the introduction, https://www.ft.com/content/b972f5e3-4f03-4986-890d-5443878424ac

Key Takeaways

1. **Post-COVID Recovery Challenges:** The economic recovery after COVID-related shocks, particularly in the absence of fiscal policies, can be extremely slow or even lead to a deep recession, underscoring the necessity for proactive fiscal measures.
2. **Monetary Policy and Inflation:** The effectiveness of monetary policy is largely driven by expectation anchoring rather than direct impacts of interest rate adjustments. However, strong anchoring can negatively affect consumption and employment.
3. **Policy Trade-offs:** There is a narrow window for optimal policy response. Excessively strong monetary policy leads to high unemployment, while overly generous fiscal stimulus fuels inflation, highlighting the importance of balancing interventions.
4. **Model Sensitivity:** Wage and price indexation parameters are the most influential in determining inflation outcomes, emphasizing the critical role of bargaining power and market dynamics.
5. **Implications for Policymakers:** The study provides a scenario-based framework to anticipate and mitigate the unintended consequences of fiscal and monetary policies, aiding policymakers in navigating complex economic environments.

Exploration of the Parameter Space in Macroeconomic Models

4.

This chapter is taken from Naumann-Woleske et al. (2024) [2] *Exploration of the Parameter Space in Macroeconomic Agent-Based Models*, with minor modifications to the figures and text to align with the context of this thesis. Section 2 from the original work has been relocated to Section 2.2 of this thesis. The original publication was a collaborative effort with *K. Naumann-Woleske*, *M. Benzaquen* and *J.-P. Bouchaud* and was later extended by *K. Naumann-Woleske* in his thesis [58]. My contributions include assisting in the initial framing of the method and conducting the analysis of the Mark0 agent-based model. The development of the algorithm and drafting of the manuscript were carried out in collaboration with the co-authors. Section 4.5 on the extension of the method to probabilistic models is entirely new.

To build on the insights from Chapter 3, where we explored the power of Agent-Based Models (ABMs) as computational scenario generators, we now delve into the robustness of the predictions these models generate. Understanding the full spectrum of possible phenomena that an ABM can produce is essential for evaluating the robustness of its predictions. However, due to the high dimensionality of parameter spaces, exploring these possibilities can be computationally expensive. In this chapter, we leverage the sloppiness approach introduced in Section 2.2 to systematically investigate the phase space of various ABMs.

The global economy can be described as a complex adaptive system where large numbers of heterogeneous agents interact in parallel [147, 148]. The evolution of this system gives rise to emergent phenomena such as business cycles, fat tails in company sizes/agent wealth, or even crises and crashes. Agent-based models (ABMs) provide tools to generate and understand these emergent phenomena (for deeper introductions to agent-based modelling see Haldane and Turrell (2019) [16], Dosi and Roventini (2019) [22], LeBaron and Tesfatsion (2008) [148], Epstein (1999) [149], Kirman and Gallegati (2022) [150], Arthur (2022) [151]). Indeed, an agent-based model is a computational device for simulating artificial economies of many heterogeneous interacting agents. The richness of phenomena that an ABM can generate makes them excellent *scenario-generators* [152–154] that can be used to understand the spectrum of possible paths for the economy, and how these future scenarios may respond to policy interventions. Unfortunately, the versatility of ABMs means that they suffer from the *wilderness of bounded rationality* [155] both due to the large number of parameters and the different choices of behavioral rules, and have been described as *black boxes* able to generate any phenomena (see Fagiolo and Roventini (2017) [156] for a discussion). Consequently, it is imperative to understand the full set of possible generated phenomena how robust

they are to changes in parameters.

Our research provides a first answer to the question of *what are the different types of dynamics that an agent-based model can generate?* In the present chapter, we first establish which parameter combinations are most influential for the observed dynamics. Interestingly, we find that across models, only few parameter combinations matter, as already alluded to in Gualdi et al. (2015) [152]. Using this feature, we develop an algorithm to systematically uncover the set of possible dynamics that an agent-based model can generate, and apply it to a medium-sized ABM. Our method is able to recover the full set of dynamics in only a few steps.

A model is a mapping from a parameter space to an observable output space. Changing the location in the parameter space changes the observed output. Research by Gutenkunst et al. (2007) [93] has shown that there are generically only a few *stiff* directions in the parameter space that significantly change the observable dynamics[1]. Conversely, there are many *sloppy* directions in which there are no significant changes in the model dynamics. The stiff and sloppy directions are defined as the eigenvectors of the Hessian matrix of a loss function at a point in parameter space. We apply this idea to six macroeconomic models including a medium-sized ABM and five Dynamic Stochastic General Equilibrium (DSGE) models, which represent economists' current *workhorse* models for monetary policy [162]. We find that each of these models is sloppy: their observable dynamics only change significantly in a handful of directions. These stiff directions typically involve many parameters, and are not aligned with the bare parameter axes, suggesting they may be missed by *one parameter at a time* variations.

1: In particular we refer to the seminal work of Gutenkunst et al. (2007) [93], with follow-on work by Apgar et al. (2010) [157], Transtrum et al. (2015) [142], Transtrum et al. (2010) [158], Transtrum et al. (2011) [159], Machta et al. (2013) [160], Mannakee et al. (2016) [161], and Hsu et al. (2020) [97]

In the DSGE models, the key directions are dominated by the parameters governing the exogenous shock processes, while deeper economic parameters typically have little influence. This is a known dependence and limitation of DSGE models [163], and reflects the formulation in terms of steady-state deviations for the selected models. In the ABM, the stiffest directions are dominated by parameters relating to *phase transitions* in the model. This finding suggests that common *one parameter at a time* sensitivity analysis may not be a sufficiently comprehensive tool to assess model sensitivity. More sophisticated multi-parameter sensitivity analysis methods exist (see Ligmann-Zielinska et al. (2020) [164] for a recent review). However, they have not permeated the macroeconomic ABM literature, and are focused only on specific set policy variables. Our approach complements these methods by considering the sensitivity of the dynamics at a given point in parameter space with respect to all parameters, while remaining computationally efficient.

We study the implications of the sloppy on the medium-scale Mark-0 ABM of Gualdi et al. (2015) [152], Gualdi et al. (2017) [165], which is a reduced form of the model by Gatti et al. (2011) [166] and has several phase transitions between different dynamics in the model's output.[2] Beginning with the phase where unemployment is constant and high, we find that the stiffest parameter directions represent the direction towards the nearest phase transitions, in this case, the realisation of the

2: A phase transition describes the change in state of a medium, such as the conditions at which water freezes (liquid → solid). In this case we refer to tipping points, where small changes in parameters drastically alter system dynamics. For example, changing from full employment to full unemployment.

model that has a periodic fluctuation around low unemployment. Exploiting this idea, we build an algorithm to step along these stiff directions and thus systematically explore the phase space within a given computational budget. In the case of the Mark-0 model, we show that this approach can efficiently recover all possible dynamics for the unemployment rate for both the low-dimensional and high-dimensional cases. Our method therefore provides a useful tool to navigate the space of parameters to find abrupt changes of behavior and their parameter dependencies.

The issue of searching the parameter space of Agent-based Models has received considerable attention in recent years. For instance, [167, 168] have combined elementary effects and kriging for a global sensitivity analysis. A prominent method and the closest to ours is meta-modelling (see Lamperti et al. (2018) [91], van der Hoog (2019) [169], Zhang et al. (2020) [170], ten Broeke et al. (2021) [171], with applications to parameter calibration in Bargigli et al. (2020) [172] and Chen and Desiderio (2021) [173]). In this approach, observable outputs are sampled at a selection of different points in the parameter space. A simpler model is then constructed that mimics the relationship between the parameters and a given set of outputs, or classifies the outcomes based on some criteria (e.g. using a neural network, as in Lamperti et al. (2018) [91]). The benefit of this approach is that once a surrogate model has been constructed, further exploring the parameter space through predicting the class of unseen parameters becomes computationally cheap. The crux is that the approach is exploratory only with respect to the classification criteria, such as the existence of a fat tailed distribution of GDP growth, that the modeler chooses to implement. This requires some ex-ante knowledge of the possible outcomes (or at least the desired outcomes) of the model by which to separate different phases.

Our approach instead is exploratory and agnostic, with no prior assumptions about the dynamics we are seeking, such as to provide an understanding of the full set of possible phenomena that can be generated by the model. We do this through intelligent sampling of the parameter space by exploiting the stiffest directions. An agnostic approach is crucial when some emergent behavior is truly unexpected, as is indeed often the case (see the detailed discussion of this point in [152]). Our method complements the meta-modelling approach, as the discovery of the different phenomena can facilitate a more precise criteria for the classification of parameters by the meta-model. In this way, one could identify new phases using our algorithm, and determine the phase transitions using a meta-model.

To develop the insights mentioned above, we use the numerical method described in Section 2.2 for estimating the Hessian matrix from which the parameter directions are extracted. We then apply this to the ABM and five DSGE models introduced in Section 4.1 to assess their sloppiness (Section 4.2). Finally, we use these insights to develop our algorithm (Section 4.3) and test it on the Mark-0 ABM (Section 4.4), for which we unveil some behavior that had escaped previous scrutiny.

4.1. Macroeconomic Models Considered

A simple Agent-Based Model: Mark-0

The first agent-based model examined in this chapter is the Mark-0 model developed by Gualdi et al. (2015) [152], which was further extended in Chapter 3. For the analysis presented here, we use the *Inactive CB* configuration without shocks (see Section 22). In this configuration, the Central Bank does not respond to inflation, with parameters set as $\pi^\star = \phi_\pi = 0$ and expectation anchoring disabled ($\tau^T = 0$). Additionally, energy prices are excluded, and the pricing function follows the specification from Sharma et al. (2020) [117],[3] which omits the demand-output ratio. The model can replicate several *macro-states* depending on the parameters (including, for example, full employment and unemployment, endogenous crises, high-inflation high-output, and low-inflation low-output scenarios). While Mark-0's aggregate behaviour is most likely not quantitatively precise, it generates qualitatively plausible and generic dynamical behaviours. In the baseline model, Gualdi et al. (2015) [152] have identified four distinct phases for the unemployment rate: full employment (FE), full unemployment (FU), residual unemployment (RU), and endogenous crises (EC). We chose the Mark-0 model because its' phase diagram has been well-studied in previous work. This allows us to compare our analysis of key parameter directions and the phase-space exploration algorithm (see Section 4.3) to previously studied and explored phases. Table 4.1[4] provides an overview of the $P = 14$ parameters that we consider in the Mark-0 model, in decreasing order of relevance to this chapter. These parameters were predetermined by Gualdi et al. (2015) [152], and we adopt them as the baseline parameters on which we focus our exploration, unless otherwise stated.

3: In this configuration, the pricing function does not consider the demand-output ratio $D_i(t)/Y_i(t)$.

4: The parameters are only slightly different as in chapter 3 (see Table A.1) for the sensitivity of loans interest rate on hiring-firing policy α_Γ, firing propensity η_0^- and the price adjustment size γ.

Table 4.1.: Parameters of the Mark-0 Model without Central Bank sorted in order of their relevance. The notation follows that of Sharma et al. (2020) [117].

Symbol	Default	Definition
$\rho^\star$	1%	Baseline interest rate
Θ	2.5	Default threshold
α_Γ	50	Loan rate effect on R
c_0	0.5	Baseline propensity to consume
R	2	Hiring/firing rate
$\frac{\gamma_w}{\gamma_p}$	1.0	Adjustment ratio
γ_p	0.1	Price-adjustment size
η_0	0.1	Baseline firing propensity
f	0.5	Bankruptcy effect on bank interest rates
δ	2%	Dividend share
β	2.0	Household intensity of choice
τ^R	0.5	Weight π^{ema} on $\hat{\pi}$
ϕ	0.1	Revival frequency
ω	0.2	EWMA memory

For the purposes of this chapter, we focus our attention on only the unemployment rate in the Mark-0 model (equivalent to output in Mark-0) because the phases identified by Gualdi et al. (2015) [152] are most visible in the dynamics of the unemployment rate.[5] For the approximation of the Hessian we use $S = 20$ seeds. We set a duration of $T = 30,000$ time steps with an initial cutoff of $T_{\text{eq}} = 10,000$, after which we find that the first two eigenvalues in the spectrum have converged to a steady state.

5: In Figure 4.1(a) we do consider all possible output series when calculating the Hessian. To ensure consistent scales when evaluating all output series, we apply the transformation $\hat{x}_t = \ln(x_t + c)$ with $c = 10^7$, to re-scale each of the observable output series x. Without this transformation, some variables (e.g. inflation) dominate as the $\|y_{s,k}(\Phi^\star)\|$ term in Eq. (A.50) becomes extremely small.

Dynamic Stochastic General Equilibrium (DSGE) Models

To contrast with the Mark-0 model, we also consider several standard DSGE models. Despite multiple shortcomings [see e.g. 156, 163], DSGE models remain the working horse models of many macro-economists [23, 174]. The first three models are borrowed to chapters 2, 3 and 8 of Galí (2015) [20], and represent the classic real business cycle and new Keynesian frameworks (chapter 8 being an open-economy model). These three models have $P = 9$, 12, and 13 parameters respectively, and have not been fitted to data explicitly. Rather they form the theoretical backbone of current DSGE models. We also consider three models that have been fitted empirically, namely the canonical Smets and Wouters (2007) [19] model with $P = 35$ parameters, the financial frictions model of Del Negro et al. (2015) [175] ($P = 50$) and the term-premium model of Carlstrom et al. (2017) [176] ($P = 51$). For the approximation of the Hessian in the DSGE models we use $S = 100$ seeds. We set a duration of $T = 5,000$ without an initial cutoff since these models are initialised to their steady state already. The outputs considered are the time-series of the log-deviations from the steady state.

4.2. Are Macroeconomic Models Sloppy?

Despite their different frameworks and levels of intricacy, all of the models display the sloppy phenomenology found in systems biology models [93]. Across all models, we observe that the eigenvalue spectrum spans several decades (Figure 4.1). In particular, for the DSGE models (b-g) the distribution is uniform across decades, which is similar to models studied in systems biology. However, the presence of phase transitions in the Mark-0 agent-based model with all observable outputs lead to a different eigenvalue spectrum (a), which is non-uniformly distributed. Instead, there are two deviating eigenvalues and a grouping of the remainder around 10^{-5}. This occurs for two reasons: First, these stiff eigenvalues correspond to the directions of the closest phase transitions, which cause extreme changes in the model dynamics, while the latter ones cause only minor shifts in the model's steady state or within the phase. The Mark-0 was constructed to be a reduced version of the Mark-1 model [166], such that only parameters affecting the phases remain. The second reason is that the amount of information, T, required to further distinguish the small eigenvalues in these models grows with their magnitude (as in Hsu et al. (2020) [97]), such that small eigenvalues may not be accurately estimated and actually be more spread out than shown in Figure 4.1.

To provide some intuition on the degree of sloppiness, the distance that can be travelled in a direction v_i given by an eigenvector without changing the observed output significantly (as measured by the loss function) is proportional to $1/\sqrt{\lambda_i}$ [93]. Thus, for the Mark-0 model (Figure 4.1(a)), one must travel roughly $\sim 10^{5/2}$ times as far in the sloppiest direction as in the stiffest direction to achieve a change in observable output that is of similar magnitude (in the absence of any phase transitions). To put this into perspective, if one wants to empiri-

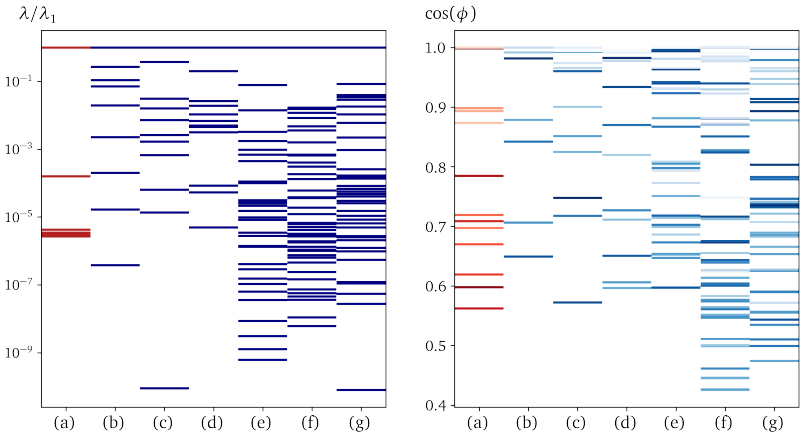


Figure 4.1.: (LHS) Eigenvalue Spectra relative to the largest eigenvalue (λ/λ_1), and (RHS) absolute cosine similarities between each eigenvector and its closest bare parameter axis. Note that the ranking of the cosine similarities does not necessarily correspond to that of the eigenvalues, see Table 4.2. The models include in red the Mark-0 ABM [152] with all variables evaluated in the full employment phase (a, $\Theta = 2.5$, $\rho^\star = 1\%$), with the remaining parameters from Gualdi et al. (2015) [152]. In blue, the Dynamic Stochastic General Equilibrium models of (b) Galí (2015) [20] chpt. 2, (c) chpt. 3, and (d) chpt. 8, as well as the estimated models of (e) Smets and Wouters (2007) [19], (f) Del Negro et al. (2015) [175], and (g) Carlstrom et al. (2017) [176] evaluated at the authors' reported parameters.

	(a)	FE	FE II	RU	RU II	FU	FU II	(b)	(c)	(d)	(e)	(f)	(g)
1:	1.00	0.81	0.81	0.91	0.91	0.93	0.98	0.98	0.75	0.98	0.99	1.00	0.74
2:	0.60	0.58	0.58	0.84	0.89	0.80	0.44	0.65	0.57	0.93	0.99	0.88	0.89
3:	0.78	0.78	0.78	0.55	0.59	0.58	0.55	0.84	0.96	0.65	0.96	0.67	0.80
4:	0.56	0.54	0.54	0.48	0.49	0.59	0.72	0.71	0.72	0.87	0.94	0.87	1.00
5:	0.71	0.58	0.58	0.45	0.59	0.49	0.55	0.88	0.99	0.73	1.00	1.00	0.91

Table 4.2.: Absolute cosine similarity between eigenvectors and the most aligned bare parameter axes for the first five eigenvectors of (a) the Mark-0 ABM (Gualdi et al., 2015) with all variables evaluated in the full employment phase (a, $\Theta = 2.5$, $\rho^\star = 1\%$), and for only the unemployment rate in the full employment phase (FE, and FE II, $\Theta = 5.0$, $\rho^\star = 1\%$), the residual unemployment phase (RU, $\Theta = 2.5$, $\rho^\star = 5\%$, and RU II, $\Theta = 5.0$, $\rho^\star = 5\%$), and the full unemployment phase (FU, $\Theta = 2.5$, $\rho^\star = 15\%$, and FU II, $\Theta = 5.0$, $\rho^\star = 12.5\%$). All remaining parameters are from Gualdi et al. (2015) [152]. The Dynamic Stochastic General Equilibrium models of (b) Galí (2015) [20] chpt. 2, (c) chpt. 3, and (d) chpt. 8, as well as the estimated models of (e) Smets and Wouters (2007) [19], (f) Del Negro et al. (2015) [175], and (g) Carlstrom et al. (2017) [176] evaluated at the authors' reported parameters.

cally fit the parameters underlying these models by constraining each direction to within a 10% confidence interval, one would require 10^5 times the amount of data for the sloppiest direction in Mark-0 as for the stiffest [160]. In the context of macroeconomics, this implies a high data requirement that is infeasible. However, it may not be necessary for Agent-based Models to be fit to such a high degree of precision in regards to each of their parameters. As noted in Gutenkunst et al. (2007) [93], *concrete* predictions can be extracted from models long before their parameters are even roughly known" because observed outputs depend only on a few stiff directions.

The importance of the stiff directions suggests the question of what the major constituents of these linear combinations are.We first quantify the degree to which the eigenvectors correspond to the bare parameter axes by considering the cosine similarity between the two vectors, where the bare axes represent the natural basis in the P-dimensional parameter space. Specifically, for each eigenvector we report the absolute cosine similarity of the bare axes with which it is

most closely aligned ($\cos\phi(v_i, v_j) = \frac{v_i \cdot v_j}{\|v_i\|\|v_j\|}$).[6] A value of one implies that the given eigenvector contains only a single non-zero parameter entry, rather than a combination of parameters. Figure 4.1 (right panel) shows the distribution of the cosine similarities for each eigenvector and its closest bare parameter axis. Note that the ranking here does not correspond to the ranking of eigenvalues; Table 4.2 shows instead the cosine similarities for the stiffest five directions. Many observed directions do not correspond to bare parameter axes, though there is a significant portion that do. We consider these separately for DSGE and ABM models.

6: The cosine similarity between two vectors is defined by $\cos\phi(v_i, v_j) = \frac{v_i \cdot v_j}{\|v_i\|\|v_j\|}$. In our case, we compare each eigenvector v_i to an identity vector e_p representing parameter direction p, where $e_{ip} = 0\ \forall i \neq p$ and $e_{pp} = 1$. From this, we extract the cosine similarity of the eigenvector to the most aligned parameter axis: $sim_i = \max_{p\in P}|\cos\phi(v_i, e_p)|$.

Stiff Directions in DSGE models

Based on Figure 4.1, DSGE models can be described as sloppy models, with many eigenvalues orders of magnitude smaller than the largest one. This means that many combinations of parameters are in fact quite irrelevant for the dynamics of the system, and that one should be able to construct reduced models with similar performance.

The stiffest parameter directions frequently correspond to individual parameter axes, with cosine similarities close to 1. These directions are often the parameters related to the exogenous shock processes, such as their persistence or co-variances. For the estimated models (f, g, and h in Figure 4.1), almost all of the first 5 stiffest eigenvectors contain large elements relating to the shock processes, with a few also containing a single additional parameter (see Table 4.2). This reflects the adiabatic nature of these models, which are formulated as deviations from a given steady state, such that shocks in a given period dominate the dynamics of the model. In light of this, once the exogenous processes have been fixed, variations in the further parameters appear to offer comparatively little discernible change in the models' output.

Stiff Parameter Directions and Phase Transitions in Agent-based Models

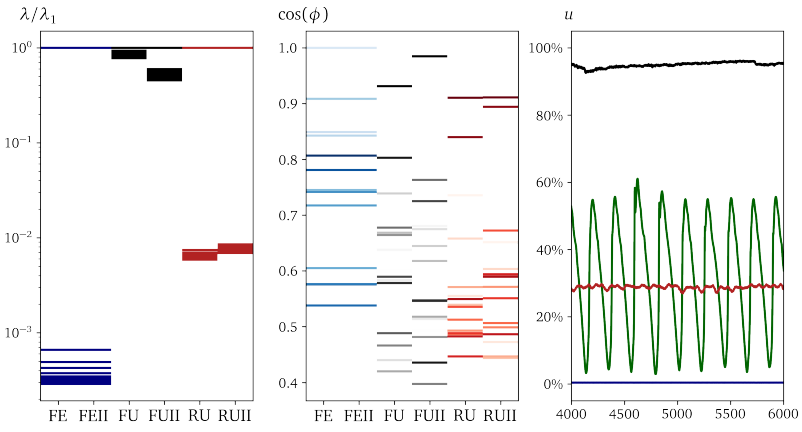


Figure 4.2.: (LHS) Eigenvalue Spectra relative to the largest eigenvalue (λ/λ_1), (Center) the absolute cosine similarities between each eigenvector and its most aligned bare parameter axis, and (Right) the unemployment rate for different phase, for the Mark-0 ABM [152]. At variance with Figure 4.1 we restrict the estimation to only the unemployment rate as a variable, with parameters in: the full employment phase (blue, with FE, $\Theta = 2.5$, $\rho^\star = 1\%$, and FE II, $\Theta = 5.0$, $\rho^\star = 1\%$), the residual unemployment phase (red, with RU, $\Theta = 2.5$, $\rho^\star = 5\%$, and RU II, $\Theta = 5.0$, $\rho^\star = 5\%$), and the full unemployment phase (black, FU, $\Theta = 2.5$, $\rho^\star = 15\%$, and FU II, $\Theta = 5.0$, $\rho^\star = 12.5\%$). The EC phase (green) is evaluated at $\Theta = 1.3$, $\rho^\star = 1\%$. All remaining parameters are from Gualdi et al. (2015) [152].

In the Mark-0 agent-based model, the stiffest directions are dominated by parameters relating to phase transitions in the model. To explore

this, we focus on the phases of the unemployment rate: full employment (FE), residual unemployment (RU), endogeneous crises (EC), and full unemployment (FU) (see Figure 4.2). The eigenspectra for two points in three phases are shown in Figure 4.2. They can generally be described by an outlier first eigenvalue, followed by a dense body of the remaining eigenvalues. The degree of separation between the first and second eigenvalue depends on the phase considered, with a large deviation in the FE phase, and a small one in the FU phase.

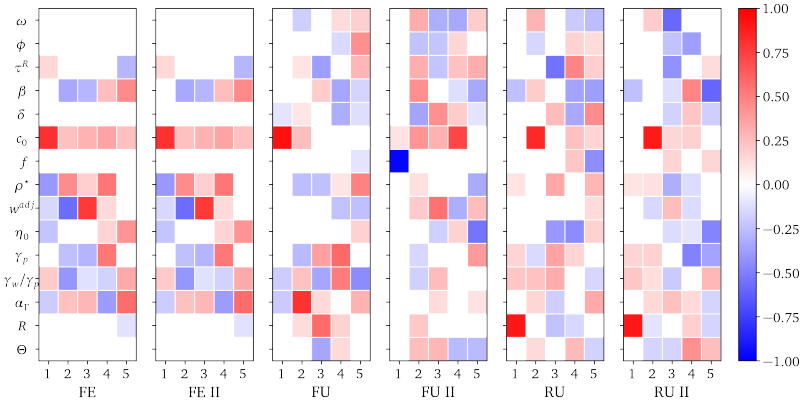


Figure 4.3.: Heatmaps of the components of the stiffest 5 eigenvectors of the Mark-0 ABM focused on the unemployment rate with all parameters. The sign of the eigenvector is such that the angle with the direction c_0 is positive. All components with a magnitude between -0.1 and 0.1 have been set to zero for exposition purposes. The directions were evaluated in the full employment phase (FE, $\Theta = 2.5$, $\rho^\star = 1\%$, and FE II, $\Theta = 5.0$, $\rho^\star = 1\%$), the residual unemployment phase (RU, $\Theta = 2.5$, $\rho^\star = 5\%$, and RU II, $\Theta = 5.0$, $\rho^\star = 5\%$), and the full unemployment phase (FU, $\Theta = 2.5$, $\rho^\star = 15\%$, and FU II, $\Theta = 5.0$, $\rho^\star = 12.5\%$). All remaining parameters are from Gualdi et al. (2015) [152].

To understand the components of the stiffest directions, Figure 4.3 shows a heatmap of the stiffest five eigenvectors for the three phases considered. In each case, the direction for the first eigenvector is dominated by one of the variables which have been found to lead to phase transitions in the *handcrafted* analysis of Gualdi et al. (2015) [152], including the parameters R and α_Γ (for a definition of these variables, see Table 4.1). In addition, we have here also identified c_0 to be a key variable in the phase transition, missed in the analysis of Gualdi et al. (2015) [152]. While the stiff directions are dominated by a single value, they are nonetheless unaligned with the bare parameter axes, as shown by the non-unity cosine similarities. In fact, the combination of parameters is generally one of multiple other phase-relevant parameters. For example, in the RU phase these mixture parameters include β, and γ_w and $\rho^\star$, each of which are related to varying phase transitions. To illustrate the relation between stiff directions and phase transitions, Figure 4.4 shows the dynamics of the unemployment rate in the Mark-0 model starting in each phase (black lines), with log step in the stiffest (green) and sloppiest (red) directions (i.e. $\log \Phi' = \log \Phi + v_i$). Note that for the full employment case (left panel), we observe a transition to residual unemployment, while for the full unemployment phase (right panel) we observe a transition to the endogenous crises phase. For the residual unemployment phase there is no explicit transition, but a strong move towards full unemployment. Note that in all cases, there is almost no change in the dynamics when traversing the sloppy direction with an equal step size. The key observation is that the variables identified in the stiffest directions do indeed point towards nearby phase transitions.[7] These results suggest that looking to the stiffest direction is a means to identify their different phases and the boundaries separating them.

7: Note that the sign of the eigenvector is arbitrary, so following such directions can lead towards the phase transition or away from it, but at the shortest path.

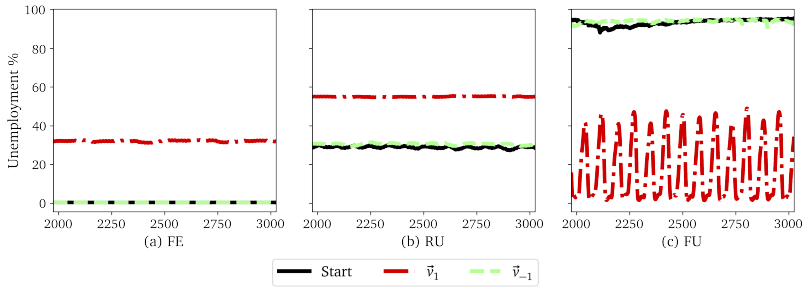


Figure 4.4.: Unemployment dynamics in different phases in their original form (black), with a unit log step in the stiffest (red) and sloppiest (green) directions identified in Figure 4.3, i.e. $\log \Phi' = \log \Phi + v_i$. The panels are: (left) the full employment phase (FE, $\Theta = 2.5$, $\rho^\star = 1\%$), (center) the residual unemployment phase (RU, $\Theta = 2.5$, $\rho^\star = 5\%$), and (right) the full unemployment phase (FU, $\Theta = 2.5$, $\rho^\star = 15\%$). All remaining parameters are from Gualdi et al. (2015) [152].

Transitions in Two Dimensions

For the unemployment rate dynamics in the two dimensional plane (bankruptcy threshold Θ, central bank's baseline interest rate $\rho^\star$), we find that the stiffest direction consistently points in the direction of a nearby phase transition. Figure 4.5 shows the approximate phase diagram in the $(\log \Theta, \log \rho^\star)$ space (dashed black lines), together with the estimation of the first (red) and second (blue) eigenvectors for a grid of points.

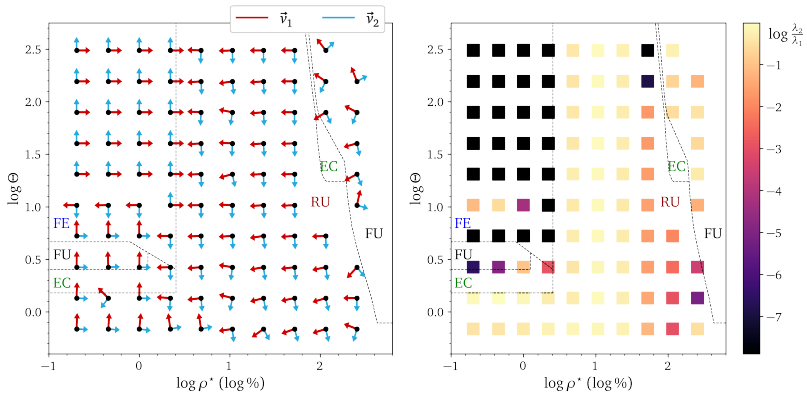


Figure 4.5.: (LHS) Phase diagram of the Mark-0 model in the $(\log \Theta, \log \rho^\star)$ plane with projected eigenvectors. The eigenvectors, v_1 (red) and v_2 (blue), are based on the 2-dimensional Hessian with only Θ and $\rho^\star$ using unemployment as output variable. Dashed black lines indicate the locations of phase transitions. (RHS) logarithm of the ratio of the first to the second eigenvalue at different points in the phase space. The black dots indicate $\ln(\lambda_2/\lambda_1) < -10$. Simulation parameters: $T = 30,000$, $S = 20$, $T_{\text{eq}} = 10,000$. The parameters of the model, other than Θ and $\rho^\star$, are from Gualdi et al. (2015) [152]

The phase diagram (Figure 4.5 left panel) suggests that the line formed by the first eigenvector is also generally the shortest path to the phase transition. The second eigenvector then points in the direction of the second closest phase transition or parameter boundary. This is clearest when considering the intersection of FE, RU and FU near $\log \Theta \approx 0.5$ and $\log \rho^\star \approx 0$. Here, at $(0, 0.4)$ the stiffest direction is the closer FU→FE transition, while for $(0.2, 0.6)$ the stiffest direction is the closer FE→RU transition. While this is true near phase boundaries, it is also true further away from them. In the FE phase, far from the boundary, the stiffest direction still points towards the closer FE→FU transition.

The proximity of a phase transition appears, as one would expect, to lead to a strong splitting in the eigenvalues. The right panel of Figure 4.5 shows the ratio of the first (λ_1) and second (λ_2) eigenvalues corresponding to the grid in the left panel. Near the FE→RU transition as well as the RU→FU transitions the ratio λ_2/λ_1 tends to zero. This is likely due to the high value $\mathcal{L}$ generated by a transition over the phase boundary, which eclipses other changes. On the other hand, it is also possible that this is simply a result of a drop in importance of all other parameters, as in the FE phase the second direction is irrelevant for almost all parameter values (except near phase transitions).

The implication of these findings is twofold: first, the stiff eigenvectors are indicators for the shortest path to the closest phase transition, second, the relationship between the eigenvalues is indicative of the distance to the phase transition, and potentially the number of different transitions.

4.3. A New Method for Systematic Phase Exploration

Using the formalism above, we now construct a probabilistic algorithm that exploits the insight that the stiffest parameter directions indicate the direction of the largest change in observed dynamics in order to generate a maximally diverse sampling of the observable output space at a relatively low computational cost. The general principle of the algorithm is to evaluate the stiffest direction at a given point, travel along them to a new point, and repeat the exercise for a given number of steps. We are constructing a new intelligent means to sample the parameter space, compared to existing methods such as Sobol sampling or Nearly Orthogonal Latin Hypercubes.

Formulation of the Algorithm

To begin with, the modeller must decide a number of steps N for which the algorithm should run, as well as a starting point $\Phi(0)$ in parameter space. Furthermore, the modeller should set the hyperparameters for the Hessian estimation: the number of random realisations S, the simulation time T and the equilibrating time T_{eq}, such that the first two eigenvalues converge to a steady ratio. The algorithm then proceeds sequentially for each step $n \in \{1, 2, \ldots, N\}$ as follows:

1. Evaluate the Hessian matrix $H^{\mathscr{L}}(\Phi(n))$ and decompose it into its eigenvalues and associated eigenvectors. For the purpose of the algorithm, we consider only the first two eigenvectors, v_1 and v_2, together with their respective eigenvalues λ_1 and λ_2. More can be considered, but an accurate estimation of their spectrum requires an exponentially increasing simulation time T (see e.g. Hsu et al. (2020) [97]).
2. Select the direction, $v(n)$ of the next step via:

$$v(n) = \begin{cases} v_1, & \text{with probability } \lambda_1/(\lambda_1+\lambda_2) \\ v_2, & \text{with probability } \lambda_2/(\lambda_1+\lambda_2) \end{cases}, \tag{4.1}$$

   such that $\lambda(n)$ is the corresponding eigenvalue. This selection generally steps in the stiffest direction, unless both directions are important, as could be the case in the presence of multiple close phase transitions for instance. We find that the first two eigenvectors are sufficient to span the search space and to find all important dynamics of the Mark-0 model (can also be seen in Figure 4.1).
3. Since the sign of the eigenvector returned by the algorithm is

ill-defined, ensure that the direction is consistent, that is

$$v(n) = \begin{cases} v(n), & \text{if } \arccos(v(n) \cdot v(n-1)) \leq 165^\circ \\ -v(n), & \text{otherwise} \end{cases}, \tag{4.2}$$

which prevents the algorithm from oscillating back and forth between two points across the same phase transition.[8]
In case this is the first step (i.e. $v(n-1)$ does not exist) we determine the direction of the first step by taking a step in both the $v(n)$ and $-v(n)$ directions, yielding parameter candidates $\Phi(n)^+$ and $\Phi(n)^-$ (see step 4 for the distance) and picking the candidate with the largest loss relative to the initial parameter point $\Phi(0)$ (see Eq. (A.50)).

8: It is possible to choose the sign of the eigenvector randomly. This non-ballistic motion might lead to a denser exploration of the phase space that is less sensitive to the initial conditions and choice of direction.

4. Pick the next point in parameter space by travelling a distance d in the $v(n)$ direction which is neither too large nor too small:

$$d = \min\left(\frac{1}{\sqrt{\lambda(n)}} \max\left(\varepsilon, \varepsilon_{\min}\sqrt{\lambda_1}\right), \varepsilon_{\max}\right) \tag{4.3}$$

$$\log \Phi^{n+1} = \log \Phi(n) + d \cdot v(n) \tag{4.4}$$

where the distance depends on three given parameters: $\varepsilon_{\min} = 0.3$, the minimum step-size to take in log-parameter space, $\varepsilon = 0.1$, the distance to travel relative to the span of the interval over which the direction can vary ($\sqrt{\lambda(n)}$), and $\varepsilon_{\max} = 1.0$, the maximum distance that can be traversed (corresponds to a factor of at most e in bare parameter space). We introduce $\varepsilon_{\min}$ and $\varepsilon_{\max}$ to counteract extreme eigenvalues, which may lead to extreme steps (unrealistically large parameters) or extremely small steps such that a large number of steps is required to generate sufficient changes. The goal is to maximise the local exploration while remaining in general proximity of the initial parameters, where our approximation in Eq. A.52 still holds. For smooth models, the eigenvectors and eigenvalues span a hyperribbon that encloses a region with the same value of the loss function, see [93]. Therefore, we make a step with our algorithm to the edge of the hyperribbon.

The most likely path followed by this algorithm is to select the top eigenvector at each step, in this way, the algorithm continually explores the direction of the largest change. Only when the first eigenvalues are similar is there a chance of bifurcating. This implies that in close proximity to a phase transition, where the eigenvalues typically diverge from one another, the direction of the phase transition is selected. Meanwhile, in areas of the parameter space where there are no immediate transitions and the eigenvalues take on a more uniform distribution, it is more likely that the algorithm explores parameter combinations along the second eigenvector as well. It is not critical that the model shows a first order phase transition, a continuous transition to another phase can also detected with the algorithm.

Computational Cost

The computational expense of agent-based models implies that there is typically an upper limit on the number of function calls that can be made. In regards to the algorithm presented, the most costly step is the calculation of the Hessian matrix, which requires $\mathcal{O}(SP)$ function calls. Consequently, the algorithm overall requires $\mathcal{O}(NSP)$ simulations of the agent-based model. For the two-dimensional Mark-0, this implies at least $8 \times 20 \times 2 \times 2 = 640$ evaluations with a duration of $T = 30,000$ time-steps are required (note one factor of two emerges from taking central difference derivatives). For reference, a Hessian matrix for the Mark-0 with $P = 14$ parameters, running $S = 20$ seeds in parallel, and $T = 30,000$, requires ~ 45 minutes of computation time on an AMD Ryzen 9 5950X processor with 32GB RAM, while a $P = 2$ Hessian requires ~ 6 minutes. A natural benchmark for our computation is the intelligent Latin Hypercube Sampling strategy [177, 178].[9] For the two-dimensional Mark-0, we find that ~ 29 samples of the parameter space are required to obtain all four phases. Considering $S = 20$ random samples to mitigate any noise effects, this corresponds to $29 \times 20 = 580$ function calls, which is on par with a seven-step estimation of the algorithm. As the dimension of the parameter-space increases, we expect the number of required parameter samples for the hypercube sampling to increase at a higher rate than the algorithm presented here. Simultaneously, the information gained from the Hessian matrices may prove more valuable than a quasi-random sampling.

9: For Latin Hypercubes, the parameter space is divided into equally probable intervals and then sample points are placed such that there is only one sample in each axis-aligned hyperplane

4.4. An Application to the Mark-0 Agent-Based Model

The testing bed for the algorithm is the identification of the phases of unemployment dynamics in the Mark-0 model, for which the number of parameters is $P = 14$. We find that in two, three and sixteen dimensions, our algorithm can recover all the major phases of unemployment (full employment, full unemployment and residual unemployment) within eight steps only. Due to the small parameter region for which the EC phase exists, it is often the case that it is missed in the two-dimensional case. Thus, the recovery of the endogenous crisis phase is dependent on the starting point. Despite dependence on initial conditions, in almost all cases our algorithm recovers at least one phase that is different to the phase of the initial parameter choice.

In the two dimensional $(\Theta,\ \rho^\star)$-space we find that the algorithm covers at least two phases irrespective of the starting point or the choice of path taken. Figure 4.6 shows two walks originating from each of three different starting points in the $(\Theta,\ \rho^\star)$ phase diagram. The solid lines represent the most likely walk in which the algorithm exclusively follows the stiffest direction, v_1, at each point. The dashed lines represent an alternative random walk, which contains at least one deviation in the direction of the second stiffest direction. The sequence of choices, and the probabilities of choosing the stiffest direction at each point are shown in Table 4.3.

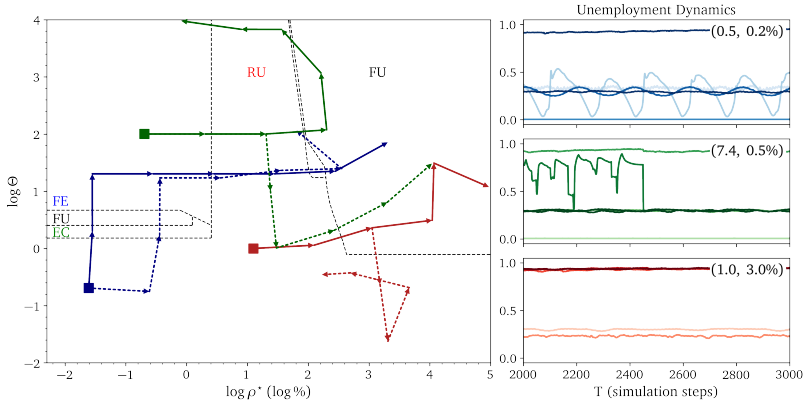


Figure 4.6.: (LHS) Phase diagram of the Mark-0 model without central bank in the $(\log \Theta, \log \rho^\star)$ plane with two likely algorithm paths for three different starting values. Solid lines show the most likely path following only v_1, while dotted lines indicate an alternate path with mixed steps. Dashed black lines indicate the locations of phase transitions. (b) Dynamics of the unemployment rate for different steps along the v_1 path for each starting point. Simulation parameters: $T = 30,000$, $S = 20$, $T_{\text{eq}} = 10,000$. The parameters of the model, other than Θ and $\rho^\star$, are those of Gualdi et al. (2015) [152]. Algorithm parameters are $\varepsilon_{\min} = 0.3$, $\varepsilon = 0.1$, $\varepsilon_{\max} = 1.0$.

		$\mathbb{P}(\vec{v}^n = \vec{v}_1)$						
Start $(\Theta, \rho^\star)$	Path	1	2	3	4	5	6	7
(0.5, 0.002)	1111111	0.53	0.70	1.00	1.00	0.98	0.62	0.64
(0.5, 0.002)	2111111	0.53	0.52	0.86	1.00	0.65	0.68	0.57
(1.0, 0.030)	111111	0.51	0.91	0.68	0.94	0.60	0.91	-
(1.0, 0.030)	1121111	0.51	0.91	0.68	0.60	0.63	0.57	0.94
(7.4, 0.005)	1111111	1.00	1.00	0.56	0.69	0.66	1.00	0.52
(7.4, 0.005)	1122112	1.00	1.00	0.56	0.58	0.62	0.87	0.75

Table 4.3.: Probability of choosing eigenvector v_1 at each step in three paths taken from a starting point in the $(\Theta, \rho^\star)$-space. Simulation parameters: $T = 30,000$, $S = 20$, $T_{\text{eq}} = 10,000$. The parameters of the model, other than Θ and $\rho^\star$, are those of Gualdi et al. (2015) [152]. Algorithm parameters are $\varepsilon_{\min} = 0.3$, $\varepsilon = 0.1$, $\varepsilon_{\max} = 1.0$.

The effectiveness of the algorithm is dependent on its initial position within the parameter space.[10] The most effective walk begins near the parameter boundaries at $\Theta = 0.5$ and $\rho^\star = 0.2\%$, and covers each of the unemployment phases (see the dynamics in the top right panel of Figure 4.6). The walk beginning in FE (green) covers the three major phases of FE, RU, and FU within three steps, while missing out on the EC phase. This is likely due to the fact that within this two-dimensional space, the EC phase is only present in small pockets, such that it is easily missed. Finally, the second RU walk (red) is the least effective, as it recovers only the RU and FU phases before terminating at extreme $\rho^\star$ values.

The probability of following the stiffest direction (Table 4.3) reflects the observations made in Section 4.2, that the eigenvalues are drawn apart in the presence of a discontinuous phase transition. In particular, the probability of taking the stiffest direction tends to one when there is a singular close phase transition, while it tends to 50% when there is either two proximate phase transitions or no proximate phase transitions. We note that the FE phase may be an exception to this, as the second eigenvalue becomes vanishingly small throughout the phase. This may be due to the low-fluctuation steady-state nature of this phase.

While the full recovery of phases may depend on starting values, each walk identified at least two phases. This suggests that a potentially optimal exploration may require a combination of parameter space sampling combined with short iterations of the algorithm. Alternatively, since the direction of the eigenvectors v_i is indeterminate, an alternative option (when not close to parameter boundaries) would be to start the algorithm in both possible directions. In that case, starting points in the central RU phase would explore both the FE and FU phases.

10: This could potentially mitigated by randomising the first direction, or simply walking in both directions

A Tentative Extension to Higher Dimensions

The principal aim of our algorithm is to facilitate phase exploration of agent-based models in high-dimensional parameter spaces. We apply the algorithm to both the three dimensional and fourteen dimensional (all parameters) formulation of the Mark-0 model to identify the unemployment phases. Figure 4.7 shows the unemployment rate dynamics for two walks following exclusively the stiffest direction in the $P = 3$ (left) and $P = 14$ (right) cases. The starting points correspond to those of the two dimensional exploration (Figure 4.6).

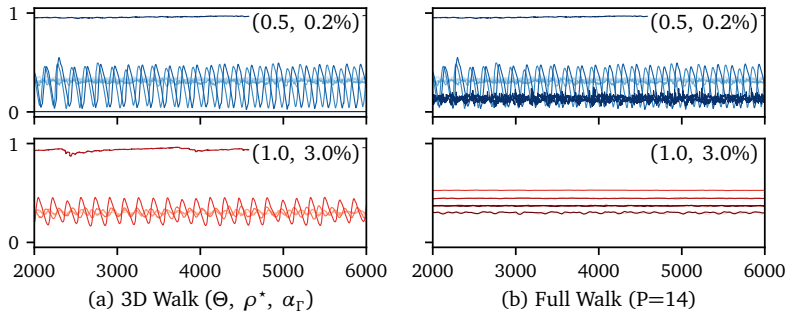


Figure 4.7.: Dynamics of the unemployment rate for different steps in the 3D case (Θ, $\rho^\star$, α_Γ) and the multi-dimensional ($P = 14$) case. We consider again the three most likely paths with 5 steps. The initial parameters of the algorithm in the (Θ, $\rho^\star$)-space are shown in the top right-hand corner of each panel and correspond to those of Figure 4.6). Simulation parameters: $T = 30,000$, $S = 20$, $T_{\text{eq}} = 10,000$. All other parameters are those of Gualdi et al. (2015) [152].

For the previously successful walk beginning at $\Theta = 0.5$, $\rho^\star = 0.2\%$ we again find that in the $P = 3$ case all phases have been recovered, while the full employment phase was missed in the $P = 14$ case. For the $\Theta = 1.0$, $\rho^\star = 3.0\%$ case we find that in the $P = 3$ case we obtain the same results as in the $P = 2$ case, while in the $P = 14$ case there is only a series of level shifts within the residual unemployment phase.[11] In particular, the lower right panel suggests that no different phases are identified (such as the large level shifts between RU and FE, or different oscillatory behaviours between EC and RU). The reason behind this is the agnostic mean-squared loss function (Eq. (A.50)) which directly compares the unemployment time-series. This means that qualitatively small changes such as level shifts may dominate. One example is the small shift in periodicity of the EC phase (Figure 4.7 top panels) that would lead to a large loss as T becomes large, though it is qualitatively still an EC phase. This suggests that our algorithm should be improved along several dimensions: i) considering an alternative agnostic loss function such as the Kullback-Leibler divergence, ii) including more observables in the loss function, allowing one to better discriminate between different phases; iii) adding an element of randomness in the amplitude and direction of the random walk would remove the (deterministic) dependence on the initial condition. We leave these extensions for further investigations.

11: Note that increasing the step size ε to 0.3 leads to a wider array of levels explored between the RU and FU phases, at the cost of missing the EC and FE phases in the $\Theta = 0.5$, $\rho^\star = 0.2\%$ case

The introduction of further dimensions leads to a different picture on key parameters, as well as the effective dimensions of the model. The introduction of α_Γ ($P = 3$) leads to a strong representation of α_Γ in the stiffest direction, in place of $\rho^\star$. This is because α_Γ controls the vertical boundary between the FE and RU phases shown in Figure 4.6, such that the distance to the phase transition in the $\log \alpha_\Gamma$ direction is likely shorter than in the $\log \rho^\star$ direction. A similar situation occurs also in the $\Theta = 1.0$, $\rho^\star = 3.0\%$ case, where the introduction of α_Γ leads to less frequent changes in $\rho^\star$. This shift is also reflected in the $P = 14$ case. The parameters with the largest changes (R, α_Γ, β, Θ, γ_w/γ_p) are those

that have previously been identified to induce changes in phases [152, 165, 179]. Meanwhile, for both starting points (Figure 4.7) at least five parameters were left unperturbed, suggesting that the model phases vary in a $Q = 9$-dimensional space, suggesting one potential avenue for the reduction of the model.

4.5. Extension with KL-Divergence

We started from the assumption that our models are not chaotic, meaning that a small change in the parameters Φ results in only a small change in the dynamics of the model. Under this assumption, we can define the distance between dynamics using the mean squared error Eq. A.50. However, if we transition to a probabilistic view of the system, instead of analyzing the trajectories $\vec{x}$, we can consider the probability distribution $p(\vec{x};\Phi)$. For probabilistic models, the natural distance between two distributions is the Kullback-Leibler (KL) divergence[12], which provides a measure of how much the perturbed distribution $p(\vec{x};\Phi+\delta\Phi)$ differs from the original distribution $p(\vec{x};\Phi)$. The Hessian of the KL divergence can be expressed as:

$$\mathcal{H}_{\mu,\nu} = \left\langle \partial_\mu \log p(\vec{x};\Phi) \partial_\nu \log p(\vec{x};\Phi) \right\rangle_{p(\vec{x};\Phi)} \tag{4.6}$$

where $\partial_\mu = \frac{\partial}{\partial \Phi_\mu}$ represents the derivative with respect to the parameters, and the expectation is taken over the probability distribution $p(\vec{x};\Phi)$.

At this point, we can make two different assumptions about the structure of the probability distribution $p(\vec{x};\Phi)$.

Either we assume that $p(\vec{x};\Phi)$ is sharply concentrated around the most likely states, similar to statistical mechanics or in the Schelling model, then the distribution is given by:

$$p(\vec{x};\Phi) = \frac{1}{Z(\Phi)} e^{-ME(\vec{x};\Phi)}, \tag{4.7}$$

where $E(\vec{x};\Phi)$ is the energy function that measures the *cost* of a state $\vec{x}$ given parameters Φ, M is a scaling factor (e.g., system size or inverse temperature) and $Z(\Phi)$ is the partition function that normalizes the distribution:

$$Z(\Phi) = \int e^{-ME(\vec{x};\Phi)} dx. \tag{4.8}$$

For large M, the probability $p(\vec{x};\Phi)$ concentrates around the state $\vec{x}(\Phi)$, which minimizes $E(\vec{x};\Phi)$. With this, one can approximate the distribution around the most likely state $\vec{x}^\star$:

$$-\frac{1}{M}\log p(\vec{x};\Phi) = \frac{1}{M}\log(Z(\Phi)) + E(\vec{x};\Phi) \tag{4.9}$$

$$\approx E(\vec{x}^\star;\Phi) + V(\Phi)\delta\vec{x} + \frac{1}{2}\delta\vec{x}\mathbb{H}(\Phi)\delta\vec{x} + \frac{1}{M}Z(\Phi) \tag{4.10}$$

with $\delta\vec{x} = \vec{x} - \vec{x}^\star$.[13] Now depending on the cases where the most likely state $\vec{x}^\star$ is within or at the boundary, one can further approximate $\log Z(\Phi)$ or the first order term vanishes.

12: Note that the distance metric needs to be symmetric, therefore we use the symmetrized Kullback-Leibler divergence:

$$D_{symKL} = \frac{1}{2}(D_{KL}(p,q) + D_{KL}(q,p)) \tag{4.5}$$

with $D_{KL}(p,q) = \sum_{x\in X} p(x)\log\left(\frac{p(x)}{q(x)}\right)$.

13: Note that $\mathbb{H}(\Phi)$ is the positive definite Hessian of the energy landscape E, not the Hessian associated with sloppiness.

Within the boundaries we can write Eq. 4.10 as

$$\mathcal{H}_{\mu,\nu} = \left\langle \partial_\mu \log p(\vec{x};\Phi) \partial_\nu \log p(\vec{x};\Phi) \right\rangle_{p(\vec{x};\Phi)} \tag{4.11}$$

$$= \left\langle \partial_\mu \Psi(\Phi) \partial_\nu \Psi(\Phi) \right\rangle_{p(\vec{x};\Phi)} \tag{4.12}$$

where $\Psi = \left(-\frac{M}{2}\log(\det(\mathbb{H}(\Phi))) + \frac{M}{2}\delta\vec{x}^T\,\mathbb{H}(\Phi)\,\delta\vec{x}\right)$ and we used that the expansion of the log-partition function around $x^\star$ is $\log(Z(\Phi)) \approx -ME(\vec{x}^\star;\Phi) - \frac{1}{2}\log(\det(\mathbb{H}(\Phi)))$.[14]

14: Note that the free energy $E(\vec{x}^\star;\Phi)$ vanishes because for the log-partition function and the free energy the expansion of the 0th order is $E(\vec{x}^\star;\Phi)$ with different signs.

We are going to look at the different terms separately. The quadratic term from the partition function [15]

$$\mathcal{H}^0_{\mu,\nu} = \frac{M^2}{4}\left\langle \partial_\mu \left(\log(\det(\mathbb{H}(\Phi)))\right) \partial_\nu \left(\log(\det(\mathbb{H})(\Phi))\right)\right\rangle_{p(\vec{x};\Phi)} \tag{4.13}$$

$$= \frac{M^2}{4}\mathrm{Tr}[\mathbb{H}^{-1}\partial_\mu\mathbb{H}(\Phi)]\mathrm{Tr}[\mathbb{H}^{-1}\partial_\nu\mathbb{H}(\Phi)] \tag{4.14}$$

15: We use the matrix derivative identity for positive definite square matrices $\partial_\zeta \log\det A = \partial_\zeta \mathrm{Tr}[\log A] = \mathrm{Tr}[A^{-1}\partial_\mu A]$

Note that none of the terms are dependent on $\vec{x}$ so we can apply directly the expectation value. And note as well that in the highest order of M we again find a rank 1 matrix.

We can continue with the quadratic term from the second order expansion in Eq. 4.12. This term can be written as

$$\mathcal{H}^{1,2}_{\mu,\nu} = \frac{M^2}{4}\left\langle \partial_\mu \left(\delta\vec{x}\,\mathbb{H}(\Phi)\,\delta\vec{x}\right) \partial_\nu \left(\delta\vec{x}\,\mathbb{H}(\Phi)\,\delta\vec{x}\right)\right\rangle_{p(\vec{x};\Phi)} \tag{4.15}$$

$$= \frac{M}{2}\left\langle \sum_{i,j}[\mathbb{H}^{-1}_{ij}\mathbb{H}_{ij}\partial_\mu\vec{x}^\star_i\partial_\nu\vec{x}^\star_j, i\leftrightarrow j]\right\rangle_{p(\vec{x};\Phi)} \tag{4.16}$$

$$= \frac{M}{2}\,\mathrm{Tr}[\mathbb{H}(\Phi)\mathbb{K}_{\mu,\nu}], \tag{4.17}$$

where

$$\left(\mathbb{K}_{\mu,\nu}\right)_{ij} := \partial_\mu\vec{x}^\star_i\partial_\nu\vec{x}^\star_j + i\leftrightarrow j. \tag{4.18}$$

Finally, we can write the cross term

$$\mathcal{H}^{1,1}_{\mu,\nu} = \left\langle \partial_\mu\left(-\frac{M}{2}\log(\det(\mathbb{H}(\Phi)))\right)\partial_\nu\left(\frac{M}{2}\delta\vec{x}^T\,\mathbb{H}(\Phi)\,\delta\vec{x}\right) + \mu\leftrightarrow\nu\right\rangle_{p(\vec{x};\Phi)} \tag{4.19}$$

$$= -\frac{M}{2}\mathrm{Tr}[\mathbb{H}^{-1}\partial_\mu\mathbb{H}(\Phi)]\mathrm{Tr}[\mathbb{H}^{-1}\partial_\nu\mathbb{H}(\Phi)] \tag{4.20}$$

where we used that

$$\left\langle \partial_\nu\left(\delta\vec{x}^T\,\mathbb{H}(\Phi)\,\delta\vec{x}\right)\right\rangle_{p(\vec{x};\Phi)} = \frac{1}{M}\mathrm{Tr}[\mathbb{H}^{-1}\partial_\mu\mathbb{H}(\Phi)] \tag{4.21}$$

To sum up, the hessian of a model with a distribution given by Eq. 4.7

can we calculated as:

$$\mathscr{H}_{\mu,\nu} = \mathscr{H}^{0}_{\mu,\nu} + \mathscr{H}^{1,1}_{\mu,\nu} + \mathscr{H}^{1,2}_{\mu,\nu} \tag{4.22}$$

$$\mathscr{H}^{0}_{\mu,\nu} = \frac{M^2}{4}\mathrm{Tr}[\mathbb{H}^{-1}\partial_\mu \mathbb{H}(\Phi)]\mathrm{Tr}[\mathbb{H}^{-1}\partial_\nu \mathbb{H}(\Phi)] \tag{4.23}$$

$$\mathscr{H}^{1,1}_{\mu,\nu} = -\frac{M}{2}\mathrm{Tr}[\mathbb{H}^{-1}\partial_\mu \mathbb{H}(\Phi)]\mathrm{Tr}[\mathbb{H}^{-1}\partial_\nu \mathbb{H}(\Phi)] \tag{4.24}$$

$$\mathscr{H}^{1,2}_{\mu,\nu} = \frac{M}{2}\,\mathrm{Tr}[\mathbb{H}(\Phi)\mathbb{K}_{\mu,\nu}] \tag{4.25}$$

For an application to the Schelling model within the boundaries see Appendix A.1 and Appendix A.2 the calculation within the boundaries.

Alternatively, we can be more general about the probability distribution and only assume that $p(\vec{x};\Phi)$ is a smooth function of Φ. In this case, small changes in Φ can be approximated by the first-order approximation for the change in $p(\vec{x};\Phi)$:

$$\partial_\mu p(\vec{x};\Phi) = \nabla \cdot \left[V_\mu(\vec{x})p(\vec{x};\Phi)\right], \tag{4.26}$$

where $V_\mu(x) = \langle \frac{\delta \vec{x}}{\delta \Phi_\mu} \rangle_{\vec{x}}$ is a field of *velocity changes* that describes how $p(x;\Phi)$ shifts locally in response to parameter changes. This field cannot be computed directly but must be estimated by smoothing the possible dynamics of the model with a kernel density estimation. With this, one can then estimate the Hessian in Eq. 4.6.

These assumptions can be further used to extend the sloppiness approach to probabilistic models.

4.6. Conclusions

We have shown that both agent-based models and DSGE models have a sloppy phenomenology, where their dynamics depend only on a few key stiff parameter directions, much like many models in all fields of science [180]. In the case of the Mark-0 agent-based model, the stiff parameter directions point towards close phase transitions. Exploiting these key directions, we developed an algorithm to systematically explore the phase space of agent-based models, which should apply universally (perhaps at the cost of tweaking the exploration parameters defined in Eq. (4.3)). Applying this algorithm to the Mark-0 model, we have recovered the four phases that were identified in the prior literature, both in low and high-dimensional cases. This method of exploration also appears to be more computationally efficient as the number of parameter dimensions grows.

Since agent-based models are able to generate a rich phenomenology, this tool may aid in understanding the possible scenarios in an agent-based model, as well as their robustness to changes in the underlying parameterisation. This suggests several avenues of further exploration, including extending our analysis to a wider variety of observable outcomes and more complex agent-based models, as well as potentially combining it with a meta-modelling approach to not only infer the set of phases but also their boundaries. Another direction, sug-

gested to us by J. Sethna, would be to follow sloppy directions instead of stiff directions, with the aim of removing irrelevant (combination of) parameters and constructing minimal models in a systematic way once the different behaviors have been classified, along the lines of Quinn et al. (2021) [180]. Building *reduced form* models using such a systematic procedure appears to us as a particularly exciting perspective.

Key Takeaways

1. **Sloppiness in Macroeconomic Models:** Both Agent-Based Models (ABMs) and Dynamic Stochastic General Equilibrium (DSGE) models exhibit sloppiness, where observable dynamics are primarily influenced by a few stiff parameter directions, while many others are relatively insignificant.
2. **Phase Transitions in ABMs:** In the Mark-0 ABM, stiff directions correspond to nearby phase transitions. Exploiting these directions enables efficient exploration of phase boundaries and different dynamic behaviors.
3. **Algorithm for Parameter Space Exploration:** A probabilistic algorithm leveraging stiff directions is introduced, which systematically and efficiently explores the phase space of ABMs, revealing all major dynamic regimes with minimal computational cost.
4. **Agnostic Approach to Dynamics:** Unlike meta-modelling methods, the proposed algorithm is agnostic, requiring no prior assumptions about the system's dynamics, making it suitable for discovering unexpected emergent phenomena.

Part II: Occupational Mobility & the Structure of Worker Transitions

Labor Markets as Complex Systems: A Quantitative Framework

5.

The intersection of complex systems and socioeconomic analysis has opened a new frontier in understanding labor markets, shifting the perspective from simple aggregations of employment transactions to dynamic, interdependent networks. Leveraging principles from complex systems, researchers have begun to unravel the intricate structure of economic networks, uncovering insights that transcend traditional economic models [181, 182]. In this view, labor markets are seen as complex, adaptive systems, where micro-level interactions—such as employment transitions and wage negotiations—combine to produce macro-level phenomena like income inequality, unemployment cycles, and workforce polarization [45, 183, 184].

The constant flux of job openings, hires, and separations mirrors the *statistical life* of particle dynamics in physical systems. Here, the labor movement is influenced by factors such as skill levels, industry demand, and economic conditions, while collectively forming discernible patterns [48, 76].

Applying methods from statistical mechanics and network theory researchers study employment dynamics, including probabilities associated with job transitions and wage evolution [57, 185]. For example, network models represent industries as interconnected nodes, illustrating how shocks in one sector propagate through others, influencing job availability and wage distributions in complex and often nonlinear ways [47, 48, 186]. These studies reveal that labor markets, much like ecosystems, exhibit emergent properties: small, localized changes can cascade into significant economic shifts, akin to phase transitions in physical systems.

This chapter leverages these principles to quantitatively analyze labor market dynamics, emphasizing the stochastic and structural properties of occupational transitions. The French labor market serves as the case study, using detailed transition data to explore occupational sizes, mobility patterns, and the structural imbalances that indicate the presence of hierarchies and bottlenecks within the French labor market—a finding consistent with similar observations in other countries [186].

The chapter starts by introducing the dataset and methods, investigating power-law distributions in occupation sizes and transition flows (Section 5.2 & 5.3).

To deepen our understanding of labor market dynamics, we focus on the statistical properties of the occupational transition matrix. This matrix provides a stochastic perspective on the labor market, with each entry representing the probability of a worker in a given occupation transitioning from one occupation to another (Section 5.4).

Then we model the labor market transitions as a markov chain, identifying discrepancies such as underestimated occupational *stickiness* in theoretical predictions. This leads to the adoption of a Mover-Stayer

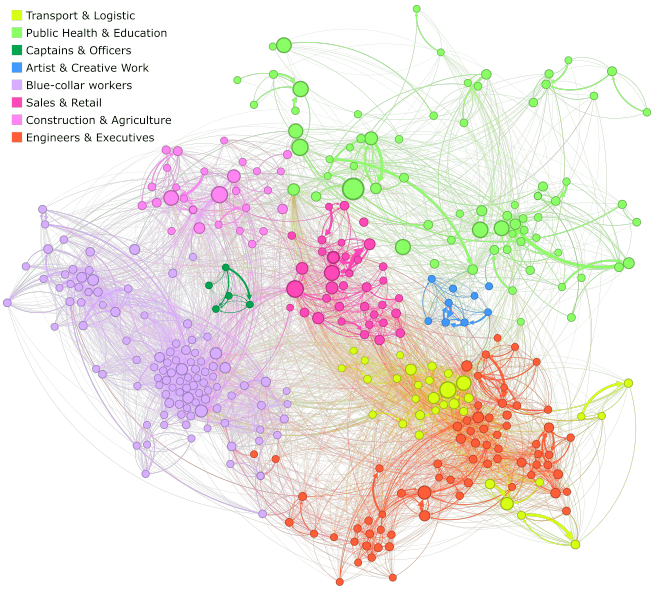


Figure 5.1.: Occupational Transition Network of France: Graph visualization of the weighted and directed labor market network derived from the transition probability matrix $\mathcal{P}_{i,j}$ (see Eq. 5.8), computed from data spanning the years 2012 to 2020. Each node symbolizes an occupation, with links that illustrate transitions between them. Node sizes correspond to the occupation's workforce size, while node color indicates its BRIM community [187] (see Section 3). Line widths are proportional to the transition probability $\mathcal{P}_{i,j}$. The layout is based on the OpenOrd algorithm [188], and the graph was generated using Gephi [189].

model, which better captures heterogeneous worker behaviors and aligns predictions with empirical data (Section 5.6).

The transition matrix serves as a tool for examining how labor markets evolve toward equilibrium. By analyzing its eigenvalues and eigenvectors, we gain insights into both the rate at which this equilibrium is reached and the dominant modes of occupational transitions. The steady-state distribution, given by the leading eigenvector, reveals the long-term allocation of workers across occupations, while subdominant eigenvalues reflect the temporal dynamics of workforce adjustments (Section 5.7).

Building on the foundation of detailed balance, we explore the statistical and spectral properties of the occupational transition matrix, offering a quantitative lens through which to interpret labor market flows and their equilibrium states (Section 5.8 & 5.9).

Finally, we interpret the occupational transition matrix as a weighted, directed network where nodes represent occupations and edges capture transitions. Clustering techniques like Louvain, Infomap, and BRIM uncover community structures, revealing clusters of closely connected occupations (Section 5.11). This analysis highlights the labor market's modular and hierarchical organization, identifying key occupations that act as bridges or bottlenecks in workforce mobility.

5.1. French Labor Market Data

We use the BTS-POSTES dataset provided by the National Institute of Statistics and Economic Studies of France (INSEE) for the years 2010 to 2021.[1] The dataset covers all employees in France for a sequence of two-year snapshots, such that one can track an individual through the years $T-1$ and T.[2] We consider an individual's change in *primary* occupation within a $(T-1, T)$ year period. The primary occupation is defined as occupation with the highest net salary provided that at least 30 days have been worked in the job and the salary is at least the mini-

1: Access to this confidential data was gained through the Centre d'accès sécurisé aux données (Ref. 10.34724/CASD). The persistence identifier for the dataset of the year 2020 is https://doi.org/10.34724/CASD.21.5012.V1

2: We cannot track individuals over more than two years because their unique identifiers are renewed and change each year.

mum wage.[3] In case of a tie, the job with the highest number of hours worked is taken.[4] The minimum salary threshold refers to the minimum wage for full-time, full-month employment in France, ensuring that primary occupations considered in the analysis meet a baseline for economic activity. Individuals without a primary occupation are classified as being *out of the occupational network*, which includes those without an active working contract or primary occupation. Note that individuals classified as out of the occupational network may not strictly align with the International Labour Organization (ILO) definition of unemployment, which requires active job-seeking and availability to start work within two weeks [190]. Due to data limitations, we cannot determine whether these individuals are actively seeking employment. An occupation is defined by the nomenclatures of Professions and Socio-professional Categories of Salaried Jobs of Private and Public Employers (PCS-ESE). Here we use the 4-digit PCE-ESE nomenclature. Based on the classification, we track year-on-year occupational transitions by analyzing the number of individuals moving from occupation i at time $T-1$ to occupation j at time T , denoted as N_{ij}^{T}.[5] We naturally exclude job changes where an individual has an extended unemployment spell, whether this be due to actively searching for a job, training, leaving or entering the workforce, or entering extended periods of leave such as maternity leave. Furthermore, we do not account for changes in companies or location if the individuals remain within the same occupational code. This ensures a focus on meaningful transitions that reflect changes in occupation rather than lateral shifts within the same job classification.

3: The indicator is defined by INSEE and is called PPS (poste principal du salarié).

4: This ensures that each individual is assigned only one job transition per year, eliminating the possibility of accounting for multiple occupation changes within the same year.

5: For simplicity, we often omit the superscript t in the text.

Understanding labor mobility requires examining realized transitions between occupations in ways that capture both the scale and structure of movement. Two complementary approaches are commonly used to represent these transitions: one focuses on absolute transition counts, while the other uses transition probabilities.

The absolute transition counts matrix, used in Sections 5.2 and 5.3, captures the raw number of workers moving between occupations. This approach highlights the total volume of flows along each pathway, making it particularly valuable for assessing the aggregate impact of labor flows, identifying high-demand sectors, and understanding workforce dynamics at scale. By emphasizing the magnitude of transitions, it provides insights into the relative importance of specific occupations in absorbing or supplying labor.

On the other hand, the transition probability matrix, introduced later in Section 5.4 and Chapter 6, normalizes transitions by the total outflows from each occupation, representing the likelihood of moving from one role to another. This probabilistic approach offers a more nuanced view of worker mobility patterns, especially when comparing occupations of varying sizes or transition volumes. By isolating the structural dynamics of transitions, it captures patterns independent of workforce scale.

Building on these representations, another perspective comes from the skill-based network, which maps potential transitions between occupations based on shared competencies. This network identifies pathways deemed feasible due to skill similarity, offering insights into mo-

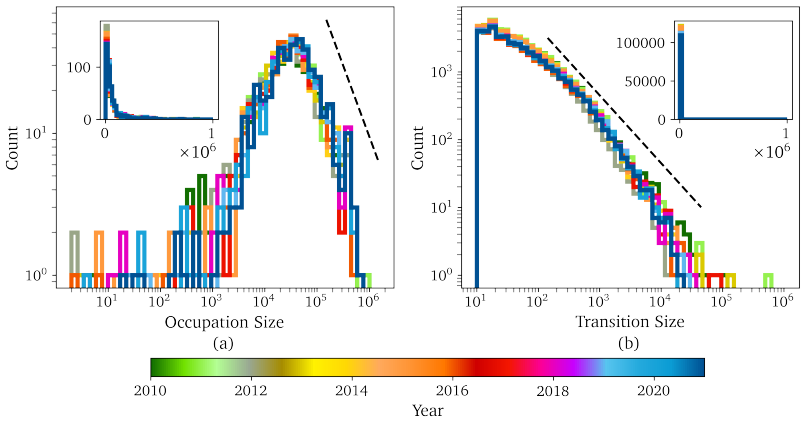


Figure 5.2.: Power-law distributions in the French Labor Market: (a) Log-log plot of the empirical probability density function for occupation sizes. The inset displays the same data with linear-linear axes. (b) Log-log plot of the empirical probability density function for transition sizes between occupations. The inset shows the data with linear-linear axes. Black dotted lines represent fitted power-law trends (Eq. 5.2) as a guidance for the eye. Different colors correspond to distributions for various years.

bility opportunities independent of realized transitions. While the observed transition network captures actual labor market flows, the skill-based network highlights ex-ante possibilities, and we investigate its structure and implications further in Chapter 7.

5.2. Power Law Distributions in the Labor Market

Power-law distributions are ubiquitous across diverse systems, ranging from phase transitions in physics to scaling laws in biological systems and the well-known examples of income distributions and city size scaling laws. Income distributions, first analyzed by Vilfredo Pareto [191], revealed the unequal distribution of income, where a small fraction of the population controls a disproportionately large share of resources, while the majority earns considerably less. This phenomenon is similarly observed in the scaling of city sizes [192] and the distribution of company sizes [193]. In this chapter, we explore distributions within the labor market to determine whether they exhibit comparable patterns to those found in other systems.

Mathematically, the probability density function (PDF) of a power-law distribution is approximated as follows

$$\log p(x) = \alpha \log x + c \tag{5.1}$$
$$\Rightarrow p(x) = Cx^{-\alpha} \tag{5.2}$$

with a constant $C = \mathrm{e}^c$ and the exponent α.

We study two distributions from labor markets that shed light on the structural imbalances: (1) the distribution of employers across different occupations $\sum_i N_{ij}$ and (2) the number of people transitioning from one occupation to another N_{ij} (Figure 5.2). The power-law distribution for both quantities is evident in the plot, particularly in the linear-linear representation shown in the inset of Figure 5.2. To illustrate this more clearly, we analyze the distributions in a double logarithmic plot, where the linear decay of the tail confirms the expected power-law behavior (see Figure 5.2 and Eq. 5.1). In comparison to the

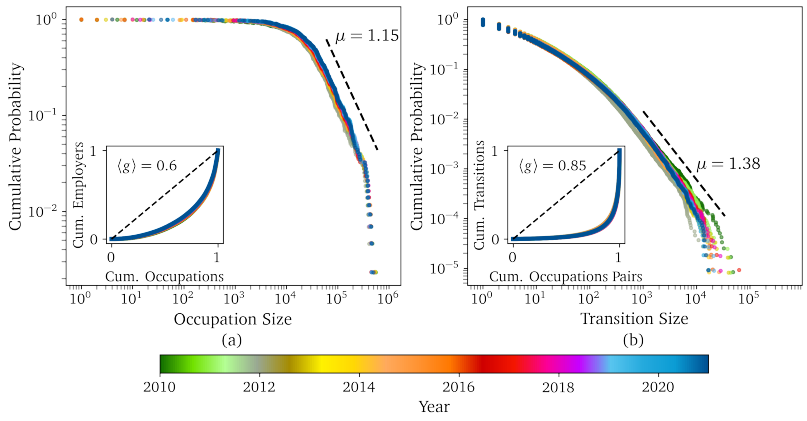


Figure 5.3.: Cumulative distributions of the French Labor Market: (a) Log-log plot of the empirical cumulative density function for occupation sizes, with a dotted line representing the fit to Equation 5.4 and an exponent $\mu = 1.15$. The inset shows the corresponding Lorenz curve, with an average Gini coefficient of $\langle g \rangle = 0.6$ across years. (b) Log-log plot of the empirical cumulative density function for transition sizes between occupations, with a dotted line representing the fit to Equation 5.4 and an exponent $\mu = 1.38$. The inset displays the Lorenz curve, with an average Gini coefficient of $\langle g \rangle = 0.85$ across years.

transition size distribution, the occupation size distribution exhibits a left tail, indicating that certain occupations have only a small number of employers. We could characterize the right tails by fitting a linear function to them, however for higher x values the number of points in each histogram bin decreases, leading to empty bins causing a non-linear appearance, which makes the fitting difficult. To address this issue, one typically shifts to the cumulative distribution function, which avoids the binning problem and facilitates the derivation of the power-law exponent. The cumulative distribution function is defined as

$$P_>(x) = \int_x^\infty p(x')\mathrm{d}x' \tag{5.3}$$

$$\Rightarrow P_>(x) \propto x^{-\mu} \tag{5.4}$$

with this, the power law exponents are linked through $\alpha = \mu + 1$. The cumulative distribution function (CDF) intuitively represents the probability of observing a value greater than or equal to x . For a power-law distribution, the CDF decays as $x^{-\mu}$, making it a useful tool for visualizing the tail behavior and directly linking the exponent μ to the underlying power-law relationship.

By fitting a linear function to the double-logarithmic plot of the cumulative distribution function (see Figure 5.3) we estimate the power-law exponent that characterizes the tail behavior of the distributions. Intuitively, smaller values of μ correspond to a slower decay of the tail, indicating greater inhomogeneity in the system. In other systems, power-law exponents are typically observed around $\mu = 1$. For instance, city size distributions exhibit $\mu \approx 1$ [194], company sizes follow $\mu \approx 1$ [195], stock market returns normalized by volatility yield $\mu \approx 3$ [196], and income distributions typically fall within μ values of 1.5 to 3 [191].

The occupation size distribution is characterized by an exponent $\alpha = 1.15$, indicates that a small number of occupations employ a large share of workers, while many other occupations have relatively few employers. This reflects an inherent asymmetry driven by factors such as economic demand, and the versatility of broadly applicable skills and tasks across companies. These *dominant occupations* serve as central hubs of labor market activity, attracting a disproportionate share of employers. Conversely, occupations with a niche or specialized employer base are represented in the left tail of the distribution.

The distribution of transitions between occupations follows a power-law pattern with an exponent $\alpha = 1.38$. This implies that certain roles act as nodes or *bridges* in the labor network, allowing individuals to move more freely between jobs. Occupations with a high number of transitions may either require transferable skills that facilitate movement across roles or be associated with industries that experience rapid growth and turnover, such as technology, finance, or sales. These occupations essentially act as central connectors, enabling fluidity and flexibility for workers who transition across occupations in response to economic shifts, skill demands, or career progression.

On the other hand, many occupations see minimal transitions, often due to specialized skills or industry isolation. For instance, highly specialized technical or niche occupations may have fewer pathways to other occupations, leading to a relatively static position within the labor market. This structure highlights bottlenecks in the labor market, where a few occupation pairs facilitate high levels of transitions, while the majority of occupations experience low mobility and limited exchange.

Both distributions remain stable over several years, with no significant variation in the power-law exponents observed across different time periods for either distribution.

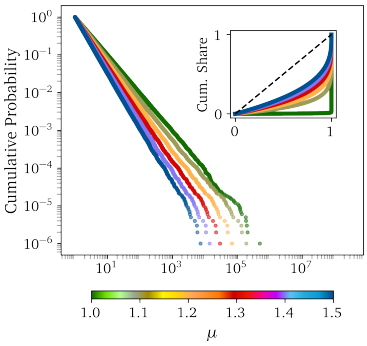
The distributions show an inhomogeneous and interdependent labor market, where dominant roles and transition-heavy occupations underpin the workforce network, mirroring hierarchies and imbalances observed in complex systems.

To further illustrate the degree of inhomogeneity[6] in labor market distributions, we can examine the *Lorenz curve*. This graphical representation, often used to depict income or wealth inequality, is equally effective for analyzing other forms of distributional imbalance. The Lorenz curve plots the cumulative share of a resource (e.g., employment or occupational transitions) against the cumulative share of units (e.g., individuals or occupations), offering a visual tool to measure how concentrated or dispersed a variable is across a population.

Assuming there are L occupations sorted by the number of people employed in each occupation, N_l, we calculate the proportion of the smallest k occupations as $S_k = \frac{\sum_{l=0}^{k} N_l}{\sum_{l=0}^{L} N_l}$. The Lorenz curve is then defined by the points $(\frac{k}{L}, S_k)$ which forms a smooth curve for large L. When compared against a 45-degree line representing perfect equality, the Lorenz curve reveals the extent of deviation from homogeneity across occupations. This deviation can be quantified using the Gini coefficient g, which is calculated as the fraction of the area between the line of equality and the Lorenz curve relative to the total area under the line of equality.

The relationship between the power-law exponent and the Lorenz curve is shown in Illustration 1. Lower values of μ correspond to a slower decay of the power-law tail, which increases inequality as reflected by a higher Gini coefficient. For $\mu \leq 1$, the Gini coefficient reaches $g = 1$, indicating a state of extreme inequality, or oligarchy, where a finite number of entities control a significant portion of the system's resources.

6: We intentionally use the term *homogeneity* instead of the more common *equality* in this context, as we aim to avoid implicitly assigning a positive or negative value judgment to the structure of the labor market.



Ill. 1: Relationship Between μ and the Gini Coefficient: Cumulative distribution functions for $\mu \in [1.0, 1.5]$, with the corresponding Lorenz curves shown in the inset.

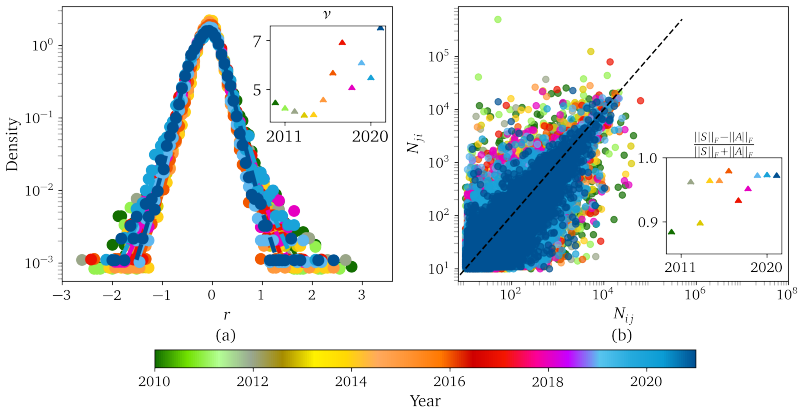


Figure 5.4.: Distribution of occupational flow ratios: (a) Empirical probability density function of the occupational flow ratio r (Eq. 5.5). The dotted lines represent the fit using a Student's t-distribution (Eq. 5.6), with the parameter ν of the fit shown in the inset. (b) Joint distribution of the back-and-forth transitions N_{ij} and N_{ji} as a scatter plot. The inset depicts the symmetry index for the distribution (Eq. 5.7). Different colors correspond to different years.

We find in the occupation size and transition size distribution, that the inhomogeneity of remains relatively stable over time, with a slight trend toward reduced inhomogeneity (see Figure 5.3, insets). Although the power-law exponent for the transition size distribution is higher than that for the occupation size distribution, the Gini coefficient is lower (average Gini of 0.6 and 0.85, respectively). This discrepancy arises because the Gini coefficient accounts for the entire distribution, while the power-law exponent primarily characterizes the tail behavior.

5.3. Occupational Flow Ratios in the Labor Market

Let N_{ij} represent the number of people who transitioned from occupation i to occupation j and N_{ji} represent the number of people who transitioned from occupation j to i. To capture the directional bias of these transitions, we define an *occupational flow ratio* r as:

$$r_{ij} = \log_{10}\left(\frac{N_{ji}}{N_{ij}}\right). \tag{5.5}$$

This flow ratio r provides a quantitative measure of the relative balance in occupational transitions. When $r \approx 0$, the transitions between occupations i and j are nearly balanced, with roughly the same number of people moving in each direction. When $|r| \gg 0$, however, it indicates a directional bias in the flow of workers, with one occupation drawing more individuals from the other.[7]

Figure 5.4 (a) shows the distribution of r for each pair of i and j. The distribution is *heavy tailed* and can be well fitted by a student's t-distribution.[8] Interestingly, this pattern can be observed also in other systems for example in company growth rates [197]. While most occupation pairs experiencing relatively balanced flows $r \approx 0$, a small number exhibiting strong directional biases. The occupation flow ratio is in the range of $r \in [-2, 2]$ and means intuitively that at the extremes, the flow between two occupations can differ by a factor of 100 in one direction compared to the other.

This heavy-tailed nature of the distribution has the following implica-

7: Note that we count each occupation pair (i, j) only once. If we were to include both (i, j) and (j, i), the flow ratio would be symmetric by definition, $r_{ij} = \log\left(\frac{N_{ji}}{N_{ij}}\right) = -\log\left(\frac{N_{ij}}{N_{ji}}\right) = -r_{ji}$.

8: The probability density function of the Student's t distribution is

$$t_\nu = \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\sqrt{\nu\pi}\,\Gamma\left(\frac{\nu}{2}\right)}\left(1 + \frac{x^2}{\nu}\right)^{-\frac{\nu+1}{2}} \tag{5.6}$$

.

tions

1. **Occupational Hierarchies and Attractiveness:** Occupations with high positive or negative r values reflect an asymmetry in worker flows, suggesting that one occupation consistently attracts or loses workers relative to the other. Such imbalances may result from differences in occupational desirability, evolving skill acquisition, or time-dependent sectoral demand. For instance, a high positive r for transitions from occupation i to j suggests that workers view occupation j as more desirable or stable, while occupation i may function as a stepping stone or transitional role.
2. **High-Transition Occupations as Labor Market Channels:** The distribution's heavy tail indicates that certain occupations play a pivotal role in the labor market, either acting as *hubs* that attract substantial inflows from other occupations or as *diffusers* that facilitate significant outflows to a wide range of roles. These roles, often marked by highly transferable skills, create central nodes within the labor network, enabling workers to transition fluidly to various other positions.

The occupational flow ratio distribution provides a snapshot of labor market dynamics, highlighting both stable and asymmetric flows between occupations. Over the years, the distribution becomes more centered around zero, reflecting increasingly balanced flows, as indicated by a rising ν parameter in the Student's t-distribution fit (see Figure 5.4(a) inset).

Symmetry in Labor Market Transitions

Instead of looking at the directional flow ratio r, we can directly look at the joint distribution of N_{ij} and N_{ji} in form of a scatter plot (see Figure 5.4(b)). We can observe a symmetry of points around the diagonal line $y = x$, or in mathematical terms, the density of points within the rectangle spanning x_1 to $x_1 + \mathrm{d}x_1$ and x_2 to $x_2 + \mathrm{d}x_2$ remains unchanged if the axes were swapped. This does not imply that the number of people transitioning from occupation i to j are exactly equal to the number transitioning from j to i. Instead, this symmetry indicates that for every pair of occupation where the flow from i to j exceeds the flow from j to i by a certain amount Δ , there is an equivalent pair where the flow ratio has the same magnitude Δ , but in the opposite direction. In other words, while individual flows may differ, the overall distribution of flow ratios is balanced, reflecting a statistical equilibrium rather than exact equality in transitions between specific pairs.

Although our empirical analysis does not show statistical significance of symmetry at the 5% level[9], the symmetry in transitions remains remarkably close. The inset of Figure 5.4(b) displays the symmetry index[10], ranging from -1 (perfect asymmetry) to 1 (perfect symmetry), with values exceeding 0.9. This near-balance suggests that, for many occupation pairs, the exchange of workers is relatively stable and reciprocal over time across the whole labor market. Microscopically, transitions between many occupations appear to be bidirectional ($r \neq 0$), reflecting a dynamic system, but macroscopically, the overall labor

9: To test statistical significance, we perform a 2D Kolmogorov-Smirnov test under the hypothesis that the joint distributions $p(N_{ij}, N_{ji})$ and $p(N_{ji}, N_{ij})$ are drawn from the same underlying distribution [198]. Note that we count the pairs $i \to j$ and $j \to i$ only once, otherwise the distributions are symmetric per definition. In other words, the first object only includes $i < j$ and the second object includes $i > j$ to avoid double counting.

10: A matrix M can be decomposed into its symmetric part, $S = \frac{M+M^T}{2}$, and its asymmetric part, $A = \frac{M-M^T}{2}$. The symmetry index is then defined as

$$s(M) = \frac{\|S\|_F - \|A\|_F}{\|S\|_F + \|A\|_F}, \qquad (5.7)$$

where $\|.\|_F$ represents the Frobenius norm $\|M\|_F := \sqrt{\Sigma_{i=0}^{N} \Sigma_{j=0}^{N} |m_{ij}|^2}$, which measures the overall magnitude of the matrix entries. If the matrix is fully symmetric, $s(M) = 1$, while if the matrix is full asymmetric, $s(M) = -1$.

market is nearly balanced such that the distribution of flow imbalances is symmetric, with as many transitions favoring one direction as the other across all occupation pairs. This balance, while not absolute, underpins the stability of the labor market, highlighting a structure where local hierarchies and bottlenecks exist, yet the occupational mobility remains globally cohesive and interconnected.

5.4. Occupational Transition Matrix

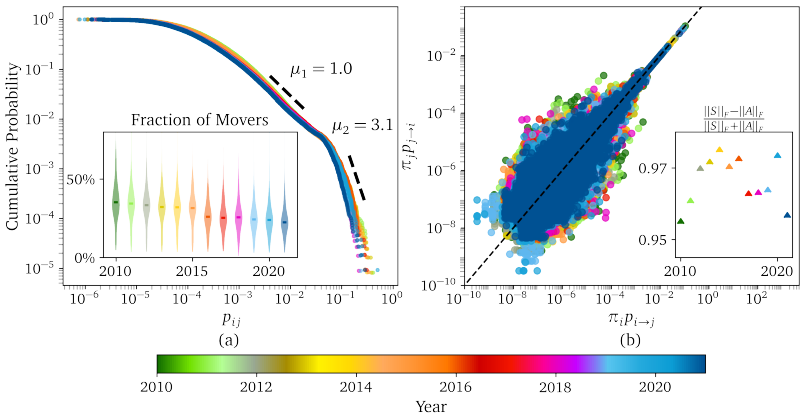


Figure 5.5.: Distribution of the Occupational Transition Matrix $\mathcal{P}^t_{ij}$: (a) Empirical cumulative distribution function of transition probabilities p_{ij}. The dotted lines represent power-law fits (Eq. 5.4) for two distinct regimes, with exponents $\mu_1 = 1.0$ and $\mu_2 = 3.1$. Inset shows a violin plot of the fraction of movers ($\sum_{i\neq j} N_{ij}$) for each year. (b) Scatter plot of the joint distribution for the detailed balance condition (Eq. 5.24). The inset displays the symmetry index of the spectral decomposition V_t (see eqs. 5.7 and 5.25). Different colors correspond to different years.

We construct the occupational transition probability matrix $\mathcal{P}_{i,j}$, which captures the likelihood of workers transitioning between occupations over consecutive years. Each entry $\mathcal{P}^T_{i,j}$ in the matrix is defined as:

$$\mathcal{P}^T_{i,j} = \mathbb{P}(X^T_i = j | X^{T-1}_j = i) = \frac{N_{ij}}{\sum_i N_{ij}}. \tag{5.8}$$

where N_{ij} is the number of individuals transitioning from occupation i at time $T-1$ to occupation j at time T. Intuitively, $\mathcal{P}^T_{i,j}$ represents the realized transition rate from occupation i to j within a given year. By normalizing each row of the matrix ($\sum_j \mathcal{P}^T_{i,j} = 1$), it can be interpreted as an empirical estimation of the underlying probability of transitioning from i to j. This year-on-year structure reflects observed labor market flows, offering a probabilistic view of workforce dynamics.

Figure 5.5(a) presents the cumulative distribution function of the off-diagonal entries in $\mathcal{P}^T$.[11] The CDF reveals a characteristic power-law behavior, with two distinct regimes. The first regime, corresponding to lower probabilities ($\mathcal{P}^T_{i,j} \lesssim 0.1$), exhibits a tail with an exponent $\mu_1 = 1$, indicating a broad range of *distant* transitions with lower probabilities. The second regime, observed for higher probabilities ($\mathcal{P}^T_{i,j} \gtrsim 0.1$), has a steeper decay with an exponent $\mu_2 = 3.1$, reflecting a smaller number of *close* transitions with significantly higher probabilities. This bifurcation highlights the dual structure of the labor market, where certain occupation pairs are strongly connected while most remain weakly linked.

The inset of Figure 5.5(a) shows the fraction of movers, defined as $\frac{\sum_{i\neq j} N_{ij}}{\sum_i N_{ij}}$. On average, approximately 25% of individuals transition out of their current occupation each year, while the remaining 75% stay

11: Diagonal elements, which represent the probability of remaining in the same occupation, are excluded from the analysis.

within the same occupation. Notably, the fraction of movers varies widely across occupations, as reflected in the substantial deviation from the mean. This variation underscores heterogeneous mobility rates, with some occupations exhibiting much higher or lower mobility rates than others.

5.5. Markov Chain

Labor market transitions can be analyzed through the lens of a discrete-time Markov chain, based on the year-on-year transition probability matrix defined in Equation 5.8. This approach assumes that the probability of moving from one occupation to another depends solely on the current occupation and not on prior occupations, as formalized by the memorylessness property:

$$\mathbb{P}(X_{n+1} = x | X_1 = x_1, X_2 = x_2, ..., X_n = x_n) = \mathbb{P}(X_{n+1} = x | X_n = x_n) \tag{5.9}$$

This framework provides a first-order probabilistic approximation of labor market dynamics, allowing us to analyze occupational flows and derive long-term behaviors such as the steady state. However, adopting the Markov assumption involves several *strong* assumptions, some of which are realistic while others simplify the complexity of real-world labor markets. We outline these assumptions, their limitations, and their relevance to this work.

- **Memorylessness:** The Markovian property (Eq. 5.9) assumes that a worker's future transition depends only on their current occupation, not on their prior career trajectory. While this simplifies the analysis, it overlooks well-documented path dependencies in career development, where previous roles influence skill accumulation, experience, and preferences (see e.g. [199]). Despite this limitation, the assumption serves as a useful first-order approximation for capturing immediate transition dynamics, enabling aggregate-level insights into labor market flows.
- **Stationarity of Transition Probabilities:** Stationarity assumes that transition probabilities remain constant over time, i.e., $p_{i,j}(t) = p_{i,j}\ \forall i, j$. While this assumption may not always hold, especially during periods of disruption, it is a reasonable approximation during stable periods (see Appendix B.2). For example, the decade preceding the COVID-19 pandemic shows relatively consistent labor market dynamics. Empirical stationarity during such periods justifies the framework.
- **Complete State Space:** The completeness assumption requires that all relevant occupations are included in the transition matrix. Using the four-digit PCS occupational codes, which have remained stable since 2009, we ensure that all occupations are represented. While these codes may not capture the full granularity of job roles or evolving interpretations, they provide a comprehensive classification suitable for modelling labor market dynamics.
- **Strong Connectedness:** This assumption posits that every occu-

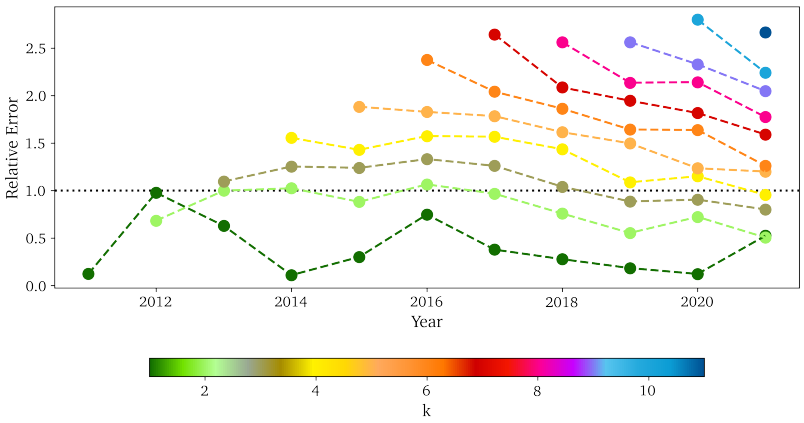


Figure 5.6.: Comparison of next-state predictions with $\mathcal{P}^t$ and $\hat{\mathcal{P}}$: Plot of the relative error Eq. 5.14 for different forecasting horizons k across various years. Each point corresponds to the relative error of the prediction $\vec{x}'_{t+k} = \left(\Pi_{\tau=t}^{t+k-1}\mathcal{P}^\tau\right)\vec{x}_t$, plotted at the year $t+k$. The dotted line represents the line of equality, where predictions using $\mathcal{P}^t$ are equal to those using $\hat{\mathcal{P}}$.

pation is reachable from any other, either directly or through intermediate steps. Our data show that, over long horizons, all occupations are eventually connected. While this does not imply seamless transitions in practice—skill gaps and barriers exist—it ensures that the framework remains valid for modelling long-term mobility.

- **Exclusion of Exogenous Shocks:** Markov chains assume a stable environment, excluding the influence of external shocks like economic crises, technological disruptions, or policy changes. Although this limits applicability during periods of crisis, the framework is robust for analyzing endogenous labor market dynamics during stable periods.

In summary, while the assumptions underlying the Markov chain framework—such as memorylessness and stationarity—are strong and may not fully reflect the complexity of real-world labor markets, they provide a simplified yet practical approach for studying occupational transitions. These assumptions enable us to compare transition matrices over time, particularly in terms of their dynamical properties, offering valuable insights into aggregate patterns. Similarly, although the explicit steady state may not carry direct economic meaning, it remains a useful tool for analyzing and comparing the structural properties of transition matrices across different years, helping to uncover persistent trends and deviations caused by external factors.

5.6. Markovian Assumption and Empirical Data

A common method for testing the validity of the Markov chain framework is to compare the empirical k-step transition matrix—representing the observed probabilities of transitioning from occupation i to j over k years—with the calculated k-step transition matrix, $\mathcal{P}^k$, derived by iteratively applying the one-year transition matrix. This approach is based on the Markov property, which states:

$$\vec{x}_{t+1} = \mathcal{P}^{t+1}\vec{x}_t, \tag{5.10}$$

where $\vec{x}_t$ is the state vector representing the labor market at time t. Using this property, the k-step transition can be calculated recursively

as:

$$\vec{x}_{t+k} = \mathcal{P}^{t+k} \dots \mathcal{P}^{t+2}\mathcal{P}^{t+1}\vec{x}_t = \left(\Pi_{\tau=t}^{t+k}\mathcal{P}^{\tau+1}\right)\vec{x}_t \tag{5.11}$$

Prior research [49, 200] has revealed a significant discrepancy between the empirical k-step transition matrix and the theoretically computed one. Specifically, the diagonal elements of the theoretical transition matrix—representing the probabilities of workers staying in the same occupation—are consistently underestimated. This suggests that the theoretical Markov chain model predicts far more transitions than are observed in reality. For example, as shown in Illustration 2, the theoretical k-step transition matrix for 10 steps predicts a staying probability of approximately 10%, meaning only 1 in 10 individuals would remain in their occupation. In contrast, empirical data shows a much higher staying probability, highlighting the *stickiness* of occupations in real labor markets [49].

To address this discrepancy, Blumen et al. (1955) [200] developed the Mover-Stayer model, which introduces a more nuanced framework to account for heterogeneous worker behaviors. This model assumes the existence of two distinct groups within the workforce:

1. Stayers: Workers who never leave their original occupations.
2. Movers: Workers who are open to transitioning between occupations.

The revised transition matrix under this framework is expressed as:

$$\tilde{\mathcal{P}} = S + (I - S)\mathcal{P}, \tag{5.12}$$

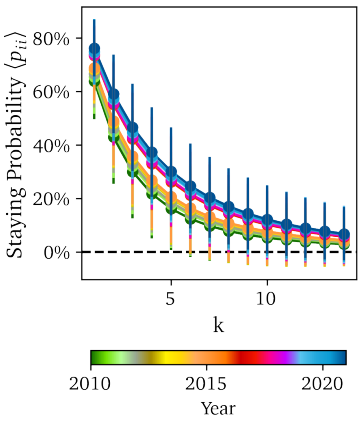
where S is a diagonal matrix representing the proportion of stayers, I is the identity matrix, and P is the original transition probability matrix. The k-step transition matrix for this revised model becomes:

$$\tilde{\mathcal{P}}_{T,k} = S + (I - S)\mathcal{P}_T^k. \tag{5.13}$$

This formulation better aligns with empirical data, as it accounts for the high probability of workers remaining in their current occupations. By incorporating the distinction between movers and stayers, the model effectively captures the *stickiness* observed in labor market transitions. The parameters of this model, including the proportions of stayers and movers, can be estimated using techniques such as maximum likelihood estimation [201] or Bayesian inference [202].

However, in our dataset, this test is not feasible because we cannot track individuals across more than two years due to changes in unique identifiers.

Instead, we adapt this framework to evaluate the predictive accuracy of the Markov model by comparing two approaches for forecasting future labor market states: time-dependent transition matrices $\mathcal{P}^t$ and the average transition matrix $\hat{\mathcal{P}} = \frac{1}{T}\sum_t \mathcal{P}^t$, which aggregates yearly transition matrices over the observation period (2010–2021). While not identical to the original method, this adaptation still provides insights into short-term dynamics captured by $\mathcal{P}^t$ and long-term structural trends revealed by $\hat{\mathcal{P}}$, allowing us to explore both immediate fluctuations and stable patterns in labor market transitions.



Ill. 2: Staying Probability as a function of k: The staying probability is represented as the mean of the diagonal elements $\left(\Pi_{\tau=t}^{t+k}\mathcal{P}^{\tau+1}\right)_{ii}$. Error bars indicate the standard deviation across different occupations, with different colors corresponding to different years. The dotted line marks a staying probability of 0%.

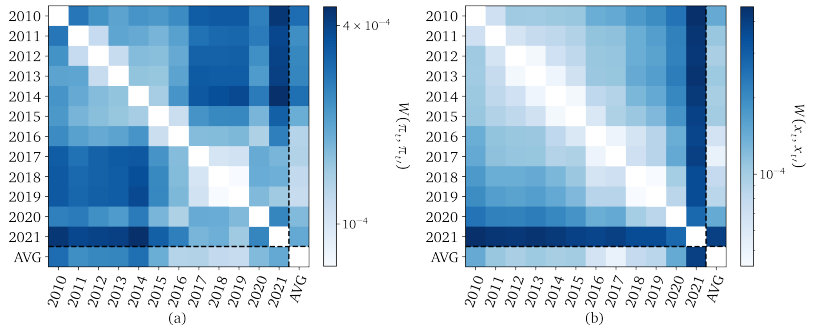


Figure 5.7.: Pairwise distance matrix: (a) Pairwise Wasserstein distance matrix between the steady states $W(\pi_t, \pi_{t'})$ (Eq. 5.23). The dotted line provides a visual delimitation between the steady states of $\mathcal{P}^t$ and the steady state of the average matrix $\hat{\mathcal{P}}$. (b) Pairwise Wasserstein distance matrix between the actual states $W(x_t, x_{t'})$ (Eq. 5.23). The dotted line provides a visual delimitation between the actual states of $\mathcal{P}^t$ and the steady state of the average matrix $\hat{\mathcal{P}}$

Starting from the observed state of the labor market at time t, we predict the state at $t+1$ using $\vec{x}'_{t+1} = \mathcal{P}^t \vec{x}_t$.[12] For longer forecast horizons, the k-step prediction is recursively calculated as outlined above. However, the predicted states $\vec{x}'_{t+k}$ often deviate from the actual observed states $\vec{x}_{t+k}$, particularly over extended horizons. To evaluate whether these predictions improve by using the average transition matrix $\hat{\mathcal{P}}$, we compare the predictive accuracy of $\mathcal{P}^T$ and $\hat{\mathcal{P}}$ using the root-mean-square error (RMSE).

12: The true state $\vec{x}_{t+1}$ is computed as $\mathcal{P}^{t+1}\vec{x}_t$. However, since this requires knowledge of $\vec{x}_{t+1}$, introducing a look-ahead bias, we use $\mathcal{P}^t$ from the prior year for predictions.

The relative error is defined as:

$$\text{RMSE}_{rel} = \frac{\text{RMSE}(\mathcal{P}^T)}{\text{RMSE}(\hat{\mathcal{P}})} = \frac{\sqrt{\frac{1}{N}\sum_{i=1}^{N}(x'_{\mathcal{P}^T,i} - x_{t+k,i})^2}}{\sqrt{\frac{1}{N}\sum_{i=1}^{N}(x'_{\hat{\mathcal{P}},i} - x_{t+k,i})^2}}, \tag{5.14}$$

where $x'_{\mathcal{P}^T,i}$ and $x'_{\hat{\mathcal{P}},i}$ are the predicted states for occupation i using $\mathcal{P}^T$ and $\hat{\mathcal{P}}$, respectively, and $x_{t+k,i}$ represents the actual state. A value $\text{RMSE}_{rel} > 1$ indicates that $\hat{\mathcal{P}}$ provides more accurate predictions, while $\text{RMSE}_{rel} < 1$ suggests that $\mathcal{P}^T$ performs better.

Figure 5.6 shows the relative RMSE (Eq. 5.14) for different forecasting horizons k. We find that $\mathcal{P}^T$ generally provides better predictions for short horizons ($k = 1, 2$), as it captures year-specific fluctuations influenced by immediate economic conditions or short-term policy changes. Conversely, for longer horizons ($k > 2$), $\hat{\mathcal{P}}$ outperforms $\mathcal{P}^T$, suggesting that long-term labor market dynamics are governed by stable structural trends rather than transient year-to-year variations.

The averaging process used to construct $\hat{\mathcal{P}}$ smooths out annual noise and idiosyncrasies, offering a more reliable representation of persistent trends in occupational transitions. These findings indicate that while the time-dependent matrices are effective for short-term forecasting, the average transition matrix better captures the underlying regularities that shape long-term labor market behavior.

While the Markov framework provides a useful first-approximation for predicting labor market dynamics, its performance is limited by the inherent simplifications of the model.

5.7. Steady State of the Labor Market

The steady state of a transition probability matrix provides insight into the long-term behavior of occupational flows. From a physics perspective, it represents the equilibrium distribution of a system, where the flux between states balances out over time. Mathematically, the steady state π_t of a markov chain is the eigenvector[13] associated with the eigenvalue $\lambda_1 = 1$, satisfying the equation:

$$\mathcal{P}\pi_t = \pi_t \tag{5.15}$$

the stationary distribution vector π_t and the transition probability matrix $\mathcal{P}_{i,j}$.

Intuitively, this eigenvector represents the distribution the system converges to after many iterations of applying $\mathcal{P}$. This convergence occurs because all other eigenvectors, associated with eigenvalues $|\lambda| < 1$, decay to zero with repeated applications of the transition matrix. This can be demonstrated by diagonalizing the transition matrix as $\mathcal{P} = Q\Lambda Q^{-1}$, with the matrix of eigenvectors Q and the diagonal matrix Λ.[14] To observe how the system converges to the steady state, we apply the transition matrix $\mathcal{P}$ repeatedly to an arbitrary initial state $\vec{u}_0$. As $k \to \infty$, this process yields:

$$\vec{u}_k = (\mathcal{P})^k \vec{u}_0 = Q\Lambda^k Q^{-1}\vec{u}_0. \tag{5.18}$$

The initial state $\vec{u}_0$ can be expressed as a linear combination of the eigenvectors of $\mathcal{P}$:

$$\vec{u}_0 = Q\vec{c} = c_1\vec{v}_1 + c_2\vec{v}_2 + \cdots + c_N\vec{v}_N, \tag{5.19}$$

with coefficients c_i of $\vec{u}_0$ in the eigenvector basis. Substituting this decomposition into the equation for $\vec{u}_k$:

$$\vec{u}_k = (\mathcal{P})^k \vec{u}_0 = Q\Lambda^k Q^{-1}\left(Q\vec{c}\right) \tag{5.20}$$

$$= c_1\lambda_1^k\vec{v}_1 + c_2\lambda_2^k\vec{v}_2 + \cdots + c_N\lambda_N^k\vec{v}_N. \tag{5.21}$$

As $k \to \infty$, all terms associated with eigenvalues $|\lambda_i| < 1$ decay to zero, leaving:

$$\vec{u}_k \underset{k\to\infty}{\to} c_1\vec{v}_1 \tag{5.22}$$

Since the steady-state distribution in a Markov chain must sum to 1, the coefficient c_1 becomes irrelevant after normalization. This demonstrates that regardless of the initial state, the system converges to the steady state governed by the eigenvector associated with the eigenvalue $\lambda_1 = 1$.

From a physical perspective, this steady state represents a configuration where the system no longer evolves under the given dynamics. A particle transitioning between states according to $\mathcal{P}$ will eventually settle into a predictable distribution.

In the context of labor markets, this particle corresponds to workers transitioning between occupations. The steady state π_t is a vector where each entry reflects the fraction of workers in a specific occu-

13: Given a $N \times N$ matrix M, then the non-zero vector $\vec{v}$ is called eigenvector of M, if $M\vec{v} = \lambda\vec{v}$, with the corresponding eigenvalue λ.

14:

$$Q = \begin{bmatrix} v_1^{(1)} & v_2^{(1)} & \cdots & v_N^{(1)} \\ v_1^{(2)} & v_2^{(2)} & \cdots & v_N^{(2)} \\ \vdots & \vdots & \ddots & \vdots \\ v_1^{(N)} & v_2^{(N)} & \cdots & v_N^{(N)} \end{bmatrix}. \tag{5.16}$$

and

$$\Lambda = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_N \end{bmatrix} \tag{5.17}$$

pation. Occupations with higher steady-state probabilities attract substantial transitions or retain workers for extended periods, emphasizing their stability and central role in the labor market. In contrast, occupations with lower probabilities are less connected or serve as transitional roles, highlighting their peripheral nature within the occupational network.

To compare steady-state distributions across different years, we use the Wasserstein distance, a metric from optimal transport theory that measures the minimal *effort* needed to transform one probability distribution into another [203]. This makes it particularly suitable for quantifying differences between steady-state vectors. Formally, the Wasserstein distance between two distributions π_1 and π_2 is given by:

$$W(\pi_1, \pi_2) = \inf_{\gamma \in \Gamma(\pi_1, \pi_2)} \int |x - y| \, d\gamma(x, y) \tag{5.23}$$

where $\Gamma(\pi_1, \pi_2)$ represents the set of all joint distributions with marginals π_1 and π_2. Intuitively, this equation calculates the minimum *cost* required to move probability mass between states to align one distribution with the other.

The Wasserstein distance between steady states from different years quantifies differences in long-term equilibrium workforce distributions. A small Wasserstein distance indicates high similarity between steady-state distributions, suggesting minimal structural changes in the labor market. In contrast, a large Wasserstein distance signals significant shifts, potentially driven by changes in sectoral demand, evolving skill requirements, or broader economic transformations.

We analyze the pairwise Wasserstein distances between the steady states for each year, as shown in Figure 5.7(a). These distances reveal that steady-state distributions are highly similar within two- to three-year intervals, forming a clear block-diagonal structure in the distance matrix. This pattern highlights periods of relative stability in the labor market's equilibrium. Over time, steady states shift gradually, with noticeable changes occurring approximately every two to three years. Despite these shifts, the system stabilizes within new three-year blocks, illustrating the slow but steady evolution of long-term labor market dynamics. The concept of a steady state assumes sufficient time for the system to fully converge. However, the observed shifts every two to three years indicate that this convergence is not achieved within such short periods

Next, we analyze the pairwise Wasserstein distances between the actual states of the labor market, shown in Figure 5.7(b). [15] The distance matrix for actual states exhibits a similar block structure to the steady states but with larger blocks and slower changes over time. This pattern indicates that while the actual state moves in the direction of the changing steady state, it never fully aligns before the steady state itself evolves again. Additionally, we observe a gradual, directed change in actual states over time, indicating a consistent direction of evolution rather than an oscillation around a fixed point. The disruption caused by the COVID-19 pandemic stands out distinctly, with the Wasserstein metric reflecting significantly higher values compared to all other periods, highlighting the unprecedented shift in state of the labor market

15: The actual state $\vec{x}_t$ represents the share of worker in each occupation at time t.

during that time.

While steady-state distribution is not an optimal or economically desirable outcome it acts as a baseline for comparison across different years, offering insights into the pace and nature of structural changes in long-term labor market dynamics.

5.8. Detailed Balance of the Labor Market

The steady state does not imply the absence of transitions between occupations but rather that, after all transitions are accounted for, the overall distribution of workers remains unchanged. A stricter condition, known as *detailed balance*, applies to symmetric transition matrices and ensures an even deeper level of equilibrium [204]. This condition characterizes reversible Markov processes, where there exists a stationary distribution $\vec{\pi}$ that satisfies the detailed balance equation:

$$\pi_i \mathcal{P}_{i,j} = \pi_j \mathcal{P}_{j,i} \tag{5.24}$$

where π_i and π_j are the steady-state probabilities of occupations i and j, and $\mathcal{P}_{i,j}$ represents the transition probability from occupation i to j.

Detailed balance means that, in equilibrium, the probability flows between any two occupations are exactly balanced. This is a stronger condition compared to the stationary distribution criterion (Eq. 5.15), which ensures only that the overall share of workers in each occupation remains constant. In contrast, detailed balance requires that the direct inflow and outflow between every pair of occupations are precisely conserved. In other words, while the stationary distribution allows the possibility of workers moving in cycles while maintaining constant occupation shares, detailed balance enforces a stricter symmetry: the probability flow of workers from occupation i to j must exactly match the flow from j to i.

Figure 5.5(b) shows the symmetry of flows by displaying the scatter plot of $(\pi_i p_{i\to j}, \pi_j p_{j\to i})$ (Eq. 5.24). A two-dimensional Kolmogorov-Smirnov test confirms symmetry at the 5% significance level.

It remains an open question why the labor market self-organizes in a way that obeys detailed balance. There is no obvious theoretical or empirical reason for this to occur, making it an intriguing subject for further research. Future studies could explore the underlying social or economic mechanisms that drive this behavior.

Since the transition matrix obeys detailed balance, we can decompose it to enable a direct comparison of the dynamical properties across the years, including those beyond the first eigenvector.

Using the detailed balance condition, we perform a similarity transformation to $\mathcal{P}$, resulting in its spectral decomposition:

$$V = \Pi \mathcal{P} \Pi^{-1}, \quad \text{where } \Pi = \begin{bmatrix} \sqrt{\pi_1} & 0 & \cdots & 0 \\ 0 & \sqrt{\pi_2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sqrt{\pi_N} \end{bmatrix} \tag{5.25}$$

If $\mathcal{P}$ obeys detailed balance, than the decomposed matrix V is symmetric.[16] The inset of Figure 5.5(b) shows the symmetry index Eq. 5.7. The index is remarkably close to 1, indicating a high degree of symmetry in the spectral decomposition V and, consequently, a close adherence to detailed balance of $\mathcal{P}$.

The eigenvalues of the transition matrix $\mathcal{P}$ remain unchanged under this transformation. Since eigenvalues represent the rates at which the system approaches equilibrium, their preservation ensures that the key timescales of labor market dynamics are maintained. However, in the original transition matrix, the eigenvectors are not necessarily orthogonal, complicating the interpretation of their contributions to the system's dynamics. By applying this transformation, the matrix becomes symmetric, and the eigenvectors become orthogonal. This orthogonality simplifies the analysis of transient dynamics, as each eigenvector now represents an independent mode of variation, eliminating overlap or redundancy. In the transformed basis, the spectral decomposition clearly separates the steady-state mode (associated with the eigenvalue $\lambda = 1$) from the transient modes, providing a more intuitive understanding of how deviations from equilibrium decay over time.

16:

$$V_{ij} = \frac{\sqrt{\pi_i}}{\sqrt{\pi_j}}\mathcal{P}_{ij} \underbrace{=}_{\text{detailed balance}} \frac{\sqrt{\pi_i}}{\sqrt{\pi_j}}\frac{\pi_j}{\pi_i}\mathcal{P}_{ji} \tag{5.26}$$

$$= \frac{\sqrt{\pi_j}}{\sqrt{\pi_i}}\mathcal{P}_{ji} = V_{ji} \tag{5.27}$$

5.9. Dynamical Properties of the Labor Market

Building on the analysis of steady states, we now investigate the dynamical properties of labor market transitions by examining the eigenvalues and eigenvectors of the transition matrices. While the steady state reflects the long-term equilibrium distribution of workers across occupations, the eigenvalues and eigenvectors associated with $|\lambda_i| < 1$ provide insights into the rates and patterns of convergence toward equilibrium and the transient dynamics that govern labor market flows.

The eigenvalues of the transition matrix determine the timescales over which deviations from the steady state decay. Specifically, the second-largest eigenvalue, $|\lambda_2|$, dominates the rate of convergence, with the spectral gap, $1 - \lambda_2$, serving as an indicator of convergence speed—a larger gap implies faster convergence.[17] On average, the relaxation time $T_{relax} \propto \frac{1}{1-\lambda_2}$ of the system is approximately 31 years, highlighting the slow adjustment process inherent in labor market transitions (see Figure 5.8 inset).

Figure 5.8 illustrates the eigenvalue spectrum of $\mathcal{P}$.[18] The largest eigenvalue is $\lambda_1 = 1$, while all other eigenvalues lie strictly within the unit circle. A clear spectral gap between the first and second eigenvalues is visible, emphasizing the system's dominant mode of convergence. For certain years, small non-zero imaginary components appear in the eigenvalues, reflecting that the transition matrix is not fully symmetric but has some small asymmetric components. However, when transitions are aggregated into the average matrix $\hat{\mathcal{P}}$, these imaginary components nearly vanish, indicating that the aggregated system is effectively symmetric (colored in dark blue in Figure 5.8).

17: The spectral gap is defined as $\lambda_1 - \max(\{\lambda_i\}) = 1 - \lambda_2$). Starting from Eq. 5.21

$$\vec{u}_k = \pi + \sum_{i=2}^{N} c_i \lambda_i^k \vec{v}_i \tag{5.28}$$

$$\|\vec{u}_k - \pi\| \leq \sum_{i=2}^{N} |\lambda_i|^k \| c_i \vec{v}_i \| \tag{5.29}$$

Since $|\lambda_2| > |\lambda_i|$ for $i > 2$, we can bound the above term

$$\|\vec{u}_k - \pi\| \leq |\lambda_2|^k \sum_{i=2}^{N} \| c_i \vec{v}_i \| = C|\lambda_2|^k \tag{5.30}$$

since the sum is a constant independent of k. We can see that $|\lambda_2|$ dominates the convergence behavior and determines how quickly the system approaches π. With this the relaxation time can be approximated by $T_{relax} \propto \frac{1}{\log|\lambda_2|} \approx \frac{1}{1-\lambda_2}$.

18: Note that the eigenvalues of V and $\mathcal{P}$ are the same.

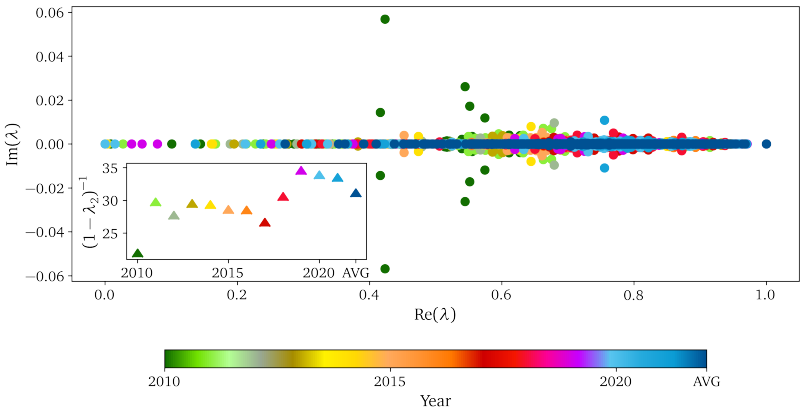


Figure 5.8.: Eigenvalue spectrum of Transition Matrices: Scatter plot of the eigenvalues of the transition matrices $\mathcal{P}_t$ and $\hat{\mathcal{P}}$ in the complex plane, with the real and imaginary components shown. Different colors correspond to different years. The inset depicts the relaxation time $T_{relax} \propto \frac{1}{1-\lambda_2}$ in units of years of the system for each year and the average matrix.

To analyze how the dynamical properties of the transition matrix evolve over time, we compare the eigenvectors of the average transition matrix, to those of the time-dependent transition matrices. The eigenvectors provide insights into the dominant patterns and modes of variation within the system. Specifically, we assess the overlap between the eigenvectors of $\hat{V}$ (derived from $\hat{\mathcal{P}}$) and V_t (derived from $\mathcal{P}^t$ for each year). This comparison allows us to explore how the dominant dynamics of the system—governed by the leading eigenvectors—shift over time.

A starting point for quantifying the similarity between eigenvectors is the cosine similarity, which measures the angle between two vectors:

$$\cos(\theta) = \frac{\hat{w}_i \vec{w}_{t,i}}{|\hat{w}_i||\vec{w}_{t,i}|} \underbrace{=}_{|\hat{w}_i|,|\vec{w}_{t,i}|=1} \hat{w}_i \vec{w}_{t,i}. \tag{5.31}$$

with $\hat{w}_i$ and $\vec{w}_{t,i}$ representing the eigenvectors of $\hat{V}$ and V_t, respectively. A cosine similarity of 1 indicates perfect alignment, while a value of 0 indicates orthogonality. However, while effective for pairwise comparisons, the cosine similarity does not capture the overall similarity of sets of eigenvectors, as it evaluates only individual pairs.

To address this limitation, we introduce the generalized cosine similarity, which assesses the alignment of an entire set of eigenvectors. This method considers the structure of the subspace spanned by the leading k eigenvectors. We first compute the overlap matrix:

$$G_k = \hat{W}_k^T W_{t,k} \tag{5.32}$$

where $\hat{W}_k$ and $W_{t,k}$ are matrices containing the first k eigenvectors of $\hat{V}$ and V_t, respectively.[19] The overlap matrix G_k encapsulates how well the subspaces defined by the two sets of eigenvectors align.

The generalized cosine similarity is then defined as:

$$d_k := -\frac{\sum_i^k \ln(s_i)}{k} \tag{5.35}$$

where s_i are the singular values of G_k.[20] Unlike standard cosine similarity, which indicates similarity through values close to 1, the gen-

19: $\hat{V}$ and V_t have in general N eigenvectors, therefore $\hat{W}$ and $W_{t,k}$ as the dimension of $N \times k$

$$W_{t,k} = \begin{bmatrix} \vec{w}_{t,0} & \vec{w}_{t,1} & \cdots & \vec{w}_{t,k} \end{bmatrix} \tag{5.33}$$

$$= \underbrace{\begin{bmatrix} | & | & & | \\ \vec{w}_{t,0} & \vec{w}_{t,1} & \cdots & \vec{w}_{t,k} \\ | & | & & | \end{bmatrix}}_{\vec{w}_{t,i} \text{ are eigenvectors of } V_t}. \tag{5.34}$$

20: This expression stems from the determinant being expressed in terms of the singular values. For a square matrix M, the singular value decomposition (SVD) is $M = U\Sigma V^\star$, where U and V are unitary matrices and Σ is a diagonal matrix containing the singular values σ_i. Since $\det(M) = \det(U\Sigma V^\star) = \det(U)\det(\Sigma)\det(V^\star) = \det(\Sigma)$, the determinant is simply the product of the singular values, $\prod_i \sigma_i$. Thus, d_k directly relates to these singular values.

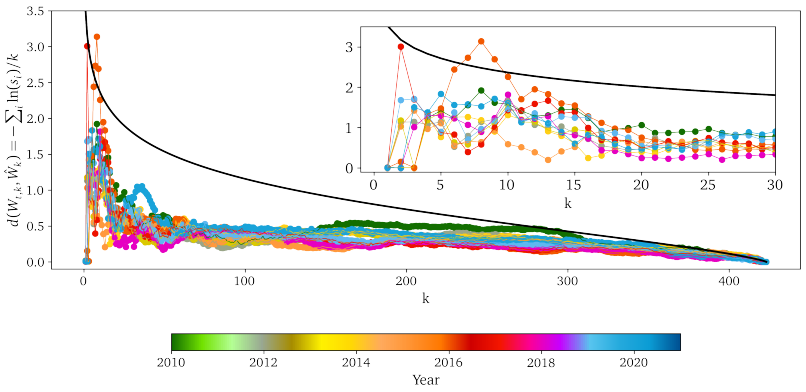


Figure 5.9.: Generalized cosine similarity: Generalized cosine similarity between the subspaces spanned by the leading k eigenvectors of V_t and $\hat{V}$ across different years. The black solid line represents the generalized cosine similarity for two random matrices [206]. The inset provides a zoomed view of the plot for the first $k \geq 30$ eigenvectors.

eralized cosine similarity is constructed such that higher values of d_k indicate greater divergence, while lower values signify stronger alignment between the eigenvectors.

This metric evaluates the *volume* spanned by the overlap between the subspaces of $\hat{W}_k$ and $W_{t,k}$. A small d_k suggests a high degree of overlap, while a large d_k implies divergence.[21] For $k = 1$, d_k simplifies to a form similar to the standard cosine similarity for two vectors.[22] As $k \to N$ (the matrix dimension), $d_k \to 0$, as the entire space is eventually spanned, albeit with potentially rotated eigenvectors.[23]

By systematically varying k, we focus on the alignment of the leading k eigenvectors, which dominate the system's dynamics. This approach provides a nuanced view of how the most influential patterns in labor market transitions evolve over time.

To contextualize the observed eigenvector overlap, we compare the empirical generalized cosine similarity to the expected similarity between eigenvectors of two random matrices [206]. This comparison provides a baseline, enabling us to distinguish meaningful deviations in eigenvector alignment from random variations. Deviations from this baseline highlight the extent to which the dynamical properties of the labor market are structured and how they change over time.

Figure 5.9 illustrates the generalized cosine similarity across years, comparing the eigenvectors of the time-dependent transition matrices V_t with those of the average transition matrix $\hat{V}$. For $k \lesssim 50$, the similarity is strong, as indicated by low d_k values well below the random matrix benchmark. This demonstrates that the leading eigenvectors, which dominate the system's dynamics, exhibit overlap between V_t and $\hat{V}$. These eigenvectors drive the system's convergence to the steady state and reflect stable, persistent dynamics across years.

Between $k \approx 50$ and $k \approx 300$, the similarity stabilizes, indicating that these eigenvectors continue to capture meaningful dynamics but with diminishing distinctiveness from the random benchmark. This plateau suggests a transition to less dominant but still structured dynamics.

Beyond $k > 350$, the similarity converges to the random matrix expectation, signifying that the eigenvectors associated with smaller eigenvalues are less structured and exhibit minimal alignment with those of the average matrix. These eigenvectors represent transient or noisy

21: This definition has the opposite intuition compared to standard cosine similarity, where values close to one indicate similarity. This design allows d_k to be compared directly with theoretical predictions derived from random matrix theory.

22: In this case, $d_1 = -\ln(s_1) = -\ln(\det(\hat{W}_1^T W_{t,1})) = -\ln(\hat{w}_0 w_{t,0})$.

23: Due to empirical irregularities, the occupational transition matrices do not perfectly follow detailed balance, meaning V is not fully symmetric. As a result, the eigenvectors of V do not form a perfectly orthonormal basis. To address this, we perform a QR decomposition on V_t and $\hat{V}$ to obtain orthonormal eigenvectors. QR decomposition expresses a real square matrix as $M = QR$, where Q is an orthogonal matrix and R is an upper triangular matrix. We use the Gram-Schmidt process, a numerical algorithm, to generate orthonormal vectors of Q [205].

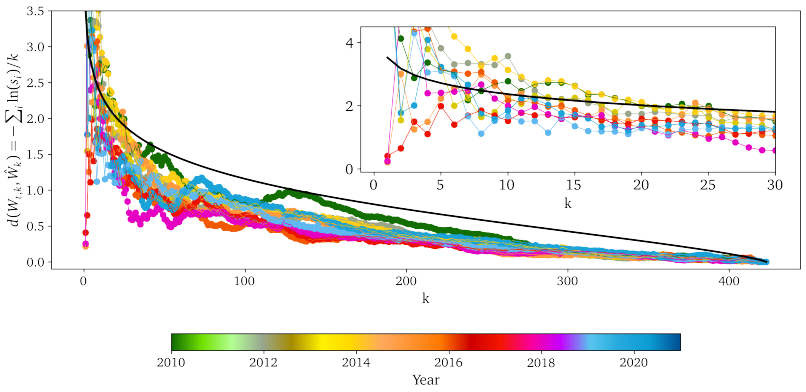


Figure 5.10.: Reversed generalized cosine similarity: Generalized cosine similarity between the subspaces spanned by the k eigenvectors starting from the smallest eigenvalue of V_t and $\hat{V}$ across different years. The black solid line indicates the generalized cosine similarity for two random matrices [206]. The inset provides a detailed view of the plot for the first $k \geq 30$ reversed eigenvectors.

aspects of the labor market.

To further show the randomness of eigenvectors associated with smaller eigenvalues, we reverse the process and compare eigenvectors starting from the smallest eigenvalues. Figure 5.10 shows that for $k < 30$, the generalized cosine similarity across years is indistinguishable from that of two random matrices. This indicates that the smallest eigenvectors, tied to smaller eigenvalues, are highly variable and do not align with the average matrix.

These findings underline a clear distinction: the leading eigenvectors associated with larger eigenvalues are stable across years, reflecting stable long-term dynamics, while trailing eigenvectors reflect transient or year-on-year variations. This divergence highlights the relative importance of dominant modes in describing consistent labor market dynamics across years, with diminishing contributions from eigenvectors associated to smaller eigenvalues.

5.10. On Measurement Noise of Transition Matrices

While the analysis of eigenvalues and eigenvectors provides valuable insights into the dynamical properties of labor market transitions, these findings are inherently dependent on the accuracy of the empirically measured transition matrix. The occupational transition matrix is derived from observed data, which is subject to various sources of noise. Understanding the impact of this noise is crucial, as it can influence the stability of eigenvalues, the alignment of eigenvectors, and, ultimately, the interpretation of the system's dynamics. Additionally, as occupational classifications become more detailed, the data requirements for reliably estimating each transition pair increase quadratically. To ensure the robustness of our conclusions, we now turn to an examination of measurement noise and its implications.

To simulate the effects of measurement noise, we construct fictitious transition matrices by sampling synthetic datasets from the actual mean transition matrix $\hat{\mathcal{P}}$.[24] For each synthetic dataset, we sample N_s transitions, varying N_s across $[0.5M, 1M, 3M, 10M, 30M]$ to reflect different dataset sizes[25]. We then reconstruct fictitious transition matrices and analyze how noise from finite sampling impacts their spectral

24: The distribution p from which we sample is $p_{ij} = \frac{N_{ij}}{N}$ where N_{ij} is the number of observed transitions from state i to j, and $N = \sum_{ij} N_{ij}$ is the total number of transitions

25: We sample each N_s 200 times.

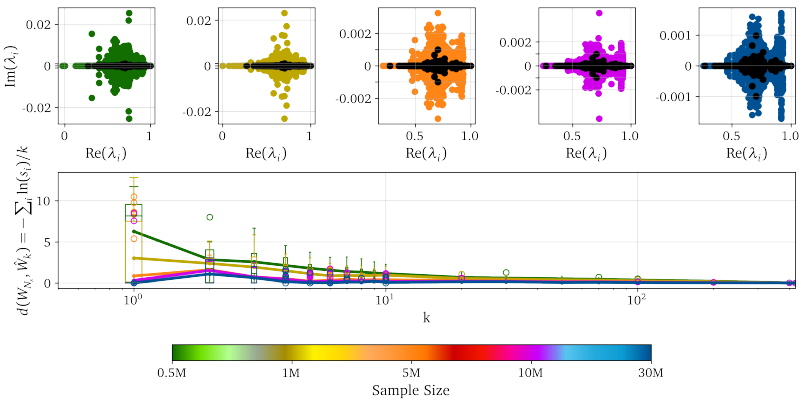


Figure 5.11.: Effect of measurement noise: (Top row) Eigenvalue spectrum of the sampled transition matrices (for all random seeds), where different sample sizes N_s are shown in different colors. The black points represent the true eigenvalue spectrum of $\hat{\mathcal{P}}$. The bottom row presents a box plot of the generalized cosine similarity between the sampled matrices (for all random seeds) and $\hat{\mathcal{P}}$. Different colors represent different sample sizes N_s.

properties.

To quantify the accuracy of the sampled matrices, we compare their spectral properties—eigenvalues and eigenvectors—to those of the true matrix $\hat{\mathcal{P}}$. First, we compare eigenvalue spectra of the sampled and true matrices, expecting that as $N_s \to \infty$, the spectra will converge (Figure 5.11, top row). Additionally, we compute the generalized cosine similarity between the eigenvectors of the sampled and true matrices to evaluate when the eigenvectors align (Figure 5.11, bottom row). These metrics allow us to assess how the eigenvalue spectra and eigenvectors are affected by finite sampling and how much data is required to minimize noise.

The top row of Figure 5.11 displays the eigenvalue spectra of the sampled matrices (colored points) compared to the eigenvalues of true matrix (black points). For small sample sizes ($N_s < 10$M), the eigenvalue spectra of the sampled matrices differ significantly from the true spectra. These sampled matrices exhibit substantial contributions in the imaginary space, indicating asymmetries in their structure. However, as N_s increases, the spectra begin to stabilize with lower imaginary part and more similar structure to the true eigenvalue spectrum. At $N_s \geq 10$M, the eigenvalue spectra closely align with those of the true matrix, and the imaginary components nearly vanish, reflecting a symmetric and accurate approximation.

The bottom row of Figure 5.11 shows the generalized cosine similarity d_k (see Eq. 5.35) between the eigenvectors of the sampled and the true matrix. Each box plot at each k is for 200 random samples. Small values indicate that the overlap of the eigenvectors is high. For $N_s \lesssim 10$M, the leading eigenvectors of the sampled matrices behave as if they were random, showing little similarity to those of the true matrix. Beginning at $N_s = 5$M, the generalized cosine similarity converges to zero for all k, indicating that the eigenvectors of the sampled matrices align with those of the true matrix and stabilize.

These results demonstrate that measurement noise significantly impacts the accuracy of the transition matrix when the sample size is small, introducing deviations in both eigenvalues and eigenvectors. However, as the sample size reaches $N_s \gtrapprox 10$M, the spectral properties converge, and the effects of noise are minimized. This suggests that at least 10 million observations are necessary to construct a transition matrix that reliably approximates the true dynamics of labor

market transitions. Ensuring sufficient data is therefore critical for robust and meaningful analysis of transition matrices in labor market studies.

5.11. Network Properties of the Labor Market

Building on our analysis of steady-state dynamics and spectral properties, we now adopt a network perspective, interpreting the transition matrix as a weighted, directed network $\mathcal{G} = \{\mathcal{N}, \mathcal{E}\}$. In this framework, each node $\mathcal{N}$ represents an occupation, while the edges $\mathcal{E}$ capture the probabilities of transitioning between them, with edge weights defined as $w_{ij} = \hat{\mathcal{P}}_{ij}$. This network representation provides a powerful lens for examining the structure of the labor market, enabling the identification of key occupations and their roles in connecting or segmenting the network.

Using methods from network science, we delve into centrality metrics to pinpoint influential occupations and clustering coefficients to assess the density of interconnections within the network. These analyses not only complement our previous findings but also uncover how the labor market organizes into interconnected sectors.

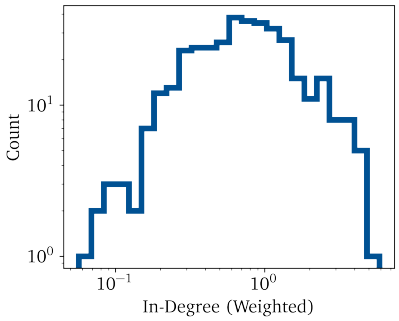


Figure 5.12.: Indegree distribution: Distribution of weighted in-degree centralities (Eq. 5.36) for each node in the labor market network $\mathcal{G}$.

- **Weighted In-Degree Centrality:** Weighted in-degree centrality quantifies the total strength of incoming transitions to an occupation

$$C_{w,j} = \sum_{i \in \text{adj}(j)} w_{ij} \tag{5.36}$$

where $i \in \text{adj}(j)$ represents nodes connected (or *adjacent*) to j [207]. This metric captures an occupation's role as a *sink* for labor mobility. The distribution, shown in Figure 5.12 suggests that while most occupations receive moderate inflows, a few serve as major hubs attracting transitions from diverse roles. Conversely, occupations with low in-degree centrality are less connected and contribute minimally to workforce redistribution. [26]

26: The transition matrix normalization ensures that out-degree centrality is constant $C_{w,i} = \sum_{j \in \text{adj}(i)} w_{ij} = 1$, for all nodes.

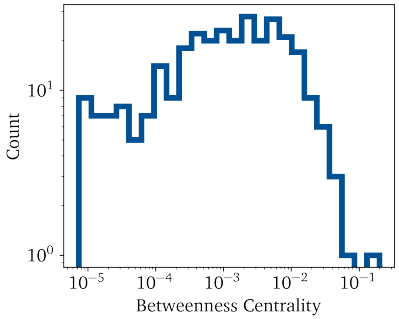


Figure 5.13.: Betweenness centrality distribution: Distribution of betweenness centralities (Eq. 5.37) of each node in the labor market network $\mathcal{G}$.

- **Betweenness centrality:** Betweenness centrality reflects how often an occupation serves as a bridge along the shortest paths between other occupations:

$$C_{b,i} = \sum_{i \neq j \neq k \in G} \frac{\sigma_{jk}(i)}{\sigma_{jk}} \tag{5.37}$$

where σ_{jk} is the number of shortest paths between j and k, and $\sigma_{jk}(i)$ is the number of those paths passing through i [208]. Occupations with high betweenness centrality connect disconnected parts of the network, facilitating transitions between otherwise isolated clusters. As shown in Figure 5.13, most occupations have moderate bridging abilities, but a few serve as critical intermediaries, representing versatile occupations connecting specialized occupations.

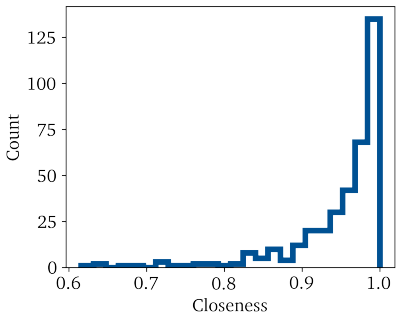


Figure 5.14.: Closeness centrality distribution: Distribution of closeness centralities (Eq. 5.38) for each node in the labor market network $\mathcal{G}$.

- **Closeness centrality:** Closeness centrality measures how *close* an occupation is to all others in terms of shortest paths:

$$C_{C,i} = \frac{n-1}{\sum_j^{n-1} d_{ij}} \tag{5.38}$$

where d_{ij} is the shortest path distance between i and j [209]. Occupations with high closeness centrality are well-integrated and easily reachable within the labor market. The distribution (Figure 5.14) is highly skewed, with most occupations having high closeness values, indicating that the labor market network is strongly connected with a small subset of isolated or niche roles.

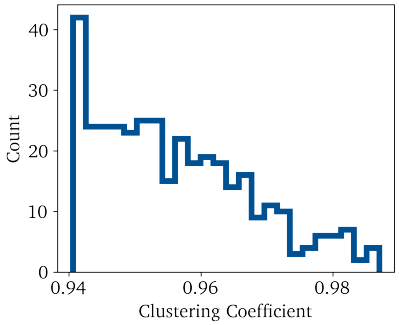


Figure 5.15.: Clustering coefficient distribution: Distribution of local clustering coefficients (Eq. 5.39) for each node in the labor market network $\mathcal{G}$.

- **Local Clustering Coefficient:** The local clustering coefficient quantifies the interconnectedness of an occupation's neighbors:

$$c_i = \frac{\sum_{j \neq k \in \text{adj}(i)} (w_{ij} w_{ik} w_{jk})^{\frac{1}{3}}}{k_i(k_i - 1)} \tag{5.39}$$

  with the degree of node $k_i = \sum_{j \in G} \delta_{ij}$ [210].[27] High clustering coefficients indicate tightly-knit communities where transitions frequently occur within related occupations. As shown in Figure 5.15, the clustering coefficient distribution is high, with most values between 0.94 and 0.98. This highlights the labor market's organization into dense, community-like structures, where transitions are concentrated within closely related groups.

27: Note that we use the weighted and undirected version of the clustering coefficient, with the weight defined as the average of the back-and-forth transition weights $\frac{w_{ij}+w_{ji}}{2}$.

The results of these network metrics reveals a labor market that is highly connected yet exhibits clear community-like structures. While some occupations act as critical hubs or bridges, facilitating widespread mobility, most transitions occur within dense, tightly-knit clusters of related roles.

Community Detection in the Labor Market Network

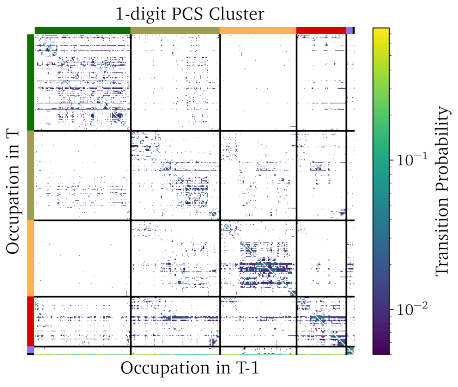


Figure 5.16.: Communities by 1-digit PCS codes: Visualization of the occupational transition matrix $\hat{\mathcal{P}}$, where color intensity represents transition probabilities. Solid lines indicate the clustering given by the 1-digit PCS codes.

We aim to detect and analyze the communities in the labor market network. Community detection allows us to uncover tightly connected groups of occupations, reflecting clusters where worker transitions are concentrated. A natural clustering into communities is already given by the 4-digit PCS occupation classification when aggregated to the 1-digit level (see Figure 5.16). However, this classification is not accurate, as it does not with the realized flows, therefore our aim is to identify

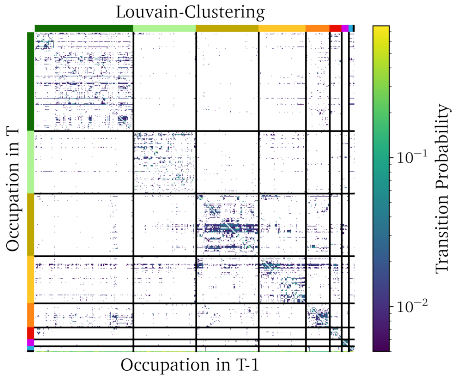


Figure 5.17.: Community detection with Louvain: Visualization of the occupational transition matrix $\hat{\mathcal{P}}$, where color intensity represents transition probabilities. Solid lines indicate the clustering results with the Louvain Clustering algorithms. The clustering reflects the best partition based on modularity from 500 random seeds. Color bars adjacent to the communities serve as visual guides.

these communities in a data-driven manner using network-based algorithms.

We compare four clustering algorithms: Louvain [211], BRIM [212], Infomap [213], and Spectral Clustering [214]. Each algorithm is applied across multiple random seeds to generate diverse clusterings. To evaluate their quality, we rely on two complementary metrics: modularity and the Gini coefficient. Modularity quantifies the strength of community structures by comparing the density of connections within clusters to the density of connections between clusters[28]. High modularity indicates that clusters are cohesive internally and well-separated from each other. This ensures that the detected communities capture meaningful divisions in the labor market. Gini coefficient assesses the balance of cluster sizes (see Section 5.2). A high Gini coefficient indicates significant inhomogeneity in the distribution of cluster sizes, with one or a few dominant clusters and many smaller ones. By minimizing the Gini coefficient alongside maximizing modularity, we aim to ensure that the detected clusters are not only strong but also balanced, avoiding trivial solutions where a single cluster dominates. Our goal is to identify clustering solutions that achieve both high modularity and low Gini coefficient, balancing structural integrity and interpretability.

1. **Louvain [211]:** The Louvain algorithm follows a two-phase iterative approach to optimize modularity (see Eq. B.4). In the first phase, each node is initially assigned to its own community, and nodes are iteratively moved to neighboring communities if this results in a positive modularity gain; the process continues until no further improvement is possible, achieving a local modularity maximum. Within the definition of modularity a resolution parameter γ is included which allows fine-tuning of the community size by adjusting the balance between internal density and inter-community connections.[29] The second phase creates a new network where each community from the first phase is treated as a single node, with edge weights between these new nodes reflecting the total weights of connections between the original communities. These two steps are repeated on the newly constructed network until no further modularity changes occur, progressively building a hierarchical structure of communities [216].[30]
2. **BRIM [212]:** The BRIM algorithm is used for community detec-

28: Modularity is defined as [215]:

$$Q = \frac{1}{E}\sum_{i=1}\sum_{j=1}(M_{i,j} - P_{i,j})\delta(a_i, a_j) \tag{5.40}$$

with the total number of links E, $M_{i,\alpha}$ the biadjacency matrix, $P_{i,j} = \frac{k_i k_j}{E}$ the probability that a link exists by chance and the membership variable that defines the block that node i belongs to a_i.

29: The resolution parameter is included such that

$$Q' = \frac{1}{E}\sum_{i=1}\sum_{j=1}(M_{i,j} - \gamma\frac{k_i k_j}{E})\delta(a_i, a_j). \tag{5.41}$$

With this, the actual number of edges within a community is compared to the expected number of edges, scaled by the resolution parameter γ. A higher γ increases the weight of the expected edge term, making it harder for larger groups of nodes to satisfy the modularity condition. At higher resolution values, smaller communities are favored because they more easily show a significant difference between actual and expected edge counts, maximizing modularity. Larger communities, by contrast, may not demonstrate sufficient internal density relative to their expected structure and are split into smaller groups.

30: We use for our analysis the python package *python-louvain* [217]. The number of clusters can be adjusted by the parameter *resolution* (γ in Eq. 5.41).

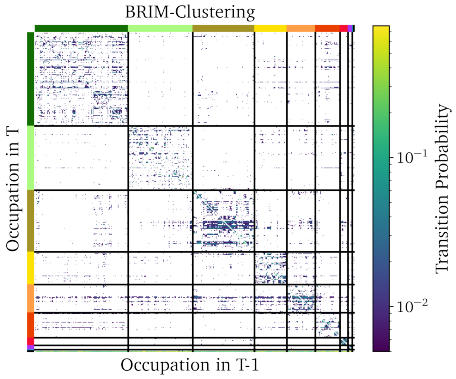


Figure 5.18.: Community detection with BRIM: Visualization of the occupational transition matrix $\hat{\mathcal{P}}$, where color intensity represents transition probabilities. Solid lines indicate the clustering results with the BRIM Clustering algorithms. The clustering reflects the best partition based on modularity from 500 random seeds. Color bars adjacent to the communities serve as visual guides.

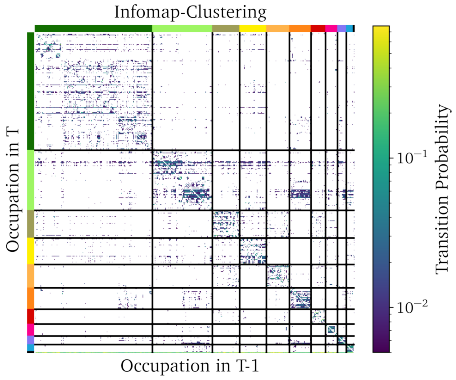


Figure 5.19.: Community detection with Infomap: Visualization of the occupational transition matrix $\hat{\mathcal{P}}$, where color intensity represents transition probabilities. Solid lines indicate the clustering results with the Infomap Clustering algorithms. The clustering reflects the best partition based on modularity from 500 random seeds. Color bars adjacent to the communities serve as visual guides.

tion within bipartite networks[31] by maximizing modularity (see Eq. B.4). It alternates between optimizing the community assignments of the two node types in the network. First, the algorithm fixes the communities for one type of node and updates the communities for the other type to maximize modularity. This process is then reversed, fixing the updated communities and optimizing for the first node type. These steps are repeated iteratively until no further improvement in modularity is achieved.[32]

3. **Infomap [213]:** Infomap is a community detection algorithm based on minimizing the map equation[33], an information-theoretic measure that compresses the description of random walks on a network. It identifies communities as regions where a random walker remains trapped for extended periods, capturing modular patterns in network flows. The algorithm works in two phases: first, it groups nodes into communities by iteratively minimizing the map equation's codelength[34], and then it builds a hierarchical structure by recursively optimizing at aggregated levels. A resolution parameter can be used to adjust the granularity of detected communities, allowing for the identification of fewer large or many small communities.[35]
4. **Spectral Clustering [214]:** Spectral clustering is a graph-based algorithm that uses the eigenvalues and eigenvectors of the transition matrix to group data points. It begins by constructing the Laplacian matrix [36] from the transition matrix and computes eigenvectors corresponding to its smallest non-zero eigenvalues. These eigenvectors provide a low-dimensional representation

31: Note that the labor market can be represented as a bipartite network, where one set of nodes corresponds to all occupations in the previous year $(T-1)$ and the other set represents the occupations in the current year (T).

32: For our analysis, we utilize the Python implementation provided by the *pycondor* package [218], which is based on the methodology developed by Platig et al. (2016) [212]. The number of clusters can be adjusted by the parameter *resolution* (γ in Eq. 5.41).

33: The map equation to be minimized is:

$$L(M) = q^{\curvearrowright} H(Q) + \sum_{m \in M} p_m^{\circlearrowright} H(P_m), \tag{5.42}$$

where $q^{\curvearrowright}$ is the rate of transitions between modules, $H(Q)$ is the entropy of for inter-module transitions, $p_m^{\circlearrowright}$ is the rate of within-module transitions for module m, and $H(P_m)$ is the entropy of the codebook for within-module transitions. Minimizing $L(M)$ balances intra- and inter-community flows for optimal compression.

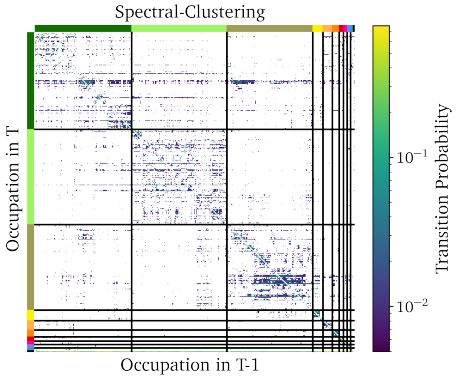


Figure 5.20.: Community detection with Spectral Clustering: Visualization of the occupational transition matrix $\hat{\mathcal{P}}$, where color intensity represents transition probabilities. Solid lines indicate the clustering results with the Spectral Clustering algorithms. The clustering reflects the best partition based on modularity from 500 random seeds. Color bars adjacent to the communities serve as visual guides.

(spectral embedding) of the data. A k-means clustering is then applied to this representation to partition the data into meaningful clusters.[37]

Both the BRIM and Louvain algorithms aim to maximize Barber's modularity, making them powerful tools for community detection. However, the Louvain algorithm has limitations when applied to complex networks. Specifically, its focus on optimizing modularity through individual node movements can lead to disconnected or weakly connected communities, as it does not explicitly account for the overall cohesion of a community [216]. Once these disconnected communities are aggregated into single nodes during the second phase, their disconnection becomes permanent, hindering further refinement and potentially resulting in suboptimal solutions. To address these shortcomings, we also employ the BRIM algorithm, which is specifically designed for bipartite networks and iteratively refines community assignments for both node sets. Additionally, we use Infomap and Spectral Clustering, which are entirely different clustering methods not based on modularity optimization, providing a diverse set of algorithms to compare against for.

To compare the clustering results, Figures 5.17, 5.18, 5.19, 5.20 show the occupational transition matrix with the communities identified by each algorithm.[38] All four methods successfully capture the block-diagonal structure in the transition matrix, better than the 1-digit PCS cluster (see Figure 5.16), indicating clusters with strong internal connectivity. However, notable differences emerge in the size distribution and modularity of the clusters. For example, BRIM tends to produce more balanced clusters, whereas Spectral Clustering generates larger clusters.

To compare the clustering algorithms quantitatively, Figure 8.12 (a) shows the modularity scores for all algorithms as box plots. Each box summarizes the performance of a clustering method at various number of communities, based on 500 random realizations using different resolution parameters. BRIM consistently achieves the highest modularity, demonstrating its effectiveness in identifying strongly connected community structures. Infomap also performs well, with modularity scores only slightly below BRIM, despite not explicitly optimizing for modularity. In contrast, the Louvain algorithm performs slightly worse than BRIM and Infomap, while Spectral Clustering

34: The codelength refers to the average number of bits needed to describe the trajectory of a random walker on the network. By assigning codewords to nodes based on their visiting rates, shorter codelengths indicate that the community structure efficiently captures the flow dynamics.

35: We employ the Python implementation of the Infomap algorithm provided by the *infomap* package [219]. The number of clusters can be adjusted by the parameter *preferred_number_of_modules*.

36: The laplacian matrix is defined as $L = D - M$ with the degree matrix D that is a diagonal matrix with the degree of each node and an adjacenency matrix M. With this, the Laplacian's eigenvalues and eigenvectors have clear interpretations in terms of graph connectivity. The smallest eigenvalue (always 0) corresponds to the graph's connected components. The second smallest eigenvalue and its eigenvector are key to partitioning the graph into two balanced clusters.

37: We utilize the *SpectralClustering* method as implemented in the *scikit-learn* Python package [220]. The number of clusters can be adjusted by the parameter *n_clusters*.

38: Each clustering method is applied with 500 random seeds, and the realization with the highest modularity is shown here.

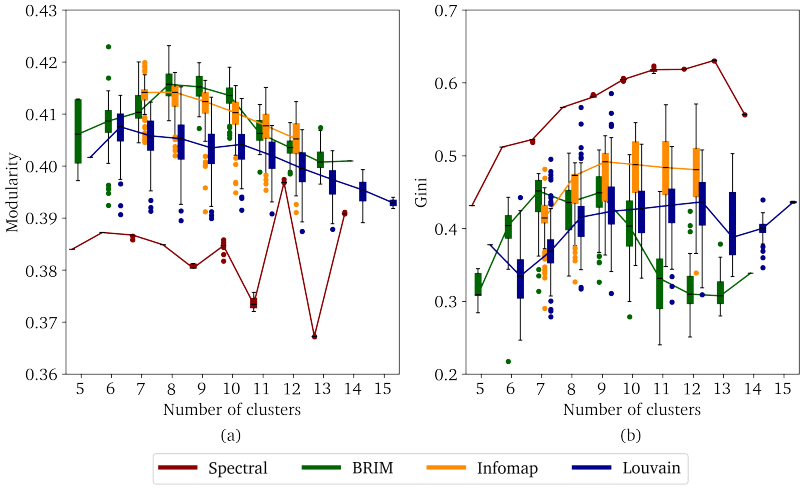


Figure 5.21.: Comparison between community detection algorithms: (a) Box plot illustrating the modularity (Eq. B.4) achieved by each clustering algorithm as a function of the number of clusters. (b) Box plot depicting the Gini coefficient (see Section 5.2) for each clustering algorithm as a function of the number of clusters. Each box plot represents results from 500 random seeds.

shows the lowest modularity scores, indicating weaker cluster cohesion.

Figure 8.12 (b) shows the Gini coefficients highlighting differences in cluster homogeneity across algorithms. Spectral Clustering produces the least balanced communities, with the highest Gini coefficient, characterized by a few dominant communities alongside many small ones. For fewer number of communities, the Louvain algorithm achieves the lowest Gini coefficient, while for larger numbers of communities, BRIM has the most balanced cluster size distribution. Infomap maintains a relatively stable modularity across varying number of clusters, with Gini coefficients ranging between 0.4 and 0.5.

These results highlight a tradeoff between modularity and cluster size homogeneity. BRIM emerges as the best-performing algorithm, achieving the highest modularity and nearly the lowest Gini coefficient, making it the most suitable choice for further analysis and is used in Chapter 6.

Key Takeaways

1. **Directional Flow Ratios:** A heavy-tailed distribution in flow ratios reveals structural imbalances, with certain occupations acting as central hubs or diffusers, while others show directional biases.
2. **Steady-State Dynamics:** The steady-state distribution, derived from the leading eigenvector of the transition matrix, reveals stable and central occupations in the labor market. Shifts in steady states occur approximately every 2–3 years, with periods of stability punctuated by structural changes.
3. **Detailed Balance:** Worker transitions between occupations exhibit a striking balance—each inflow into an occupation is matched by a corresponding outflow. While the underlying mechanisms driving this self-organization remain unclear, this symmetry allows for spectral decomposition, providing a powerful tool to analyze how labor market dynamics evolve over time beyond just their long-term equilibrium.
4. **Slow Adjustment to Shocks:** The labor market takes decades to fully adjust to disruptions, with a relaxation time of around 31 years. This suggests that while some changes happen quickly, deeper structural shifts unfold over much longer timescales.
5. **Network Perspective:** Interpreting the labor market as a network, where occupations are represented as nodes and real transitions as weighted edges, reveals the dual role of certain occupations as hubs or bridges that facilitate labor mobility, while also highlighting the inherent community structure.

The Structure of Occupational Mobility in France

6.

This chapter is taken from Knicker et al. (2024) [4] *The Structure of Occupational Mobility in France*. To align with the objectives of this thesis, I modified and extended several parts to provide additional context. The original work was a collaborative effort with *K. Naumann-Woleske* and *M. Benzaquen*.

We observed in Chapter 5 that occupational transitions reveal hierarchies, with certain occupations acting as hubs due to substantial inflows or as diffusers facilitating significant outflows. These dynamics manifest in the asymmetry of occupational flow ratios. The transition probability matrix can be interpreted as a stochastic process, enabling the quantification of long-term steady states and the associated relaxation times. Additionally, network metrics highlighted the clustered nature of the labor market network, with a power-law distribution in betweenness centrality indicating the presence of bottlenecks that constrain occupational mobility. Building on these findings, this chapter explores specifically the role of bottlenecks in impeding occupational mobility. By developing new metrics and leveraging insights from the structural properties of occupational transitions, we aim to quantify these bottlenecks and explore their implications for labor market dynamics.

Reallocating workers across occupations is crucial for resilient labor markets, allowing economies to adapt to technological innovations and economic fluctuations. However, this process often encounters challenges. Structural barriers, such as skill mismatches and occupational rigidities, can impede the smooth transition of workers [221], and the presence of bottleneck occupations further contributes to sluggish reallocation [47]. As labor markets evolve with automation and digital transformation, understanding the mechanisms that hinder worker transitions becomes increasingly important. For example, innovations like Generative Artificial Intelligence (GenAI) are diminishing the need for repetitive tasks, potentially affecting 44% of working hours in impacted industries [222]. At the same time, there is a growing demand for workers in technology-related roles, such as information security specialists [223], which underscores the need for effective worker reallocation.

We investigate the question of how to identify occupations that may impede rapid reallocation of workers across occupations using a fitness and complexity-inspired method. This complements the prior exploration of the descriptive properties in Chapter 5. Our main contribution is the development of novel, data-driven metrics for accessibility and transferability that quantify occupational mobility and ef-

fectively identify bottleneck occupations in the labor market. Transferability is a measure of the diversity of potential destinations when leaving a given occupation, while accessibility is a measure of the diversity of different occupations from which a given occupation can be accessed. Using these metrics, we identify *condenser* occupations, with low transferability and high accessibility, that act as bottlenecks in occupational transitions.

It is important to note that when we refer to *bottlenecks* or *condenser* occupations, we do so in a descriptive manner. These terms are used to characterize structural features of the observed transition network, highlighting areas of low inter-community mobility or high condensation, without implying any normative judgment about the labor market's optimal functioning.

We apply our framework to the French labor market using employer tax data from 2012 to 2020 to construct year-on-year occupational transition matrices (Section 6.2). The transition matrix exhibits an inherent block structure, with, on average, only one out of six occupations per block exhibiting significant inter-block transitions (Section 6.2). The implied steady state of this occupational transition matrix is far from the current distribution of workers across occupations. Over half of the occupations deviate more than 10% from the implied steady state distribution (Section 6.2). Furthermore, transitions into and out of the employed labor force account for $\sim 40\%$ of all inter-community transitions, suggesting the existence of important barriers hindering direct movement between occupational communities (Section 6.2).

To identify the occupations acting as bottlenecks to reallocation we introduce two metrics: *transferability* and *accessibility*. Transferability quantifies the diversity of outside options for a given occupation, while accessibility quantifies the diversity of pathways to a given occupation (Section 6.3). This enables us to classify occupations into four categories: (1) Hub, (2) Diffuser, (3) Channel, and (4) Condenser occupations. Overall, 93% of occupations are condensers, which emerge as bottleneck occupations due to their low transferability but high accessibility: many workers can transition into these occupations but they offer limited follow-on opportunities. This aligns with human capital theory, which posits that as workers gain experience and accumulate occupation-specific skills, their incentives to transition to different occupations decrease, resulting in lower mobility over time [224, 225]. Our framework complements this theory by introducing metrics that quantify the structural properties of actual transitions, offering a novel perspective on the constraints in labor reallocation while remaining consistent with the underlying principles of human capital theory.

Furthermore, we find that transferability is negatively correlated with the number of employees in a given occupation. In Section 6.3, we show that our accessibility metric successfully identifies occupations with high betweenness centrality, a metric commonly used in the literature to determine bottlenecks [47]. In particular, occupations with high accessibility show high skill similarity with occupations inside *and* outside of the community, making them the entry-points for inter-community transitions (see Chapter 7).

Our main contribution to the literature is the additional information

gained from the transferability metric. Specifically, occupations with high transferability have a high skill similarity to occupations outside of the community, thus serving as the origins of inter-community occupational transitions (Section 7.5). Together with the negative relationship between transferability and the number of employees in an occupation, it is these occupations that may be at the heart of slow labor reallocation. To demonstrate how these metrics can effectively inform policy interventions, Section 6.4 illustrates that leveraging both transferability and accessibility metrics can help identify occupations for targeted policies aimed at reducing frictions in the labor market.

6.1. Literature Review

At the macro-level, structural shifts may be attributed to technological advancements [226] or macroeconomic events such as the *China Shock* [227] or the COVID-19 pandemic [228]. Empirical evidence suggests that worker reallocation following structural shocks is a slow process [227, 229]. Understanding the reasons for slow worker reallocation is crucial, as it may influence policymakers' decisions on labor market interventions, particularly in the context of pandemic-induced job market shifts, the green transition, and the impacts of automation.

Several factors have been identified as contributing to slow worker reallocation. First, skills are often imperfectly transferable between occupations, leading to mismatches that impede transitions [47, 221, 230]. Adão et al. (2022) [221] find that, in occupations with strong skill specificity, the pace of labor market adjustments slows down because the transition relies less on reallocating older workers and more on gradually integrating younger workers into these roles. Bocquet (2022) [47] was the first to highlight and quantify the role of bottleneck or *bridge occupations* on the speed of reallocation. He finds that reallocation is significantly slowed down due to the structure of the occupation network and emphasizes that the impact of skill frictions can be effectively modeled within a search and matching framework, making the underlying effects more tractable.

Second, local labor markets are often imperfectly connected on an occupational and sectoral level [229, 231]. Kline and Moretti (2013) [232] emphasize geographical barriers, suggesting that labor markets in highly competitive and productive cities are not well-connected, leading to a scarcity of occupational transitions. Schmutz and Sidibé (2020) [231] develop a general equilibrium search and matching model to incorporate these geographical frictions.

Third, recent approaches emphasize analyzing labor flows and their network patterns to understand occupational mobility, which tracks workers' transitions through a complex network based on various attributes [46–48, 183].

Incorporating skill data into analyses of labor flows has revealed that skill-relatedness not only facilitates labor transitions between occupations but also promotes regional diversification and local industry growth [45, 184, 186, 233]. A recent study by Aufiero et al. (2024) [185] investigates the relationship between job fitness, skill coherence, and

wages in the US labor market by employing network-based tools from the Economic Complexity framework. Going beyond these studies, we here relate network properties to slow worker reallocation by developing a set of metrics to measure and understand the patterns that lead to slow reallocation.

6.2. Occupational Transition Network of France

Using micro-data covering all employees in France, we construct an occupational transition network for the period 2012-2020 to identify bottlenecks in labor flows by examining the labor market's implied steady state, community structure, and entry and exit dynamics. We observe a scarce transition matrix with a strong community structure, where only 9% of all transitions occur across communities, and with over half of all occupations deviating by over 10% from their implied steady state. Furthermore, we show that, on average, 40% of the total inflow to each occupational community originates from outside the labor market. These results indicate that there are significant constraints in the reallocation of workers in the French labor market, which appears to be out of its equilibrium state.

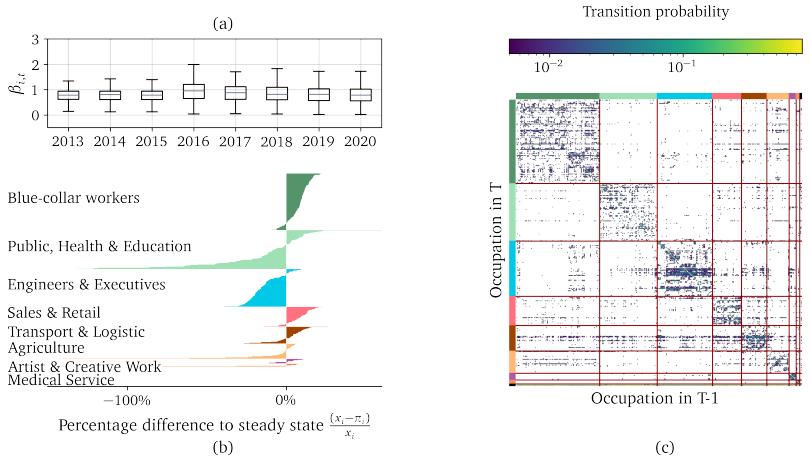


Figure 6.1.: Occupational Transition Matrix and steady state distribution (a) Temporal stability analysis from 2012-2020 where each box corresponds to the distribution of standardized coefficient $\beta_{i,t}$ (see Appendix B.2) estimated by the regression in Eq. B.6. The grey line in each box corresponds to the median of the distribution. (b) Bar plot of percentage difference between the actual share of workers in each occupation x and the share of workers implied by the steady state π. (c) Transition probability matrix $\mathcal{P}_{i,j}$, where each entry represents the fraction of individuals in $(T-1)$ (columns) who transitioned to occupation in (T) (rows). Self-loops are excluded. Red lines and colored bars indicate the BRIM communities.

Constructing the Occupational Transition Network

We follow the construction procedure described in Section 5.4, where $\mathcal{P}^T_{i,j}$ represents the occupational transition network, capturing the probability of transitioning from occupation i at time $T-1$ to j at time T. Over the period from 2012-2020 we account for 253.3 million person-year observations, among which 41.2 million (16.2%) undergo primary-to-primary occupational transitions and 184.9 million (73.0 %) remain in the same occupation, 10.9 million (4.3%) leave the network and 16.3 million (6.5%) enter the occupational network. The box-plot in Figure 6.1(a) shows the regression coefficients of the year-specific transition matrices in comparison to 2012, implying that it is reasonable to assume temporal stability for the purposes of this paper (see Appendix B.2).[1] To mitigate noise effects, links between two oc-

1: We perform the regression $\mathcal{P}^t_{i,j} = \alpha_{i,t} + \beta_{i,t}\mathcal{P}^{t'=2012}_{i,j} + \epsilon_{i,t} \quad \forall t \in \{2013, \ldots, 2020\}$ and assume temporal stability if $\alpha_{i,t} \to 0$ and $\beta_{i,t} \to 1$ (for more details see Appendix B.2 Figure B.1)

cupations are removed if the flow is below a fraction $\Theta = 1\%$ from the originating occupation.[2] For $\Theta = 0\%$ each occupation connects to 278 other occupations on average, with many connections due to a few individuals transitioning. With the filtering process we remove 5.3% of observations from the data which results in a total of 6900 existing links, with 18 links per occupation on average. This magnitude aligns with estimates provided by France Travail's ROME job classification system,[3] where an average of 13 transitions per occupation is deemed possible, considering similar skill and work environments.

2: The results provided in Section 6.3 remain robust with small changes in the order of a few percent in Θ.

3: See the official website https://www.data.gouv.fr/en/datasets/repertoire-operationnel-des-metiers-et-des-emplois-rome/.

Occupational Communities

Both the occupational transition network (Figure 5.1) and the occupational transition matrix (Figure 6.1) reveal a community structure indicating barriers to the free reallocation of workers among jobs (see Section 5.11). As expected, workers are most often *trapped* within an occupational community: on average, 91% of transitions happen within a given community and only 1 out of 6 occupations per community (on average 17%) has over 10% of their outgoing transitions to different communities.[4]

To identify the occupational communities, we utilize the BRIM algorithm (see Section 5.11), which is designed to maximize the modularity of Barber (2007) [215], due to its rapid convergence and its ability to incorporate interdependence between communities across the network [212]. To quantify the intra-community density compared to inter-community density of flows we use the network modularity introduced in Section 5.11 Eq. B.4).[5]

4: We define an occupation with an inter-community transition as one where over 10% of transitions occur outside its community. If we set the threshold at 5%, this would result in 2 out of 5 occupations per community exhibiting inter-community transitions.

5: Modularity serves as a metric for assessing the clustering of networks or graphs, quantifying the effectiveness of partitioning a network into distinct communities. Networks exhibiting high modularity contain tightly connected nodes within communities while showing limited connections between nodes belonging to different communities. The equation to calculate modularity (see Eq. B.4) is a bipartite expansion of the modularity proposed by Newman and Girvan (2004) [234].

Employing the BRIM algorithm results in 8 communities and a modularity of 0.42. The communities identified through the maximization of modularity broadly correspond to (1) Blue-collar workers, (2) Public, Health & Education, (3) Engineers & Executives, (4) Sales & Retail (5) Transport & Logistic (6) Agriculture, (7) Artist & Creative Work (8) Medical Service (see Appendix B.4 for tables with exact community membership). Although communities could be given by the INSEE's 1-digit PCS codes, we choose the BRIM communities over the INSEE partition since they have a higher modularity of 0.42, compared to 0.36 when using the 1-digit PCS.[6] The community membership structure and community size distribution determined by BRIM is robust across different clustering algorithms (see Section 5.11).

6: The 1-digit PCS occupational types are grouped into six broad subgroups: (1) Agriculture (Agriculteurs), (2) Craftsmen, traders, and entrepreneurs (Artisans, commerçants et chefs d'entreprises), (3) Managers and higher intellectual professions (Cadres et professions intellectuelles supérieures), (4) Intermediate Occupations (Professions intermédiaires), (5) Employees (Employés), and (6) Workers (Ouvriers). Note that within the Agriculture category there is only one single occupation.

The low inter-community mobility represents *bottlenecks* in occupational flows in the French labor market (see [47]). Similarly, Neffke and Henning (2013) [45] also find an inherent community structure within the occupational transition network for Sweden, suggesting frictional reallocation. They attribute this phenomenon primarily to the skill-relatedness of different jobs: individuals are more likely to transition to jobs that require similar skill-sets, whereas different skill-sets represent a higher barrier to making a given transition thus potentially requiring additional training for some transitions.

Stationary Distribution of Occupations

Assuming that the right-stochastic transition matrix $\mathcal{P}_{i,j}$ remains constant (consistent with our findings), it can be thought of as the transition matrix of a Markov chain, which implies a steady state distribution of occupational employment shares to which the system converges. As already discussed in Section 5.7, the implied steady state distribution $\vec{\pi}$ over the different occupations is given by the eigenvector associated with the eigenvalue $\lambda = 1$ of the transition matrix, while the rate of convergence for individual occupations is given by the eigenvector associated with the second largest eigenvalue.

While the steady-state distribution does not necessarily represent an optimal or desirable outcome from an economic perspective, it serves as a useful benchmark for understanding the structural dynamics of the labor market. The steady-state distribution provides a baseline that reflects the equilibrium implied by the observed transition matrix. It allows us to quantify deviations from this baseline and identify occupation communities where flows appear constrained or misaligned with the tendency of the network. By benchmarking against the steady-state distribution, we can identify patterns of excess or underemployment, informing where interventions might be necessary to facilitate smoother labor market transitions.

We analyze the implied steady-state distribution for each BRIM community and find that the current state of the labor market deviates from the implied steady state. This deviation can be explained by the block structure of the occupational transition matrix due to limited inter-community flows (see Figure 6.1).

Figure 6.1(b) shows the percentage difference between the actual state $\vec{x}$ and the implied steady state $\vec{\pi}$ for each occupation. Deviation of the current distribution from the implied steady state suggests either a slower convergence rate compared to structural changes or insufficient time for the system to reach equilibrium. These observations indicate labor market constraints that impede the system from reaching a steady state. Moreover, the majority of transitions occurs within communities, which explains the over and under-representation of workers in different communities in comparison to the occupation's implied steady state.

Over half of all occupations (223 out of 372 occupations) exhibit deviations of more than 10% from the steady state distribution. Specifically, 88.1% of occupations classified as Blue-collar workers, 84.2% of Sales & Retail, and 69.7% of Transport & Logistics occupations experience excess employment. In contrast, the communities of Engineers & Executives and Public, Health & Education are underemployed, with 86.1% and 64.0% of their occupations, respectively, having fewer workers than the steady-state distribution suggests. In other communities, the number of occupations with excess employment and underemployment are balanced (see Figure 6.1(b)). The half-time until convergence of the system to its implied equilibrium is given by the spectral gap of $\mathcal{P}_{i,j}$, which yields $\frac{1}{1-\lambda_2} \approx 31$ years (see Section 5.9). Putting this timescale into perspective is challenging, but if worker reallocation were fast, we would expect an order of magnitude of few years rather than few decades. For example, Bocquet (2022) [47] finds that

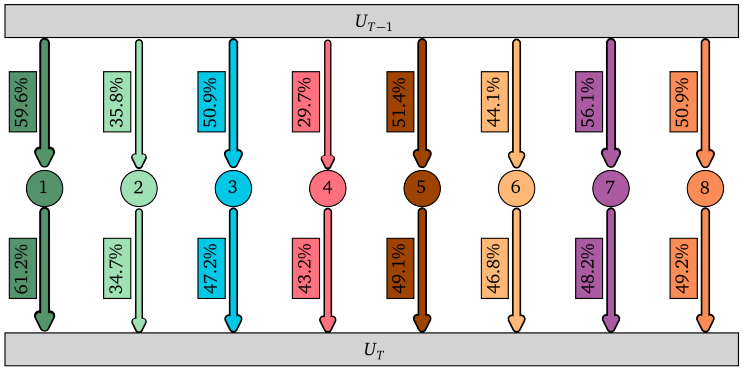


Figure 6.2.: Inflow and outflow dynamics of the Labor Market: Each node in the network represents an occupational BRIM community: (1) Blue-collar workers, (2) Public, Health & Education, (3) Engineers & Executives, (4) Sales & Retail (5) Transport & Logistic (6) Agriculture, (7) Artist and Creative Work (8) Medical Service and (U) Out of the occupation network. Flows indicate transitions to and from outside the labor market normalized by the total number of transitions from and to the community.

the speed of labor market reallocation after an asymmetric shock is around 1 year. This is because they assumes a new transition matrix due to the shock and measures transition time using a normalized cumulative impulse response. In contrast, we assume a stable transition matrix and measure the timescale using the spectral gap between the first and second eigenvalues of the transition matrix. The convergence timescale of 31 years, as discussed in Section 5.7, indicates that reallocation of workers to the implied steady state is slower than the pace of small structural shifts in the labor market. This suggests that the steady state is unlikely to be fully realized before it evolves again.

Entry and Exit from the Occupation Network

Given the low inter-community mobility within our network, the primary alternative for entering or leaving a community is through transitions into or out of the occupational network itself. Over the period of 2012-2020, 16.3 million (2.0 million p.a.) individuals entered the occupational network, while 10.9 million (1.4 million p.a.) exited, resulting in a net inflow of 600,000 workers per year on average.

Figure 6.2 shows the inflow and outflow dynamics for each occupational community from and to outside of the labor market. The width of link from node i to j in Fig. 6.2 corresponds to the proportion of people transitioning from i to j out of the total number of individuals leaving i and the total number of individuals transitioning into j.

On average 47.3% of transitioning workers enter an occupational community from outside the occupational network, while 47.5% of those leaving an occupational community enter a state of unemployment (see Figure 6.2). This underscores the substantial influence of unemployment on inter-community transitions. The prevalence of an interim unemployment phase preceding inter-community transitions suggests considerable bottlenecks for direct transitions between communities. We leave the study of the unemployment dynamics to future work, as our sample only includes observations in the period starting $T-1$ to the end of T, implying that we cannot ascertain where those unemployed at the end of T potentially re-enter the occupational network in $T+1$.

When analyzing the occupational network across different age groups, we find that for individuals under 30, the number of workers entering

the network is 25% higher than those exiting. In contrast, for individuals over 60, the number of workers leaving the network exceeds the number entering by 100%. (see Appendix B.3 Figure B.2). This implies a strong role for job-entry and retirement exit dynamics across the unemployment transitions. To adjust for these effects, we note that excluding individuals under the age of 18 entering the occupational network, and those above the age of 64 leaving the occupational network, the ratio between transitions into and out of the labor force with the total number of direct occupation-to-occupation transitions shrinks from 14.2% to 6.3%, such that we can safely neglect these dynamics in the analysis.

6.3. Identifying Bottlenecks: A Data-driven Taxonomy

The occupational transition network developed in Section 6.2 shows evidence of a strongly clustered structure, with a majority of transitions occurring within a community. In this Section we apply methods from the fitness and complexity literature to develop a taxonomy of occupations, classifying them either as (i) Hubs, (ii) Condensers, (iii) Diffusers, or (iv) Channels, based on their accessibility and transferability. Accessibility (fitness) describes the diversity of occupations from which a given occupation can be reached, while transferability (inverse complexity) describes the diversity of occupations that can be reached from a given occupation. We show that the majority of occupations in the French labor market exhibit a higher accessibility than transferability, resulting in workers becoming *condensed* into these occupations, with limited opportunities for transitioning elsewhere, thus Condenser and Channel occupations emerge as bottleneck occupations. We find that the accessibility metric correctly identifies bottlenecks as it shows high correlation with betweenness centrality. Additionally transferability correlated with closeness centrality showing that entering high transferability occupation helps workers to diffuse into different communities of the network. Furthermore, we investigate the effect of gender and in Chapter 7 we show that our metrics provide information not otherwise contained in factors such as skill similarities across occupations (Section 7.5). We do not find any age related effect on our metrics (see Appendix B.7). Future research could be directed towards understanding how the pressure originating from high accessibility but low transferability might, to some extent, explain the occupational wage-distribution.

Accessibility and Transferability of Occupations

We quantify the accessibility and transferability of an occupation by measuring the diversity of transitions to and from that occupation using economic fitness and complexity metrics that describe a hierarchy as nested patterns in a system [235]. Economic fitness and complexity has previously been used to quantify the productivity and technology composition of economies [235–237], with a particular focus on economic and technological growth and development using patent and export data [238–240]. Fitness and complexity have only recently been

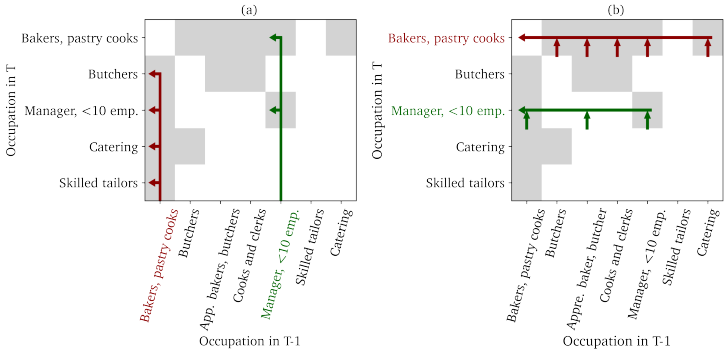


Figure 6.3.: Schematic visualization of the degree of Transferability and Accessibility. An illustrative representation of a transitional occupation matrix $\tilde{\mathcal{P}}_{i,j}$, where a grey square signifies a connection between two occupations ($\tilde{\mathcal{P}}_{i,j} > 10^{-2}$), while a white square denotes no connection. (a) Transferability of an occupation is linked to the number of occupations a worker can transition to from a given occupation in time T-1. (b) Accessibility of an occupation is associated with the number of occupations from which a worker could have transitioned to a given occupation at time T.

related to labor context [185], yet in a different manner (see Section 7.7).

The fitness, $F_i^{(n)}$, of an occupation i is an extensive measure defined as the sum over the complexities of the occupations j that lead to the given occupation i. The fitness can be interpreted as an occupation's *accessibility*, quantifying the ease and extent of workers transitioning from other occupations to occupation i.[7] The complexity, $Q_j^{(n)}$, of occupation j indicates the difficulty of transferring from occupation j to other occupations. The inverse complexity $1/Q_j^{(n)}$ of an occupation thus serves as an indicator of the occupational opportunities offered by j, and we refer to it as the *transferability* of an occupation.

Following Tacchella et al. (2012) [235], these two metrics are determined as the fixed point solution to

$$\tilde{F}_i^{(n)} = \sum_j \tilde{\mathcal{P}}_{i,j} Q_j^{(n-1)} \qquad \tilde{Q}_j^{(n)} = \left(\sum_i \tilde{\mathcal{P}}_{i,j} \frac{1}{F_i^{(n-1)}} \right)^{-1} \tag{6.1}$$

$$F_i^{(n)} = \frac{\tilde{F}_i^{(n)}}{\sum_i \tilde{F}_i^{(n)}} \qquad Q_j^{(n)} = \frac{\tilde{Q}_j^{(n)}}{\sum_j \tilde{Q}_j^{(n)}} \tag{6.2}$$

with $i \in \mathbb{V}^{T-1}$, $j \in \mathbb{V}^T$ where $\mathbb{V}^T$ is the space of occupations in year T, and n is the iteration index.[8] The transition probability matrix $\tilde{\mathcal{P}}_{i,j}$ is conditioned on transitions only, excluding self-loops such that $\mathrm{Tr}(\tilde{\mathcal{P}}_{i,j}) = \sum_i \tilde{\mathcal{P}}_{i,i} = 0$. Note that the fitness of an occupation after one iteration $F_i^{(1)}$, is the weighted number of occupations from which individuals transitioned into occupation i. Meanwhile, the complexity of an occupation after one iteration $Q_j^{(1)}$, is determined by the inverse of the weighted count of occupations to which individuals transition starting from occupation j. It is important to note that these metrics only have an interpretation in relative terms, but not in absolute terms as there is no unit attached.

Figure 6.3 shows a schematic example of transferability (a) and accessibility (b) based on $\tilde{\mathcal{P}}_{i,j}$. A grey square indicates a link between two occupations, i.e. $\tilde{\mathcal{P}}_{i,j} > 10^{-2}$. In Figure 6.3 bakers and pastry cooks exhibit higher transferability than manager as they have more opportunities for transitions. Similarly, the occupation of bakers and pastry

7: Note that an occupation is more accessible when it is reached by workers from occupations that have limited options for transition. Thus, an occupation accessible to a wide range of other occupations contributes less to the transferability of that occupation. Instead, higher accessibility is associated with occupations that are predominantly transitioned into by workers from other occupations without dispersing elsewhere.

8: The initialization is done via $\tilde{F}_i^{(0)} = 1 \ \forall i$ and $\tilde{Q}_j^{(0)} = 1 \ \forall j$. We use 200 iterations and confirm that that the algorithm converges (the fixed point solution is reached).

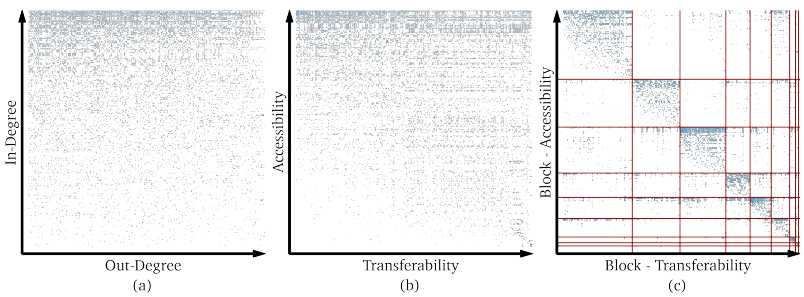


Figure 6.4.: Sorted, Transition Probability Matrix. X-axis corresponds to occupations in (T-1) and y-axis to occupations in (T). (a) Diversification and Ubiquity ordering (b) Accessibility and Transferability ordering (c) In-Block Accessibility and Transferability ordering. Each Block, detected by the BRIM algorithm (red lines), has a higher NODF than the raw matrix. The intensity of the coloring corresponds to higher probability flow from (T-1) to (T).

cooks is more accessible than the occupation of a manager, i.e. more different occupations lead to bakers and pastry cooks than to a manager.

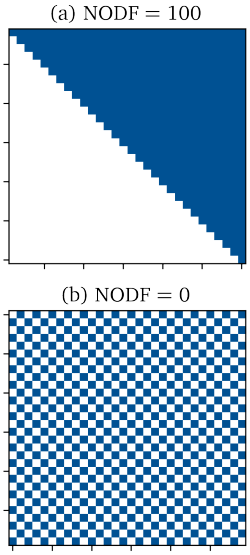
Nestedness in Occupational Transitions

The convergence of the fitness and complexity algorithm depends on the transition matrix exhibiting a specific structural property known as nestedness [235]. Nestedness refers to a hierarchy in the connections of a system, where less-connected nodes are subsets of better-connected ones (see Appendix B.1 and Illustration 3). This property ensures that the iterative process for calculating Accessibility and Transferability reaches a stable fixed point.

Ordering the axes of the binary transition matrix based only on the in- and out-degree of the respective occupation does not reveal a nested structure (Figure 6.4 (a)). However, when sorting the x-axis by transferability and the y-axis by accessibility, an inherent nested structure within the labor market becomes apparent (see Figure 6.4 (b)). In this arrangement, the bottom-right corner of the matrix contains on the x-axis occupations characterized by the highest transferability and on the y-axis occupations with lowest accessibility (see Figure 6.4 (b)). This nested hierarchy reflects the labor market's flow dynamics: occupations higher on the y-axis are broadly accessible from a diverse array of occupations, while those along the x-axis offer more diverse opportunities for outgoing transitions (see also Section 5.3).

To quantify this nestedness, we use the NODF metric (Nestedness based on Overlap and Decreasing Fill), which measures the extent to which the connections of less-central nodes are subsets of those of better-connected nodes. As visualized in Illustration 3, NODF values range from 0 (indicating no nestedness, akin to a checkerboard pattern) to 100 (indicating perfect nestedness). For the French labor market, we find a NODF value of 10.01, which is comparable to other economic systems [241], such as the company-product space, where NODF values are around 12 [242]. This level of nestedness supports the convergence of the fitness and complexity algorithm, allowing us to reliably compute Accessibility and Transferability metrics.

However, these metrics do not yet account for the community structure of the labor market. To incorporate this, we evaluate the in-block accessibility and transferability for communities identified using the BRIM algorithm (Figure 6.4(c)). Within these communities, the nestedness increases, with a mean NODF value of $\langle \text{NODF} \rangle = 12.56$, exceed-



Ill. 3: Nestedness of binary matrices: (a) Binary matrix that is fully nested, NODF=100. (a) Checkerboard matrix that is not nested, NODF=0.

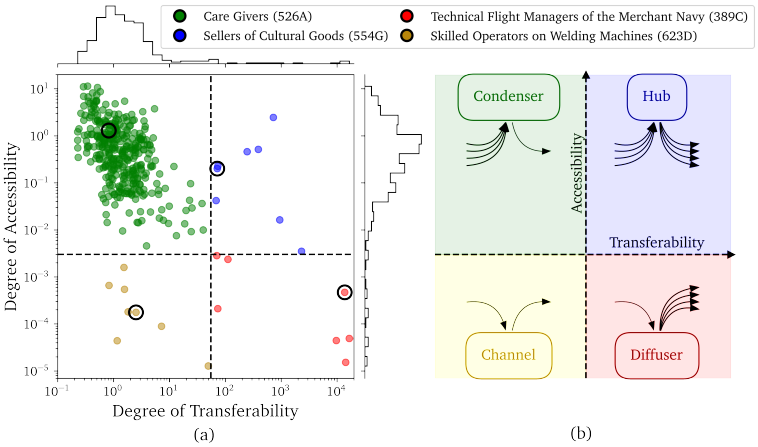


Figure 6.5.: Accessibility and Transferability Space. (a) Each point represents a single occupation and is located in the space by its corresponding accessibility (y-axis) and transferability (x-axis) metric. The dashed lines denote the mean-shift clustering boundaries at $\Theta^A = 0.0031$ and $\Theta^T = 49.9$. The histograms show the distribution along the accessibility (horizontal) and transferability (vertical) dimensions. (b) Proposal of a new taxonomy based on the degree of accessibility and transferability of an occupation into four categories. The arrows schematically represent the potential number of incoming and outgoing transitions for the different categories.

ing the overall matrix's nestedness. As already found in Section 6.2 most transitions occur within these communities, but some extend beyond them (see Figure 6.4 (c)). Transitions outside the community are mainly associated with occupations having high transferability, suggesting a bottleneck effect for individuals in low-transferability occupations aspiring to transition beyond their community. These findings already suggests a potential avenue for increasing reallocation speed since one can identify precisely occupations that cause frictions (see Section 6.4).

Accessibility and Transferability of the French Labor Market

Figure 6.5(a) shows the distribution of accessibility (y-axis) and transferability (x-axis) for all occupations in the French labor market (defined in Section 6.2). For the interpretation of accessibility and transferability we have to focus on relative terms, i.e. the comparison between two occupations. From Figure 6.5(a) we note that the accessibility of occupations spans seven orders of magnitude, while the transferability spans six. The majority of occupations show a high degree of accessibility relative to their degree of transferability, suggesting that they *condense* the flow of labor: moving between these occupations slowly reduces the outside options of the workers. In contrast to this, there is only a handful of occupations where relative transferability remains high while relative accessibility is also high. Let us stress again that our analysis focuses solely on an average year-on-year transition matrix due to data limitations. Specifically, we do not directly observe individuals' residence times in a given occupation, which may affect the diversity of outside options (transferability) of the occupation, thus potentially affecting the dispersion of transferability and accessibility of the occupational space.

To study this phenomena of condensation, and identify associated bottlenecks in labor reallocation, we propose a taxonomy that classifies occupations into four groups (see Figure 6.5(b)):

- **Hubs** are characterized by a high transferability and accessibility. These occupations are broadly accessible to individuals from many occupations while also providing access to a variety of

other occupations and often do not require specialized training or education. For example, sellers of cultural goods (554G).

- **Condensers** are characterized by a high accessibility but low transferability. While accessible from a broad range of occupations, these occupations allow for transitions only to a few other occupations, thus funnelling individuals, e.g. care givers (526A).
- **Diffusers** are characterized by a low accessibility but high transferability. As the inverse of the Condenser, these occupations are hard to reach, often due to specialized training, but exhibit a diversity of opportunities, such as officers and technical flight managers of the merchant navy (389C).
- **Channels** are characterized by both a low accessibility and a low transferability, representing occupations with only a few occupations to transition from and into. For example, skilled operators on welding machines (623D).

The classification of occupations into this taxonomy depends on two thresholds, Θ^A and Θ^T, that separate the low and high regimes for the accessibility and transferability of occupations.[9] Figure 6.5(a) shows that the defined taxonomy clusters are well partitioned in the data of the French labor market.[10] Appendix B.5 provides a comprehensive table of occupations within each cluster (Table 9-12).

The cluster of Condenser occupations contains the largest number of occupations (348 out of 371), describing 94% of the total person-year datums in the dataset and 94% of estimated total wages in 2020, thus indicating the economic significance of this phenomenon.[11] The categories of Channel (8 occupations), Hub (8 occupations), and Diffusers (7 occupations) make up the remaining set of occupations. The large share of condensers could be interpreted as a general trend of specialization: over the course of a career one specializes in terms of experience and skill, therewith inadvertently reducing the transferability to other occupations. In fact, for 3 out of 5 occupations (59.6%) the expected transferability of the next occupation is lower than a person's current occupation. This aligns with human capital theory, which posits that as workers accumulate specific human capital, the returns to staying in their current role increase, making transitions less attractive [224].

However, it is important to note that while human capital theory provides one explanation, our analysis focuses solely on occupational transitions between the current and subsequent occupations across all workers, without tracing individual career paths through the labor market.

Figure 6.6 shows the relation between the number of employees in a given occupation to the occupation's accessibility and transferability metrics. As expected, highly accessible occupations have a higher number of employees given the diversity of occupations, both from within and across communities, that transfer into these occupations. By contrast, more transferable occupations, i.e. those with a more diverse set of outside options, are generally smaller. This highlights the condensation effect: within a community of occupations, there are inroads via highly-accessible occupations, but only a smaller fraction of individuals transition into highly transferable occupations that allow

9: Thresholds are determined by a mean-shift algorithm [243]. Specifically, we use the MeanShift implementation from the scikit-learn package in Python with a bandwidth of 3 [220]. To ensure robustness, we employed both spectral clustering and k-means clustering (for each dimension separately), which resulted in similar threshold values [220].

10: We perform a significance testing with a random network, preserving both in- and out-degree distributions as in the labor market, to confirm that the cluster sizes of the French labor market are significant. This suggests that the findings are intricacies specific to the labor market's flows, and not merely due to the distribution of in- and out-degrees of occupations.

11: We compute estimated total wages as the sum over occupations' median wage multiplied by the number of employees in the occupation

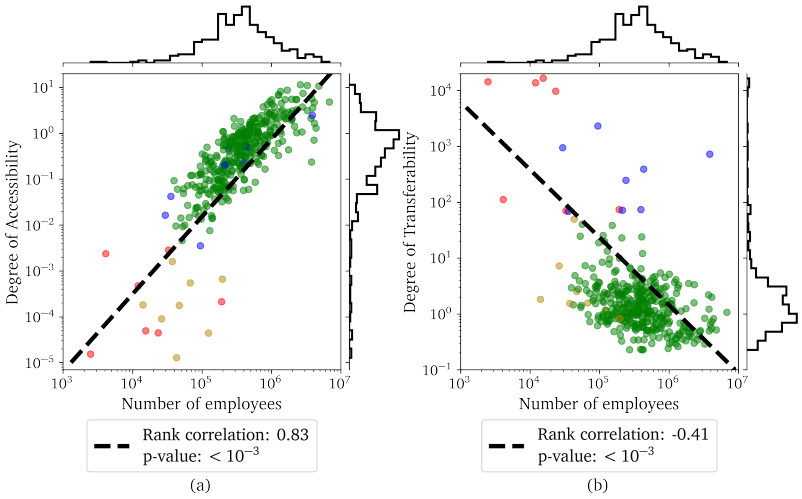


Figure 6.6.: Occupation size effect on Accessibility and Transferability. Scatter plot depicts the relationship between the accessibility (a), and transferability (b) and the number of employees in a given occupation. A dashed line serves as a visual guide and represents the linear fit of the points. There is a significant and positive rank correlation of 0.83 between accessibility and number of employees and a significant and negative rank correlation of -0.41 observed between number of employees and transferability.

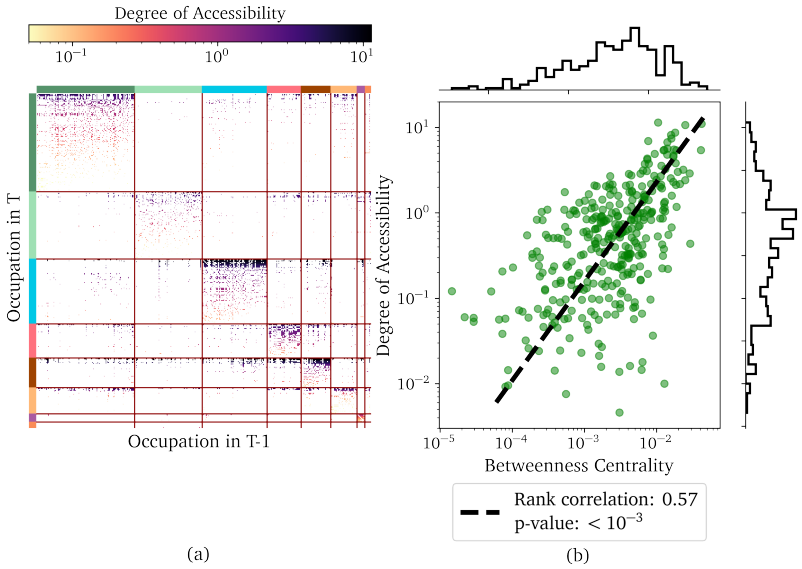


Figure 6.7.: Correlation of Accessibility and betweenness centrality. (a) Heatmap of the occupational transition matrix, with rows colored according to the accessibility of each respective occupation. (b) Scatter plot depicting the relationship between the accessibility of an occupation and its betweenness centrality, a measure of bottlenecks, showing a significant positive rank correlation of 0.57. Only condenser occupations are shown. A dashed line serves as a visual guide and represents the linear fit of the points.

for a higher degree of inter-community mobility.

Identifying Bottlenecks in the French Labor Market

The nested and clustered structure of the occupational transition matrix implies the presence of bottlenecks—occupations that account for a disproportionately large share of inter-community transitions compared to intra-community transitions. The accessibility of an occupation captures this particular phenomenon. Figure 6.7(a) shows the occupational transition matrix colored by the degree of accessibility (dark having high accessibility) and split into the 8 BRIM communities (Section 3). One can clearly see that, intuitively, those occupations who receive a higher share of inflow from inter-community transition are also those with a higher accessibility, thus these occupations represent the handful of entry-points into different occupational communities. By contrast, those occupations with a low accessibility receive workers almost exclusively from within their given community (See Figure B.3 in Appendix B.6). Prior research has used the betweenness centrality [244] (Eq. 5.37), i.e. the fraction of shortest paths between

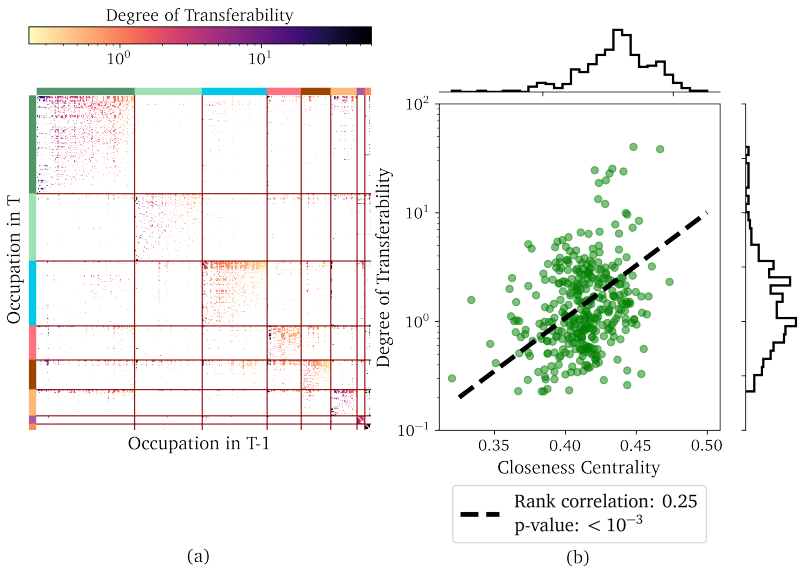


Figure 6.8.: Correlation of Transferability and closeness centrality. (a) Heatmap of the occupational transition matrix, with columns colored according to the transferability of each respective occupation. (b) Scatter plot depicting the relationship between the transferability of an occupation and its closeness centrality, a measure of a nodes ability to spread information in a graph, revealing a significant positive rank correlation of 0.27. Only condenser occupations are shown. A dashed line serves as a visual guide and represents the linear fit of the points.

any two nodes that pass through a given node, to identify bottlenecks in network structures (see Bocquet (2022) [47] for an application to labor markets). Focusing on the condenser occupations, comprising 94% of the employed workforce, Figure 6.7(b) shows that our accessibility metric captures the betweenness centrality quite well, with a significant rank correlation of 0.57.[12] Accessibility and betweenness centrality identify bottlenecks from an ex-post perspective by focusing on the diversity of incoming transitions, highlighting occupations within a community that serve as *targets* for workers transitioning from other communities (i.e., considering both incoming and outgoing links). However, this approach does not address the ex-ante perspective of the worker—namely, the potential opportunities they can access from a given occupation. To capture this additional dimension, we use the transferability metric, which allows us to identify bottlenecks not only from the perspective of where workers arrive in T (accessibility) but also from the perspective of where they start in T-1 (transferability).

12: We use rank correlation instead of Pearson correlation because, for accessibility and transferability metrics derived from the economic fitness and complexity framework, the relative ordering of occupations is more important than the absolute differences in their values.

Figure 6.8(a) shows the occupational transition matrix color-coded by the transferability of a given occupation. The figure shows that it is those occupations with a high transferability that have a higher dispersion in possible destinations, both within and across communities. As with high-accessibility occupations, there are only a few occupations per community that show a high degree of transferability, and, as highlighted in Section 6.3, high-transferability occupations tend to have a smaller number of employees. Furthermore, higher transferability is positively correlated (although in weaker fashion than for accessibility) with the ratio of inter- to intra-community transitions (see Figure B.3 in Appendix B.6). Thus, occupations with high transferability are more likely to facilitate transitions between communities. This role is further supported by the positive correlation between transferability and closeness centrality (Eq.5.38), as shown in Figure 6.8(b). The consequence of this being that low transferability occupations reflect the here-presented definition of bottlenecks: they limit the degree to which there is free mobility between occupations. This is not intended

to suggest any occupation should be able to transition to any other occupation (e.g. in the sense that some occupations require specialized training), but rather that this is, empirically, the identified structure of the labor market.

Gender Effect on Accessibility and Transferability

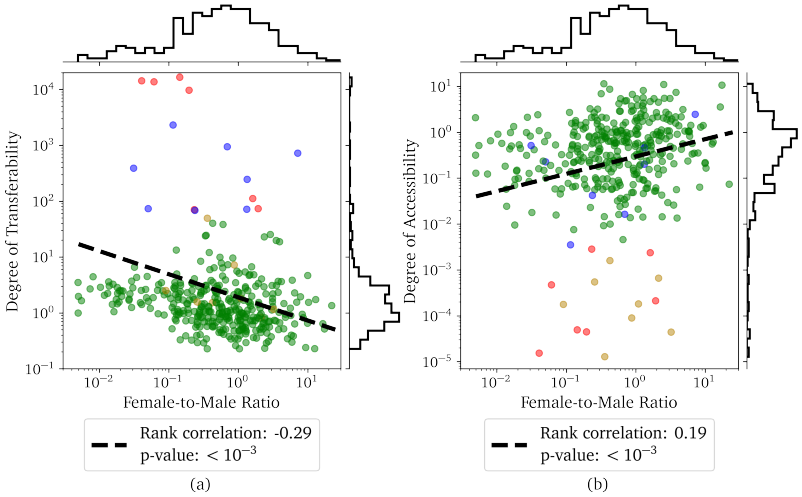


Figure 6.9.: Gender effect on Accessibility and Transferability. Scatter plot depicts the relationship between Female-to-Male ratio (a) the degree of transferability and (b) the degree of accessibility for each occupation. Color coding corresponds to the introduced occupation taxonomy. A dashed line serves as a visual guide and represents the linear fit of the points. There is a significant and negative rank correlation of -0.29 observed between Female-to-Male ratio and transferability and a significant and positive rank correlation of 0.19 between Female-to-Male ratio and accessibility (legend of (a,b)).

The female-to-male ratio of employees in an occupation has a significant influence on the accessibility and transferability metrics. The rank correlation between transferability and the gender ratio is at -0.29, while for accessibility the correlation is 0.19, both with significant p-values of $< 10^{-3}$ (see Figure 6.9 (a,b)). Occupations with lower transferability tend to exhibit higher female-to-male ratios, whereas those with higher accessibility tend to have higher female-to-male ratios. This discrepancy between the transferability and accessibility and gender ratios in different occupations is captured by our metrics.

To assess gender differences, we construct separate male and female networks from the data and independently calculate their accessibility and transferability metrics. Figure B.5(a) in Appendix B.7 shows that within the female network, 52 occupations that are classified as Condensers in the male network are classified as Hub, Diffuser, and Channel in the female network. This means that the dispersion of transferability and accessibility of occupations is much wider, and the female labor flow network reveals more occupations with fewer transitions between them and more occupations that are challenging to access. Since only the relative difference between the accessibility and transferability matters, when it comes to reallocating workers in response to shocks, female workers face higher constraints than male workers.

6.4. Informing Policy using Accessibility and Transferability

In the previous sections, we have shown that there are several occupations that can act as bottlenecks between occupational communities,

and we can identify these using accessibility and transferability metrics. These occupations may be part of the reason we observe a significant deviation between the realized distribution of occupations and the steady state implied by the occupational transition matrix.[13] While the steady-state distribution does not necessarily represent an optimal labor market outcome, it allows us to quantify deviations from this baseline applying labor market interventions. In this section, we show how to improve the speed of convergence, i.e. to reduce frictions, through a small rebalancing of the occupational flows. In particular, for a given occupation rebalancing outgoing flows away from high accessibility and towards high transferability targets leads to faster convergence towards the implied steady state.

13: We here explicitly do not make a statement about whether the implied steady state is desirable.

For the occupational transition matrix, the rate of convergence to the steady state is determined by the second eigenvalue λ_2[14], with a lower eigenvalue indicating a faster rate of convergence. Therefore, our objective is to minimize $\Delta\lambda_2 = \lambda_2^\star - \lambda_2$, where $\lambda_2^\star$ is the second eigenvalue of the transition matrix following a small perturbation $\mathcal{P}^\star = \mathcal{P} + \epsilon V$. Here V is a $N \times N$ matrix indicating the location of the interventions. Specifically, we consider the simple case where for a chosen occupation k, we increase the flow to occupation l^+ in equal magnitude to a decrease of the flow to occupation l^-, thus preserving the normalization of $\mathcal{P}_{i,j}$. Mathematically, one can write

14: Note that the rate of convergence is determined by the spectral gap between the first and second eigenvalues. Since the transition probability matrix is stochastic, the first eigenvalue is $\lambda_1 = 1$.

$$(V)_{i,j} = \begin{cases} -1, & i = l^-, j = k \\ 1, & i = l^+, j = k \\ 0, & \text{else} \end{cases} \tag{6.3}$$

The constant $\epsilon \ll 1$ gives the size of the intervention, which is assumed to be small so as to keep fixed the implied steady state distribution and the distributions of the accessibility and transferability metrics. We consider here the case of $\epsilon = 0.01$, whose interpretation is that we reallocate 1% of the outgoing labor flows for occupation k.[15]

15: To assess the effects of the implemented strategies on these metrics, we compute the rank correlation between the metrics of the original and perturbed networks. The transferability metrics show a significant average rank correlation of 0.961, while the accessibility metrics exhibit a significant average rank correlation of 0.975 across all policies. These findings indicate that even with minor perturbations, the relative values of the accessibility and transferability metrics remain stable. Furthermore, the mean absolute percentage difference between the steady states of the perturbed matrices and the unperturbed matrix is below 10^{-4}.

Using matrix perturbation theory, the first-order change in the second eigenvalue can be approximated by[16]

$$\Delta\lambda_2 = \vec{v}_2^T V \vec{v}_2 = (\vec{v}_2^T)_k \cdot ((\vec{v}_2)_{l^+} - (\vec{v}_2)_{l^-}), \tag{6.4}$$

where we have used the second eigenvector, $\vec{v}_2$, of the unperturbed transition matrix (see Appendix B.8). As shown in Section 5.9, the second eigenvector $\vec{v}_2$ describes the slowest mode by which the system approaches its stationary distribution. States with large magnitudes, $|(\vec{v}_2)_i| \to 1$, dominate the system's slow dynamics, as they correspond to states with the highest contribution to this relaxation mode. Sign differences in the entries indicate partitioning of states, as they represent opposite directions in the evolution of the system's slow mode, with states in different partitions exhibiting divergent behavior as the system converges to equilibrium.

If the value of the eigenvector of the target occupation, $(\vec{v}_2)_k$, has a different sign than the difference in eigenvector values of the source and destination occupations, $((\vec{v}_2)_{l^+} - (\vec{v}_2)_{l^-})$, the change in second

16: Note that, to apply first-order matrix perturbation, the matrix must be symmetric. The occupational transition matrix has a symmetry value of $\frac{\|S\|_F - \|A\|_F}{\|S\|_F + \|A\|_F} = 0.9$ (see Section 5.8 and Eq.5.7). This indicates that the occupational transition matrix is almost symmetric. Furthermore, our numerical results demonstrate that the predictions from first-order perturbation theory align well with the numerical findings (see Figure 6.10(a) and Appendix B.8, Figure B.6).

eigenvalue, $\Delta\lambda_2$, is negative:

$$\Delta\lambda_2 < 0 \text{ if } \begin{cases} (\vec{v}_2)_{l^+} > (\vec{v}_2)_{l^-} \wedge (\vec{v}_2)_k < 0 \\ (\vec{v}_2)_{l^+} < (\vec{v}_2)_{l^-} \wedge (\vec{v}_2)_k > 0 \end{cases} \tag{6.5}$$

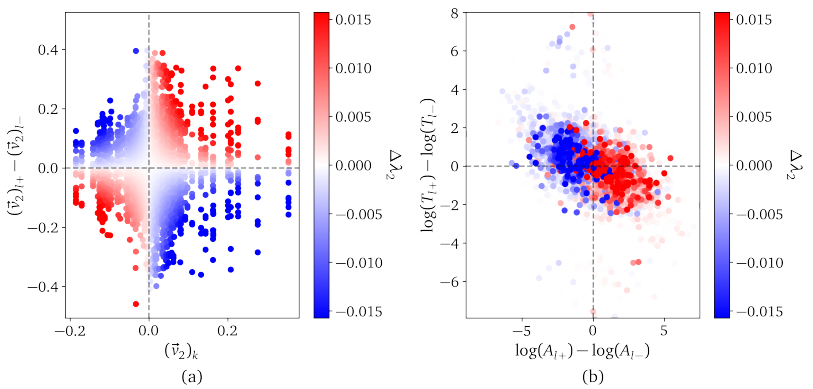


Figure 6.10.: Reducing constraints through effective policy making (a) Scatter plot showing the eigenvector value difference between the destination and source occupation $(\vec{v}_2)_{l^+} - (\vec{v}_2)_{l^-}$ against the eigenvector value of the target occupation $(\vec{v}_2)_k$ (see Eq.6.4). The color coding indicates the change in the second eigenvalue, $\Delta\lambda_2$. (b) Scatter plot showing the relation between log-difference of source and target occupation in transferability $\log(T_{l^+}) - \log(T_{l^-})$ and accessibility $\log(A_{l^+}) - \log(A_{l^-})$. The color coding indicates the change in the second eigenvalue, $\Delta\lambda_2$.

Figure 6.10(a) shows the distribution of $(\vec{v}_2)_k$ versus $(\vec{v}_2)_{l^+} < (\vec{v}_2)_{l^-}$ and the consequent distribution of eigenvalue changes $\Delta\lambda_2$. Positive and negative changes in the adjustment speed $\Delta\lambda_2$ align with calculated regimes (see Eq. 6.5). While the bulk of the distribution is concentrated around 0, there are several significant outliers in the $(\vec{v}_2)_{l^+} < (\vec{v}_2)_{l^-} \wedge (\vec{v}_2)_k > 0$ quadrant, hinting at room for policy interventions to improve labor reallocation. Improvements in adjustment speed (measured by $\Delta\lambda_2$) can be achieved by considering accessibility and transferability metrics. Figure 6.10(b) illustrates the relationship between the log difference in accessibility and transferability metrics of the flow-increase occupation l^+ and the flow-decrease occupation l^- on, along with the difference in the second eigenvalue, $\Delta\lambda_2$. Each point represents a policy that constitutes a perturbation V for all possible combinations of target k and destination l^+ occupation.[17] The eigenvalue difference is more likely to be negative when the accessibility of the flow-increase occupation is lower than that of the flow-decrease occupation, with a significant correlation coefficient of 0.26 (Appendix B.8). Simultaneously, we find that $\Delta\lambda_2 < 0$ when the transferability of the destination occupation is higher than the transferability of the source occupation, with a significant correlation coefficient of -0.13 (Appendix B.8). In terms of our taxonomy, this implies an improvement in the speed of convergence to the implied steady state when we shift the target of outgoing individuals to diffuser occupations in place of condenser occupations. Practically, this implies fewer individuals transitioning into already large occupations (see Section 6.3) with a low transferability and limited further mobility. For example, increasing the flow from *Executives in charge of economic, financial and commercial studies (372A)* to *Banking Operations Executives (376B)*, while reducing the flow to *IT project managers, IT managers (388C)*, has a strongly positive impact on the convergence speed.

Historically, the French state's labor market policies have focused on on-the-job training together with technical and vocational education [245], which links to the latter objective: increasing the transferability of currently low-transferability occupations. While we do not

17: For each link in the occupational transition matrix from occupation k to l^+ where $\mathcal{P}_{i,j} > 0.01$ a perturbation V is applied. With this, the target and destination occupations are given. The source occupation is then chosen at random from existing links of the target occupation k with a flow greater than epsilon, ensuring that $\mathcal{P}_{i,j} \geq 0 \ \forall i, j$.

focus here on the *how* of implementing labor market policy, the above results present an intuition about which occupations to target with policy, presenting a set of channels that may optimally reduce frictions under the assumption that the implied steady state is indeed a desirable target. At this point, we should however note that while we categorize different occupations by their accessibility and transferability, our dataset does not distinguish between the supply (unemployed and their characteristics) and the demand (vacancies) drivers of the transition rates between occupations. Nonetheless, integrating information with accessibility and transferability measures can enhance the effectiveness of a policy aimed at alleviating bottlenecks in the labor market.

6.5. Conclusions

In this chapter, we analyzed the labor flow network of all economically active workers in the French labor market using micro-data from French employers (BTS-POSTES). Here the observed flows represent realized transitions within the labor market, as opposed to potential transitions based on skill similarity. Realized transitions provide a more accurate reflection of the complex dynamics emerging from the interaction of workers' preferences, employers' hiring decisions, and the prevailing economic conditions. The observed transition network, therefore, provides a richer empirical perspective, capturing structural constraints and revealing patterns of mobility that extend beyond what can be inferred from skill similarity alone. We demonstrated the presence of significant constraints in the labor market evidenced by strong communities and a lack of inter-community transitions, hindering reallocation of workers (Section 6.2). We then employed a quantitative method to assess the accessibility and transferability of occupations, using an economic fitness and complexity framework, to identify precise bottlenecks in the labor market (Section 6.3). We showed that the emergence of bottlenecks can be explained by a condensation effect due to occupations with high accessibility but low transferability, that is in line with the principles of human capital theory, where workers' accumulation of occupation-specific skills reduces their incentives to transition to other roles. Finally, in Section 6.4, we introduced a quantitative approach with the aim of evaluating policy consequences using the rate of convergence to the stationary distribution, that might aid policymakers in assessing potential policy initiatives.

While our approach provides valuable insights into the structure of the labor market, it has certain limitations. The analysis relies on year-to-year snapshots of transitions without longitudinal panel data to capture the full trajectory of workers over time. This could introduce potential biases, as residence times within different occupations vary, which could slightly influence the degree of transferability of an occupation. Future research with panel data could extend this work by analyzing the time-dependent condensation effect of individual workers transitioning to more specialized occupations, and link this to labor market outcomes such as lifetime earnings and wage progressions. Additionally, it would be valuable to explore whether our metrics

could be used to explain wage differences in occupations as a mismatch between transferability and accessibility affecting workers' and firms' bargaining power.

Key Takeaways

1. **Deviation from Steady State:** Many occupations are not in their expected long-term equilibrium—over half deviate significantly from their steady-state distribution. This suggests that structural barriers, such as skill mismatches or limited job mobility, slow down worker reallocation.
2. **Community Structure:** The labor market exhibits a strongly clustered structure, with limited inter-community transitions and a majority of flows occurring within communities.
3. **Identifying Bottlenecks:** A new framework based on accessibility and transferability metrics reveals bottleneck occupations. These *condenser* jobs attract workers but offer limited pathways to other careers, restricting upward mobility and reinforcing labor market rigidity.
4. **Policy Implications:** By identifying mobility bottlenecks and key transition pathways, this approach offers a data-driven way to design labor market policies. Targeted interventions, such as retraining programs focused on low-transferability occupations, could help reduce labor frictions and enable a smoother transition toward a more balanced employment landscape.

Skill Network of the Labor Market

7.

Section 7.5 is adapted from Knicker et al. (2024) [4] *The Structure of Occupational Mobility in France*, as its content aligns more with this chapter. All other sections in this chapter are entirely new contributions.

When studying occupational transitions, two types of networks are commonly considered: the observed transition network and the skill-based network (see, e.g. [47]). In the previous chapters, we focused on labor market dynamics purely through the lens of realized flows, examining the structure and properties of transitions as they occur. In this chapter, we shift our attention to the role of skills in shaping these transitions, investigating the extent to which the skill requirements of specific occupations explain the observed patterns of worker mobility.

Skill similarity between occupations has been shown to significantly influence occupational transitions [45, 246] and has been used to reconstruct labor flow networks [48]. A skill-based network uses measures of skill similarity to map potential pathways between occupations based on shared competencies. While skill-based networks can only partially explain mobility structures and reflect an ex-ante perspective—unaffected by realized worker and firm decisions—they offer a key advantage: they identify potential mobility opportunities grounded in skill relatedness. This provides insights into the set of occupations that workers might feasibly transition into based on their skill profiles [45].

In this chapter, we delve into the role of skills in influencing labor market flows and dynamics. By incorporating skill data into our framework, we aim to understand how similarities in skill requirements shape occupational connectivity, transition patterns, and bottlenecks. Specifically, we explore two distinct skill similarity metrics to construct skill-based networks, investigate how skills explain the patterns of transferability and accessibility introduced in Section 6.3, and examine the alignment between skill-driven clusters and task-based occupational classifications.

7.1. French Skill Data

The ROME dataset (Répertoire Opérationnel des Métiers et des Emplois) is a comprehensive data framework developed by France Travail[1] to enhance professional mobility and align labor market supply and demand in France. By adopting a skill-based approach, the dataset addresses recruitment challenges and supports workforce matching. It encompasses 10,840 technical skills, and 10,244 knowl-

1: France Travail, formerly known as Pôle emploi, is a French government agency responsible for registering unemployed individuals, assisting them in finding employment, and providing financial support.

edge items associated to more than 500 occupations organized into five categories and 27 subcategories [247]. Serving as a dynamic and evolving tool, the ROME dataset standardizes the language of recruitment, providing insights into the competencies required for various occupations. This makes it particularly well-suited for studying skill-based occupational flows.

To analyze the relationships between occupations based on their skill requirements, we construct a weighted skill network using two distinct similarity metrics: cosine similarity and θ-similarity. Each occupation j is represented by a binary skill vector V_j,

$$V_{j,s} = \begin{cases} 1, & \text{if occupation } j \text{ requires skill s} \\ 0, & \text{else} \end{cases} \tag{7.1}$$

where each entry is 1 if the occupation requires a specific skill and 0 otherwise. The similarity metrics allow us to quantify the degree of alignment between two occupations based on their technical skill and knowledge profiles.

Cosine similarity (see 5.9, Eq. 5.31) captures the extent to which occupations share overlapping skill sets and was used previously by researchers such as Gathmann and Schönberg (2010) [246] to measure empirically transferability of skills. However, it can overemphasize the influence of ubiquitous skills (e.g. "Communicating with Supervisors") that are common across many roles. As a result, occupations with fundamentally different skill requirements may appear similar, because they share these widely used skills.

In contrast, the θ-similarity metric focuses on the *effective use* of skills, emphasizing those that are particularly relevant or distinctive to a given occupation [248]. By discounting ubiquitous skills, θ-similarity provides a more nuanced view of skill complementarity, better reflecting the unique characteristics of each occupation's requirements.

First, the Revealed Comparative Advantage (RCA) is computed to identify skills that are disproportionately important to each occupation. The RCA for skill s in occupation j is calculated as:

$$\text{RCA}(j,s) = \frac{\frac{V_{j,s}}{\sum_{s \in S} V_{j,s}}}{\frac{\sum_{j \in J} V_{j,s}}{\sum_{j \in J, s \in S} V_{j,s}}} \tag{7.2}$$

where $V_{j,s}$ represents the presence of skill s in occupation j, J is the set of all occupations, and S is the set of all skills. If $\text{RCA}(j,s) > 1$, the skill s is considered to be *effectively used* by the occupation j.[2] Using this RCA normalization, the θ-similarity between two occupations is defined as the minimum of the conditional probabilities that a pair of occupations effectively uses the same skill, which as introduced by Alabdulkareem et al. (2018) [248]:

$$\theta(i,j) = \frac{\sum_{s \in S} e(i,s) \cdot e(j,s)}{\max\left(\sum_{s \in S} e(i,s), \sum_{s \in S} e(j,s)\right)} \tag{7.3}$$

where $e(j,s) = 1$ if $\text{RCA}(j,s) > 1$ and $e(j,s) = 0$ otherwise. This metric accounts for the relative importance of skills and their co-

2: The RCA compares the proportion of a skill s in the total skill requirements of a specific job j ($\frac{V_{j,s}}{\sum_{s \in S} V_{j,s}}$) with the proportion of that skill in the total skill requirements across all jobs ($\frac{\sum_{j \in J} V_{j,s}}{\sum_{j \in J, s \in S} V_{j,s}}$). For instance, if j represents Data Scientist and s represents Python, an $\text{RCA}(j,s) > 1$ would suggest that Python is significantly more crucial for Data Scientists than for the average occupation. Conversely, if j represents Teacher and s represents Communication, an $\text{RCA}(j,s) < 1$ would indicate that, although Communication is important, its relevance to Teachers is not disproportionately higher than its relevance across other occupations.

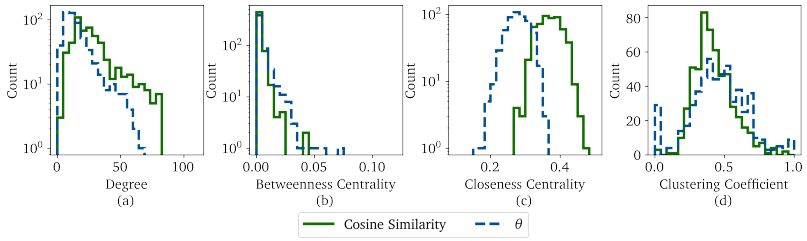


Figure 7.1.: Network properties of the Skill Network: Distribution of network properties for the skill networks constructed using cosine similarity (green) and θ-similarity (blue), including (a) degree, (b) betweenness centrality, (c) closeness centrality, and (d) clustering coefficient.

occurrence, offering a robust, data-driven measure of job similarity.[3]

Using either cosine similarity or θ-similarity, we construct a weighted skill network $\mathcal{G}$, where the nodes $\mathcal{N}$ represent occupations and edges $\mathcal{E}$ represent the similarity between two occupations, with weights $w_{ij} = w_{ji}$ corresponding to the similarity metric.[4] This network captures the degree of overlap or complementarity between occupations based on their skill profiles. Nodes with stronger edge weights are closely related in terms of their skill requirements, while weaker weights indicate less similarity. By analyzing this network, we gain insights into the structure of skill relationships in the labor market and how these relationships influence occupational transitions.

7.2. Network Properties of Skill Network

Figure 7.1 shows the distribution of the network metrics described in Section 5.11, providing a comparison between the cosine similarity and θ-similarity networks.

In terms of degree distribution (Figure 7.1(a)), the θ-similarity network naturally exhibits a lower mean compared to the cosine similarity network. This difference arises from the RCA filtering applied in the θ-similarity metric, which prioritizes distinctive skills while discounting ubiquitous skills that are common across many occupations. Although this results in a sparser network, it emphasizes more specific and meaningful relationships between occupations.

Figure 7.1(b) displays the betweenness centrality distribution, a measure of bottlenecks in the network that capture the importance of occupations in bridging different clusters. The θ-similarity network has a wider distribution of betweenness centrality compared to the cosine similarity network. This suggests that the θ-similarity network better highlights occupations that serve as critical intermediaries, facilitating transitions between otherwise disconnected communities.

The closeness centrality distribution (Figure 7.1(c)) reveals higher values for the cosine similarity network. This is expected, as the greater number of connections in the cosine similarity network results in shorter average path distances, making occupations appear more closely interconnected overall.

Finally, the local clustering coefficient distribution (Figure 7.1(d)), which measures the tendency of neighboring nodes to form tightly connected triangles, is largely similar between the two networks. However, the θ-similarity network shows a small subset of occupations with notably lower clustering coefficients, indicating some degree of

3: Aufiero et al. (2024) [185] propose a similar metric, referred to as job relatedness, defined as:

$$B^{Jobs}_{jj'} = \frac{1}{D} \sum_{s \in S} \frac{e(i,s)e(j,s)}{u_s} \tag{7.4}$$

with the skill ubiquity $u_s = \sum_j e(j,s)$, representing the total number of jobs that effectively use skill s and $D = \max\left(\sum_{s \in S} e(i,s), \sum_{s \in S} e(j,s)\right)$. This additional scaling by u_s adjusts the similarity based on how widespread or specialized a skill is across all occupations. We choose to stick to the θ-similarity (Eq. 7.3) definition because it focuses purely on the co-occurrence of effectively used skills between jobs without introducing an additional layer of normalization. This makes it simpler to interpret while retaining its robustness in capturing meaningful relationships between jobs.

4: Note that the network is undirected, as both similarity metrics are symmetric.

isolation for these roles.

7.3. Relationship of Skills and Transition Rates

Skill similarity between occupations is expected to correlate with the transition rates between them, as shared skills facilitate worker mobility. To quantify this relationship, we employ a Poisson regression, a statistical method designed for modelling count data [249]. This approach allows us to model the flow of workers N_{ij} between two occupations i and j, normalized by the total inflow into i. The regression model is specified as follows:

$$\log\left(\frac{N_{ij}}{\sum_j N_{i,j}}\right) = \alpha + \beta \cdot d_{skill_{ij}} \tag{7.5}$$

Here, $d_{skill_{ij}}$ represents the skill similarity (calculated using either cosine similarity or θ-similarity) between occupations i and j . The parameter α is the intercept, reflecting the baseline flow when skill similarity is $d_{skill_{ij}} = 0$, and β measures the sensitivity of worker transitions to changes in skill similarity. An additive increase in skill similarity by δ corresponds to an exponential increase in transition rates of $\exp(\delta \cdot \beta)$.

By applying this regression model, we can assess the strength and significance of the relationship between skill similarity and occupational flows.

The regression results, summarized in Table 7.1, demonstrate consistent parameters for both skill similarity networks. For the cosine similarity network, $\beta = 4.3$ (p-value $< 10^{-3}$), indicating that an additive increase of $\delta = 0.1$ in skill similarity results in an approximate $\exp(0.43) = 1.54$-fold increase in transition rates. For the θ-similarity network, $\beta = 4.5$ (p-value $< 10^{-3}$), corresponding to a $\exp(0.45) = 1.57$-fold increase in transitions for the same $\delta = 0.1$ increase in skill similarity. When skill similarity is $d_{skill_{ij}} = 0$, the baseline transition rates are on the order of magnitude of $\approx 10^{-3}$, which can be attributed to noise in both the transition rates and skill data.

The relationship between skill similarity and worker transitions is robust and highly significant across both networks. However, the pseudo R^2, which measures the improvement of the fitted model over a null model[5], reveals differences. The θ-similarity network achieves a higher pseudo R^2 value (0.164) compared to the cosine similarity network (0.145), indicating a better fit to the data and a stronger ability to explain the observed transition patterns.[6] This suggests that,

Model	Parameter	Coefficient	Pseudo R-squared
Cosine Similarity	α	-6.27**	0.145
	β	4.3**	0.145
θ	α	-6.1**	0.164
	β	4.5**	0.164

5: We use here the likelihood ratio index [250] that is defined as

$$R^2 = 1 - \frac{\log L_{\text{fitted}}}{\log L_{\text{null}}}, \tag{7.6}$$

where $\log L_{\text{fitted}}$ and $\log L_{\text{null}}$ are the log-likelihoods of the fitted and null models, respectively. The null-model is a poisson model with only an intercept.

6: Unlike R^2 in linear regression, which measures the proportion of variance explained, pseudo R^2 evaluates the improvement in the model's likelihood over the null model. Pseudo R^2 values are lower in Poisson regression due to the discrete nature of count data and the inherent variance-mean relationship of the Poisson distribution. Values between 0.2-0.4 indicate a good model fit [251].

Table 7.1.: Summary table of Poisson Model: Summary table for the Poisson model specified in Eq. 7.5, applied to the two skill networks constructed using cosine similarity and θ-similarity.

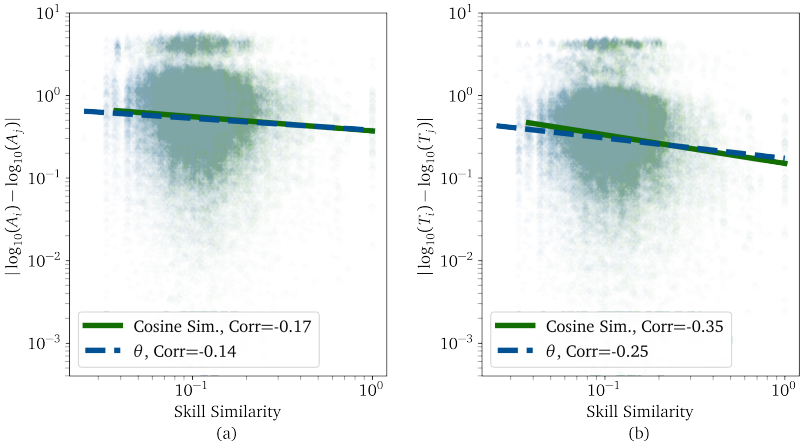


Figure 7.2.: Correlation between skill similarity and differences in Accessibility and Transferability: Scatter plots showing (a) the accessibility difference $|\log(A_i) - \log(A_j)|$ and (b) the transferability difference $|\log(T_i) - \log(T_j)|$ as functions of skill similarity (measured using cosine similarity and θ-similarity) between each pair of occupations i and j. Lines represent linear fits, with the corresponding correlation values displayed in the legend.

while both metrics serve as strong predictors of occupational flows, θ-similarity captures a slightly broader range of transitions, likely due to its emphasis on distinctive skill relationships through RCA filtering.

The cosine similarity network, by considering all skill overlaps, captures a wider array of transitions, including those influenced by ubiquitous skills. However, this reduces its ability to identify finer distinctions in flow patterns.

7.4. Relationship of Skills and Accessibility and Transferability

To better understand how skill similarity translates into occupational flow patterns, we analyze its relationship with the accessibility and transferability metrics (see Section 6.3). Specifically, we examine the correlation between skill similarity and the absolute log differences of accessibility $|\log(A_i) - \log(A_j)|$ and transferability $|\log(T_i) - \log(T_j)|$ between occupations i and j.

Figure 7.2 (a) shows the relationship for accessibility and panel (b) for transferability. We find a negative and significant correlation[7] for both skill similarity metrics. This indicates that as skill similarity increases, the log difference in accessibility and transferability decreases. Occupations with high skill similarity exhibit highly similar flow patterns in terms of where workers come from (accessibility) and where they go to (transferability).

7: Significant at the 5% level.

This finding demonstrates that skill similarities between occupations are directly reflected in their transition dynamics. Occupations that share similar skill sets not only overlap in competencies but also fulfill analogous roles within the labor market, exhibiting comparable patterns of workforce mobility. These results reinforce the idea that skill similarity is a driver of occupational flow dynamics, linking the structural properties of skills to the functional behavior of the labor market.

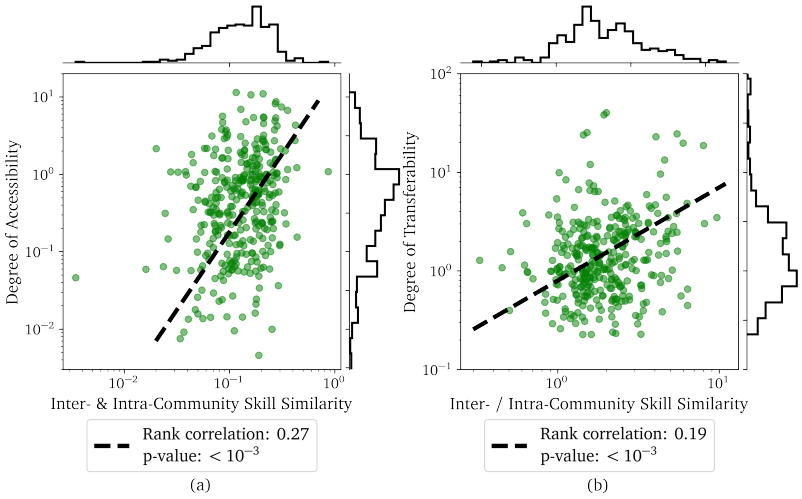


Figure 7.3.: Skill effect on Accessibility and Transferability. Scatter plot depicts the relationship between (a) d^p the product of inter- and intra-community skill similarity and the degree of accessibility and (b) d^r the ratio of inter- and intra-community skill similarity and the degree of transferability for each occupation. A dashed line serves as a visual guide and represents the linear fit of the points. There is a significant and positive rank correlation of 0.27 observed between the product of inter- and intra-community of skill similarity and accessibility and a significant and positive rank correlation of 0.19 between the ratio of inter- and intra-community skill similarity and the degree of transferability for each occupation.

7.5. Skill Effects on Accessibility and Transferability

Using skill data enables us to better understand why highly accessible occupations act as bottlenecks and why highly transferable occupations facilitate inter-cluster movement (see Section 6.3).

To investigate this, we calculate the average skill similarity of an occupation i to other occupations within its community $d_{\text{Intra},i}$ and to all occupations outside its community $d_{\text{Inter},i}$ (see Appendix B.1, Eq. B.3). These measures allow us to derive two skill scores: the product $d^p = d_{\text{Intra}} \cdot d_{\text{Inter}}$, and the ratio $d^r = d_{\text{Inter}}/d_{\text{Intra}}$. Occupations with high scores in d_{Intra} and d_{Inter} exhibit a balanced skill similarity profile across and within communities, whereas those with unique or specialized skills score lower.

Figure 7.3 illustrates the relationship between accessibility and the product of inter- and intra-community skill similarity d^p in panel (a), and between transferability and the ratio d^r in panel (b). Occupations with higher accessibility scores are positively correlated with higher d^p (rank correlation = 0.27, p-value$< 10^{-3}$), indicating that these occupations share substantial skill overlap both within their own community *and* with other occupations in other communities. This suggests that highly accessible occupations act as entry points into new occupational communities due to their skill similarity with a broad range of roles, making them pivotal in facilitating transitions from outside the community.

Conversely, occupations with higher transferability scores are positively correlated with d^r (rank correlation = 0.19, , p-value$< 10^{-3}$), reflecting their strong skill similarity with occupations outside their own community. These occupations serve as inter-community connectors, enabling transitions that bridge disparate clusters. This suggests that high transferability occupations act as springboards to inter-community transitions.

While accessibility highlights entry points into communities based on shared skills, transferability emphasizes connectors that enable inter-cluster transitions, providing a richer understanding of occupational

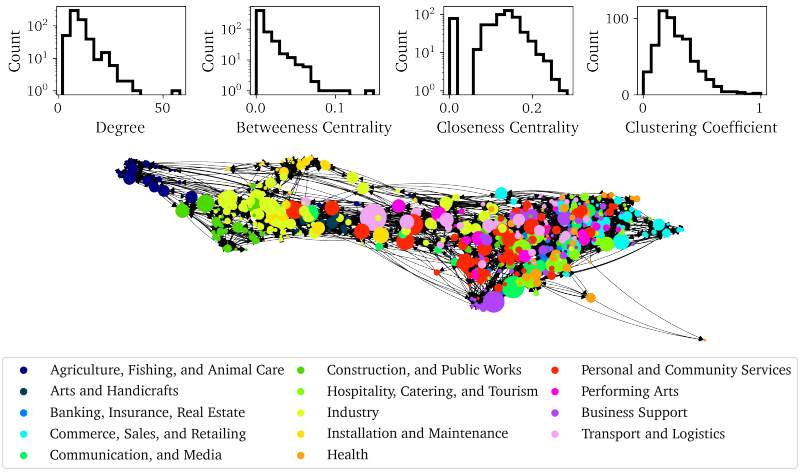


Figure 7.4.: Expert-Assessed Skill Network: Visualization of the expert-assessed skill network, where each node represents an occupation in the ROME classification. Node size corresponds to its betweenness centrality, and node color indicates its 1-digit ROME cluster. Edge weights are determined by the relevance index α: $w_{ij} = \frac{1}{2}$ for $\alpha = 1$, $w_{ij} = \frac{1}{3}$ for $\alpha \in [2, 5]$, and $w_{ij} = 0$ otherwise. The top row displays the distributions of network metrics, including degree, betweenness (Eq. 5.37), closeness (Eq. 5.38) centralities, and the clustering coefficient (Eq. 5.39).

flows than traditional network metrics alone.

7.6. Expert-Assessed Skill Network

To complement the skill networks based on calculated similarity metrics, we construct a skill network using expert-assessed connections from the ROME framework. This dataset provides detailed information on occupational transitions deemed feasible by experts, specifying a source occupation, a target occupation, and an index α that ranks the relevance or likelihood of the transitions. Transitions with $\alpha = 1$ represent very close and highly feasible transitions, while $\alpha = 10$ indicates more distant transitions.

Using this data, we create a weighted, directed skill network where nodes represent occupations and edge weights correspond to the expert-assessed likelihood of transitions. Specifically, transitions with $\alpha = 1$ are assigned a weight of $\frac{1}{2}$, reflecting their high relevance, while transitions with $\alpha \in [2, 5]$ are assigned a weight of $\frac{1}{3}$, capturing moderate but meaningful likelihood. Transitions with $\alpha > 5$ are excluded, as the number of such transitions sharply decreases, allowing us to focus on the most relevant connections. This approach follows the methodology of Bocquet (2022) [47], who first explored the emergence of bottleneck occupations using skill data of French administrative records.

The expert-assessed network provides advantages over networks based solely on calculated skill similarities. It incorporates practical feasibility, as the expert assessments account for real-world constraints and opportunities, such as the ease of reskilling or training. By drawing on expert judgment, these connections are firmly rooted in labor market realities, rather than *abstract* mathematical relationships. Additionally, the expert network minimizes biases introduced by ubiquitous skills—those shared by many occupations but lacking meaningful differentiation.

Figure 7.4 illustrates the expert-assessed skill network, revealing a distinct ribbon-like structure where occupations classified under the same 1-digit ROME code are grouped closely together (nodes in the network from the same 1-digit ROME classification have the same

color). The network exhibits a gradient from left to right, with two dominant coarse-grained clusters. On one side, occupations related to agriculture, construction, and industry dominate, reflecting a focus on manual skills. On the other side, occupations classified under banking, insurance, and sales prevail, representing a cluster of occupations with cognitive skills. Bridging these two extremes are intermediate occupations such as performing arts, transport and logistics, and personal and community services, which exhibit a mix of both manual and cognitive elements.

This segmentation of the labor market into manual and cognitive clusters aligns with similar findings from the ONET database, which contains administrative data on the U.S. economy. Using the θ-similarity metric (Eq. 7.3), Alabdulkareem et al. (2018) [248] identified two distinct groups in the ONET dataset: one associated with cognitive skills and the other with physical skills.[8] Similarly, Aufiero et al. (2024) [185], employing a variation of the θ-similarity normalized by skill ubiquity (Eq. 7.4), found two primary communities in their analysis: one dominated by technical occupations and the other by humanities and communication occupations.

8: Note that Alabdulkareem et al. (2018) [248] compute $\theta(s, s')$ instead of $\theta(i, j)$, constructing a skill network where nodes represent individual skills rather than occupations. In this case, $\theta(s, s')$ reflects the complementarity of two skills being effectively used by the same occupation.

The network metrics, shown in the top-row of Figure 7.4, highlight again the presence of bottlenecks within the skill network. These are evident from the power-law distribution of betweenness centrality, with a few occupations emerging as critical intermediaries that facilitate transitions between otherwise disconnected clusters.

To better understand the relationship between bottlenecks and labor market dynamics, we analyze the correlation between betweenness centrality and the flow-based metrics of accessibility and transferability. Accessibility shows a positive correlation with betweenness centrality ($r = 0.26$, p-value $< 10^{-3}$), indicating that bottleneck occupations tend to attract diverse incoming flows, drawing workers from a wide range of origins. In contrast, transferability is negatively correlated with betweenness centrality ($r = -0.25$, p-value $< 10^{-3}$), suggesting that occupations with high bottleneck properties are less likely to facilitate diverse outgoing transitions.

7.7. Routine, Non-Routine, and Cognitive-Manual Classification

To further explore the structure of the skill network and validate the detected clusters, we analyze occupations using the routine, non-routine, and cognitive-manual classification (see [252] for more details). This widely recognized framework categorizes occupations based on the nature of their tasks. It distinguishes between tasks that are highly automatable (routine) and those requiring human judgment, creativity, or interpersonal skills (non-routine). Additionally, the cognitive-manual dimension differentiates between tasks that demand mental effort and problem-solving (cognitive) and those involving physical effort and tangible outputs (manual) [252].

This classification is particularly relevant as it highlights how technological advancements and automation impact different segments

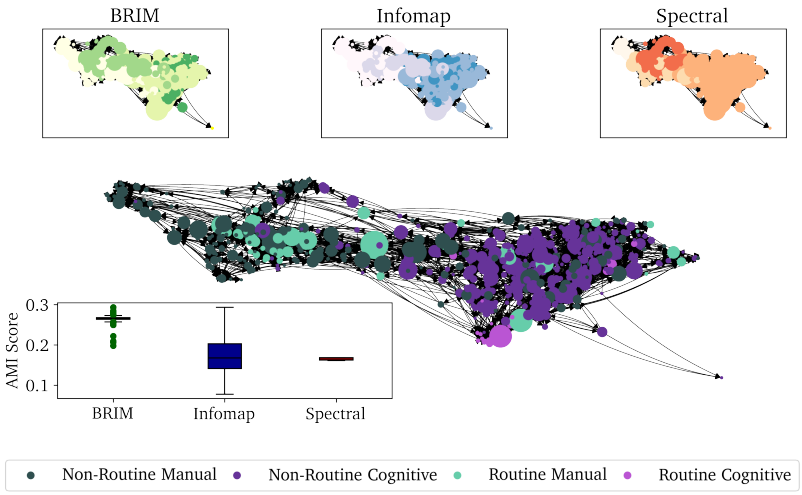


Figure 7.5.: Expert-Assessed Skill Network with task-based classification: Expert-assessed skill network constructed as in Figure 7.4, with nodes color-coded according to the task-based classification. The top row displays the clustering with the highest modularity achieved by BRIM [212], Infomap [213], and Spectral Clustering [214] algorithms. The box plot illustrates the Adjusted Mutual Information (AMI) scores (Eq. 7.7) for all random seeds, comparing the clustering results with the task-based classification.

of the labor market [43]. By comparing this task-based classification with the community clusters detected through the skill network, we can assess the extent to which the clusters align with established occupational categories. Such an analysis provides insights into whether the community clusters capture meaningful divisions within the labor market and how accessibility and transferability dynamics relate to these categories.

Since the ROME dataset does not provide this classification directly, we manually assign occupations to the four categories—routine cognitive, routine manual, non-routine manual, and non-routine cognitive. This assignment is based on a detailed review of the information available in the ROME dataset, including job descriptions, access to employment, competencies, working environments, tasks, and skills. The complete classification for each occupation are documented in Appendix B.9. In total, we find 321 non-routine cognitive, 192 non-routine manual, 46 routine manual and 17 routing cognitive occupations.

Figure 7.5 shows the same skill network as in Figure7.4 but here the nodes are color coded according to the task-based classification. First we see that within the network the nodes with the same task-based classification are clustered together. Second, indeed, on the left side of the network, we observe a concentration of manual occupations, encompassing both routine manual (light blue) and non-routine manual (gray blue). On the right side of the network, occupations shift toward cognitive tasks, with a cluster of routine cognitive occupations (light purple) and non-routine cognitive (dark purple). The middle section of the network serves as a transition zone, primarily connecting non-routine cognitive and non-routine manual occupations. This transition zone highlights the versatility and interconnectivity of non-routine tasks, which bridge the cognitive and manual domains.

The alignment between this classification and the structure of the skill network is remarkable, as it demonstrates that the skill network effectively captures the inherent divisions and overlaps within the labor market. The spatial organization of the categories reflects real-world task-based similarities, with occupations clustering based on their cognitive or manual focus and their routine or non-routine nature.

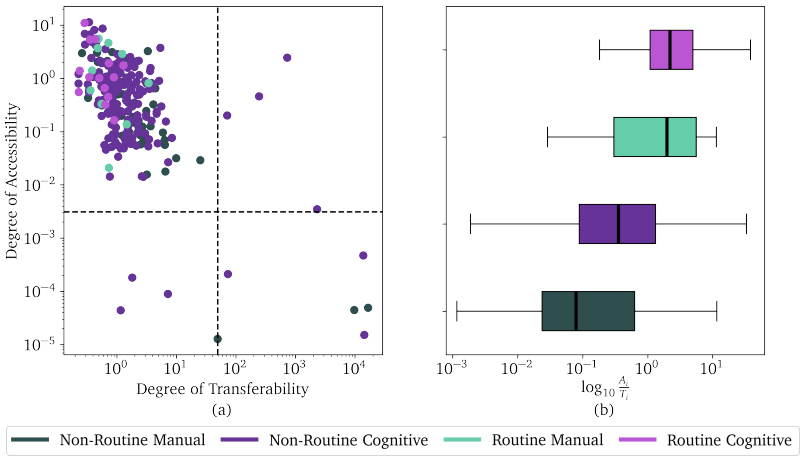


Figure 7.6.: Accessibility and Transferability with task-based classification: (a) Accessibility-Transferability space, as defined in Section 6.3, with each point color-coded according to the task-based classification. (b) Box plot of the bottleneck ratio $\log_{10} \frac{A_i}{T_i}$ for occupation i within each task-based cluster.

Evaluating Community Clustering Algorithms Against the Routine-Non-Routine and Cognitive-Manual Classification

To evaluate how well community detection algorithms align with the routine, non-routine, and cognitive-manual classification of occupations, we compare the performance of three clustering methods: Spectral Clustering, BRIM, and Infomap.[9] The comparison uses the Adjusted Mutual Information (AMI) score, which quantifies the similarity between the detected clusters and the true occupational classification.[10]

Each clustering algorithm is run 50 times with different random seeds, and the AMI scores are summarized in the boxplot in Figure 7.5. Additionally, Figure 7.5 visualizes the best clustering result for each algorithm in the color-coded skill network. The results reveal clear differences in performance across the methods. BRIM achieves the highest mean AMI score across all runs, indicating that it most accurately reflects the task-based classification. Infomap, while showing higher variance in AMI scores, performs well in certain runs but less consistently. Spectral clustering, though less variable, generally underperforms compared to BRIM and Infomap.

Although none of the algorithms perfectly capture all four clusters corresponding to the routine-non-routine and cognitive-manual classification, both BRIM and Infomap excel at capturing specific aspects. BRIM reliably distinguishes between routine and non-routine occupations, while Infomap is effective at separating cognitive and manual roles. This highlights that the task-based classification is inherently encoded in the directed transitions of the expert-assessed skill network.

9: Louvain clustering was excluded from this analysis, as it could not produce exactly four clusters corresponding to the classification categories.

10: The Adjusted Mutual Information (AMI) is a measure of the similarity between two clusterings, adjusted for chance. It is defined as

$$\frac{\mathrm{MI}(U,V) - \mathrm{E}[\mathrm{MI}(U,V)]}{\langle(\mathrm{H}(U),\mathrm{H}(V))\rangle - \mathrm{E}[\mathrm{MI}(U,V)]}, \quad (7.7)$$

where U and V are the two clusterings being compared, $\mathrm{MI}(U,V)$ is the mutual information, $\mathrm{H}(U)$ and $\mathrm{H}(V)$ are the entropies of the clusterings, $\langle\cdot\rangle$ denotes the average, and $\mathrm{E}[\mathrm{MI}(U,V)]$ is the expected mutual information under random labeling. Intuitively, AMI measures how much information is shared between two clusterings while accounting for the similarity that could occur by chance. It ranges from 0 to 1, where 0 indicates no agreement beyond random chance, and 1 represents perfect agreement.

7.8. Accessibility-Transferability Ratios and Routine Occupations

To explore how the task-based classification aligns with the accessibility and transferability metrics, we analyze the relationship between these metrics and occupational categories. Figure 7.6(a) illustrates the accessibility-transferability space, with each occupation color-coded

according to the task-based classification. Figure 7.6(b) presents a box plot of the bottleneck ratio, defined as $\log_{10}\left(\frac{A_i}{T_i}\right)$.

Routine occupations—both routine manual and routine cognitive—have significantly higher bottleneck ratios, indicating that these roles act as bottlenecks in the labor market. This finding aligns with the challenges routine jobs face in enabling workers to transition into other roles, largely due to the narrow and repetitive nature of their associated tasks [40]. Conversely, the bottleneck ratio is lowest for non-routine occupations, suggesting that these roles facilitate transitions more effectively. This adaptability likely stems from the versatile, transferable skills required in non-routine manual jobs, which are applicable across a wide range of tasks and sectors.

These findings underscore the value of integrating community detection methods, flow-based metrics like accessibility and transferability, and skill data. While clustering algorithms capture the structural divisions within the skill network, the transferability-to-accessibility ratio offers a complementary functional perspective, shedding light on how different task-based occupational categories interact within the broader labor market. Together, these approaches provide a comprehensive understanding of the interplay between skills, tasks, and labor market dynamics.

Key Takeaways

1. **Skills Drive Mobility:** Workers tend to move between occupations with similar skill requirements, reinforcing the idea that skill compatibility is a fundamental factor shaping job transitions.
2. **Structural Imbalances Persist Across Perspectives:** Network properties of the skill-based labor network highlights similar structural imbalances as the occupational flow based network.
3. **Accessibility and Transferability:** High-accessibility occupations serve as entry points, while high-transferability roles connect communities due to their inter and intra skill similarities. Routine occupations show higher bottleneck ratios, restricting mobility.
4. **Routine Jobs Limit Mobility:** Occupations with routine cognitive or manual tasks create mobility bottlenecks, restricting workers' ability to transition into new roles and making them more vulnerable to labor market disruptions.

Part III: The Interplay of Skill, Luck & Competition in Success

Beyond Luck: Measuring Skill and Competition in Venture Capital and Analyst Forecasts

8.

8.1. Introduction

Financial decision-making, whether in venture capital (VC) or public markets, is shaped by a complex interplay of skill, randomness, and competition. Venture capitalists operate in high-stakes environments, seeking startups with outsized growth potential while navigating uncertainty and external factors such as market trends and macroeconomic conditions. Their success hinges on a combination of industry expertise, network connections, and timing—essentially picking the right startups under the right conditions. In public markets, a similar challenge unfolds as analysts forecast financial metrics for companies. Here, skill materializes as the ability to distill market data into accurate predictions, balancing insight with the inherent unpredictability of financial systems.

While skill and luck are often considered independently, real-world outcomes frequently arise from their intricate interplay. A decision-maker's success is not solely determined by being *skillful* or *lucky* but rather by their position on a two-dimensional continuum of both factors. For instance, a highly skilled investor may still rely on favorable market conditions to succeed, while even an unskilled participant can benefit from sheer luck. Conversely, adverse random events can undermine even the most skilled decision-makers. This dual perspective underscores the need for a framework that quantifies skill and luck simultaneously, offering a more nuanced understanding of competitive success.

This chapter introduces the skill-luck-depth framework, previously applied to sports, games, and social dynamics [253], to analyze competition in VC and public markets. Venture capitalists make long-term decisions, often waiting years to realize returns, with competition driven by access to high-potential startups. Public markets, on the other hand, function on much shorter time horizons, with analysts continuously evaluated based on forecasting accuracy. These contrasting conditions make VC and public markets ideal for examining how skill, luck, and competition sensitivity manifest in different environments.

I seek to answer the question of *how much skill and luck is involved in Venture Capital and Analyst Forecasts?* For venture capital, this involves analyzing how sectors and funding stages influence these factors. For public markets, I examine how forecasting horizons affect the interplay of skill and randomness.

Methodologically, I extend the skill-luck-depth framework [253] from network science to finance, refining *depth of competition* as the skill importance factor and introducing benchmark *player* to connect skill to well-defined outcome variables. Empirically, it provides new insights into the dynamics of VC and analyst predictions, offering a skill met-

ric that accounts for persistence and removes the influence of luck or one-off events. On an interdisciplinary level, it bridges complexity-inspired modelling with investing behavior, showcasing the framework's versatility and paving the way for future applications in other competitive domains.

The main finding is that both venture capital and analyst forecasts involve minimal influence of luck. However, these domains are highly competitive, meaning that small differences in skill among investors or analysts result in only slight changes to their returns or accuracies.

This chapter begins by a review of the existing literature, then introducing the skill-luck-depth framework developed by Jerdee and Newman (2024) [253], followed by an analysis of the venture capital market and analyst forecasts using this framework, and closes with a discussion and limitations of the framework.

8.2. Literature Review

The question of whether success is primarily driven by skill or luck has long been debated across various disciplines. As introduced in Chapter 2, Section 2.1 Pluchino et al. (2018) [66] address this issue using an agent-based model, demonstrating that success is not solely a function of talent but is significantly influenced by luck. Their findings challenge the traditional meritocratic view, showing that even in environments where talent follows a normal distribution, success follows a highly skewed pattern, with moderately talented but exceptionally lucky individuals dominating the top ranks.

This debate is particularly relevant in the context of financial markets and venture capital, where investment decisions and forecasting abilities are often attributed to expertise and skill. However, a long-standing discussion persists: Do venture capitalists succeed due to superior investment selection and value creation, or are their returns largely dictated by market timing and external conditions? Similarly, do financial analysts and fund managers possess genuine forecasting ability, or are their success streaks indistinguishable from randomness?

In light of this, understanding the balance between skill and luck in venture capital investing and financial markets is crucial for effective capital allocation and investment strategies. In this chapter, we combine the two interconnected research streams—the persistence of performance in financial markets and the role of skill in venture capital—to assess whether outperformance in these fields is attributable to genuine ability or randomness.

A large body of research investigates whether fund managers exhibit persistent skill. Hartzmark (1991) [254] analyzed futures traders and found no evidence of sustained outperformance, implying that short-term trading success is largely driven by luck. Similarly, Malladi (2020) [255] applied the False Discovery Rate (FDR) method to hedge funds and found that only around 3% of managers exhibit true skill, while 1 out of 3 are unskilled, and the majority have returns indistinguishable from random chance. This aligns with Mauboussin and Arbesman

(2011) [256] who demonstrated that streaks of outperformance among fund managers often mimic random sequences rather than reflecting skill. In contrast to this, when it comes to forecasting earnings, Leone and Wu (2007) [257] found that top-ranked financial analysts demonstrate genuine forecasting skill, with earnings forecasts and stock recommendations showing persistent accuracy over time.

Evidence from private equity suggests a similar pattern. Korteweg and Sorensen (2017) [258] found that buyout fund managers exhibit persistent skill, with top-quartile PE firms generating 7–8% higher annual returns than bottom-quartile firms. Yet, their findings also indicate that venture capital fund performance is much less persistent, suggesting that luck plays a more dominant role in early-stage investing compared to leveraged buyouts.

Gompers et al. (2006) [259] extended this question to venture-backed firms, finding that serial entrepreneurs who succeeded once are significantly more likely to succeed again, implying that entrepreneurial and investment skill are real and measurable. However, Liechti et al. (2014) [260] showed that entrepreneurs tend to underestimate the role of luck in their success, even when random factors like timing and market trends significantly impact outcomes.

Additionally, Soto-Simeone et al. (2021) [261] challenged the idea that new venture survival follows a random walk process, arguing that entrepreneurial experience, adaptability, and strategic decision-making have measurable effects on long-term success. This suggests that while venture capital outcomes appear luck-driven in the short term, over longer horizons, skill effects emerge.

Khanna and Mathews (2022) [262] introduced a game-theoretical model showing that new, unestablished VCs often outcompete experienced VCs because startups value reliability over skill. This leads to a crowding-out effect, where VCs with genuine skill may lose deals to those who are simply more aggressive in offering funding commitments. This mechanism may partially explain why venture capital outperformance is difficult to sustain.

The existing research suggests that both skill and luck play a role in financial markets and venture capital, but their influence depends heavily on the investment context and the observed period of time. Our study builds on this literature by applying the skill-luck-depth model to measure the extent to which VC and financial forecasting performance is truly skill-driven or randomness-dominated.

By integrating insights from financial analysts and venture capital, we contribute to a more comprehensive understanding of how skill manifests across different investment disciplines and can be measured alongside luck.

8.3. A Framework for Analyzing Competitive Performance

Skill-Luck-Depth Model

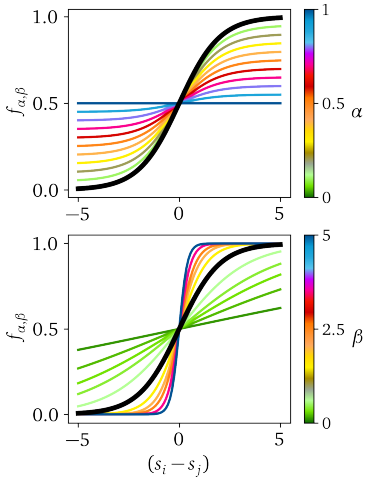
To analyze competitive dynamics in venture capital (VC) and public markets, I employ the skill-luck-depth model, a probabilistic framework that quantifies the interplay of skill, randomness, and competition sensitivity in pairwise contests. This model was developed by Jerdee and Newman (2024) [253] in the context of sports, games, and social situations and predicts the probability of one competitor outperforming another based on differences in their skill-levels.

At the core of the model is the winning probability function $f_{\alpha,\beta}(s_i - s_j)$ between a pairwise competition of two players with the skill-levels s_i and s_j:

$$f_{\alpha\beta}(s_i - s_j) = \frac{1}{2}\alpha + (1-\alpha)\frac{1}{1+e^{-\beta(s_i - s_j)}}, \tag{8.1}$$

The parameter α captures the behavior of the winning probability when there are large skill differences. It can be interpreted as a measure of *luck*, as it reflects the non-zero probability that a less skilled competitor can still win, even when facing an opponent with significantly higher skill. Conversely, the parameter β, referred to here as the *skill importance factor*, governs how sensitive outcomes are to differences in skill. Higher β values lead to sharper stratification in competition, where even small differences in skill result in significant differences in winning probabilities. Lower β values produce a flatter curve, where outcomes are less dependent on skill differences and randomness plays a greater role (see Illustration 4).

This framework accommodates a wide range of competitive environments. In systems dominated by skill, pairwise competition outcomes align closely with differences in $s_i - s_j$. When luck dominates ($\alpha \approx 1$), all competitors have roughly equal chances of success, regardless of skill.



Ill. 4: Sensitivity of $f_{\alpha,\beta}$ to α, β: Black line corresponds to $\alpha = 0$ and $\beta = 1$. (a) Behavior of $f_{\alpha,\beta}$ for varying α. α determines the behavior for $(s_i - s_j) \to \infty$. (b) Behavior of $f_{\alpha,\beta}$ for varying β. β determines the steepness of $f_{\alpha,\beta}$. Illustration taken from Jerdee and Newman (2024) [253] but slightly adjusted.

Bayesian Parameter Estimation

The parameters of the skill-luck-depth model are estimated using a Bayesian inference approach, which not only provides parameter estimates but also accounts for uncertainty in those estimates.[1]

Bayesian inference begins with prior assumptions about the parameters. For the skill s_i of player i, I use a Gaussian prior with mean zero and variance of $\frac{1}{2}$ which sets the scale of the skill s.[2] Luck α is assigned a uniform prior of $P(\alpha) = 1$, reflecting no initial bias about the degree of randomness. For the skill importance factor β, a Cauchy prior is used to accommodate both steep and flat competitive hierarchies.[3]

To fit the parameters, I aim to estimate the values of α, β, and s that best align with the observed data A.[4] Here, the observed data consist of the outcomes from pairwise comparisons in VC investments and financial analyst forecasts, which I will describe in detail in the following subsection. The joint posterior probability can be decomposed, assuming

1: The priors outlined in this section closely follow the setup used in Jerdee and Newman (2024) [253], ensuring comparability of the results.

2: The Gaussian prior on s has the following form:

$$P(s) = \prod_{i=1}^{P} \frac{1}{\sqrt{\pi}} e^{-s_i^2} \tag{8.2}$$

with the number of players P.

3: The cauchy prior has the form of

$$P(\beta) = \frac{2\omega/\pi}{\beta^2 + \omega^2} \tag{8.3}$$

where ω determines the decay. Here I use $\omega = 4$ as in Jerdee and Newman (2024) [253].

4: Assume there a P players and m matches, then the observed data A is the $P \times P$ matrix where each entry A_{ij} is the number of times player i won over player j.

independence between all parameters, as:

$$P(s, \alpha, \beta|A) = P(A|s, \alpha, \beta)\frac{P(\alpha)P(\beta)P(s)}{P(A)} \tag{8.4}$$

with the data likelihood $P(A|s, \alpha, \beta)$.[5] The prior $P(A)$ is unknown but can be disregarded as it is independent of the parameters and only determines the scaling.

5: The likelihood of the data is given by

$$P(A|s, \alpha, \beta) = \prod_{ij} f_{\alpha,\beta}(s_i - s_j)^{A_{ij}} \tag{8.5}$$

To estimate posterior distributions, I use Hamiltonian Monte Carlo (HMC) [263], a computationally efficient sampling technique well-suited for this type of inference to find the best fitting parameters.[6] The resulting posterior distributions provide insights into the relative skill-levels of competitors s, the degree of randomness influencing outcomes α, and the impact of outcomes to skill differences β. The estimation procedure and the underlying code are adapted from Jerdee and Newman (2024) [253] and the Python package provided therein.

6: I sample each run 20 000 points based on the HMC method.

Construction of Pairwise Games

To apply the skill-luck-depth model, I construct pairwise games that compare competitors based on their observed outcomes. Instead of generating pairwise comparisons for all possible combinations of competitors, which can become computationally infeasible and difficult to interpret, I introduce 3 *benchmark* players.[7] These players serve as specific benchmarks for outcome variables in the competition, defined as quantiles of the outcome distribution: portfolio multiples for venture capital and prediction accuracy for analyst forecasts, respectively.

7: Note that the benchmark players are fictitious and only represent ex-post benchmarks as quantiles of the observed data.

Each VC (or analyst) i is associated with a portfolio P_i of N_i companies, either through investment (for VCs) or forecasting (for analysts).

To create the benchmark players, I use the ex-post distribution of portfolio outcomes:

- For VCs, this is the distribution of investment multiples (see Section 8.4 for more details).
- For analysts, this is the distribution of forecasting accuracy (see Section 8.5 for more detail).

A benchmark player is assigned a single company in their portfolio, where the company's multiple (or forecast accuracy) is chosen as a specific quantile of the entire distribution.

Each real player i competes with each of their portfolio companies against the three benchmark players through pairwise contests. A player wins a contest

- For VCs, it their company's multiple is higher than the benchmark's.
- For analysts, if their forecast accuracy is higher than the benchmark's.

With this the data matrix A has dimension $P \times 3$ where each entry A_{ij} is the number of times player i won against the benchmark players. The total number of games is then $\sum_{ij} A_{ij} = 3 \cdot \sum_i N_i$. By doing so, I not only capture the relative performance of competitors against one another

but also evaluate their performance against absolute benchmarks. This dual perspective provides a richer understanding of the competitive landscape, revealing not just how good a competitor is relative to their peers, but also how their performance measures up to a desirable or expected standard.

8.4. Skill and Randomness in Venture Capital Investments

Venture capital investing is inherently competitive, with success depending on the ability to identify and invest in startups that can generate exceptional returns. To analyze this competitive landscape, I utilize the Crunchbase dataset, which offers detailed information on publicly announced funding rounds for startups and their investors. This dataset provides key insights into the investment behavior and performance of VCs across different funding stages.

The dataset includes detailed records of funding rounds in which a VC participated, the total funding amount for each round, and subsequent funding events for the same startups. I focus on start-ups α, founded between 2010 and 2022, that received investments from VCs with at least two investments during this period and are geographically located in the United States. For each VC i, I construct a portfolio P_i, comprising all startups they have backed. The average portfolio multiple for a VC is then defined as

$$M_i = \langle M_\alpha \rangle_{\alpha \in P_i}, \tag{8.6}$$

where M_α is the investment amount multiple of an individual startup α in the portfolio. Similarly $N_i = \sum_{\alpha \in P_i} \mathbb{1}$ is the number of investments per VC i.

The multiple M_α for a portfolio company is calculated as the ratio of the funding amount raised in the next round to the funding amount raised in the current round. This serves as a proxy for the startup's growth in investment amount between rounds.[8] If a startup fails to raise a subsequent round within three years, the multiple is set to

8: It is important to note that I do not have access to company valuations, only the funding amounts. To interpret the multiple as a true measure of growth, I assume that the proportion of shares sold remains consistent across consecutive funding rounds.

Table 8.1.: Table of Venture Capital data: The dataset is divided into funding rounds (Series A, B, and C) across six sectors. N denotes the number of startup companies, and P_i represents the number of investors.

		Series A	Series B	Series C
Tech & Software	N	666	243	88
	P_i	124	55	21
Financial & Legal	N	344	75	28
	P_i	70	19	8
Business & Services	N	1631	589	246
	P_i	283	125	54
Consumer Tech	N	1654	447	144
	P_i	292	105	37
Health & Environment	N	1344	931	387
	P_i	253	178	84
Communication & Ed.	N	235	66	8
	P_i	47	16	2
Σ	N	5874	2351	901
	P_i	641	346	156

$M_\alpha = 0$. If multiple VCs invest in the same company during the same round, the company α is included in the portfolios P_i of all participating VCs.

We focus on the rounds between Seed and Series C. Across all rounds and sectors, the dataset encompasses 9,126 startups and 908 unique investors. The median multiple for individual startups is approximately $\text{Med}(M_\alpha) \approx 2.5$, while the median multiple for investors, representing the average performance of their portfolios, is at $\text{Med}(M_i) \approx 1.3$ (see full breakdown dataset in Table 8.1).

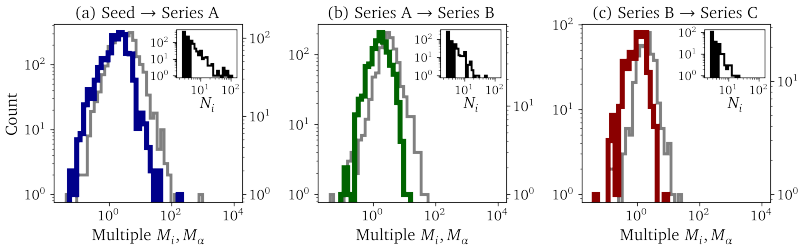


Figure 8.1.: Distributions of Venture Capital data: Distribution of startup multiples (grey) and VC multiples (colored) across funding stages. The distributions are shown for (a) Seed to Series A, (b) Series A to B, and (c) Series B to C. The inset displays the distribution of the number of startups invested per VC. All distributions are shown on a log-log scale.

Figure 8.1 illustrates the distribution of multiples for both startups (grey) and VCs (colored) across funding stages. The colored distributions represent the VC multiples, calculated as the mean multiple of all startups within a VC's portfolio, while the grey distributions represent individual startup multiples. The inset provides the number of startups invested per VC at each stage: Seed to Series A (blue), Series A to B (green), and Series B to C (red). Both distributions exhibit power-law behavior, with a small number of startups achieving significant growth ($M_i \approx 10$) and a median VC multiple only slightly above 1. This skewed distribution highlights the concentration of success in a few high-performing investments.

Portfolio Distributions & Random Selection Process

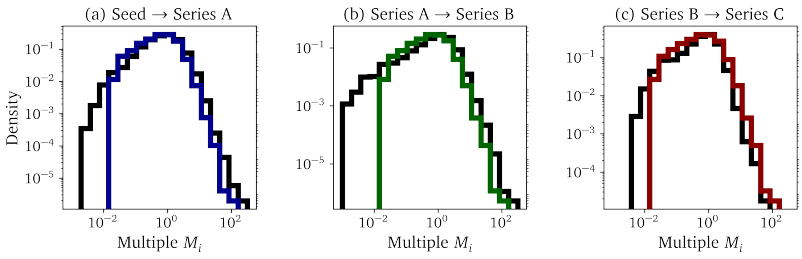


Figure 8.2.: Comparison of VC multiples to random sampling: Distribution of actual VC multiples (colored) compared to the random sampling of portfolios (black) across funding stages. The distributions are shown for (a) Seed to Series A, (b) Series A to B, and (c) Series B to C. The random sampling preserves the number of startups invested per VC but assigns startups randomly to portfolios. All distributions are normalized for comparability.

We first examine whether venture capitalists exhibit skill in selecting high-performing startups by comparing their actual multiple distributions to those generated by a random selection process (Figure 8.2). In this step, I simulate random portfolios for each funding stage (Seed to Series A, Series A to B, and Series B to C)[9] by preserving the number of startups invested in by each VC while randomly assigning startups to these portfolios. This random sampling serves as a baseline for understanding whether VCs, as a group, deviate from what would be expected under a purely random selection process.

The comparison reveals two key patterns. First, on the right tail, rep-

9: To simulate the random selection process, I sample the entire system using 1,000 random seeds, resulting in a total of 1,000 times the number of portfolios in each round. Startups are assigned to portfolios using a random process without replacement, ensuring that each startup is only included once per random portfolio set. This approach mirrors realistic constraints, preserving the total number of portfolios and their sizes while introducing randomness.

resenting the most successful investors, the actual VC multiple distribution aligns closely with the random sampling distribution (see Figure 8.2). This indicates that, across all VCs, there is little systematic evidence of outperformance in selecting top-performing startups beyond what would be achieved through a random process. However, for Series B to C investors, there is a slight outperformance compared to random sampling, suggesting some ability to identify high-growth startups at later stages.

Second, on the left tail, representing the worst-performing investors, the actual VC distribution has significantly less weight than the random sampling. This suggests that VCs collectively exhibit a selection process that avoids extremely poor outcomes, systematically steering away from catastrophic failures in their portfolios. Together, these findings highlight that while VCs as a group may not assemble more portfolios with the highest-performing startups compared to random, they demonstrate skill in minimizing exposure to the lowest-performing investments.

Luck-Skill Impact Landscape of Venture Capital Investments

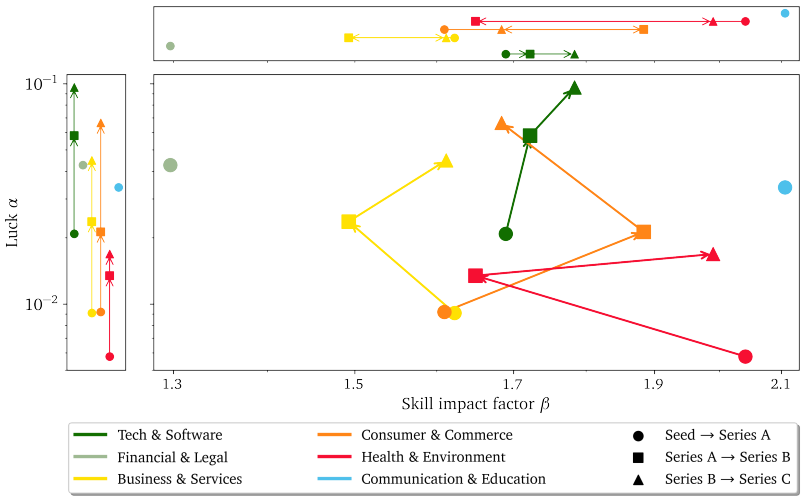


Figure 8.3.: Luck and skill impact factor plane: Estimated parameters for luck (α) and skill impact factor (β) for each sector and funding round. Sectors are distinguished by color, while marker styles indicate funding rounds: circles for Seed to Series A, squares for Series A to B, and triangles for Series B to C. Arrows connect consecutive funding rounds within the same sector, pointing in the direction of higher rounds.

To construct the pairwise games required for applying the skill-luck-depth framework, I introduce three benchmark investors. Each portfolio company of every VC is compared against these benchmarks. The benchmarks represent the ex-post 25th, 50th, and 75th percentiles of the multiple distribution for all VCs, calculated separately for each vintage, funding round, and sector.

For each VC, a pairwise game is created by comparing the return multiple of each portfolio company against these benchmarks. A *win* is recorded if the multiple exceeds the benchmark's threshold, while a *loss* is recorded otherwise.[10]

Once the model is fitted, the skill parameter s_i represents a VC's ability to identify startups that consistently generate high returns. This reflects the quality of their decision-making and their ability to recognize potential successful companies ahead of time. The luck parameter α captures the influence of external factors that drive startup suc-

10: In total, there are $3 \cdot N_\alpha$ games for each funding round and sector, due to the 3 benchmark players.

cess but lie beyond the VC's control, such as market dynamics, sector-specific trends, and macroeconomic conditions. The skill importance factor β reflects the steepness of the competitive hierarchy, determining how sharply performance outcomes differentiate based on skill in the respective sector and funding round.

Figure 8.3 illustrates the results of the skill-luck-depth analysis, plotting the estimated parameters for luck α and skill impact factor β across different sectors and funding rounds. Each color represents a sector, while marker styles distinguish the funding rounds: circles for Seed to Series A, squares for Series A to B, and triangles for Series B to C. Arrows connect consecutive funding rounds within the same sector, pointing in the direction of higher rounds. The subplots show the one-dimensional change of parameters for each sector.

The results reveal two key findings. First, most points are clustered in a region of low luck α values between 0 and 0.1 and moderate skill impact factor β values around 1.7. This indicates that, for the majority of sectors and funding rounds, randomness plays a minimal role in determining outcomes, while skill has a measurable but not overwhelming impact on performance differentiation. Compared to sport and board games, the luck and skill impact factors are on the order of magnitude as chess or tennis [253].

Second, as funding rounds progress from early (Seed to Series A) to later stages (Series B to C), luck α increases while the skill impact factor β fluctuates in both directions.

Financial and Legal and Communication and Education for the Series B and C rounds are removed from the plot due to limited data coverage, which results in a poor fitting of the parameters.

In early rounds (e.g., Seed to Series A), where luck plays a minimal role, VCs must focus heavily on startup selection. At this stage, relying on chance would most likely result in low-performing portfolios, emphasizing the need for careful and informed decision-making. In the second round, skill becomes either more or less impactful depending on the sector, meaning that, for example for Business & Service or Health & Environment, changes in skill advantages over competitors do lead to lesser gains in multiples compared to competitors.

Impact of Benchmark Investor

Building on the insights from the luck-skill impact factor plane, where I examined trends across sectors and funding rounds, I now focus on individual VCs to better understand how skill correlates with their portfolio outcomes. Figure 8.4 provides a detailed view of this relationship, plotting each VC's skill value against their portfolio multiple, allowing us to zoom in at the VC level and explore how skill influences performance beyond sector-level trends.

Figure 8.4 plots skill values against portfolio multiples, with different funding rounds distinguished by color and marker styles. [11] The color intensity, reflects the number of companies in each VC's portfolio. The dotted horizontal line highlight the multiple of 1 and the vertical dashed lines, representing average benchmarks across all vintages

11: Circles for Seed to Series A, squares for Series A to B, and triangles for Series B to C.

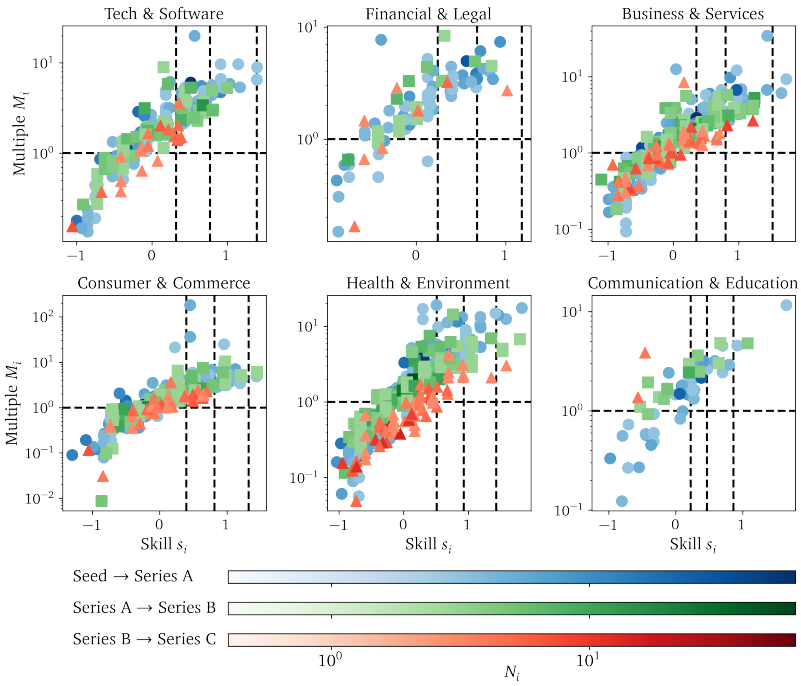


Figure 8.4.: Skill vs. multiple at the VC level: The skill values of individual VCs (x-axis) are plotted against their portfolio multiples (y-axis). Different funding rounds are distinguished by color and marker style: blue circles for Seed to Series A, green squares for Series A to B, and red triangles for Series B to C. The color intensity, as indicated by the color bar, reflects the number of companies in each VC's portfolio. The dotted horizontal line represents a portfolio multiple of 1, and the vertical dashed lines indicate the mean benchmark positions across all vintages and funding rounds, providing visual guidance.

and funding rounds to offer visual reference points for interpreting the results.[12]

The analysis reveals three main findings. First, skill values are approximately homogeneously distributed between -1 and 1, regardless of the number of companies in a VC's portfolio. This indicates that skill is evenly spread across VCs and is not inherently tied to the number of their investments.

Second, as expected, skill is positively correlated with portfolio multiples. However, skill provides more nuanced insights than multiples alone by accounting for the role of luck in portfolio outcomes. VCs achieve high skill values only if they consistently outperform across many portfolio companies, rather than relying on one or two outlier successes. This highlights the value of skill as a measure of repeatable, systematic success, rather than being driven by *lucky shots* that might inflate portfolio averages. Prominent venture capitalists, such as Sequoia Capital and Andreessen Horowitz, unsurprisingly exhibit high skill scores, but the analysis also highlights less well-known investors, such as CRV and Kleiner Perkins, who achieve equally high skill scores.

Finally, while skill values are relative—depending on the performance of VCs within the same group—the inclusion of benchmark investors provides an absolute measure of performance. The vertical benchmark lines link skill values directly to obtainable multiples, allowing for interpretation in real-world terms. This connection bridges the gap between relative and absolute measures, offering a clearer picture of what specific skill-levels translate to in terms of portfolio performance.

12: Benchmarks are calculated as percentiles for each funding round, sector, and vintage, resulting in a large number of specific benchmarks. To simplify visualization, the vertical lines show the average benchmark positions across all funding rounds, sectors, and vintages. These lines provide an approximate reference for where the benchmarks lie relative to their skill values, aiding in interpretation without overcrowding the figure.

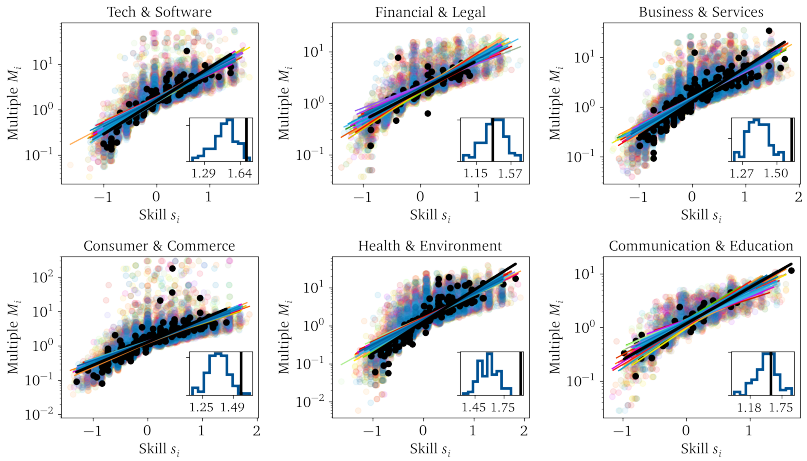


Figure 8.5.: Skill vs. multiple with randomly sampled portfolios: The scatter plots show the relationship between skill values and portfolio multiples for both real and randomly sampled VC portfolios. Black points represent real VCs, while colored points correspond to randomly assembled portfolios, with each color indicating a different random seed. The solid lines depict the linear fits for each dataset. The inset displays the distribution of slopes from the linear fits of the random portfolios, with the solid black line indicating the slope of the real system for comparison.

Measuring of True Skill

To evaluate whether real VCs exhibit true skill beyond what could be attributed to randomness, we combine the random sampling process with the skill-luck-depth framework. The objective is to determine whether the skill distribution of real VCs differs significantly from that of randomly assembled portfolios.

To do this, we first construct randomized portfolios following the process described above. We then apply the skill-luck-depth model to these portfolios, treating them as if they were real investment decisions. This allows us to assign skill values to purely random portfolios and compare them with the skill estimates from real VC portfolios.

To quantify this difference, we analyze the slope of the relationship between skill and investment multiples across different sectors for the Seed to Series A funding round. Figure 8.5 shows the scatter plots for the random seeds (colored) and the true data (black), with corresponding linear fits. The inset in each panel displays the distribution of slopes, where the black line indicates the slope of the real system.

The results reveal a clear difference between real and random portfolios. The slope of the real system is significantly higher than the mean slope of the randomly generated portfolios. This suggests that real VCs exhibit a true skill effect within the skill-luck-depth framework. The effect is particularly pronounced in sectors such as Tech & Software, Business & Services, Consumer & Commerce, and Health & Environment where the correlation between skill and multiple is much stronger in real portfolios than in their randomized counterparts. Additionally, we observe that the dispersion of skill values for a given multiple is higher in real portfolios, reinforcing the denoising effect of the skill-luck-depth framework. This suggests that the model successfully isolates persistent skill signals in real data, while random portfolios tend to produce more uniform skill estimates.

However, two sectors—Financial & Legal and Communication & Education— do not exhibit a statistically significant difference between real and random portfolios. This is likely due to insufficient data, making it challenging to obtain a reliable fit.

This analysis confirms that real VCs demonstrate a measurable skill ef-

fect, as their portfolios consistently show a stronger skill-multiple correlation than randomly assembled ones. The skill-luck-depth model effectively distinguishes real investment performance from chance, though its effectiveness depends on the availability of sufficient data within each sector.

8.5. Skill and Randomness in Analyst Forecasts

To evaluate analyst performance, I utilize the Institutional Brokers' Estimate System (IBES) dataset, which provides detailed forecasts for annual financial reports across various time horizons. Both venture capitalists and analysts share the core challenge of forming expectations about company performance, making their comparison meaningful. However, the key distinction lies in their roles: VCs can actively and operationally influence the companies they invest in but they rely heavily on intangible factors for evaluation. In contrast, analysts focus on more mature companies, leveraging extensive financial data to inform their forecasts.

Analysts are *still* optimistic

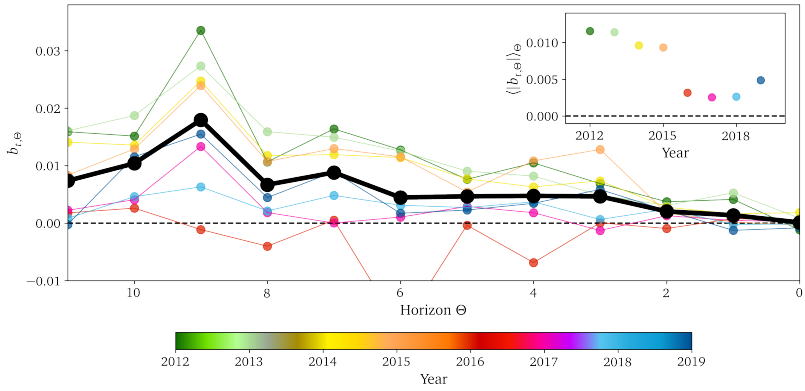


Figure 8.6.: Forecast bias across horizons: The ex-post forecast bias $b_{\alpha,t-\theta}$ is shown for different years (colored lines) across forecasting horizons θ, with the black line representing the average bias across all years. Positive values indicate overoptimistic forecasts, also found by Guedj and Bouchaud (2005) [264] for the period of 1987-2004. The inset illustrates average absolute bias across horizons.

For each company, I use the analysts prediction of Earnings per Share (EPS) values at horizons θ ranging from 11 months before the annual financial statement to 1 month before. To ensure comparability across companies with different stock prices, I normalize EPS by the company's share price, obtaining an earnings yield as the primary metric. I follow the same data cleaning method as in Guedj and Bouchaud (2005) [264] by only keeping stocks that are rated by at least 10 analysts and have an EPS between $[-10, 10]$. Additionally, I focus only on US stocks between the years 2012-2019. In total, the dataset contains 460743 rated companies by 2810 unique analysts (see Table 8.2).

Let the earnings yield for stock α at the time of the annual financial statement t be denoted as $\epsilon_{\alpha,t}$. The forecast made by analyst i for $\epsilon_{\alpha,t}$ at time $t-\theta$, where θ is the time before the earnings announcement, is represented as $f^i_{\alpha,t-\theta}$. I first examine the distribution the ex-post forecast bias, that is, following Guedj and Bouchaud (2005) [264], as:

$$b_{\alpha,t-\theta} = f_{\alpha,t-\theta} - \epsilon_{\alpha,t} \tag{8.7}$$

$\theta = 0$	N	30149
	P_i	1145
$\theta = 1$	N	43773
	P_i	1741
$\theta = 2$	N	59952
	P_i	1896
$\theta = 3$	N	4301
	P_i	231
$\theta = 4$	N	38263
	P_i	1579
$\theta = 5$	N	56343
	P_i	1821
$\theta = 6$	N	12958
	P_i	737
$\theta = 7$	N	35694
	P_i	1525
$\theta = 8$	N	52441
	P_i	1780
$\theta = 9$	N	9499
	P_i	487
$\theta = 10$	N	90272
	P_i	2511
$\theta = 11$	N	26643
	P_i	1034
Σ	N	460743
	P_i	2810

Table 8.2.: Table of analyst data: The dataset is divided into forecasting horizons Θ. N denotes the number of companies that are rated, and P_i represents the number of analysts. The sum over analyst is the count of unique analysts.

Here, $f_{\alpha,t}$ is the average forecast over all analysts, $\langle f^i_{\alpha,t-\theta}\rangle_i$. Guedj and Bouchaud (2005) [264] found that between 1987 and 2004, analyst forecasts were positively biased, meaning analysts tended to be overoptimistic.

Figure 8.6 illustrates the forecast bias $b_{t,\theta} = \langle b_{\alpha,t-\theta}\rangle_\alpha$ for different years (colored lines) across forecasting horizons θ. The black line represents the average bias across all years. The results show a consistent pattern of positive bias across forecasting horizons, with predictions becoming more accurate (closer to zero bias) as the earnings announcement approaches ($\theta \to 0$). The inset in Figure 8.6 shows the average absolute bias across horizons. This results is consistent, though with a slightly lower order of magnitude, with the findings of Guedj and Bouchaud (2005) [264], which indicates that the overoptimistic bias can still be shown in their predictions. Interestingly, from 2016 inwards the absolute prediction became significantly better (see inset in Figure 8.6).

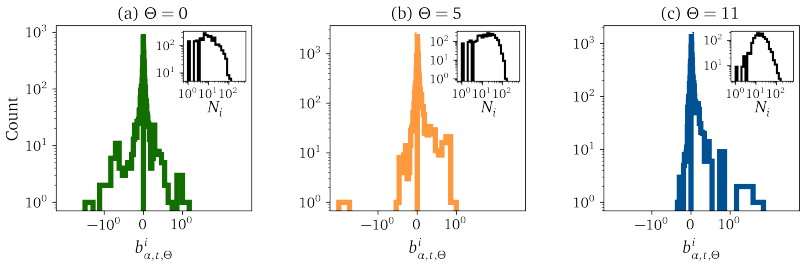


Figure 8.7.: Distribution of forecast biases: The distribution of individual analyst forecast biases $b^i_{\alpha,t,\theta}$ is shown for horizons $\theta \in \{0, 5, 11\}$ months before the earnings announcement date. Inset shows the distribution of the number of companies rated by analysts N_i.

Additionally, Figure 8.7 shows the distribution of forecast biases for individual analysts $b^i_{\alpha,t,\theta}$ for the horizons $\theta \in \{0, 5, 11\}$ months before the earning announcement date. These distributions are heavy-tailed, with the majority of predictions clustered close to the actual earnings yield, but a few outliers exhibiting large deviations. The skewness toward positive values is clearly visible in the distributions for larger horizons, reflecting the persistent overoptimism of analysts. Interestingly, the overall magnitude of biases—both positive and negative—remains similar across forecasting horizons. However, as the forecasting horizon shortens (θ decreases), the weight of the positive bias shifts closer to zero, indicating that predictions made closer to the earnings announcement are more accurate. The inset shows the distribution of number of companies rated by analyst i, denoted as N_i which has a median of ≈ 20 companies.

Prediction Error of Analysts

As an error measure I calculate the absolute percentage difference between an analyst's forecast and the true earnings yield reported in the financial statement as

$$\Delta^i_{\alpha,t-\theta} = \frac{|f^i_{\alpha,t-\theta} - \epsilon_{\alpha,t}|}{\epsilon_{\alpha,t}} \tag{8.8}$$

The percentage difference serves as a measure of forecasting accuracy, with lower values indicating better performance.[13] Each analyst's forecasts, across multiple companies and horizons, are aggregated to form a *portfolio* of predictions. The overall prediction error of an analyst i at the period $t - \theta$ is evaluated as the mean percentage difference, calculated as $\Delta^i_{t-\theta} = \langle \Delta^i_{\alpha,t-\theta} \rangle_\alpha$, where the average is taken over all stocks α in the analyst's portfolio.

13: Note that this follows the inverse logic of the VC analysis: while a high multiple indicates good performance for VCs, here, lower values indicate better performance for analysts.

Forecast Accuracy & Random Process

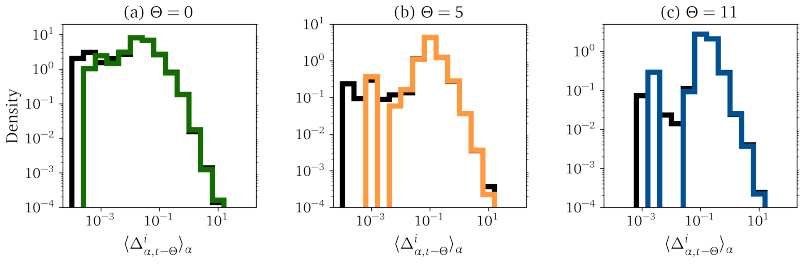


Figure 8.8.: Random shuffling of analyst predictions: The distribution of the percentage difference of predictions ($\Delta^i_{t-\theta}$) is shown for actual analyst performance (colored) compared to a random predictor baseline (black) for three forecasting horizons: (a) $\theta = 0$, (b) $\theta = 5$, and (c) $\theta = 11$ months before the earnings announcement.

To further evaluate analyst performance, similarly for Venture Capitalists (see Section 8.4), I implement an ex-post random shuffling approach to assess deviations from aggregated analyst accuracy. For each fixed stock, horizon θ, and year, I randomly shuffle the prediction error of analysts while maintaining these conditions constant. This process is repeated across 100 random seeds to simulate a baseline of randomly assigned predictions for each analyst. By aggregating the mean error ($\Delta^i_{t-\theta}$) across all stocks predicted by a given analyst i, this method allows us to compare actual analyst performance against the expected performance of random forecaster.[14]

14: Note that we use the term *random forecaster*, even though we do not generate predictions from a probability distribution. Instead, we randomly shuffle existing predictions across stocks, forecasting horizons, and years. This approach introduces a slight bias, as shown by Guedj and Bouchaud (2005) [264], who demonstrated that analysts tend to exhibit herding behavior in their forecasts.

Figure 8.8 shows the actual distributions of $\Delta^i_{t-\theta} = \langle \Delta^i_{\alpha,t-\theta} \rangle_\alpha$ for $\theta \in 0, 5, 11$ compared to the random shuffling (black lines). On the right tail, representing poorly performing analysts with high prediction errors, the distribution of actual performance aligns closely with the expectations of the random predictor. This indicates that the least accurate analysts perform as poorly as would be expected by chance, suggesting that their predictions contribute little added value relative to randomness.

In contrast, on the left tail, corresponding to highly accurate analysts with low prediction errors, actual performance deviates from the random baseline. The random benchmark suggests there should be more analysts with near-zero average prediction errors than observed in the actual data. This deviation is particularly evident for longer horizons (θ), where analysts tend to excel with predictions of some stocks but struggle with others, ultimately lowering their mean accuracy.

These findings indicate a performance asymmetry: while the worst predicting analysts perform no better than random, the most accurate analysts fail to reach their potential based on the random benchmark. Analysts appear to achieve high accuracy for a subset of assets but encounter significant challenges with others, mirroring patterns observed in venture capital.

Skill and Luck in Analyst Predictions

To construct the pairwise games required for applying the skill-luck-depth framework, I introduce three benchmark analysts. Each forecasted company of every analyst is compared against these benchmarks. The benchmarks represent the ex-post 10th, 25th, and 50th percentiles of the error distribution for all analysts, calculated separately for each year.[15]

15: Note that I take here lower percentiles as compared to the VC case as a *good* analyst achieves lower errors where as *good* VCs achieve higher multiples.

A *win* is recorded if an analyst's prediction error is lower (better) than the benchmark, and a *loss* otherwise. By introducing these benchmarks, I evaluate not only how analysts perform relative to each other but also how their accuracy aligns with desirable performance standards.

The skill parameter s_i represents an analyst's ability to consistently provide accurate forecasts across their portfolio of companies and horizons. High skill values correspond to lower errors and higher forecasting accuracy. The luck parameter α captures randomness in forecast outcomes, reflecting market volatility, company-specific surprises, and other unforeseen factors that affect earnings.

Figure 8.9 illustrates the results of fitting the skill-luck-depth model in the two-dimensional plane of luck α and skill impact factor β for different forecasting horizons θ. I conduct the fitting on a randomly selected subset of 250 analysts rather than the entire dataset to ensure time-efficient computations, as fitting the full dataset would be computationally intensive. To confirm the robustness of the results, I tested varying subset sizes and multiple random selections, ensuring that the findings remain consistent across different configurations.

The luck parameter α in public markets is $\alpha \approx 0.1$. The skill impact fac-

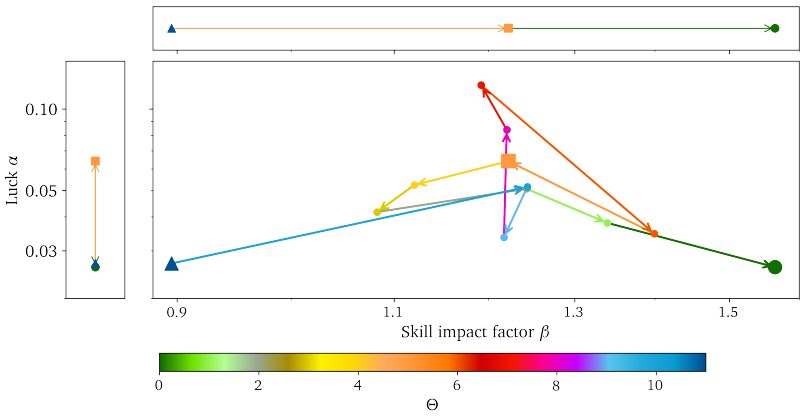


Figure 8.9.: Luck vs. skill impact factor for public markets: The estimated parameters of luck (α) and skill impact factor (β) are plotted for different forecasting horizons (θ), where each color represents a different horizon. The arrows indicate the progression across horizons, moving closer to the earnings announcement date.

tor β in public markets is consistently below 1.6 for different horizons. This suggests a flatter competitive landscape compared to venture capital, where small differences in skill have a less pronounced impact on outcomes.

Across forecasting horizons, no clear monotonic trend emerges for luck. Instead, luck appears to increase slightly for mid-range horizons ($\theta = 6$–8) before decreasing again as the horizon shortens closer to the earnings announcement date. This variation may reflect differences in uncertainty levels for predictions made farther from or nearer to the actual earnings release.

For the skill impact factor, there is an increase in β at lower horizons ($\theta \to 0$), indicating that as the earnings announcement approaches, skill becomes more significant in determining prediction accuracy (see top panel in Figure 8.9. This aligns with the expectation that analysts rely on more precise and timely data when closer to the announcement date, leading to sharper performance differentiation based on skill.

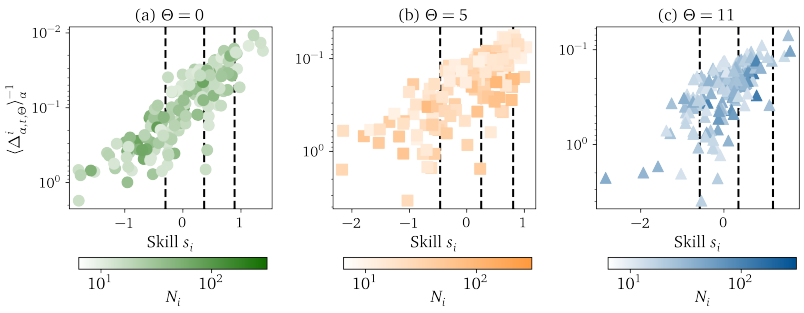


Figure 8.10.: Skill vs. accuracy for analysts: Scatter plots show the relationship between skill (x-axis) and accuracy (y-axis) for individual analysts, with each panel representing different forecasting horizons: (a) $\theta = 0$ (blue circles), (b) $\theta = 5$ (green squares), and (c) $\theta = 11$ (red triangles). Dashed vertical lines represent benchmarks corresponding to the 10th, 25th, and 50th percentiles of the accuracy distribution, linking skill values to fixed levels of accuracy for comparison.

Figure 8.10 presents a scatter plot showing the relationship between the skill parameter of each analyst and their respective accuracy $\langle \Delta^i_{\alpha,t-\theta} \rangle^{-1}_{\alpha}$. As expected, a strong positive correlation is evident: higher accuracy corresponds to higher skill. However, skill is not merely a direct reflection of accuracy—it is a denoised measure that also incorporates the consistency of performance across predictions, providing a more robust assessment of an analyst's ability.

The color intensity in the figure represents the number of companies rated by each analyst. No clear relationship emerges between the number of companies rated and either skill or accuracy, indicating that the breadth of coverage does not inherently influence these metrics.

To contextualize the skill values, dashed vertical lines represent benchmarks corresponding to the 10th, 25th, and 50th percentiles of the accuracy distribution. These benchmarks allow for a direct comparison of analysts, linking their skill scores to fixed levels of accuracy. By doing so, the figure provides a reference framework for evaluating the relative and absolute performance of analysts. Analysts with skill values above the benchmarks demonstrate superior accuracy and consistency compared to their peers and the defined percentiles.

Overall, this analysis highlights the value of skill as a more nuanced metric that extends beyond raw accuracy. It provides a means to compare analysts fairly, both within the dataset and against fixed benchmarks, offering a clear perspective on their relative and absolute performance.

Measuring of True Skill

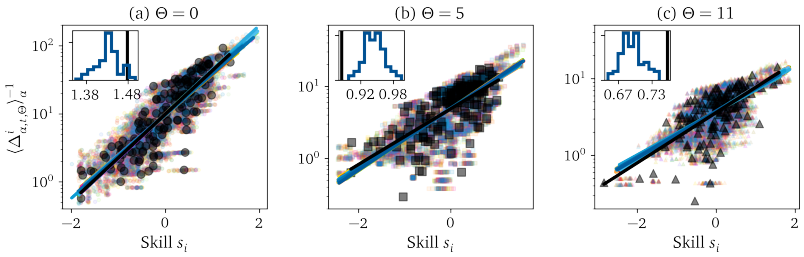


Figure 8.11.: Skill vs. multiple with randomly sampled portfolios: The scatter plots show the relationship between skill values and prediction accuracy. Black points represent real analysts, while colored points correspond to randomly assembled portfolios. The solid lines depict the linear fits for each dataset. The inset displays the distribution of slopes from the linear fits of the random portfolios, with the solid black line indicating the slope of the real system for comparison.

Building on the approach used for venture capital, we apply the skill-luck-depth framework to public market analysts by comparing the skill distributions of real analysts to those of randomly sampled prediction portfolios.

We analyze the relationship between prediction accuracy and skill, comparing the regression slopes obtained from real analyst portfolios with those from randomly generated ones. Figure 8.11 visualizes this comparison, where the colored points represent randomly sampled predictions, while the black points correspond to real analyst portfolios. The solid lines indicate the linear fits for each dataset, and the inset shows the distribution of regression slopes, with the solid black line marking the slope of the real system.

The results reveal that for short-term forecasts ($\theta = 0$) and long-term forecasts ($\theta = 11$), the real system exhibits a significantly steeper slope than the randomly sampled portfolios. This indicates that even though the mean portfolio accuracy is the same, the skill is slightly higher for the true portfolios, indicating a measurable effect with the skill-luck-depth framework of real portfolios.

However, an interesting pattern emerges at $\theta = 5$, where the effect is reversed. Here, the slope of the real system remains statistically significant but is actually lower than that of the random portfolios. One possible explanation is that analysts exhibit herding behavior, relying more on consensus estimates rather than independent analysis at this stage, which could reduce the variation in skill [264]. Additionally,

data constraints may play a role—if analysts do not have significantly new information at this horizon, their forecasts may become more uniform, making random portfolios appear similarly skilled.

Overall, these findings indicate that true skill is most apparent in short-term and long-term forecasts, where analysts can either leverage the most recent information effectively or apply deep expertise to long-term projections.

8.6. Skill & Competition: A Comparative Analysis

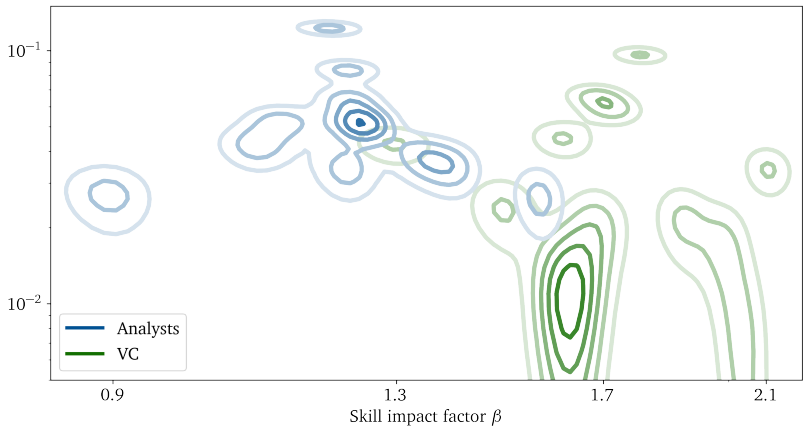


Figure 8.12.: Luck vs. skill impact factor VCs vs. Analysts: Density plot of estimated parameters of luck (α) and skill impact factor (β) are plotted for each funding round and sector for VCs in green and for each horizon for Analysts in blue.

As shown in Figure 8.12, the luck parameter α exhibits the same order of magnitude for both VCs and analysts. This suggests that the likelihood of fully unskilled participants achieving success—whether by finding a high-performing startup or making an accurate prediction—is nearly zero in both domains, highlighting the skill-driven nature of these competitive industries.

However, the skill impact factor β differs between the two systems. In venture capital, β is higher, indicating that differences in skill among investors translate into more substantial variations in returns. A highly skilled VC can achieve higher portfolio gains compared to their less skilled peers. By contrast, analysts exhibit a lower β, meaning that differences in skill lead to relatively smaller improvements in prediction accuracy. This implies a flatter competitive landscape in public markets, where even highly skilled analysts experience less pronounced gains from their expertise compared to the sharp differentiation seen among VCs.

These findings underscore the distinct dynamics of the two industries. While both are driven by skill rather than luck, the competitive advantage afforded by skill is more pronounced in venture capital, making it a higher-stakes environment for those with superior abilities. Conversely, public markets appear less stratified, offering a narrower range of differentiation based on skill.

One can compare these findings to the findings of Jerdee and Newman (2024) [253] who applied the skill-luck-depth framework to sports and games. The skill impact factor and luck in VC and analyst prediction

is of the order of magnitude as games like Basketball ($\alpha = 0.12,\ \beta \approx 0.8$), Chess ($\alpha = 0.08,\ \beta \approx 1$), Tennis ($\alpha = 0.04,\ \beta \approx 1.4$) or Soccer ($\alpha = 0.04,\ \beta \approx 1.6$)[253].

8.7. Conclusions

This chapter demonstrates the adaptability and potential of the skill-luck-depth framework in analyzing competitive dynamics in financial contexts, offering advancements in the understanding of decision-making in venture capital and public markets. By disentangling skill from luck, the framework provides actionable metrics for evaluating performance. For example, the skill parameter enables investors to develop clear investment strategies. An outside investor seeking to allocate capital to the VC asset class can assess fund managers based on skill, with the additional parameters of luck and skill importance factor offering valuable context. These parameters highlight the varying levels of diligence required, depending on the sector or funding round, to identify top-performing VCs. Similarly, analyst skill scores provide a new lens for identifying consistent outperformers in public markets.

The ability to denoise performance metrics is a critical contribution. Traditional metrics, such as raw returns or forecast accuracy, often conflate skill with the effects of luck or random events. By isolating the skill component, this framework offers a more robust measure of persistent and systematic decision-making. Moreover, skill as a dependent variable opens the door to regression analyses that can uncover the factors driving success, enabling deeper insights into how expertise manifests in different contexts.

Another key aspect is the use of pairwise comparisons with benchmark investors and analysts. This approach anchors skill to interpretable, absolute metrics, allowing for normative statements about performance. For instance, introducing benchmarks enables outside investors to evaluate whether a VC's skill is likely to yield returns above a certain threshold, providing a clear linkage between skill and expected outcomes. This advancement is particularly valuable in venture capital, where absolute returns are a critical consideration for comparing the asset class to other investment opportunities.

Despite these achievements, several limitations must be acknowledged. First, Bayesian parameter estimation requires extensive data for reliable fitting. While this is less problematic in public markets, where data is abundant, it presents challenges in venture capital, where datasets are often sparse, particularly in specific sectors or funding rounds. These limitations can lead to uncertainties in parameter estimates, particularly for outliers, such as underrepresented sectors.

Another limitation is the ex-post nature of this analysis. The framework relies on historical data to estimate skill, luck, and the skill importance factor. While this provides valuable retrospective insights, its utility for forward-looking investment strategies depends on the persistence of skill scores over time. Future work should investigate whether skill metrics remain stable across years and changing market

conditions, which would establish their predictive power.

Data availability poses additional challenges, particularly in venture capital. The absence of company valuations means that funding amounts must serve as a proxy for growth, assuming consistent shares sold across rounds. This assumption may not always hold, as valuations can increase significantly even with modest funding changes. Furthermore, the dataset does not specify the amounts contributed by individual VCs within a funding round, limiting the precision of portfolio multiple calculations. Weighted averages based on investment amounts would improve the accuracy of these metrics. Additionally, our dataset captures only startups that successfully received funding, overlooking the role of skill in discarding startups. A holistic view of the market would require including startups that never received funding but still achieved success independently, as this represents an important dimension of VC decision-making skill that remains unaccounted for.

Finally, the model does not account for network or reputation effects in venture capital, which can play a pivotal role in startup success. Prominent VCs often benefit from a *self-fulfilling prophecy* where their involvement boosts a startup's perceived credibility and likelihood of success. This dimension is not captured in the framework but remains a critical factor in understanding VC performance. By contrast, such factors are less relevant for analysts, whose evaluations are primarily driven by forecast accuracy.

Looking ahead, this framework has significant implications and potential for future research. Expanding the analysis to include additional domains, such as private equity or actively managed funds, could provide valuable insights into competitive dynamics across a broader range of contexts. Incorporating temporal effects to examine how skill, luck, and competition sensitivity evolve over time would further enhance the framework's utility. Finally, addressing data limitations, such as including company valuations or the specific amounts invested by VCs, would improve the robustness of the results and enable more precise evaluations.

Key Takeaways

1. **Skill Metrics of Investors and Analysts:** The study introduces a denoised skill metric that filters out luck, providing a clearer measure of persistent performance in both venture capital investing and financial forecasting.
2. **Luck Plays a Minor Role:** Success in both venture capital and financial analysis is largely skill-driven, with little evidence that luck alone leads to consistent outperformance. However, public markets exhibit a more level playing field, where even highly skilled analysts only marginally outperform their peers.
3. **Venture Capitalists Excel at Risk Management:** While venture capitalists do not significantly outperform random selection when identifying high-growth startups, they demonstrate strong skill in avoiding poor investments, which contributes to long-term success.
4. **Persistent Optimism in Forecasting:** Financial analysts tend to overestimate company performance, exhibiting a persistent optimism bias. However, this bias has gradually declined since 2016, suggesting improved forecasting accuracy.
5. **Asymmetric Analyst Performance** While the least accurate analysts perform no better than random, even the top-performing analysts fall short of their theoretical potential. This suggests that forecasting success is not uniformly distributed across assets, with analysts excelling in certain areas while struggling in others, a pattern also observed in venture capital decision-making.

Conclusions

9. Perspectives & Conclusions

This thesis delves into the complexities of economic systems by integrating agent-based modeling, network science, and statistical physics to explore the interplay between individual decisions, structural dynamics, and emergent phenomena. By uniting these methodologies, it addresses challenges in macroeconomics, labor market dynamics, and the role of uncertainty in shaping economic outcomes.

The thesis underscores how individual decisions aggregate to form emergent patterns, such as inflationary dynamics, workforce mobility, or the interplay of skill and chance.

Part I The first part of this thesis focuses on using agent-based modeling to explore macroeconomic dynamics, emphasizing the role of individual decision-making and its aggregation into systemic outcomes. Central to this investigation is the study of post-COVID inflation, which serves as a compelling case to analyze the interplay between monetary policy, fiscal interventions, and the expectations of economic agents.

The Mark-0 ABM is a powerful tool for simulating complex macroeconomic systems. By incorporating an expectation-anchoring mechanism, the analysis demonstrates how the stability of inflation expectations significantly influences the effectiveness of monetary policy. This extension sheds light on critical mechanisms, such as the wage-price spiral, and uncovers the intricate trade-offs between controlling inflation and managing unemployment. These findings offer valuable insights for policymakers seeking to navigate the delicate balance between economic stability and growth. However, the model's reliance on simplified behavioral rules and its assumptions about agent homogeneity pose limitations when applying its insights to more nuanced real-world scenarios.

Several extensions to the Mark-0 model could enhance its ability to capture complex dynamics. For example, the model currently does not account for a central bank with a dual mandate. Incorporating such a mandate might improve the model's representation of economic reality, though it could also challenge trust anchoring, especially when balancing opposing objectives. Additionally, redistribution effects are overlooked, as the model treats households as a single aggregate. Introducing distinct household types, such as *wealthy* and *hand-to-mouth* agents, could allow for a deeper exploration of how inflation and monetary policies affect different segments of society. Including a financial sector where households can invest as well as consume would further enrich the model's insights.

The Mark-0 economy also assumes a constant workforce. The COVID pandemic, however, has shifted workers' perspectives on the work-leisure balance, likely influenced by accumulated savings. This has

contributed to phenomena such as the *Great Resignation* and, potentially, to wage inflation. Furthermore, long-term demographic trends in developed economies are shifting away from workforce growth toward stagnation or even decline, factors that could have far-reaching implications for inflationary dynamics.

The current modeling of the energy sector is simplistic. A more detailed representation, such as a firm network with advanced production functions and multiple commodity types, could reveal new dynamics, including power-law distributions of firm sizes and cascading bankruptcies. While these enhancements would make the model more realistic, they risk undermining its simplicity as an accessible and easy-to-use scenario generator for policymakers. Striking a balance between realism and usability is critical, particularly given the increased complexity and additional data requirements that come with more detailed models.

The availability of data has been improving, thanks to collaborations between researchers and central banks or government institutions. However, greater engagement with policymakers—the end users of these models—could further align model development with practical needs. Policymakers sharing their key challenges and requirements would enable researchers to design more relevant and impactful models, ultimately bridging the gap between theoretical insights and real-world applications. These efforts would ensure that ABMs remain valuable tools for understanding macroeconomic phenomena and informing policy decisions.

This thesis further addresses the inherent complexity of agent-based models (ABMs) and the challenges posed by their vast parameter spaces by introducing the *sloppiness* framework. This framework identifies the key parameters—referred to as stiff directions—that govern model behavior, allowing for efficient exploration of high-dimensional models while reducing computational costs. By narrowing the focus to influential drivers, the approach not only simplifies the analysis of macroeconomic systems but also provides policymakers with a clear understanding of the factors most critical to systemic outcomes.

The sloppiness framework is applied to systematically explore the Mark-0 ABM, revealing that only a small subset of stiff parameter combinations significantly influence the model's dynamics. This enhances both the interpretability and predictive scope of the model, allowing modelers to understand parameter sensitivities and the range of possible dynamics more deeply. However, the current application focuses on the dynamics of a single variable, such as inflation or unemployment. Expanding the approach to assess the Hessian matrix for multi-variable changes would enable a more holistic evaluation, such as the impact of a central bank's dual mandate. Additionally, the growing adoption of differential ABMs simplifies the implementation of the sloppiness framework, making it more accessible for broader use. Furthermore, shifting the analysis from time-series dynamics to predicted distributions of variables—using as a loss function the Kullback-Leibler divergence—could mitigate path dependencies introduced by the stochasticity of ABMs (see [58]).

Our framework also offers a structured pathway for improving model-data alignment. By initially focusing on the most critical parameters and progressively addressing less influential ones, the approach enhances parameter fitting efficiency and can help prioritize the acquisition of important datasets for calibration. Importantly, while the stiff parameters provide a focused lens on model behavior, there is potential in the sloppy directions. These less-sensitive parameters can inform alternative policies that achieve similar outcomes, expanding the decision-making toolkit for policymakers.

Part II The second part of this thesis investigates labor markets using the methodologies of complex systems and network science, treating them as networks of interconnected occupations linked by worker transitions. This approach uncovers structural imbalances, mobility patterns, and the pivotal role of skills in shaping occupational flows, contributing to the broader theme of understanding economic systems as dynamic and interdependent systems.

Chapter 5 examines the structural properties of occupational transitions in the French labor market, uncovering power-law distributions in occupation sizes and transition flows. These findings emphasize the hierarchical nature of the labor market, with certain occupations acting as mobility hubs while others remain isolated. The analysis also highlights a near-equilibrium symmetry in worker flows, revealing inherent bottlenecks and structural imbalances. The use of network science approaches offers a computational advantage, enabling efficient processing of large datasets and providing a different perspective on labor market dynamics. However, some insights remain descriptive and are obtained implicitly from previous economic approaches. The use of aggregated data simplifies the analysis and computation but reduces specificity, limiting its utility for designing targeted retraining programs or other direct interventions.

To address this gap, Chapter 6 introduces accessibility and transferability metrics, which quantify worker mobility to identify occupational bottlenecks. These metrics reveal that *condenser* occupations, defined by high accessibility but low transferability, act as barriers to efficient mobility, reinforcing structural imbalances. While these metrics provide a robust framework for identifying bottlenecks, they are limited by the static nature of the used dataset. Incorporating path dependencies and tracking worker transitions over time would allow researchers to construct longitudinal labor paths and would offer a more nuanced understanding of labor market dynamics. Furthermore, while the focus on identifying bottleneck occupations provides valuable insights into the *what*, equal attention should be given to the *how*—operationalizing these findings for effective policy interventions. Greater collaboration with policymakers could facilitate access to more detailed datasets and inform the design of metrics and models tailored to applicable policies. Early involvement in survey design and policy planning could ensure that research aligns more closely with the specific needs and actionable goals of policymakers. Future studies should also explore how accessibility, transferability patterns, and broader labor market dynamics evolved during the COVID crisis and recovery. Datasets covering this period were unavailable for this

research, however their future availability could provide valuable insights into labor market shocks and serve as a stress test for the applicability of existing models under unprecedented conditions.

Chapter 7 incorporates skill similarity into the labor market framework, shedding light on how skills shape occupational flows and identifying critical bottlenecks. By comparing skill-based networks with realized flow networks, the analysis highlights both their overlaps and the complementary insights they provide, emphasizing the potential of combining these approaches to address labor market inefficiencies. However, the analysis also reveals gaps in available data, particularly concerning path-dependent skill acquisition and the impact of retraining programs. Micro-level data on how retraining—both on the job and for unemployed individuals—affects workers' positioning in the labor market would enable the development of models that not only identify key areas but also answer the crucial question of *how* to implement policies effectively.

Part III The third part investigates the interplay between skill and luck in determining outcomes within complex systems, examining how these factors influence success and performance across domains. This analysis aligns with the core themes of the thesis by exploring the balance between deterministic, skill-driven mechanisms and stochastic, luck-based elements that shape emergent patterns and structures in economic systems. Using a framework grounded in complexity science, the chapter disentangles skill and luck in the contexts of venture capital investing and analysts' predictions of company metrics. The findings reveal that while luck diminishes over time, skill plays a progressively larger role, with skill improvements having a greater impact on venture capitalists' performance compared to analysts.

However, the analysis does not account for macroeconomic, vintage, or network effects, which are significant influences in these systems. Future work could extend the framework by regressing the derived skill values of individual agents on explanatory variables, such as economic conditions or network characteristics. If skill persistence and repeatable patterns emerge from such regressions, the metrics developed here could move beyond descriptive insights to actionable strategies, enabling the formulation of investment hypotheses for both public and private market investors.

Ultimately, this thesis demonstrates how the integration of complexity science and economic modeling can provide a holistic understanding of socio-economic systems. By bridging theoretical and practical approaches, it offers a toolkit to address the multifaceted challenges of our interconnected world.

Bibliography

References in citation order:

[1] Johannes Voit. *The statistical mechanics of financial markets*. Springer, 2003 (cited on page v).

[2] Karl Naumann-Woleske et al. 'Exploration of the Parameter Space in Macroeconomic Agent-Based Models'. In: *Handbook of Complexity Economics*. 2024 (cited on pages vii, 12, 25, 51, 52, 57, 192).

[3] Max Sina Knicker et al. 'Post-COVID inflation and the monetary policy dilemma: An agent-based scenario analysis'. In: *Journal of Economic Interaction and Coordination* (2024), pp. 1–55 (cited on pages vii, 23).

[4] Max Sina Knicker et al. 'The Structure of Occupational Mobility in France'. In: *arXiv preprint arXiv:2407.14179* (2024) (cited on pages vii, 105, 124).

[5] Brian Arthur et al. 'The Economy as an Evolving Complex System II'. In: *Reading, Addison-Wesley* (1997) (cited on pages 2, 6).

[6] Jamie Morgan. *What is Neoclassical Economics?*. Taylor & Francis, 2015 (cited on page 3).

[7] Léon Walras. *Éléments d'économie politique pure: ou, Théorie de la richesse sociale*. F. Rouge, 1877 (cited on page 3).

[8] Alfred Marshall. ''The Pure Theory of Domestic Values'(1879)'. In: *The Foundations of Price Theory Vol 3*. Routledge, 1879, pp. 113–152 (cited on page 3).

[9] David Vines and Samuel Wills. 'The rebuilding macroeconomic theory project'. In: *Oxford Review of Economic Policy* 34.1/2 (2018), pp. 1–42 (cited on pages 3, 4).

[10] Harold L Cole, Lee E Ohanian, et al. 'The Great Depression in the United States from a neoclassical perspective'. In: *Federal Reserve Bank of Minneapolis Quarterly Review* 23 (1999), pp. 2–24 (cited on page 3).

[11] John Maynard Keynes. 'The general theory of employment'. In: *The quarterly journal of economics* 51.2 (1937), pp. 209–223 (cited on page 3).

[12] Michel De Vroey. *A history of macroeconomics from Keynes to Lucas and beyond*. Cambridge University Press, 2016 (cited on page 4).

[13] Michel De Vroey. 'Getting rid of Keynes? A survey of the history of macroeconomics from Keynes to Lucas and beyond'. In: *A Survey of the History of Macroeconomics from Keynes to Lucas and Beyond (April 2, 2010). National Bank of Belgium Working Paper* 187 (2010) (cited on page 4).

[14] Robert Lucas. 'Econometric policy evaluation: A critique'. In: *Carnegie-Rochester conference series on public policy*. Vol. 1. North-Holland. 1976, pp. 19–46 (cited on page 4).

[15] Finn E Kydland and Edward C Prescott. 'Time to build and aggregate fluctuations'. In: *Econometrica: Journal of the Econometric Society* (1982), pp. 1345–1370 (cited on page 4).

[16] Andrew G Haldane and Arthur E Turrell. 'Drawing on different disciplines: macroeconomic agent-based models'. In: *Journal of Evolutionary Economics* 29 (2019), pp. 39–66 (cited on pages 5, 25, 57).

[17] Alan P Kirman. 'Whom or what does the representative individual represent?' In: *Journal of economic perspectives* 6.2 (1992), pp. 117–136 (cited on page 5).

[18] Federico Guglielmo Morelli et al. 'Confidence collapse in a multihousehold, self-reflexive DSGE model'. In: *Proceedings of the National Academy of Sciences* 117.17 (2020), pp. 9244–9249 (cited on page 5).

[19] Frank Smets and Rafael Wouters. 'Shocks and Frictions in US Business Cycles: A Bayesian DSGE Approach'. In: *American economic review* 97.3 (2007), pp. 586–606 (cited on pages 5, 61, 62).

[20] Jordi Galí. *Monetary Policy, Inflation, and the Business Cycle: An Introduction to the New Keynesian Framework*. Second. Princeton, N.J: Princeton University Press, 2015 (cited on pages 5, 61, 62).

[21] Takeshi Yagihashi. 'DSGE models used by policymakers: A survey'. In: *Policy Research Institute, Ministry of Finance, Japan, PRI Discussion Paper Series* (2020) (cited on page 5).

[22] Giovanni Dosi and Andrea Roventini. 'More Is Different ... and Complex! The Case for Agent-Based Macroeconomics'. In: *Journal of Evolutionary Economics* 29.1 (Mar. 2019), pp. 1–37. DOI: 10.1007/s00191-019-00609-y (cited on pages 5, 6, 57).

[23] Olivier Blanchard. 'On the Future of Macroeconomic Models'. In: *Oxford Review of Economic Policy* 34.1-2 (Jan. 2018), pp. 43–54. DOI: 10.1093/oxrep/grx045 (cited on pages 5, 61).

[24] David Vines and Samuel Wills. 'The rebuilding macroeconomic theory project part II: multiple equilibria, toy models, and policy models in a new macroeconomic paradigm'. In: *Oxford Review of Economic Policy* 36.3 (2020), pp. 427–497 (cited on page 5).

[25] William A. Branch and Bruce McGough. 'Heterogeneous Expectations and Micro-Foundations in Macroeconomics'. In: *Handbook of Computational Economics*. Ed. by Cars Hommes and Blake LeBaron. Vol. 4. Handbook of Computational Economics. Elsevier, 2018, pp. 3–62. DOI: https://doi.org/10.1016/bs.hescom.2018.03.001 (cited on page 5).

[26] Greg Kaplan et al. 'Monetary policy according to HANK'. In: *American Economic Review* 108.3 (2018), pp. 697–743 (cited on page 5).

[27] W Brian Arthur. 'Complexity economics'. In: *Complexity and the Economy* (2013) (cited on page 5).

[28] W Brian Arthur. 'Foundations of complexity economics'. In: *Nature Reviews Physics* 3.2 (2021), pp. 136–145 (cited on page 5).

[29] Jean-Philippe Bouchaud. 'Economics needs a scientific revolution'. In: *Nature* 455.7217 (2008), pp. 1181–1181 (cited on page 6).

[30] J Doyne Farmer et al. 'Economics needs to treat the economy as a complex system'. In: *Paper for the INET conference 'rethinking economics and politics*. Vol. 14. 2012 (cited on page 6).

[31] J Doyne Farmer and Duncan Foley. 'The economy needs agent-based modelling'. In: *Nature* 460.7256 (2009), pp. 685–686 (cited on page 6).

[32] Dale T Mortensen and Christopher A Pissarides. 'Job creation and job destruction in the theory of unemployment'. In: *The review of economic studies* 61.3 (1994), pp. 397–415 (cited on page 7).

[33] Peter A Diamond. 'Aggregate demand management in search equilibrium'. In: *Journal of political Economy* 90.5 (1982), pp. 881–894 (cited on page 7).

[34] Dale T Mortensen. 'The matching process as a noncooperative bargaining game'. In: *The economics of information and uncertainty*. University of Chicago Press, 1982, pp. 233–258 (cited on page 7).

[35] Christopher A Pissarides. 'Short-run equilibrium dynamics of unemployment, vacancies, and real wages'. In: *The American Economic Review* 75.4 (1985), pp. 676–690 (cited on page 7).

[36] Jens J Krüger. 'Productivity and structural change: a review of the literature'. In: *Journal of Economic Surveys* 22.2 (2008), pp. 330–363 (cited on page 8).

[37] Simon Kuznets. 'Modern economic growth: findings and reflections'. In: *The American economic review* 63.3 (1973), pp. 247–258 (cited on page 8).

[38] John R Hicks. 'Marginal productivity and the principle of variation'. In: *Economica* 35 (1932), pp. 79–88 (cited on page 8).

[39] David Card and John E DiNardo. 'Skill-biased technological change and rising wage inequality: Some problems and puzzles'. In: *Journal of labor economics* 20.4 (2002), pp. 733–783 (cited on page 8).

[40] Daron Acemoglu and David Autor. *Skills, Tasks and Technologies: Implications for Employment and Earnings*. 2011 (cited on pages 8, 134).

[41] Joseph A Schumpeter. *Capitalism, socialism and democracy*. routledge, 1962 (cited on page 8).

[42] H David. 'The "task approach" to labor markets: an overview'. In: *Journal for Labour Market Research* 46.3 (2013), pp. 185–199 (cited on page 8).

[43] Carl Benedikt Frey and Michael A Osborne. 'The future of employment: How susceptible are jobs to computerisation?' In: *Technological forecasting and social change* 114 (2017), pp. 254–280 (cited on pages 8, 132).

[44] Omar A Guerrero and Robert L Axtell. 'Employment growth through labor flow networks'. In: *PloS one* 8.5 (2013), e60808 (cited on page 8).

[45] Frank Neffke and Martin Henning. 'Skill relatedness and firm diversification'. In: *Strategic Management Journal* 34.3 (2013), pp. 297–316 (cited on pages 8, 10, 76, 107, 109, 124).

[46] Ian M Schmutte. 'Free to move? A network analytic approach for learning the limits to job mobility'. In: *Labour Economics* 29 (2014), pp. 49–61 (cited on pages 8, 107).

[47] Leonard Bocquet. 'The Network Origin of Slow Labor Reallocation'. In: (2022) (cited on pages 8, 10, 76, 105–107, 109, 110, 118, 124, 130).

[48] R Maria del Rio-Chanona et al. 'Occupational mobility and automation: a data-driven network model'. In: *Journal of The Royal Society Interface* 18.174 (2021), p. 20200898 (cited on pages 8, 17, 76, 107, 124).

[49] Ian Bradley and Ronald L Meek. *Matrices and society: matrix algebra and its applications in the social sciences*. Vol. 501. Princeton University Press, 1986 (cited on pages 9, 87).

[50] Halina Frydman et al. 'Testing the adequacy of Markov chain and mover-stayer models as representations of credit behavior'. In: *Operations Research* 33.6 (1985), pp. 1203–1214 (cited on page 9).

[51] Hie Joo Ahn et al. *The dual us labor market uncovered*. Tech. rep. National Bureau of Economic Research, 2023 (cited on page 9).

[52] E Alvarez et al. 'A locally stationary Markov chain model for labour dynamics'. In: *Journal of data science* 7 (2008), pp. 27–42 (cited on page 9).

[53] Mr Ippei Shibata. *Labor market dynamics: A hidden markov approach*. International Monetary Fund, 2019 (cited on page 9).

[54] Robert E Hall and Marianna Kudlyak. *Churn and Stability: The Remarkable Heterogeneity of Flows among Employment, Job Search, and Non-Market Activities in the US Population*. Tech. rep. Working paper, 2022 (cited on page 9).

[55] Victoria Gregory et al. 'The alpha beta gamma of the labor market'. In: *Journal of Monetary Economics* (2024), p. 103695 (cited on page 9).

[56] Denis Fougere and Thierry Kamionka. 'Econometrics of individual labor market transitions'. In: *The Econometrics of Panel Data: Fundamentals and Recent Developments in Theory and Practice*. Springer, 2008, pp. 865–905 (cited on page 10).

[57] Kevin Donovan et al. 'Labor market dynamics and development'. In: *The Quarterly Journal of Economics* 138.4 (2023), pp. 2287–2325 (cited on pages 10, 76).

[58] Karl Naumann-Woleske. 'Scenario Discovery in a Complex Economy: Exploring the Parameter space of Agent-based Models'. PhD thesis. Institut Polytechnique de Paris, 2024 (cited on pages 12, 57, 159).

[59] Andrew Ilachinski. *Cellular automata: a discrete universe*. World Scientific Publishing Company, 2001 (cited on page 13).

[60] Stephen Wolfram. *New kind of science*. 1997 (cited on page 13).

[61] Carsten Knudsen. 'Chaos without nonperiodicity'. In: *The American Mathematical Monthly* 101.6 (1994), pp. 563–565 (cited on page 13).

[62] Edgar F Codd. *Cellular automata*. Academic press, 2014 (cited on page 14).

[63] Sébastian Grauwin et al. 'Competition between collective and individual dynamics'. In: *Proceedings of the National Academy of Sciences* 106.49 (2009), pp. 20622–20626 (cited on pages 14, 15, 181).

[64] Thomas C Schelling. 'Dynamic models of segregation'. In: *Journal of mathematical sociology* 1.2 (1971), pp. 143–186 (cited on page 14).

[65] Laetitia Gauvin et al. 'Phase diagram of a Schelling segregation model'. In: *The European Physical Journal B* 70 (2009), pp. 293–304 (cited on page 15).

[66] Alessandro Pluchino et al. 'Talent versus luck: The role of randomness in success and failure'. In: *Advances in Complex systems* 21.03n04 (2018), p. 1850014 (cited on pages 15, 16, 137).

[67] Domenico Delli Gatti et al. *Macroeconomics from the Bottom-up*. Vol. 1. Springer Science & Business Media, 2011 (cited on page 17).

[68] Silvano Cincotti et al. 'Credit money and macroeconomic instability in the agent-based model and simulator Eurace'. In: *Economics* 4.1 (2010), p. 20100026 (cited on page 17).

[69] Giovanni Dosi et al. 'Micro and Macro Policies in the Keynes+Schumpeter Evolutionary Models'. In: *Journal of Evolutionary Economics* 27.1 (Jan. 2017), pp. 63–90. DOI: 10.1007/s00191-016-0466-4 (cited on pages 17, 25).

[70] Herbert Dawid and Domenico Delli Gatti. 'Chapter 2 - Agent-Based Macroeconomics'. In: *Handbook of Computational Economics*. Ed. by Cars Hommes and Blake LeBaron. Vol. 4. Handbook of Computational Economics. Elsevier, Jan. 2018, pp. 63–156. DOI: 10.1016/bs.hescom.2018.02.006. (Visited on 01/26/2021) (cited on pages 17, 25).

[71] Sebastian Poledna et al. 'Economic forecasting with an agent-based model'. In: *European Economic Review* 151 (2023), p. 104306 (cited on pages 17, 19).

[72] Samuel Wiese et al. 'Forecasting macroeconomic dynamics using a calibrated data-driven agent-based model'. In: *arXiv preprint arXiv:2409.18760* (2024) (cited on pages 17, 19).

[73] Isabelle Salle et al. 'Inflation targeting in a learning economy: An ABM perspective'. In: *Economic Modelling* 34 (2013), pp. 114–128 (cited on page 17).

[74] Francesco Lamperti and Andrea Roventini. 'Beyond climate economics orthodoxy: impacts and policies in the agent-based integrated-assessment DSK model'. In: *European Journal of Economics and Economic Policies* 19.3 (2022), pp. 357–380 (cited on page 17).

[75] Francesco Lamperti et al. 'Faraway, so close: Coupled climate and economic dynamics in an agent-based integrated assessment model'. In: *Ecological Economics* 150 (2018), pp. 315–339 (cited on page 17).

[76] Kathyrn R Fair and Omar A Guerrero. 'Endogenous labour flow networks'. In: *arXiv preprint arXiv:2301.07979* (2023) (cited on pages 17, 18, 76).

[77] Marco Bardoscia et al. 'The impact of prudential regulation on the UK housing market and economy: Insights from an agent-based model'. In: *Journal of Economic Behavior & Organization* 229 (2025), p. 106839 (cited on page 17).

[78] Cars Hommes et al. 'CANVAS: A Canadian behavioral agent-based model for monetary policy'. In: *Journal of Economic Dynamics and Control* (2024), p. 104986 (cited on page 18).

[79] Bence Mérő et al. 'A high-resolution, data-driven agent-based model of the housing market'. In: *Journal of Economic Dynamics and Control* 155 (2023), p. 104738 (cited on page 18).

[80] Marco Pangallo and R Maria del Rio-Chanona. 'Data-driven economic agent-based models'. In: *arXiv preprint arXiv:2412.16591* (2024) (cited on page 18).

[81] R Maria del Rio-Chanona et al. 'Large language models reduce public knowledge sharing on online Q&A platforms'. In: *PNAS nexus* 3.9 (2024), pgae400 (cited on page 18).

[82] Marijn J Ton et al. 'A global dataset of 7 billion individuals with socio-economic characteristics'. In: *Scientific Data* 11.1 (2024), p. 1096 (cited on page 18).

[83] A Borsos et al. 'Agentbased modeling at central banks: Recent developments and new challenges'. In: *This Book* (2024) (cited on page 18).

[84] Cars Hommes. 'Behavioral and experimental macroeconomics and policy analysis: A complex systems approach'. In: *Journal of Economic Literature* 59.1 (2021), pp. 149–219 (cited on page 18).

[85] Michael WM Roos. *Values, attitudes and economic behavior*. 777. Ruhr Economic Papers, 2018 (cited on page 18).

[86] Joshua M Epstein. *Generative social science: Studies in agent-based computational modeling*. Princeton University Press, 2012 (cited on page 18).

[87] Isabelle Salle et al. *Lifetime memories of inflation: Evidence from surveys and the lab*. Tech. rep. National Bureau of Economic Research, 2024 (cited on page 18).

[88] Joep Lustenhouwer and Isabelle Salle. 'Learning to be rational in the presence of news: a lab investigation'. In: *Available at SSRN 4739235* (2024) (cited on page 18).

[89] Emanuele Borgonovo et al. 'Sensitivity analysis of agent-based models: a new protocol'. In: *Computational and Mathematical Organization Theory* 28.1 (2022), pp. 52–94 (cited on page 19).

[90] Boxin Tang. 'Orthogonal array-based Latin hypercubes'. In: *Journal of the American statistical association* 88.424 (1993), pp. 1392–1397 (cited on page 19).

[91] Francesco Lamperti et al. 'Agent-Based Model Calibration Using Machine Learning Surrogates'. In: *Journal of Economic Dynamics and Control* 90 (May 2018), pp. 366–389. DOI: 10.1016/j.jedc.2018.03.011 (cited on pages 19, 59).

[92] Chathika Gunaratne et al. 'Generating mixed patterns of residential segregation: An evolutionary approach'. In: *Journal of Artificial Societies and Social Simulation* 26.2 (2023) (cited on page 19).

[93] Ryan N. Gutenkunst et al. 'Universally Sloppy Parameter Sensitivities in Systems Biology Models'. In: *PLOS Computational Biology* 3.10 (Oct. 2007), e189. DOI: 10.1371/journal.pcbi.0030189 (cited on pages 19, 20, 58, 61, 62, 67).

[94] Kevin S Brown and James P Sethna. 'Statistical mechanical approaches to models with many poorly known parameters'. In: *Physical review E* 68.2 (2003), p. 021904 (cited on page 20).

[95] Katherine N. Quinn. 'Chebyshev Approximation and the Global Geometry of Model Predictions'. In: *Physical Review Letters* 122.15 (2019). DOI: 10.1103/PhysRevLett.122.158302 (cited on page 20).

[96] Joshua J Waterfall et al. 'Sloppy-model universality class and the Vandermonde matrix'. In: *Physical review letters* 97.15 (2006), p. 150601 (cited on page 20).

[97] Chieh-Ting (Jimmy) Hsu et al. 'Numerical Parameter Space Compression and Its Application to Biophysical Models'. In: *Biophysical Journal* 118.6 (Mar. 2020), pp. 1455–1465. DOI: 10.1016/j.bpj.2020.01.023 (cited on pages 20, 58, 61, 66).

[98] Katherine N Quinn et al. 'Information geometry for multiparameter models: New perspectives on the origin of simplicity'. In: *Reports on Progress in Physics* 86.3 (2022), p. 035901 (cited on page 20).

[99] Stanislao Gualdi et al. 'Tipping Points in Macroeconomic Agent-Based Models'. In: *Journal of Economic Dynamics and Control* 50.C (2015), pp. 29–61 (cited on pages 23, 25, 27, 30, 33, 35, 37, 38, 41, 52, 55).

[100] Olivier Blanchard. 'Where Danger Lurks'. In: *Finance and Development (IMF)* (2014) (cited on pages 23, 26).

[101] François de Soyres et al. *Fiscal Policy and Excess Inflation during Covid-19: A Cross-Country View*. Tech. rep. July 2022. (Visited on 04/14/2023) (cited on page 24).

[102] François de Soyres et al. 'Demand-Supply Imbalance during the Covid-19 Pandemic: The Role of Fiscal Policy'. In: *International Finance Discussion Papers (IFDP)* (Aug. 2022). (Visited on 04/14/2023) (cited on page 24).

[103] Thomas Ferguson and Servaas Storm. *Myth and Reality in the Great Inflation Debate: Supply Shocks and Wealth Effects in a Multipolar World Economy*. Tech. rep. Institute for New Economic Thinking Working Paper Series, Jan. 2023. DOI: 10.36687/inetwp196. (Visited on 04/14/2023) (cited on page 24).

[104] Joseph Stiglitz and Ira Regmi. *The Causes of and Responses to Today's Inflation*. Tech. rep. Roosevelt Institute, 2022. (Visited on 04/14/2023) (cited on pages 24, 25, 34, 47, 54).

[105] Alberto Cavallo and Oleksiy Kryvtsov. 'What can stockouts tell us about inflation? Evidence from online micro data'. In: *Journal of International Economics* (2023), p. 103769 (cited on page 24).

[106] Veronica Guerrieri et al. 'Macroeconomic Implications of COVID-19: Can Negative Supply Shocks Cause Demand Shortages?' In: *American Economic Review* 112.5 (May 2022), pp. 1437–1474. DOI: 10.1257/aer.20201063. (Visited on 01/03/2024) (cited on page 24).

[107] Francesco Ferrante et al. 'The Inflationary Effects of Sectoral Reallocation'. In: *International Finance Discussion Paper* 1369 (Feb. 2023), pp. 1–57. DOI: 10.17016/ifdp.2023.1369. (Visited on 01/03/2024) (cited on page 24).

[108] Isabella Weber et al. 'Inflation in Times of Overlapping Emergencies: Systemically Significant Prices from an Input-output Perspective'. In: (Dec. 2022) (cited on page 24).

[109] Ben Bernanke and Olivier Blanchard. 'What Caused the U.S. Pandemic-Era Inflation?' In: (2023). (Visited on 05/23/2023) (cited on page 24).

[110] Andrew Glover et al. 'How Much Have Record Corporate Profits Contributed to Recent Inflation?' In: *The Federal Reserve Bank of Kansas City Economic Review* (Jan. 2023). DOI: 10.18651/ER/v108n1GloverMustredelRiovonEndeBecker. (Visited on 04/14/2023) (cited on page 25).

[111] Isabella Weber and Evan Wasner. 'Sellers' Inflation, Profits and Conflict: Why Can Large Firms Hike Prices in an Emergency?' In: *Economics Department Working Paper Series* (Jan. 2023) (cited on pages 25, 36).

[112] Alan S. Blinder. 'The Anatomy of Double-Digit Inflation in the 1970s'. In: *Inflation: Causes and Effects*. University of Chicago Press, 1982, pp. 261–282. (Visited on 04/14/2023) (cited on page 25).

[113] Alan S. Blinder and Jeremy B. Rudd. 'The Supply-Shock Explanation of the Great Stagflation Revisited'. In: *NBER Chapters* (2013), pp. 119–175. (Visited on 04/14/2023) (cited on page 25).

[114] Michael Bruno and William Easterly. 'Inflation Crises and Long-Run Growth'. In: *Journal of Monetary Economics* 41.1 (Feb. 1998), pp. 3–26. DOI: 10.1016/S0304-3932(97)00063-9. (Visited on 04/14/2023) (cited on page 25).

[115] Stanislao Gualdi et al. 'Monetary Policy and Dark Corners in a Stylized Agent-Based Model'. In: *Journal of Economic Interaction and Coordination* 12.3 (Oct. 2017), pp. 507–537. DOI: 10.1007/s11403-016-0174-z. (Visited on 12/06/2020) (cited on pages 25, 26, 28, 29, 32, 33, 55, 188).

[116] Jean-Philippe Bouchaud et al. 'Optimal Inflation Target: Insights from an Agent-Based Model'. In: *Economics: The Open-Access, Open-Assessment E-Journal* (2018). DOI: 10.5018/economics-ejournal.ja.2018-15. (Visited on 08/26/2020) (cited on pages 25, 27–29, 33, 38, 55).

[117] Dhruv Sharma et al. 'V-, U-, L-, or W-shaped Recovery after COVID? Insights from an Agent Based Model'. In: *PLoS One* 16.3 (July 2020). (Visited on 08/31/2020) (cited on pages 25–28, 33, 35, 41, 42, 49, 53, 54, 60).

[118] Giovanni Dosi and Andrea Roventini. 'More Is Different ... and Complex! The Case for Agent-Based Macroeconomics'. In: *Journal of Evolutionary Economics* 29.1 (Mar. 2019), pp. 1–37. DOI: 10.1007/s00191-019-00609-y. (Visited on 01/26/2021) (cited on page 25).

[119] Mervyn King and John Kay. *Radical Uncertainty: Decision-making for an unknowable future*. Hachette UK, 2020 (cited on pages 27, 55).

[120] Jean-Philippe Bouchaud. 'Radical complexity'. In: *Entropy* 23.12 (2021), p. 1676 (cited on pages 27, 55).

[121] Edoardo Gaffeo et al. 'Adaptive Microfoundations for Emergent Macroeconomics'. In: *Eastern Economic Journal* 34.4 (Oct. 2008), pp. 441–463. DOI: 10.1057/eej.2008.27. (Visited on 04/18/2023) (cited on page 27).

[122] Domenico Delli Gatti et al. *Macroeconomics from the Bottom-up*. First. New Economic Windows. Milano: Springer Science & Business Media, 2011 (cited on page 27).

[123] Dimitris Christelis et al. 'Trust in the central bank and inflation expectation'. In: *Available at SSRN 3540974* (2020) (cited on page 32).

[124] Ricardo Reis. 'The burst of high inflation in 2021–22: how and why did we get here?' In: (2022) (cited on page 34).

[125] Jeremy B. Rudd. 'Why Do We Think That Inflation Expectations Matter for Inflation? (And Should We?)' In: *Review of Keynesian Economics* 10.1 (Jan. 2022), pp. 25–45. DOI: 10.4337/roke.2022.01.02. (Visited on 03/15/2023) (cited on page 34).

[126] Ricardo Reis. 'Losing the Inflation Anchor'. In: *Brookings Papers on Economic Activity* Fall (Sept. 2021), pp. 307–361. (Visited on 04/14/2023) (cited on pages 34, 38, 46).

[127] Milton Friedman. *The optimum quantity of money, and other essays*. Tech. rep. 1969 (cited on page 36).

[128] Hiroshi Ugai et al. 'Effects of the quantitative easing policy: A survey of empirical analyses'. In: *Monetary and economic studies-Bank of Japan* 25.1 (2007), p. 1 (cited on page 36).

[129] Ricardo Reis and Silvana Tenreyro. 'Helicopter money: what is it and what does it do?' In: *Annual Review of Economics* 14 (2022), pp. 313–335 (cited on page 36).

[130] Lucy Chennells. 'The windfall tax'. In: *Fiscal Studies* 18.3 (1997), pp. 279–291 (cited on page 36).

[131] Isabella Weber. 'Big Oil's Profits and Inflation: Winners and Losers'. In: *Challenge* (2022), pp. 1–9 (cited on page 36).

[132] Mr Thomas Baunsgaard and Nate Vernon. *Taxing Windfall Profits in the Energy Sector*. International Monetary Fund, 2022 (cited on page 36).

[133] David Miles et al. *And yet it moves: Inflation and the great recession*. International Center for Monetary and Banking Studies/-Centre for Economic …, 2017 (cited on page 38).

[134] FRED. 'Personal Consumption Expenditures [PCE]'. In: *U.S. Bureau of Economic Analysis, Federal Reserve Bank of St. Louis* (2023). accessed February 5, 2023 (cited on page 39).

[135] FRED. 'Consumer Price Index: Total All Items for the United States [CPALTT01USM657N]'. In: *Organization for Economic Co-operation and Development, Federal Reserve Bank of St. Louis* (2023). accessed February 5, 2023 (cited on page 39).

[136] FRED. 'Spot Crude Oil Price: West Texas Intermediate (WTI) [WTISPLC]'. In: *Federal Reserve Bank of St. Louis* (2023). accessed February 5, 2023 (cited on page 39).

[137] Maria Grazia Attinasi et al. 'Supply chain disruptions and the effects on the global economy'. In: *Economic Bulletin Boxes* 8 (2022) (cited on page 39).

[138] Behnam Zakeri et al. 'Pandemic, War, and Global Energy Transitions'. In: *Energies* 15.17 (2022), p. 6114 (cited on page 40).

[139] William H Beveridge. *Full employment in a free society (Works of William H. Beveridge): a report*. Routledge, 2014 (cited on page 48).

[140] Peter A Diamond and Ayşegül Şahin. 'Shifts in the Beveridge curve'. In: *Research in Economics* 69.1 (2015), pp. 18–25 (cited on page 48).

[141] Gene Kindberg-Hanlon and Michael Girard. *What Caused the Beveridge Curve to Shift Higher in the United States During the Pandemic?* International Monetary Fund, 2024 (cited on page 49).

[142] Mark K. Transtrum et al. 'Perspective: Sloppiness and Emergent Theories in Physics, Biology, and Beyond'. In: *The Journal of Chemical Physics* 143.1 (July 2015), p. 010901. DOI: 10.1063/1.4923066 (cited on pages 51, 58).

[143] Jorge Alvarez et al. 'Wage-Price Spirals: What is the Historical Evidence?' In: (2022) (cited on page 52).

[144] A Steven Holland. 'The Changing Responsiveness of Wages to Price-Level Shocks: Explicit and Implicit Indexation'. In: *Economic Inquiry* 26.2 (1988), pp. 265–279 (cited on page 52).

[145] Robert J Shiller. *Narrative economics: How stories go viral and drive major economic events*. Princeton University Press, 2020 (cited on page 54).

[146] Théo Dessertaine et al. 'Out-of-equilibrium dynamics and excess volatility in firm networks'. In: *Journal of Economic Dynamics and Control* 138 (2022), p. 104362 (cited on page 55).

[147] Leigh Tesfatsion. 'Agent-Based Computational Economics: Growing Economies from the Bottom Up'. In: *SSRN Electronic Journal* (2002) (cited on page 57).

[148] Blake LeBaron and Leigh Tesfatsion. 'Modeling Macroeconomies as Open-Ended Dynamic Systems of Interacting Agents'. In: *American Economic Review* 98.2 (May 2008), pp. 246–250. DOI: 10.1257/aer.98.2.246 (cited on page 57).

[149] Joshua M. Epstein. 'Agent-Based Computational Models and Generative Social Science'. In: *Complexity* 4.5 (1999), pp. 41–60. DOI: 10.1002/(SICI)1099-0526(199905/06)4:5<41::AID-CPLX9>3.0.CO;2-F (cited on page 57).

[150] Alan Kirman and Mauro Gallegati. 'A Quarter of a Century of Complex Economics: How Far Have We Come?' In: *Handbook of Complexity Economics*. London, New York: Routledge, 2022 (cited on page 57).

[151] Brian Arthur. 'Some Thoughts on Agent-based Modeling and the Role of Computation in Economics'. In: *Handbook of Complexity Economics*. London, New York: Routledge, 2022 (cited on page 57).

[152] Stanislao Gualdi et al. 'Tipping Points in Macroeconomic Agent-Based Models'. In: *Journal of Economic Dynamics and Control* 50.C (2015), pp. 29–61 (cited on pages 57–60, 62–65, 69–71).

[153] Richard Bookstaber. *The End of Theory*. Princeton, N.J: Princeton University Press, 2017 (cited on page 57).

[154] Jean-Philippe Bouchaud. 'To Make Sense of Complex Systems, Send in the Agents'. In: *Risk.net* (2020) (cited on page 57).

[155] Christopher A. Sims. 'Macroeconomics and Reality'. In: *Econometrica* 48.1 (Jan. 1980), p. 1. DOI: 10.2307/1912017 (cited on page 57).

[156] Giorgio Fagiolo and Andrea Roventini. 'Macroeconomic Policy in DSGE and Agent-Based Models Redux: New Developments and Challenges Ahead'. In: *Journal of Artificial Societies and Social Simulation* 20.1 (2017), p. 1. DOI: 10.18564/jasss.3280 (cited on pages 57, 61).

[157] Joshua F. Apgar et al. 'Sloppy Models, Parameter Uncertainty, and the Role of Experimental Design'. In: *Molecular BioSystems* 6.10 (2010), pp. 1890–1900. DOI: 10.1039/B918098B (cited on page 58).

[158] Mark K. Transtrum et al. 'Why Are Nonlinear Fits to Data so Challenging?' In: *Physical Review Letters* 104.6 (Feb. 2010), p. 060201. DOI: 10.1103/PhysRevLett.104.060201 (cited on page 58).

[159] Mark K. Transtrum et al. 'Geometry of Nonlinear Least Squares with Applications to Sloppy Models and Optimization'. In: *Physical Review E* 83.3 (Mar. 2011), p. 036701. DOI: 10.1103/PhysRevE.83.036701 (cited on page 58).

[160] Benjamin B. Machta et al. 'Parameter Space Compression Underlies Emergent Theories and Predictive Models'. In: *Science* 342.6158 (Nov. 2013), pp. 604–607. DOI: 10.1126/science.1238723 (cited on pages 58, 62).

[161] Brian K. Mannakee et al. 'Sloppiness and the Geometry of Parameter Space'. In: *Uncertainty in Biology*. Ed. by Liesbet Geris and David Gomez-Cabrero. Vol. 17. Cham: Springer International Publishing, 2016, pp. 271–299. DOI: 10.1007/978-3-319-21296-8 (cited on page 58).

[162] Greg Kaplan and Giovanni L. Violante. 'Microeconomic Heterogeneity and Macroeconomic Shocks'. In: *Journal of Economic Perspectives* 32.3 (Aug. 2018), pp. 167–194. DOI: 10.1257/jep.32.3.167 (cited on page 58).

[163] Joseph E Stiglitz. 'Where Modern Macroeconomics Went Wrong'. In: *Oxford Review of Economic Policy* 34.1-2 (2018), pp. 70–106 (cited on pages 58, 61).

[164] Arika Ligmann-Zielinska et al. ''One Size Does Not Fit All': A Roadmap of Purpose-Driven Mixed-Method Pathways for Sensitivity Analysis of Agent-Based Models'. In: *Journal of Artificial Societies and Social Simulation* 23.1 (2020), p. 6 (cited on page 58).

[165] Stanislao Gualdi et al. 'Monetary Policy and Dark Corners in a Stylized Agent-Based Model'. In: *Journal of Economic Interaction and Coordination* 12.3 (Oct. 2017), pp. 507–537. DOI: 10.1007/s11403-016-0174-z (cited on pages 58, 71).

[166] Domenico Delli Gatti et al. *Macroeconomics from the Bottom-up*. First. New Economic Windows. Milano: Springer Science & Business Media, 2011 (cited on pages 58, 61).

[167] Giovanni Dosi et al. 'On the robustness of the fat-tailed distribution of firm growth rates: a global sensitivity analysis'. In: *Journal of Economic Interaction and Coordination* 13.1 (2018), pp. 173–193 (cited on page 59).

[168] Giovanni Dosi et al. 'The impact of deunionization on the growth and dispersion of productivity and pay'. In: *Industrial and Corporate Change* 30.2 (2021), pp. 377–408 (cited on page 59).

[169] Sander van der Hoog. 'Surrogate Modelling in (and of) Agent-Based Models: A Prospectus'. In: *Computational Economics* 53.3 (Mar. 2019), pp. 1245–1263. DOI: 10.1007/s10614-018-9802-0 (cited on page 59).

[170] Yi Zhang et al. *Validation and Calibration of an Agent-Based Model: A Surrogate Approach*. https://www.hindawi.com/journals/ddns/2020/6946370/. Research Article. Jan. 2020. DOI: 10.1155/2020/6946370 (cited on page 59).

[171] Guus ten Broeke et al. 'The Use of Surrogate Models to Analyse Agent-Based Models'. In: *Journal of Artificial Societies and Social Simulation* 24.2 (2021), p. 3 (cited on page 59).

[172] Leonardo Bargigli et al. 'Network Calibration and Metamodeling of a Financial Accelerator Agent Based Model'. In: *Journal of Economic Interaction and Coordination* 15.2 (Apr. 2020), pp. 413–440. DOI: 10.1007/s11403-018-0217-8 (cited on page 59).

[173] Siyan Chen and Saul Desiderio. 'A Regression-Based Calibration Method for Agent-Based Models'. In: *Computational Economics* (Feb. 2021). DOI: 10.1007/s10614-021-10106-9 (cited on page 59).

[174] Lawrence J. Christiano et al. 'On DSGE Models'. In: *Journal of Economic Perspectives* 32.3 (Aug. 2018), pp. 113–140. DOI: 10.1257/jep.32.3.113 (cited on page 61).

[175] Marco Del Negro et al. 'Inflation in the Great Recession and New Keynesian Models'. In: *American Economic Journal: Macroeconomics* 7.1 (Jan. 2015), pp. 168–196. DOI: 10.1257/mac.20140097 (cited on pages 61, 62).

[176] Charles T. Carlstrom et al. 'Targeting Long Rates in a Model with Segmented Markets'. In: *American Economic Journal: Macroeconomics* 9.1 (Jan. 2017), pp. 205–242. DOI: 10.1257/mac.20150179 (cited on pages 61, 62).

[177] M. D. McKay et al. 'A Comparison of Three Methods for Selecting Values of Input Variables in the Analysis of Output from a Computer Code'. In: *Technometrics* 21.2 (1979), pp. 239–245. DOI: 10.2307/1268522 (cited on page 68).

[178] Thomas M. Cioppa and Thomas W. Lucas. 'Efficient Nearly Orthogonal and Space-Filling Latin Hypercubes'. In: *Technometrics* 49.1 (2007), pp. 45–55 (cited on page 68).

[179] Jean-Philippe Bouchaud et al. 'Optimal Inflation Target: Insights from an Agent-Based Model'. In: *Economics: The Open-Access, Open-Assessment E-Journal* (2018). DOI: 10.5018/economics-ejournal.ja.2018-15 (cited on page 71).

[180] Katherine N. Quinn et al. 'Information Geometry for Multiparameter Models: New Perspectives on the Origin of Simplicity'. arXiv. 2021 (cited on pages 73, 74).

[181] Daron Acemoglu et al. 'The network origins of aggregate fluctuations'. In: *Econometrica* 80.5 (2012), pp. 1977–2016 (cited on page 76).

[182] Marco Bardoscia et al. 'The physics of financial networks'. In: *Nature Reviews Physics* 3.7 (2021), pp. 490–507 (cited on page 76).

[183] Charlie Joyez and Catherine Laffineur. 'The occupation space: network structure, centrality and the potential of labor mobility in the French labor market'. In: *Applied Network Science* 7.1 (2022), p. 16 (cited on pages 76, 107).

[184] Neave O'Clery and Stephen Kinsella. 'Modular structure in labour networks reveals skill basins'. In: *Research Policy* 51.5 (2022), p. 104486 (cited on pages 76, 107).

[185] Sabrina Aufiero et al. 'Mapping job fitness and skill coherence into wages: an economic complexity analysis'. In: *Scientific Reports* 14.1 (2024), p. 11752 (cited on pages 76, 107, 113, 126, 131).

[186] Frank Neffke et al. 'Inter-industry labor flows'. In: *Journal of Economic Behavior & Organization* 142 (2017), pp. 275–292 (cited on pages 76, 107).

[187] Vincent D Blondel et al. 'Fast unfolding of communities in large networks'. In: *Journal of statistical mechanics: theory and experiment* 2008.10 (2008), P10008 (cited on page 77).

[188] Shawn Martin et al. 'OpenOrd: an open-source toolbox for large graph layout'. In: *Visualization and data analysis 2011*. Vol. 7868. SPIE. 2011, pp. 45–55 (cited on page 77).

[189] Mathieu Bastian et al. 'Gephi: an open source software for exploring and manipulating networks'. In: *Third international AAAI conference on weblogs and social media*. 2009 (cited on page 77).

[190] Mito Tsukamoto et al. *Towards the Right to Work: A Guidebook for Designing Innovative Public Employment Programmes*. 2012 (cited on page 78).

[191] Vilfredo Pareto. *Cours d'économie politique*. Vol. 1. Librairie Droz, 1964 (cited on pages 79, 80).

[192] Felix Auerbach. 'Das gesetz der bevölkerungskonzentration [The law of population concentration]'. In: *Petermanns Geographische Mitteilungen* 59 (1913), pp. 74–76 (cited on page 79).

[193] Yuji Ijiri and Herbert Alexander Simon. *Skew distributions and the sizes of business firms*. Vol. 24. North-Holland, 1977 (cited on page 79).

[194] George Kingsley Zipf. *Human Behavior and the Principle of Least Effort*. Vol. 1. Cambridge:Addison-Wesley, 1949 (cited on page 80).

[195] Robert L Axtell. 'Zipf distribution of US firm sizes'. In: *science* 293.5536 (2001), pp. 1818–1820 (cited on page 80).

[196] Vasiliki Plerou et al. 'Scaling of the distribution of price fluctuations of individual companies'. In: *Physical review e* 60.6 (1999), p. 6519 (cited on page 80).

[197] Hideaki Aoyama et al. *Econophysics and companies: statistical life and death in complex business networks*. Cambridge University Press, 2010 (cited on page 82).

[198] John A Peacock. 'Two-dimensional goodness-of-fit testing in astronomy'. In: *Monthly Notices of the Royal Astronomical Society* 202.3 (1983), pp. 615–627 (cited on page 83).

[199] Günther Schmid. *Full employment in Europe: Managing labour market transitions and risks*. Edward Elgar Publishing, 2008 (cited on page 85).

[200] Isadore Blumen et al. 'The industrial mobility of labor as a probability process'. In: *(No Title)* (1955) (cited on page 87).

[201] Halina Frydman. 'Maximum likelihood estimation in the mover-stayer model'. In: *Journal of the American Statistical Association* 79.387 (1984), pp. 632–638 (cited on page 87).

[202] Denis Fougère and Thierry Kamionka. 'Bayesian inference for the mover–stayer model in continuous time with an application to labour market transition data'. In: *Journal of Applied Econometrics* 18.6 (2003), pp. 697–723 (cited on page 87).

[203] Leonid V Kantorovich. 'Mathematical methods of organizing and planning production'. In: *Management science* 6.4 (1960), pp. 366–422 (cited on page 90).

[204] Ludwig Boltzmann. *Lectures on gas theory*. Univ of California Press, 1872 (cited on page 91).

[205] YK Wong. 'An application of orthogonalization process to the theory of least squares'. In: *The Annals of Mathematical Statistics* 6.2 (1935), pp. 53–75 (cited on page 94).

[206] Romain Allez and Jean-Philippe Bouchaud. 'Eigenvector dynamics: general theory and some applications'. In: *Physical Review E—Statistical, Nonlinear, and Soft Matter Physics* 86.4 (2012), p. 046202 (cited on pages 94, 95).

[207] Mark EJ Newman. 'Scientific collaboration networks. II. Shortest paths, weighted networks, and centrality'. In: *Physical review E* 64.1 (2001), p. 016132 (cited on page 97).

[208] Ulrik Brandes. 'A faster algorithm for betweenness centrality'. In: *Journal of mathematical sociology* 25.2 (2001), pp. 163–177 (cited on page 98).

[209] L Freeman. 'Centrality in networks: I. conceptual clarifications. social networks'. In: *Social Network* 10 (1979), pp. 0378–8733 (cited on page 98).

[210] Jukka-Pekka Onnela et al. 'Intensity and coherence of motifs in weighted complex networks'. In: *Physical Review E—Statistical, Nonlinear, and Soft Matter Physics* 71.6 (2005), p. 065103 (cited on page 99).

[211] Vincent D Blondel et al. 'Fast unfolding of communities in large networks'. In: *Journal of statistical mechanics: theory and experiment* 2008.10 (2008), P10008 (cited on page 100).

[212] John Platig et al. 'Bipartite community structure of eQTLs'. In: *PLoS computational biology* 12.9 (2016), e1005033 (cited on pages 100, 101, 109, 132, 195).

[213] Ludvig Bohlin et al. 'Community detection and visualization of networks with the map equation framework'. In: *Measuring scholarly impact: methods and practice* (2014), pp. 3–34 (cited on pages 100, 101, 132).

[214] Jianbo Shi and Jitendra Malik. 'Normalized cuts and image segmentation'. In: *IEEE Transactions on pattern analysis and machine intelligence* 22.8 (2000), pp. 888–905 (cited on pages 100, 101, 132).

[215] Michael J Barber. 'Modularity and community detection in bipartite networks'. In: *Physical Review E* 76.6 (2007), p. 066102 (cited on pages 100, 109, 195).

[216] Vincent A Traag et al. 'From Louvain to Leiden: guaranteeing well-connected communities'. In: *Scientific reports* 9.1 (2019), pp. 1–12 (cited on pages 100, 102).

[217] Thomas Aynaud. *Python-Louvain: Louvain Algorithm for Community Detection*. https://github.com/taynaud/python-louvain. Version 0.16, BSD License. 2016. URL: https://github.com/taynaud/python-louvain (cited on page 100).

[218] Genís Calderer. *PyCONDOR: Python implementation of CONDOR with the BRIM algorithm*. https://github.com/genisott/pycondor. Version accessed: 2024. 2018. URL: https://github.com/genisott/pycondor (cited on page 101).

[219] Daniel Edler et al. *The MapEquation software package*. https://mapequation.org. 2025 (cited on page 102).

[220] Fabian Pedregosa et al. 'Scikit-learn: Machine learning in Python'. In: *Journal of machine learning research* 12.Oct (2011), pp. 2825–2830 (cited on pages 102, 116).

[221] Rodrigo Adão et al. 'Fast and slow technological transitions'. In: (2022) (cited on pages 105, 107).

[222] Bhaskar Ghosh et al. 'GenAI Will Change How We Design Jobs. Here's How.' In: *Harvard Business Review* (2023) (cited on page 105).

[223] Attilio Di Battista et al. 'Future of jobs report 2023'. In: *World Economic Forum, Geneva, Switzerland. https://www. weforum. org/reports/the-future-of-jobs-report-2023*. 2023 (cited on page 105).

[224] Gary S Becker. *Human capital: A theoretical and empirical analysis, with special reference to education*. University of Chicago press, 2009 (cited on pages 106, 116).

[225] Simon Marginson. 'Limitations of human capital theory'. In: *Studies in higher education* 44.2 (2019), pp. 287–301 (cited on page 106).

[226] Enghin Atalay et al. 'New technologies and the labor market'. In: *Journal of Monetary Economics* 97 (2018), pp. 48–67 (cited on page 107).

[227] David Autor et al. *On the persistence of the China shock*. Tech. rep. National Bureau of Economic Research, 2021 (cited on page 107).

[228] Guido Matias Cortes and Eliza Forsythe. 'Heterogeneous labor market impacts of the COVID-19 pandemic'. In: *ILR Review* 76.1 (2023), pp. 30–55 (cited on page 107).

[229] Rafael Dix-Carneiro et al. 'Globalization, trade imbalances, and labor market adjustment'. In: *The Quarterly Journal of Economics* 138.2 (2023), pp. 1109–1171 (cited on page 107).

[230] Francisco J Buera et al. 'Skill-biased structural change'. In: *The Review of Economic Studies* 89.2 (2022), pp. 592–625 (cited on page 107).

[231] Benoıt Schmutz and Modibo Sidibé. 'Search and Zipf'. In: (2020) (cited on page 107).

[232] Patrick Kline and Enrico Moretti. 'Place based policies with unemployment'. In: *American Economic Review* 103.3 (2013), pp. 238–243 (cited on page 107).

[233] Cristian Jara-Figueroa et al. 'The role of industry-specific, occupation-specific, and location-specific knowledge in the growth and survival of new firms'. In: *Proceedings of the National Academy of Sciences* 115.50 (2018), pp. 12646–12653 (cited on page 107).

[234] M.E.J Newman and M Girvan. 'Finding and evaluating community structure in networks'. In: *Physical Review E* 69.2 (2004) (cited on page 109).

[235] Andrea Tacchella et al. 'A new metrics for countries' fitness and products' complexity'. In: *Scientific reports* 2.1 (2012), p. 723 (cited on pages 112–114).

[236] Andrea Zaccaria et al. 'How the taxonomy of products drives the economic development of countries'. In: *PloS one* 9.12 (2014), e113770 (cited on page 112).

[237] César A Hidalgo and Ricardo Hausmann. 'The building blocks of economic complexity'. In: *Proceedings of the national academy of sciences* 106.26 (2009), pp. 10570–10575 (cited on page 112).

[238] Greg Morrison et al. 'On economic complexity and the fitness of nations'. In: *Scientific Reports* 7.1 (2017), p. 15332 (cited on page 112).

[239] Angelica Sbardella et al. 'Green technology fitness'. In: *Entropy* 20.10 (2018), p. 776 (cited on page 112).

[240] Matteo Straccamore et al. 'Urban economic fitness and complexity from patent data'. In: *Scientific Reports* 13.1 (2023), p. 3655 (cited on page 112).

[241] Zhuo-Ming Ren et al. 'Bridging nestedness and economic complexity in multilayer world trade networks'. In: *Humanities and Social Sciences Communications* 7.1 (2020), pp. 1–8 (cited on pages 114, 196).

[242] Dario Laudati et al. 'The different structure of economic ecosystems at the scales of companies and countries'. In: *Journal of Physics: Complexity* 4.2 (2023), p. 025011 (cited on pages 114, 196).

[243] Dorin Comaniciu and Peter Meer. 'Mean shift: A robust approach toward feature space analysis'. In: *IEEE Transactions on pattern analysis and machine intelligence* 24.5 (2002), pp. 603–619 (cited on page 116).

[244] Linton C Freeman. 'A set of measures of centrality based on betweenness'. In: *Sociometry* (1977), pp. 35–41 (cited on page 117).

[245] Bernard Gazier. 'Opportunities or tensions: Assessing French Labour market reforms from 2012 to 2018'. In: *International Journal of Comparative Labour Law and Industrial Relations* 35.3 (2019) (cited on page 121).

[246] Christina Gathmann and Uta Schönberg. 'How general is human capital? A task-based approach'. In: *Journal of Labor Economics* 28.1 (2010), pp. 1–49 (cited on pages 124, 125).

[247] France Travail. 'Répertoire Opérationnel des Métiers et des Emplois (ROME)'. In: *Data.gouv.fr* (2009). accessed January 1, 2025 (cited on page 125).

[248] Ahmad Alabdulkareem et al. 'Unpacking the polarization of workplace skills'. In: *Science advances* 4.7 (2018), eaao6030 (cited on pages 125, 131).

[249] JA Nelder. 'Log linear models for contingency tables: a generalization of classical least squares'. In: *Journal of the Royal Statistical Society: Series C (Applied Statistics)* 23.3 (1974), pp. 323–329 (cited on page 127).

[250] Daniel McFadden. 'Quantitative methods for analysing travel behaviour of individuals: some recent developments'. In: *Behavioural travel modelling*. Routledge, 2021, pp. 279–318 (cited on page 127).

[251] Giselmar AJ Hemmert et al. 'Log-likelihood-based pseudo-R 2 in logistic regression: deriving sample-sensitive benchmarks'. In: *Sociological Methods & Research* 47.3 (2018), pp. 507–531 (cited on page 127).

[252] Emil Mihaylov and Kea Gartje Tijdens. 'Measuring the routine and non-routine task content of 427 four-digit ISCO-08 occupations'. In: (2019) (cited on page 131).

[253] Maximilian Jerdee and MEJ Newman. 'Luck, skill, and depth of competition in games and social hierarchies'. In: *Science Advances* 10.45 (2024), eadn2654 (cited on pages 136, 137, 139, 140, 144, 153, 154).

[254] Michael L Hartzmark. 'Luck versus forecast ability: Determinants of trader performance in futures markets'. In: *Journal of Business* (1991), pp. 49–74 (cited on page 137).

[255] Rama K Malladi. 'Luck versus skill in evaluating hedge fund managers' performance'. In: *Journal of Business and Management* 26.1 (2020), pp. 22–39 (cited on page 137).

[256] Andrew Mauboussin and Samuel Arbesman. 'Differentiating Skill and Luck in Financial Markets with Streaks'. In: *Available at SSRN 1664031* (2011) (cited on page 137).

[257] Andrew J Leone and Joanna S Wu. 'What does it take to become a superstar? Evidence from institutional investor rankings of financial analysts'. In: *Evidence from Institutional Investor Rankings of Financial Analysts (May 23, 2007). Simon School of Business Working Paper No. FR* (2007), pp. 02–12 (cited on page 138).

[258] Arthur Korteweg and Morten Sorensen. 'Skill and luck in private equity performance'. In: *Journal of Financial Economics* 124.3 (2017), pp. 535–562 (cited on page 138).

[259] Paul Gompers et al. *Skill vs. luck in entrepreneurship and venture capital: Evidence from serial entrepreneurs*. 2006 (cited on page 138).

[260] Diego Liechti et al. 'Luck and entrepreneurial success'. In: *Swiss Finance Institute Research Paper* 14-51 (2014) (cited on page 138).

[261] Aracely Soto-Simeone et al. 'The role of skill versus luck in new venture survival'. In: *International Journal of Management Reviews* 23.4 (2021), pp. 549–556 (cited on page 138).

[262] Naveen Khanna and Richmond D Mathews. 'Skill versus reliability in venture capital'. In: *Journal of Financial Economics* 145.2 (2022), pp. 41–63 (cited on page 138).

[263] Radford M Neal. 'MCMC using Hamiltonian dynamics'. In: *arXiv preprint arXiv:1206.1901* (2012) (cited on page 140).

[264] Olivier Guedj and Jean-Philippe Bouchaud. 'Experts' earning forecasts: bias, herding and gossamer information'. In: *International Journal of Theoretical and Applied Finance* 8.07 (2005), pp. 933–946 (cited on pages 147–149, 152).

[265] Mário Almeida-Neto et al. 'A consistent metric for nestedness analysis in ecological systems: reconciling concept and measurement'. In: *Oikos* 117.8 (2008), pp. 1227–1239 (cited on page 196).

[266] Eurostat. 'Being young in Europe today–Labour market–Access and participation'. In: (2018) (cited on page 227).

Appendix

Part I A.

A.1. Hessian of Schelling Model within the boundaries

Now, we can use the above formulas to calculate the hessian with the Kullback Leibler divergence for the Schelling model as in Section 2 and Grauwin et al. (2009) [63]. We know that the stationary probability distribution is given by

$$p(\vec{x};\Phi) = \frac{1}{Z}\exp\frac{F(\vec{x},\Phi)}{T} \tag{A.1}$$

with the free energy $F(x) = V(x)+TS(x)$ and a normalisation constant Z. For larger number of cells in each block H, the free energy can be written as

$$F(x,\Phi) = H\sum_q f(\rho_q) = H\sum_q \left[-T\rho_q\ln\rho_q - T(1-\rho_q)\ln(1-\rho_q) + \alpha\rho_q u(\rho_q) + (1-\alpha)\int_0^{\rho'_q} u(\rho'_q)\,d\rho'_q\right] \tag{A.2}$$

With the asymmetrically peaked utility function we can find for the free energy:

$$u(\rho) = \begin{cases} 2\rho & \text{if } \rho \le \frac{1}{2} \\ m + 2(1-m)(1-\rho) & \text{if } \rho > \frac{1}{2} \end{cases} \tag{A.3}$$

$$f(\rho) = -T\rho\ln\rho - T(1-\rho)\ln(1-\rho)\begin{cases} +(1+\alpha)\rho^2 & \text{if } \rho \le \frac{1}{2} \\ -(1+\alpha)(1-m)\rho^2 + (2-m)\rho - \frac{(1-\alpha)(2-m)}{4} & \text{if } \rho > \frac{1}{2} \end{cases} \tag{A.4}$$

Note that $F(\vec{x},\Phi)$ where $\Phi = \{\alpha, m, \rho_0\}$ and $\vec{x}$ are the microscopic configurations of densities in each block. We know that the schelling model has two steady state distributions depending on the parameter set Φ where we find the system either in a homogeneous phase with block density of $\rho^\star = \rho_0$ in each block or in a heterogeneous phase where the blocks are either of density $\rho_1^\star$ or $\rho_2^\star$ with $\rho_1^\star < \rho_2^\star$.

Let us first consider the case where $T \to 0$, then the entropic terms vanish.
To calculate the hessian with the equations derived above, we need to find $\vec{x}^\star$ that maximizes the free energy $F(\vec{x}) = H\sum_q f(\rho_q)$ under the constraint $g(\vec{x}) = \sum_q \rho_q = Q\rho_0 = const$ that the sum over all densities in each block must be equal to the number of blocks times the initial density. Since the constraint is linear in ρ_q, it vanishes in the lagrangian function, and we find that in the homogeneous phase $\vec{x}^\star = \vec{x}^\star(\Phi) = \vec{x}^\star(\rho_0) = (\rho_0, \rho_0, ..., \rho_0)$.

With this we can calculate the equations 4.23

$$\mathscr{H}^0_{\mu,\nu} = \frac{M^2}{4}\mathrm{Tr}[\mathbb{H}^{-1}\partial_\mu \mathbb{H}(\Phi)]\mathrm{Tr}[\mathbb{H}^{-1}\partial_\nu \mathbb{H}(\Phi)] \tag{A.5}$$

$$\mathscr{H}^{1,1}_{\mu,\nu} = -\frac{M}{2}\mathrm{Tr}[\mathbb{H}^{-1}\partial_\mu \mathbb{H}(\Phi)]\mathrm{Tr}[\mathbb{H}^{-1}\partial_\nu \mathbb{H}(\Phi)] \tag{A.6}$$

$$\mathscr{H}^{1,2}_{\mu,\nu} = \frac{M}{2}\,\mathrm{Tr}[\mathbb{H}(\Phi)\mathbb{K}_{\mu,\nu}], \tag{A.7}$$

where

$$\left(\mathbb{K}_{\mu,\nu}\right)_{ij} := \partial_\mu \vec{x}^\star_i \partial_\nu \vec{x}^\star_j + i \leftrightarrow j. \tag{A.8}$$

With the following calculations:

$$F(\vec{x}^\star, \Phi) = F(\vec{x}^\star, \alpha, \rho_0) = H \cdot Q \cdot (1+\alpha)\rho_0^2 \tag{A.9}$$

$$(\mathbb{H})_{ij} = \partial_i \partial_j F(x, \Phi)\big|_{\vec{x}=\vec{x}^\star} = 2H(1+\alpha)\delta_{ij} \tag{A.10}$$

$$(\mathbb{H}^{-1})_{ij} = \frac{1}{2H(1+\alpha)}\delta_{ij} \tag{A.11}$$

$$\partial_\mu(\mathbb{H})_{ij} = \partial_\mu\left(2H(1+\alpha)\delta_{ij}\right) = 2H\delta_{\mu,\alpha}\delta_{ij} \tag{A.12}$$

$$\mathrm{Tr}[\mathbb{H}^{-1}\partial_\nu \mathbb{H}] = \mathrm{Tr}[\frac{1}{(1+\alpha)}\delta_{\nu,\alpha}\mathbb{1}^{QxQ}] = \frac{Q}{(1+\alpha)}\delta_{\nu,\alpha} \tag{A.13}$$

$$\left(\mathbb{K}_{\mu,\nu}\right)_{ij} = \partial_\mu \vec{x}^\star_i \partial_\nu \vec{x}^\star_j + i \leftrightarrow j = 2\partial_\mu \vec{x}^\star \partial_\nu \vec{x}^\star J^Q \tag{A.14}$$

$$\mathrm{Tr}[\mathbb{H}K_{\mu,\nu}] = \mathrm{Tr}[4H(1+\alpha)\partial_\mu \vec{x}^\star \partial_\nu \vec{x}^\star \mathbb{1}^{QxQ}] = 4H \cdot Q(1+\alpha)\delta_{\mu,\rho_0}\delta_{\nu,\rho_0} \tag{A.15}$$

where J^Q is the all-ones matrix and $J^Q\,\mathbb{1}^{QxQ} = \mathbb{1}^{QxQ}$.

Now we can write with $\Phi = (\alpha, m, \rho = 0)$

$$\mathscr{H}^0 = \frac{M^2Q^2}{4(1+\alpha)^2}\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \tag{A.16}$$

$$\mathscr{H}^{1,1} = \frac{-MQ^2}{2(1+\alpha)^2}\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \tag{A.17}$$

$$\mathscr{H}^{1,2} = 2MHQ(1+\alpha)\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{A.18}$$

$$\mathscr{H} = \begin{pmatrix} A & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & B \end{pmatrix} \tag{A.19}$$

$$\tag{A.20}$$

with $A = \frac{MQ^2}{2(1+\alpha)^2}(\frac{M}{2} - 1)$ and $B = 2MHQ(1+\alpha)$. So in the homogeneous phase for $\rho \leq \frac{1}{2}$ is not as exciting as the hessian matrix is of rank two with the eigenvalues $\lambda_1 = A, \lambda_2 = B$ and $\lambda_3 = 0$ and the eigenvectors $v_1 = (1,0,0)^T$ (pointing in the direction of α) and $v_2 = (0,0,1)^T$ (pointing in the direction of ρ_0), $v_3 = (0,1,0)^T$.
Now we can to the same calculation for the case where $\rho > \frac{1}{2}$, then

$$f(\rho) = -(1+\alpha)(1-m)\rho^2 + (2-m)\rho - \frac{(1-\alpha)(2-m)}{4} \tag{A.21}$$

We get an additional contribution due to the dependency of m:

$$(\mathbb{H})_{ij} = \partial_i \partial_j F(x,\Phi)\big|_{\vec{x}=\vec{x}^\star} = -2H(1+\alpha)(1-m)\delta_{ij} \tag{A.22}$$

$$(\mathbb{H}^{-1})_{ij} = \frac{1}{-2H(1+\alpha)(1-m)}\delta_{ij} \tag{A.23}$$

$$\partial_\mu(\mathbb{H})_{ij} = \partial_\mu\left(-2H(1+\alpha)(1-m)\delta_{ij}\right) = -2H((1-m)\delta_{\mu,\alpha} + (1+\alpha)\delta_{\mu,m})\delta_{ij} \tag{A.24}$$

$$\mathrm{Tr}\left[\mathbb{H}^{-1}\partial_\nu\mathbb{H}\right] = \mathrm{Tr}\left[\left(\frac{1}{(1+\alpha)}\delta_{\nu,\alpha}\mathbb{1}^{QxQ}\right) + \left(\frac{1}{(1-m)}\delta_{\nu,m}\mathbb{1}^{QxQ}\right)\right] = Q\left(\frac{1}{(1+\alpha)}\delta_{\nu,\alpha} + \frac{1}{(1-m)}\delta_{\nu,m}\right) \tag{A.25}$$

$$\left(\mathbb{K}_{\mu,\nu}\right)_{ij} = \partial_\mu\vec{x}_i^\star\partial_\nu\vec{x}_j^\star + i \leftrightarrow j = 2\partial_\mu\vec{x}^\star\partial_\nu\vec{x}^\star J^Q \tag{A.26}$$

$$\mathrm{Tr}[\mathbb{H}K_{\mu,\nu}] = \mathrm{Tr}[-4H(1+\alpha)(1-m)\partial_\mu\vec{x}^\star\partial_\nu\vec{x}^\star\mathbb{1}^{QxQ}] = -4H\cdot Q(1+\alpha)(1-m)\delta_{\mu,\rho_0}\delta_{\nu,\rho_0} \tag{A.27}$$

Now we can write with $\Phi = (\alpha, m, \rho_0)$

$$\mathscr{H}^0 = \frac{M^2Q^2}{4}\begin{pmatrix} \frac{1}{(1+\alpha)^2} & \frac{1}{(1+\alpha)(1-m)} & 0 \\ \frac{1}{(1+\alpha)(1-m)} & \frac{1}{(1-m)^2} & 0 \\ 0 & 0 & 0 \end{pmatrix} \tag{A.28}$$

$$\mathscr{H}^{1,1} = \frac{-MQ^2}{2}\begin{pmatrix} \frac{1}{(1+\alpha)^2} & \frac{1}{(1+\alpha)(1-m)} & 0 \\ \frac{1}{(1+\alpha)(1-m)} & \frac{1}{(1-m)^2} & 0 \\ 0 & 0 & 0 \end{pmatrix} \tag{A.29}$$

$$\mathscr{H}^{1,2} = -2MHQ(1+\alpha)(1-m)\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{A.30}$$

$$\mathscr{H} = \begin{pmatrix} A & B & 0 \\ B & C & 0 \\ 0 & 0 & D \end{pmatrix} \tag{A.31}$$

$$\tag{A.32}$$

With $A = \frac{1}{(1+\alpha)^2}\frac{MQ^2}{4}(M-2)$, $B = \frac{1}{(1+\alpha)(1-m)}\frac{MQ^2}{4}(M-2)$, $C = \frac{1}{(1-m)^2}\frac{MQ^2}{4}(M-2)$ and $D = -2MHQ(1+\alpha)(1-m)$. Hence we find for the eigenvalues and eigenvectors (with $B^2 = AC$ [1])

1: Note that in general $\mathscr{H}_{k,k}\mathscr{H}_{l,l} = \mathscr{H}^2_{k,l} = \mathscr{H}^2_{l,k}$

$$\lambda_1 = \frac{1}{2}\left(\sqrt{A^2 - 2AC + 4B^2 + C^2} + A + C\right) \tag{A.33}$$
$$= A + C \quad (\propto \mathcal{O}(M^2)) \tag{A.34}$$
$$\vec{v}_1 = \left(-\frac{-A + C - \sqrt{A^2 - 2AC + 4B^2 + C^2}}{2B}, 1, 0\right)^T \tag{A.35}$$
$$= \left(\frac{A}{B}, 1, 0\right)^T \tag{A.36}$$
$$\lambda_2 = \frac{1}{2}\left(-\sqrt{A^2 - 2AC + 4B^2 + C^2} + A + C\right) \tag{A.37}$$
$$= 0 \tag{A.38}$$
$$\vec{v}_2 = \left(-\frac{-A + C + \sqrt{A^2 - 2AC + 4B^2 + C^2}}{2B}, 1, 0\right)^T \tag{A.39}$$
$$= \left(-\frac{C}{B}, 1, 0\right)^T \tag{A.40}$$
$$\lambda_3 = D \quad (\propto \mathcal{O}(M)) \tag{A.41}$$
$$\vec{v}_3 = (0, 0, 1)^T \tag{A.42}$$

So we find that due to the structure of equation A.32, the first eigenvalue is of order M^2 to a eigenvector that is spanned in the α-m-space, the second eigenvalue is 0 and the third eigenvalue of order M to the eigenvector that is aligned with the ρ_0 parameter axis.

A.2. Extension KL divergence at the boundaries

In the case where the most probable state is at the boundary, the first term in expansion 4.10 does not vanish. So we write

$$-\frac{1}{M}\log p(\vec{x}; \Phi) = E(\vec{x}; \Phi) \tag{A.43}$$
$$\approx E(\vec{x}^\star; \Phi) + \vec{V}(\Phi) \cdot \delta\vec{x} + \frac{1}{2}\delta\vec{x}\,\mathbb{H}(\Phi)\,\delta\vec{x} + \frac{1}{M}Z(\Phi) \tag{A.44}$$

with the vector $\vec{V}$ locally normal to the manifold at point $\vec{x}^\star$ and the constraint that $\vec{V}(\Phi)\cdot\delta\vec{x} \geq 0$. This implies that in the direction of $\vec{V}$ we we have the constraint and for directions perpendicular to $\vec{V}$ our problem is unconstrained, which means that in these directions the linear term vanishes and we need to keep the second order term. Then, we can calculate the partition function by rotating with a rotation matrix R with $\det R = 1$ such that y_0 is the direction of the constraint and all

other directions $y_1, ..., y_N$ are unconstrained.

$$Z(\Phi) = \int \mathrm{d}\vec{x}\; e^{-\left(E(x^\star;\Phi)+\vec{V}(\Phi)\delta\vec{x}+\frac{1}{2}\delta\vec{x}\,\mathbb{H}(\Phi)\,\delta\vec{x}\right)}\;\Theta(\vec{V}(\Phi)\delta\vec{x}) \tag{A.45}$$

$$= C\int \mathrm{d}y_0 \int \mathrm{d}y_1, ..., \mathrm{d}y_N\; e^{-\left(\vec{V}(\Phi)\delta\vec{y}+\frac{1}{2}\delta\vec{y}\,\mathbb{H}(\Phi)\,\delta\vec{y}\right)}\;\Theta(\vec{V}(\Phi)\delta y_0) \tag{A.46}$$

$$= C\int_0^\infty \mathrm{d}y_0 \int_\infty^\infty \mathrm{d}y_1, ..., \mathrm{d}y_N\; e^{-\left(\vec{V}(\Phi)\delta\vec{y}+\frac{1}{2}\delta\vec{y}\,\mathbb{H}(\Phi)\,\delta\vec{y}\right)} \tag{A.47}$$

$$= C'\int_0^\infty \mathrm{d}y_0 \int_\infty^\infty \mathrm{d}y_1, ..., \mathrm{d}y_N\; e^{-\frac{1}{2}(\vec{y}-\vec{\mu})^T\mathbb{H}(\Phi)(\vec{y}-\vec{\mu})} \tag{A.48}$$

$$= C'\int_0^\infty \mathrm{d}y_0 \int_\infty^\infty \mathrm{d}y_1, ..., \mathrm{d}y_N\; \mathcal{N}(\vec{y},\vec{\mu},\mathbb{H}) \tag{A.49}$$

Where we completed the squares in eq. A.48 and with the multivariate gaussian distribution $\mathcal{N}(\vec{y},\vec{\mu},\mathbb{H})$ with mean $\vec{\mu}$, $\vec{\mu} = V\mathbb{H}^{-1} - \vec{x}^\star$ covariance matrix $\mathbb{H} = \sum^{-1}$.

A.3. Table of Parameters

	Parameter description	Value
R_0	Ratio of hiring-firing rate (η^+/η^-):	2.0
Θ	Maximum credit supply available to firms :	3.2
Γ_0	Financial Fragility sensitivity:	0.0
$\rho^\star$	Baseline interest rate:	0.001*
α_c	Influence of deposit rates on consumption	12
ϕ_π	Intensity of interest rate policy of Central Bank:	0.0*
π^*	Central Bank inflation target:	0.0*
τ^T	Inflation target trust parameter:	0.0*
g_w	Factor to adjust wages to inflation expectations:	0.8
g_p	Factor to adjust prices to inflation expectations:	0.8
y_0	Initial production:	0.7
γ	Parameter to set adjustment of prices:	0.01
η_0^-	Firing propensity:	0.2
α_Γ	Influence of loans interest rate on hiring-firing policy:	450
c_0	Fraction of savings in consumption budget:	0.5
β	Price sensitivity parameter:	2
ϕ	Revival probability per unit time:	0.1
ω	Moving average parameter:	0.2
δ	Dividend rate:	0.02
δ_e	Fraction of energy sector's equity redistributed:	0.04
g_e	Share of Energy Price share in GDP:	0.0325
f	Interpolation parameter of cost of default:	0.5
μ	Easy-Credit policy multiplier:	1.3
N_F	Total number of firms:	10000
N	Total number of households:	10000

Table A.1.: Inactive Central Bank parameter set for Mark-0 Model (see Section 3.3) . Parameters, which change for the Reactive Central Bank and Anchored/Floating Expectation scenarios, can be taken from sections 22 and 22 respectively and are indicated in the table with a star. All rates are given in monthly time scales.

A.4. Counterfactuals

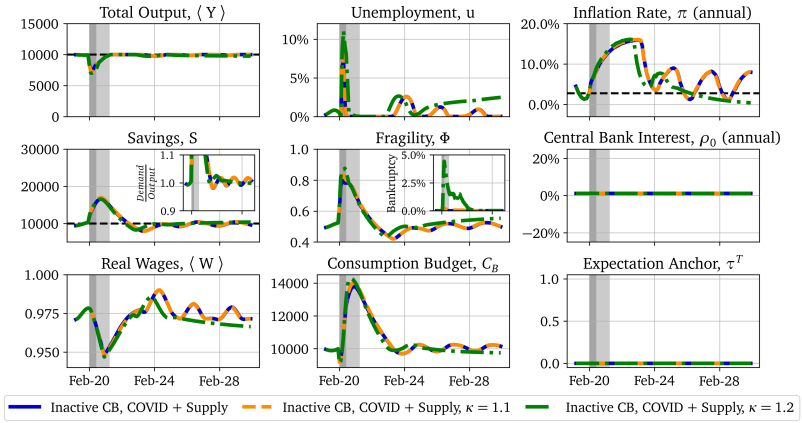


Figure A.1.: Economic dashboard for the Inactive Central Bank scenario without any policy and with a stronger COVID shock: The dynamics for a COVID shock strength of $\kappa = 1.0$ (blue), $\kappa = 1.1$ (orange) and $\kappa = 1.2$ (green) in the Inactive Central Bank scenario without any policy. Once the COVID shock gets too strong, the economy collapses in the long run.

A.5. Monetary Policy

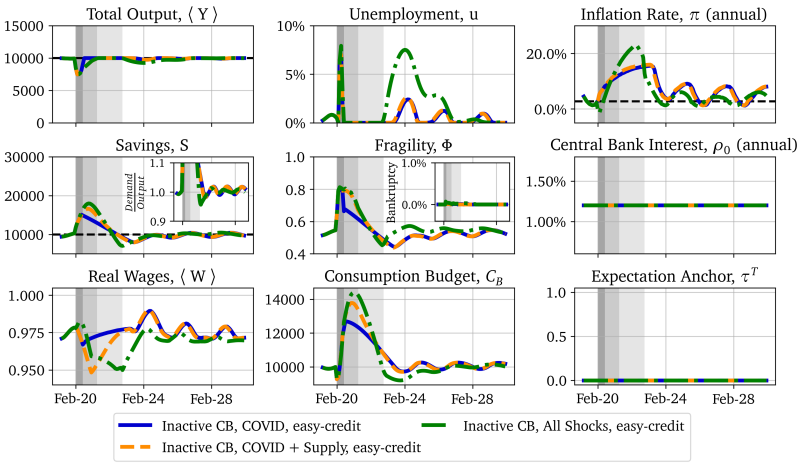


Figure A.2.: Dynamics for the three shocks for the Inactive Central Bank scenario: The dynamics for the three shocks, COVID only (blue), COVID and Supply Chain shock (orange) and all shocks (green) for the Inactive Central Bank scenario with Easy-Credit policy. The areas shaded in grey indicate the duration of the three shocks: the COVID shock lasting until the end of the dark grey area, the supply chain shock until the end of the grey area, and the price shock until the end of the light grey area.

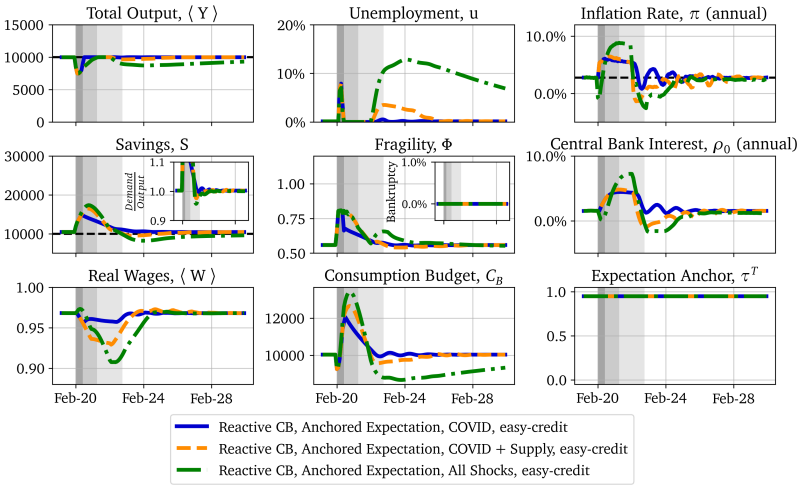


Figure A.3.: Dynamics for the three shocks for the Reactive Central Bank with Anchored Expectation scenario: The dynamics for the three shocks, COVID only (blue), COVID and Supply Chain shock (orange) and all shocks (green) to the scenario with reactive Central Bank and Anchored Expectation of economic agents with an Easy-Credit policy. The areas shaded in grey indicate the duration of the three shocks: the COVID shock lasting until the end of the dark grey area, the supply chain shock until the end of the grey area, and the price shock until the end of the light grey area.

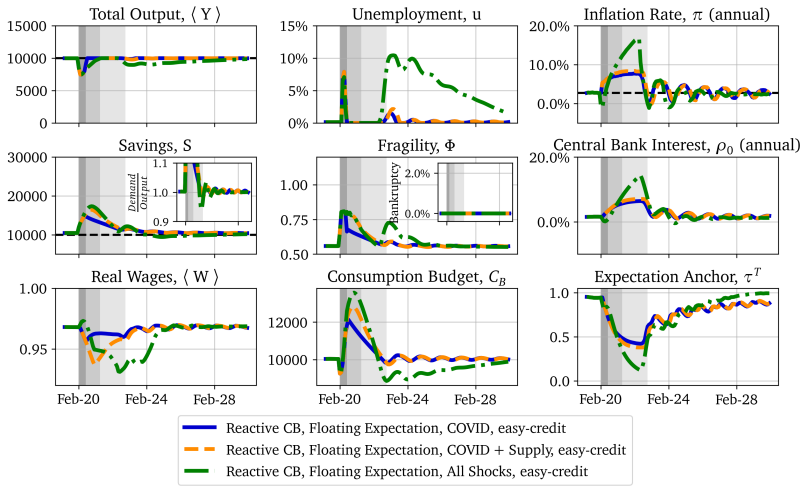


Figure A.4.: Dynamics for the three shocks for the Reactive Central Bank with Floating Expectation scenario: The dynamics for the three shocks, COVID only (blue), COVID and Supply Chain shock (orange) and all shocks (green) to the scenario with reactive Central Bank and Floating Expectation of economic agents with an Easy-Credit policy. The areas shaded in grey indicate the duration of the three shocks: the COVID shock lasting until the end of the dark grey area, the supply chain shock until the end of the grey area, and the price shock until the end of the light grey area.



Figure A.5.: Dynamics for all scenarios and all shocks: The full dashboard for Figure 3.6 All dynamics are with Easy-Credit policy in the Inactive Central Bank scenario (blue), Reactive Central Bank with Anchored Expectation scenario (orange) and Reactive Central Bank with Floating Expectation scenario (green). The areas shaded in grey indicate the duration of the three shocks: the COVID shock lasting until the end of the dark grey area, the supply chain shock until the end of the grey area, and the price shock until the end of the light grey area.

A.6. Importance of Expectation Anchoring

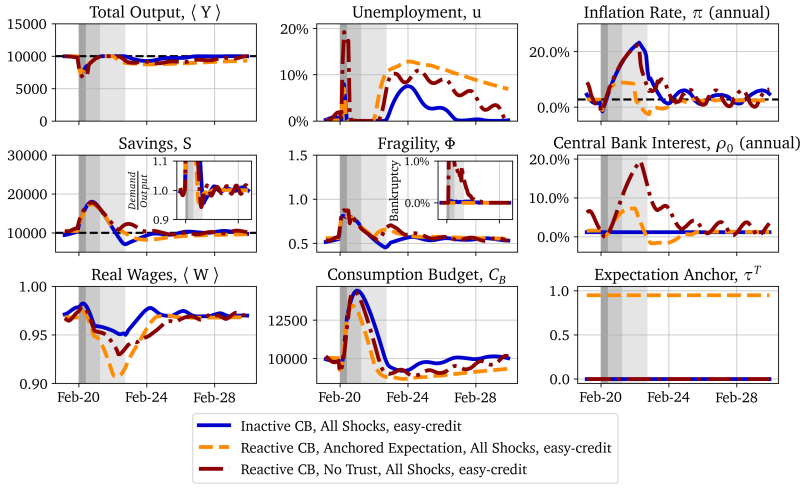


Figure A.6.: Dynamics with de-Anchored Expectation ($\tau^T = 0$): All dynamics are with Easy-Credit policy and all shocks, including the Inactive Central Bank scenario (blue), Reactive Central Bank with Anchored Expectation scenario (orange dashed) and Reactive Central Bank with no trust scenario $\tau^T = 0$ (red dash-dotted). The areas shaded in grey indicate the duration of the three shocks: the COVID shock lasting until the end of the dark grey area, the supply chain shock until the end of the grey area, and the price shock until the end of the light grey area. Note that without expectation anchoring, the inflation path is very similar with (red dashed-dotted) and without (blue) monetary policy.

A.7. Monetary Policy with Stronger Central Bank Reaction

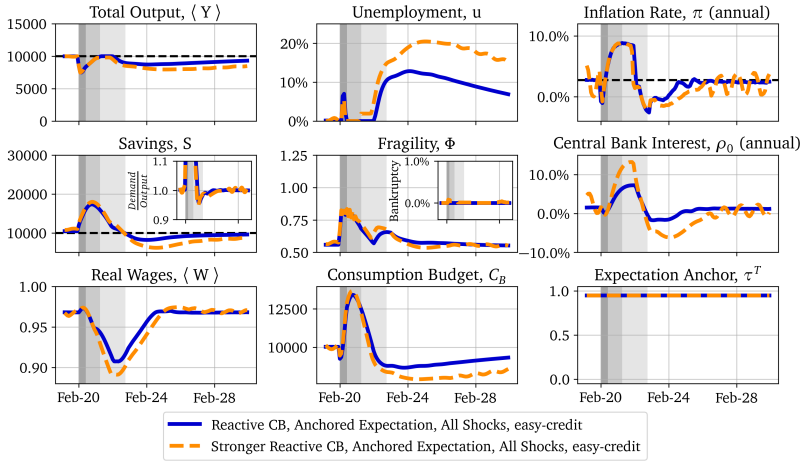


Figure A.7.: Economic dashboard for the stronger Reactive Central Bank with Anchored Expectation scenario and all shocks. The dynamics for a Central Bank strength of $\phi_\pi = 1.0$ (blue) and $\phi_\pi = 2.0$ (orange). A stronger central bank is not able to decrease inflation due to an external price shock, but with stronger policies, the consumption of households reduces, which leads to an increase in unemployment.

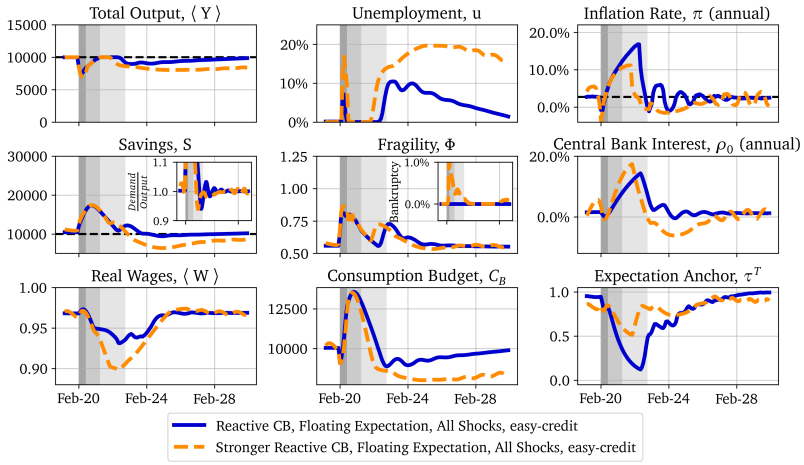


Figure A.8.: Economic dashboard for the stronger Reactive Central Bank with Floating Expectation scenario and all shocks. The dynamics for a Central Bank strength of $\phi_\pi = 1.0$ (blue) and $\phi_\pi = 2.0$ (orange) with Floating Expectation of economic agents. A stronger Central bank can reduce excess inflation at the cost of increased unemployment. The central bank must maintain a balance between falling inflation and rising unemployment.

A.8. Sensitivity of Monetary Policy to α_c and α_Γ

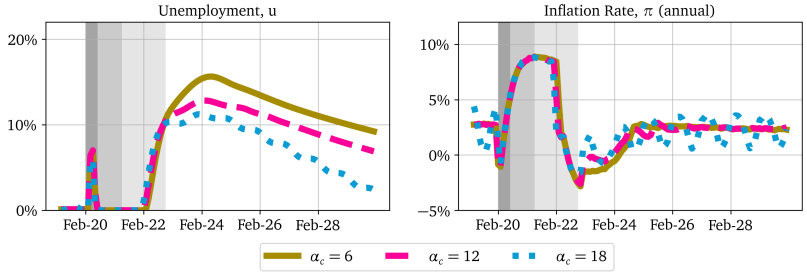


Figure A.9.: Sensitivity of α_c for the Reactive Central Bank with Anchored Expectation scenario and all shocks. As already described in [115], larger α_c lead to a greater magnification of price trends (cet. par.). This amplification, in turn, causes the fluctuations in unemployment and inflation. Furthermore, larger α_c increases consumption which reduces unemployment due to higher demand.

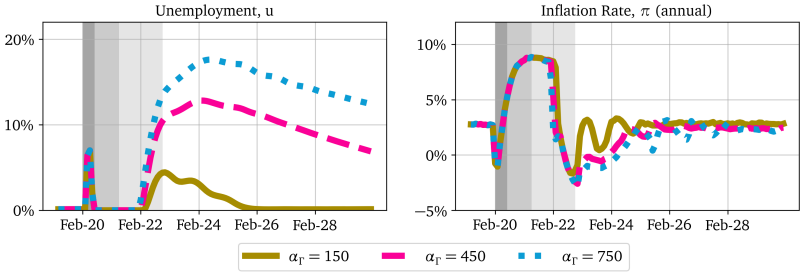


Figure A.10.: Sensitivity of α_Γ for the Reactive Central Bank with Anchored Expectation scenario and all shocks. Larger values of α_Γ increases the influence of financial fragility on the hiring/-firing policy of firms (cet. par.). This in turn leads to a larger downward adjustment of the workforce which increases unemployment.

A.9. Helicopter Money

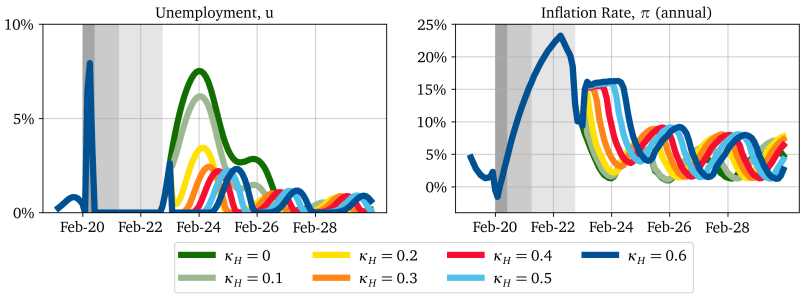


Figure A.11.: Helicopter Money in the Inactive Central Bank scenario and all shocks. Unemployment (left) and inflation (right) for a helicopter drop of size $\kappa_H S$ one month after the price shock. Already with $\kappa_H \geq 0.2$ unemployment is reduced almost to zero. A further increase of Helicopter Money only increases inflation without reducing unemployment more.

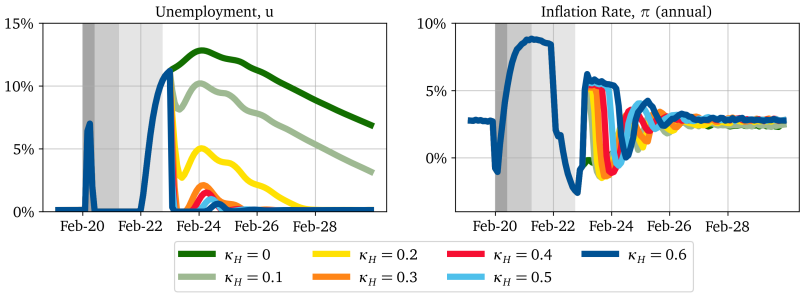


Figure A.12.: Helicopter Money in the Reactive Central Bank with Anchored Expectation scenario. Unemployment (left) and inflation (right) for a helicopter drop of size $\kappa_H S$ one month after the price shock. With $\kappa_H \geq 0.3$ unemployment is reduced almost to zero. A further increase of Helicopter Money only increases inflation without reducing unemployment further. However, the Central Bank policy manages to keep inflation under control.

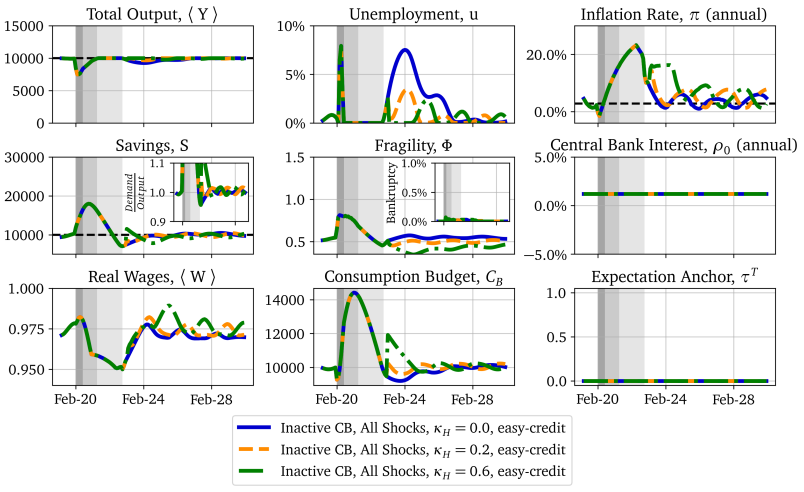


Figure A.13.: Economic dashboard for Helicopter Money in the Inactive Central Bank scenario and all shocks. Full dynamics for scenario without Helicopter Money (blue), for Helicopter Money with $\kappa_H = 0.2$ (orange) and $\kappa_H = 0.6$ (green) one month after the price shock and with Easy-Credit policy.

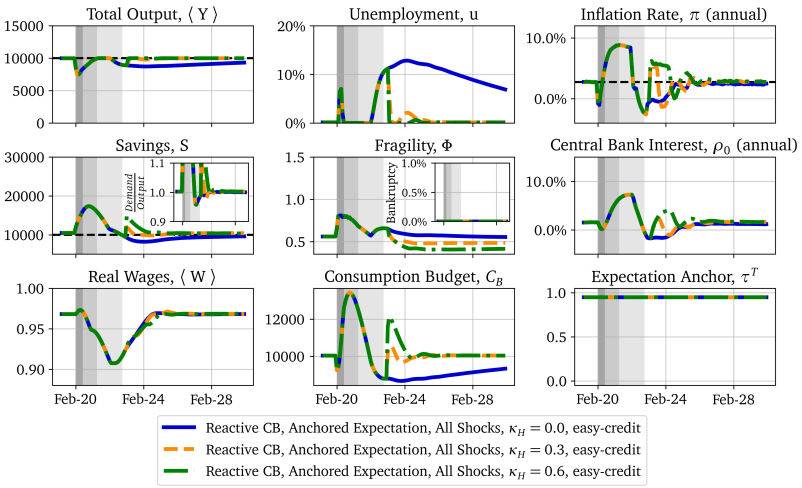


Figure A.14.: Economic dashboard for Helicopter Money in the Reactive Central Bank with Anchored Expectation scenario and all shocks. Full dynamics for scenario without Helicopter Money (blue), for Helicopter Money with $\kappa_H = 0.3$ (orange) and $\kappa_H = 0.6$ (green) one month after the price shock and with Easy-Credit policy.

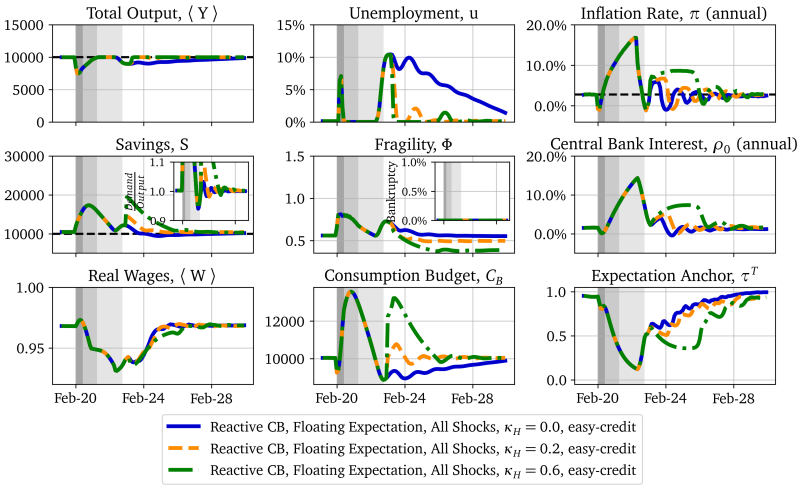


Figure A.15.: Economic dashboard for Helicopter Money in the Reactive Central Bank with Floating Expectation and all shocks. Full dynamics for scenario without Helicopter Money (blue), for Helicopter Money with $\kappa_H = 0.2$ (orange) and $\kappa_H = 0.6$ (green) one month after the price shock and with Easy-Credit policy.

A.10. Windfall Tax

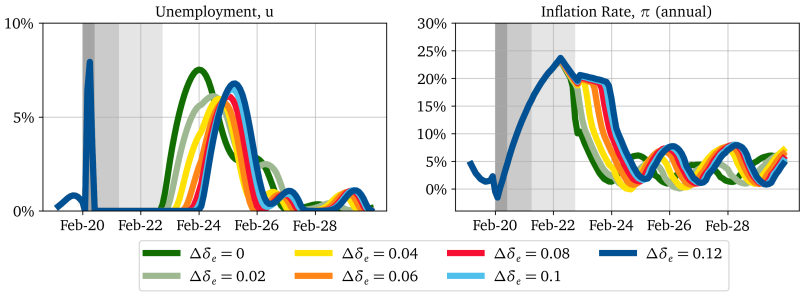


Figure A.16.: Windfall Tax in the Inactive Central Bank scenario. Unemployment (left) and inflation (right) for Windfall Tax of $\delta_e + \Delta\delta_e$ one year before the end of the price shock with a duration of two years. With $\Delta\delta_e \approx 4\%$ unemployment is reduced strongly. A further increase of tax does only increase unemployment again. For larger dividends, inflation increases because increased demand due to increased savings.

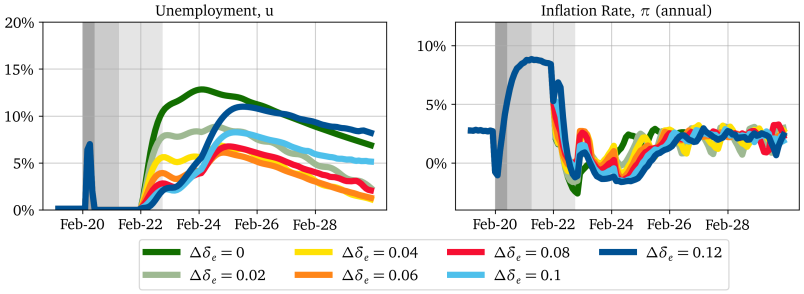


Figure A.17.: Windfall Tax in the Reactive CB with Anchored Expectation scenario. Unemployment (left) and inflation (right) for Windfall Tax of $\Delta\delta_e$ one year before the end of the price shock with a duration of two years. With $\Delta\delta_e \approx 6\%$ unemployment is reduced strongly. A further increase of $\Delta\delta_e$ leads to a significant resurgence of unemployment.

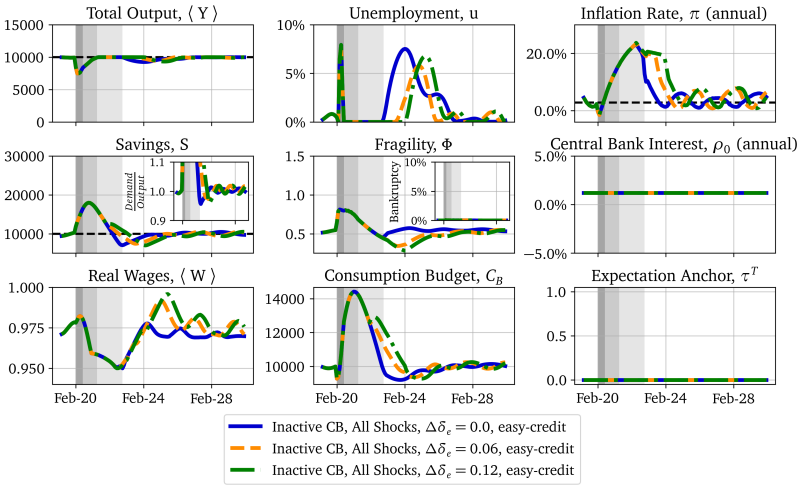


Figure A.18.: Economic dashboard for Windfall Tax in the Inactive Central Bank scenario, all shocks. Full dynamics for the Inactive Central Bank scenario without Windfall Tax (blue), for tax of $\Delta\delta_e = 6\%$ (orange) and $\Delta\delta_e = 12\%$ (green) one year before the end of the price shock with a duration of two years.

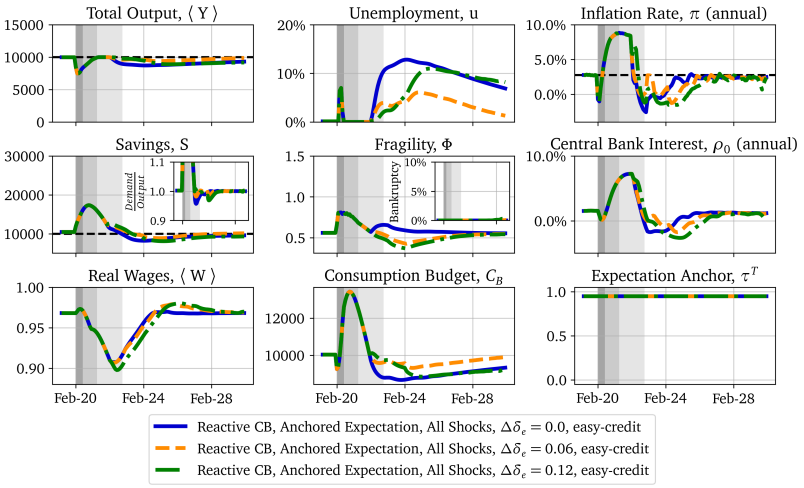


Figure A.19.: Economic dashboard for Windfall Tax in the Reactive Central Bank with Anchored Expectation scenario and all shocks. Full dynamics for scenario without Windfall Tax (blue), for tax of $\Delta\delta_e = 6\%$ (orange) and $\Delta\delta_e = 12\%$ (green) one year before the end of the price shock with a duration of two years.

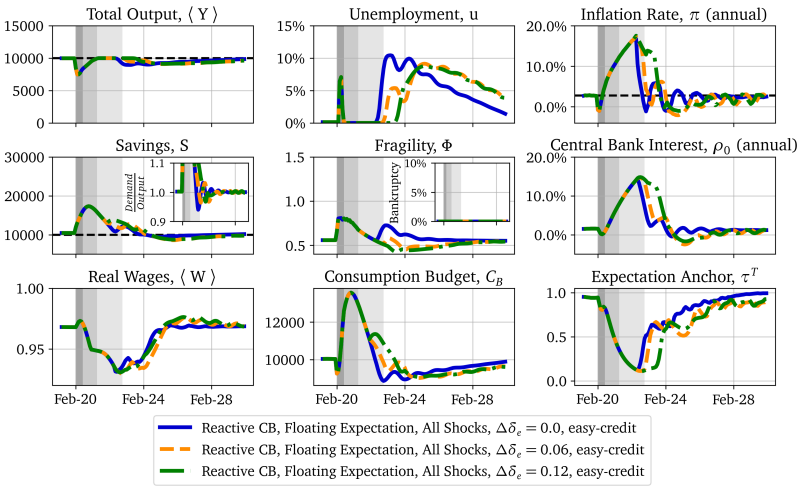


Figure A.20.: Economic dashboard for Windfall Tax in the Reactive Central Bank with Floating Expectation scenario and all shocks. Full dynamics for scenario without Windfall Tax (blue), for tax of $\Delta\delta_e = 6\%$ (orange) and $\Delta\delta_e = 12\%$ (green) one year before the end of the price shock with a duration of two years.

A.11. Sloppiness

For the sloppiness analysis we follow the same approach as in [2]. We define a mean-square loss error which is dependent on the set of parameters Φ and a small deviation δ. The loss function can be written as

$$\mathscr{L}(\Phi,\delta) = \frac{1}{2SKT}\sum_s\sum_k\sum_t\left(\frac{y_{s,k,t}(\Phi+\delta)-y_{s,k,t}(\Phi)}{\left\|y_{s,k}(\Phi)\right\|}\right)^2, \tag{A.50}$$

where $y_{s,k,t}(\Phi)$ is the realisation of output variable $k \in \{1,\dots,K\}$ at time $t \in \{1,\dots,T\}$ for random realisation $s \in \{1,\dots,S\}$. In this case, we use the output variables inflation and unemployment with 50 random realisations and we average over a period of 11 years that stars 1 year before the COVID shock in February 2020.

The Hessian matrix for the mean-squared loss at point Φ is then defined as

$$H^{\mathscr{L}}_{i,j}(\Phi) := \left.\frac{d^2\mathscr{L}(\Phi,\delta)}{d\log\Phi_i d\log\Phi_j}\right|_{\delta=0}, \tag{A.51}$$

$$H^{\mathscr{L}}_{i,j}(\Phi) = \frac{1}{SKT}\sum_s\sum_k\sum_t\frac{1}{\left\|y_{s,k,t}(\Phi)\right\|^2}\frac{dy_{s,k,t}(\Phi)}{d\log\Phi_i}\frac{dy_{s,k,t}(\Phi)}{d\log\Phi_j}, \tag{A.52}$$

Since the parameters in general have different order of magnitudes, we take the derivative with respect to the log parameters to only consider relative parameter changes. Note, that in Eq. A.52 the second derivative term vanishes as we evaluate $H^{\mathscr{L}}_{i,j}(\Phi)$ at $\delta = 0$.

The Hessian matrix can be decomposed into its eigenvalues and eigenvectors where the eigenvalues indicates the relative importance of the corresponding eigenvectors that represent the linear combination of parameters in parameter space. We are interested in the stiff directions, so the eigenvectors with largest eigenvalues that bring the biggest change to the dynamics of the system.

The eigenvalue spectrum spans several decades (Figure A.21), which

indicates the sloppiness of the model at that set of parameters. Based on the spectrum, we select the four largest eigenvectors to consider for our analysis (see Figures A.22, A.23, A.24).

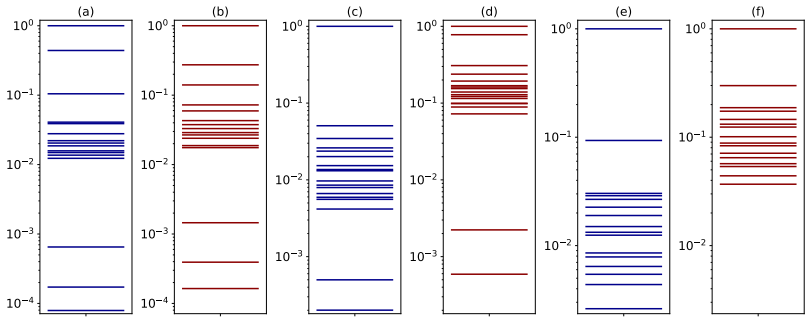


Figure A.21.: Eigenvalue Spectrum for three scenarios and all shocks. Sloppiness parameter $\epsilon = 0.01$, $n_{seeds} = 50$ target variable inflation (red) unemployment (blue), (a,b) Inactive Central Bank, (c,d) Reactive Central Bank with Anchored Expectation, (e,f) Reactive Central Bank with Floating Expectation.

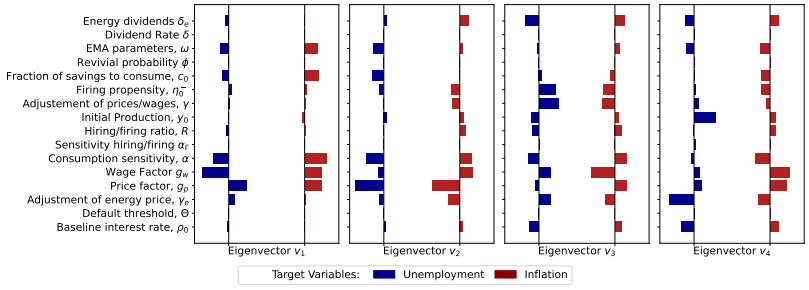


Figure A.22.: Eigenvectors of the Hessian matrix for the Inactive Central Bank Scenario and all shocks. The length of the eigenvectors is normed to 1 and the colors indicate the target variables, unemployment (blue) and inflation (red).

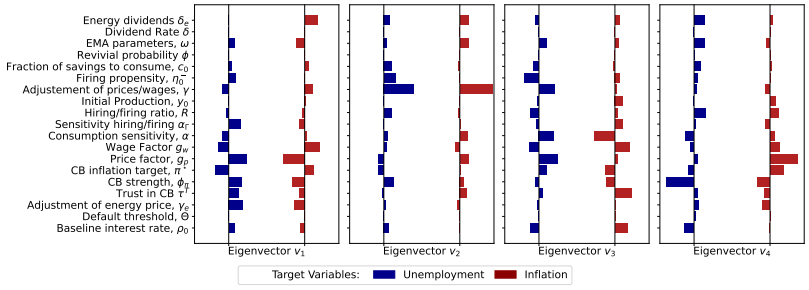


Figure A.23.: Eigenvectors of the Hessian matrix for the Reactive Central Bank with Anchored Expectation scenario and all shocks. The length of the eigenvectors is normed to 1 and the colors indicate the target variables, unemployment (blue) and inflation (red).

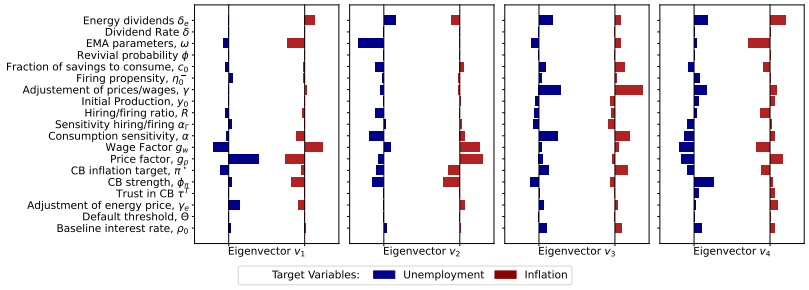


Figure A.24.: Eigenvectors of the Hessian matrix for the Reactive Central Bank with Floating Expectation scenario and all shocks. The length of the eigenvectors is normed to 1 and the colors indicate the target variables, unemployment (blue) and inflation (red).

A.12. Sensitivity to the Indexation Parameters g_p and g_w

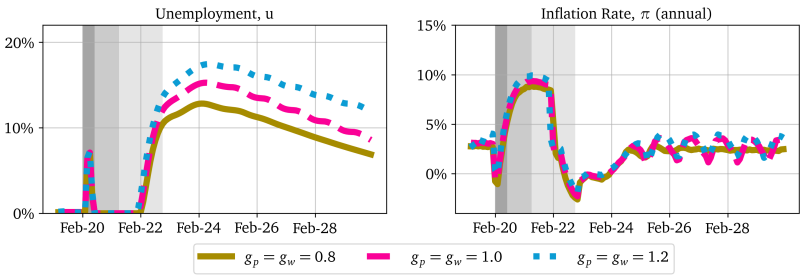


Figure A.25.: Sensitivity of $g_p = g_w = const$ for the Reactive Central Bank with Anchored Expectation scenario and all shocks. Increasing $g_p = g_w$ increases slightly inflation (cet. par.), causing slightly higher interest rates which leads to a reduction of consumption budget that reduces demand and therefore increases unemployment.

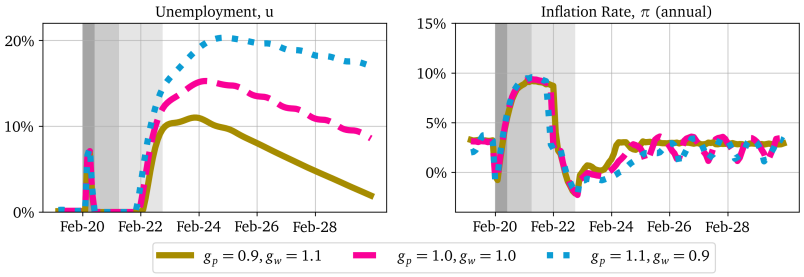


Figure A.26.: Sensitivity of $\frac{g_p + g_w}{2} = const$ for the reactive Central Bank with Anchored Expectation scenario and all shocks. (blue) When $g_w > g_w$, there is a higher bargaining power of worker which increases wages (cet. par.), therefore there is a higher consumption budget and higher demand which decreases unemployment. (yellow) For larger market power ($g_p > g_w$), there is the reverse effect of decreased wages causing a decrease in consumption budget and demand which therefore increases unemployment.

Part II B.

B.1. Cluster Methods & Labor Data

Skill Data and the ROME Classification

In addition to the annual declaration of social data, we utilize occupational skill data, specifically the Operational Directory of Trades and Jobs (Répertoire Opérationnel des Métiers et des Emplois, ROME). This directory, created as a reference system by the French Employment Center (Pôle emploi), categorizes occupations based on their skills and professional fields. The dataset encompasses skills ranging from know-how and hard skills to soft skills for over a thousand occupation labels. We convert these skills into vectors of approximately 800 dimensions, where each entry represents the extent to which a given skill is associated with that occupation. However, the ROME occupation classification comprises over 5000 occupations and differs from the 4-digit occupation classification PCS. Consequently, the mapping from ROME classification to PCS classification is not a one-to-one correspondence. In cases where a single PCS code is corresponds to multiple ROME occupations, the skill vector for that PCS occupation is calculated as the element-wise average of the skill vectors associated with the corresponding ROME occupations. Finally, as a distance metric between two PCS codes, we use a simple cosine similarity of the skill vectors for each occupation.

$$d_{i,j} = \frac{\mathbf{S}_i \cdot \mathbf{S}_j}{\|\mathbf{S}_i\| \left\|\mathbf{S}_j\right\|} \tag{B.1}$$

where $\mathbf{S}_k$ are the skill vectors of the PCS codes. In Section 7.5, we use the ratio and product of inter- and intra-community skill similarities, d_{Intra} and d_{Inter}. For a given occupation i, this metric calculates the average skill similarity to all occupations j within (intra) and outside (inter) occupation i's community.

$$d_{\text{Intra},i} = \langle \tilde{d}_{i,j} \rangle_{c^i = c^j} \tag{B.2}$$

$$d_{\text{Inter},i} = \langle \tilde{d}_{i,j} \rangle_{c^i \neq c^j} \tag{B.3}$$

where c^i is the BRIM community of occupation i.

Brim Algorithm

To identify the communities within our bipartite network, we employ the modularity maximization algorithm BRIM, which seeks to maximize the Barber modularity defined as follows Platig et al. (2016) [212], Barber (2007) [215]:

$$Q = \frac{1}{E} \sum_{i=1}^{N_R} \sum_{\alpha=1}^{N_C} (M_{i,\alpha} - P_{i,\alpha}) \delta(a_i, a_\alpha) \tag{B.4}$$

with the total number of links E, $M_{i,\alpha}$ the biadjacency matrix, $P_{i,\alpha}$ the probability that a link exists by chance and the membership variable that defines the block that node i belongs to a_i. This specific algorithm proved effective for the paper's objectives, particularly due to its consideration of the interdependence between communities in the different parts of the network, resulting in a well-defined partition for the labor flow network. Nonetheless, we also tested a partitioning by the BiLouvain algorithm to validate the robustness of our findings.

Nestedness

In this study, we employ the NODF (Nestedness based on Overlap and Decreasing Fill) measure, a widely utilized metric for nestedness analysis developed by Almeida-Neto et al. (2008) [265]. This metric is chosen due to its relevance to a structural property in complex networks, wherein the neighborhood of a given node is a subset of the neighborhoods of better-connected nodes. In the context of the fitness and complexity metric, our focus lies in ranking occupations based on their transferability and accessibility. It has been demonstrated that various social systems exhibit this hierarchical behavior in their rankings Ren et al. (2020) [241], Laudati et al. (2023) [242], and the nestedness metric serves as a quantitative measure for this phenomenon. The values span a range from 0 to 100, where 100 represents a fully nested network, and 0 reflects a non-nested pattern resembling a checkerboard like structure. This metric is based on two key properties: decreasing fill and paired overlap. Paired overlap is essentially the percentage of 1's in a given column (or row) that align with identical positions in another column. Then, NODF is the average of all paired values across columns (n) and rows (m):

$$\text{NODF} = \frac{\sum N_{\text{paired}}}{\left(\frac{n(n-1)}{2}\right) + \left(\frac{m(m-1)}{2}\right)} \tag{B.5}$$

In the realm of fitness and complexity research, where scholars examine import/export matrices of countries and their products, diverse values of the NODF measure are observed. For instance, the world trade network for Bovine exhibits a NODF value of 0.12, the typewriter network has a NODF value of 0.35, and on a company-product level, the NODF value is also 0.12 Ren et al. (2020) [241], Laudati et al. (2023) [242].

B.2. Temporal Stability

The structure of the occupation transitions in France has remained stable over the sample period of 2012-2020, despite events such as COVID-19. To test temporal stability we focus on link-stability over time, both in existence and magnitude of the transition rates. This can be captured in a simple regression comparing the current transition rate to that of 2012 as a baseline year:

$$\mathcal{P}_{i,j}^{t} = \alpha_{i,t} + \beta_{i,t}\mathcal{P}_{i,j}^{2012} + \epsilon_{i,t} \quad \forall t \in \{2013, \ldots, 2020\}, \tag{B.6}$$

such that the hypothesis of temporal stability can be captured by the joint test of $H_0 : \alpha_{i,T} = 0, \beta_{i,T} = 1$. That is, we expect the magnitude of transitions to be the same as 2012 without any general trend ($\alpha_{i,T}$).

Our findings indicate that the coefficient $\alpha_{i,t} \to 0$, and the coefficient $\beta_{i,t} \to 1$ remains stable over time (Figure B.1). This stability underscores the persistence of transition probabilities across years and allows as to combine all observations over the period of 2012-2020.

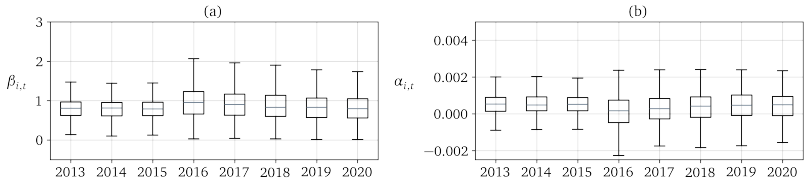


Figure B.1.: Regression Parameters of Linear Regression in eq. B.6. (a) Box plot of the coefficient $\beta_{i,t}$ for occupation i and year t. (b) Box plot of the intersection coefficient $\alpha_{i,t}$ for occupation i and year t.

B.3. Entry and Exit from the Occupation Network

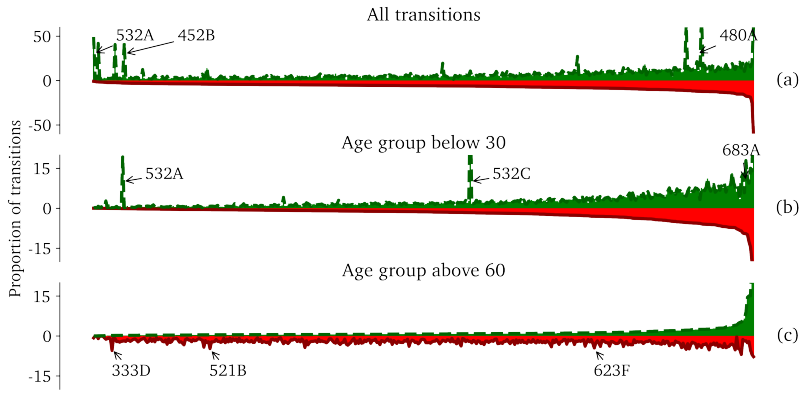


Figure B.2.: Entry and Exit Dynamics in the Labor Market: Distribution of inflows (green) and outflows (red) for each occupation (on the x-axis) over the sample period for (a) all transitions, (b) workers below 30 years old, and (c) workers above 60 years old. The y-axis represents the proportion of transitions relative to the total number of workers in each occupation. Throughout the sample period, there is a slightly higher influx of workers into the labor market, amounting to 2% more inflows. Among the age group below 30, the influx is 25% higher, while for the age group below 60, outflows exceed inflows by a factor of two.

Figure B.2 illustrates the proportion of workers transitioning into (green) and out of (red) specific occupations across different age groups. There is variability in the inflow and outflow across occupations. In panel (a), which depicts the entire network distribution, occupations like Gendarmes (532A) or Chief Warrant Officers (452B) exhibit significantly higher inflows than outflows, whereas Foremen and supervisors in agriculture, forestry (480A) experience higher outflows than inflows.

In Figure B.2 (b), displaying distributions only for workers under 30, occupations such as Apprentice bakers and butchers (683A) show more outflows than inflows, while Non-commissioned military officers have higher inflows than outflows. Overall, in the under-30 age group the inflow of workers is 25% higher than the outflow.

Conversely, for workers above 60, the outflow exceeds the inflow by a factor of two. The occupations with the highest outflow are for example Administrative executives of France Telecom (333D), France Telecom employees (521B) Qualified metal machining operators (623F).

B.4. BRIM Communities

Table 1: Occupations in the BRIM Community “Blue Collar Workers"

PCS-ESE	Occupation Name in French
477D	Techniciens de l’environnement et du traitement des pollutions
623F	Opérateurs qualifiés d’usinage des métaux travaillant à l’unité ou en petite série, moulistes qualifiés
623E	Soudeurs manuels
477A	Techniciens de la logistique, du planning et de l’ordonnancement
621A	Chefs d’équipe du gros oeuvre et des travaux publics
621B	Ouvriers qualifiés du travail du béton
623D	Opérateurs qualifiés sur machine de soudage
465C	Photographes
476A	Assistants techniques, techniciens de l’imprimerie et de l’édition
475A	Techniciens de recherche-développement et des méthodes de production des industries de transformation
475B	Techniciens de production et de contrôle-qualité des industries de transformation
476B	Techniciens de l’industrie des matériaux souples, de l’ameublement et du bois
624B	Monteurs, metteurs au point très qualifiés d’ensembles mécaniques travaillant à l’unité ou en petite série
534A	Agents civils de sécurité et de surveillance
624C	Monteurs qualifiés d’ensembles mécaniques travaillant en moyenne ou en grande série
624D	Monteurs qualifiés en structures métalliques
623G	Opérateurs qualifiés d’usinage des métaux sur autres machines (sauf moulistes)
487B	Responsables du tri, de l’emballage, de l’expédition et autres responsables de la manutention
622E	Autres monteurs câbleurs en électronique
623A	Chaudronniers-tôliers industriels, opérateurs qualifiés du travail en forge, conducteurs qualifiés d’équipement de formage, traceurs qualifiés
477C	Techniciens d’installation et de maintenance des équipements non industriels (hors informatique et télécommunications)
621G	Mineurs de fond qualifiés et autres ouvriers qualifiés des industries d’extraction (carrières, pétrole, gaz...)
477B	Techniciens d’installation et de maintenance des équipements industriels (électriques, électromécaniques, mécaniques, hors informatique)
621E	Autres ouvriers qualifiés des travaux publics
621D	Ouvriers des travaux publics en installations électriques et de télécommunications
621C	Conducteurs qualifiés d’engins de chantiers du bâtiment et des travaux publics
481A	Conducteurs de travaux (non cadres)

Continued on next page

Table 1: Occupations in the BRIM Community “Blue Collar Workers" (Continued)

PCS-ESE	Occupation Name in French
481B	Chefs de chantier (non cadres)
484A	Agents de maîtrise en fabrication : agroalimentaire, chimie, plasturgie, pharmacie.
484B	Agents de maîtrise en fabrication : métallurgie, matériaux lourds et autres industries de transformation
485A	Agents de maîtrise et techniciens en production et distribution d’énergie, eau, chauffage
485B	Agents de maîtrise en fabrication des autres industries (imprimerie, matériaux souples, ameublement et bois)
486B	Agents de maîtrise en maintenance, installation en électricité et électronique
474B	Techniciens de recherche-développement et des méthodes de fabrication en construction mécanique et travail des métaux
474A	Dessinateurs en construction mécanique et travail des métaux
473C	Techniciens de fabrication et de contrôle-qualité en électricité, électromécanique et électronique
473B	Techniciens de recherche-développement et des méthodes de fabrication en électricité, électromécanique et électronique
486D	Agents de maîtrise en maintenance, installation en mécanique
622A	Opérateurs qualifiés sur machines automatiques en production électrique ou électronique
472C	Métreurs et techniciens divers du bâtiment et des travaux publics
486E	Agents de maîtrise en entretien général, installation, travaux neufs (hors mécanique, électromécanique, électronique)
487A	Responsables d’entrepôt, de magasinage
622C	Monteurs câbleurs qualifiés en électricité
623B	Tuyauteurs industriels qualifiés
474C	Techniciens de fabrication et de contrôle-qualité en construction mécanique et travail des métaux
479B	Experts salariés de niveau technicien, techniciens divers
624F	Ouvriers qualifiés des traitements thermiques et de surface sur métaux
673B	Ouvriers de production non qualifiés travaillant par formage de métal
673A	Ouvriers de production non qualifiés travaillant par enlèvement de métal
672A	Ouvriers non qualifiés de l’électricité et de l’électronique
671C	Ouvriers non qualifiés des travaux publics et du travail du béton
655A	Autres agents et ouvriers qualifiés (sédentaires) des services d’exploitation des transports
653A	Magasiniers qualifiés
673C	Ouvriers non qualifiés de montage, contrôle en mécanique et travail des métaux

Continued on next page

Table 1: Occupations in the BRIM Community “Blue Collar Workers" (Continued)

PCS-ESE	Occupation Name in French
652A	Ouvriers qualifiés de la manutention, conducteurs de chariots élévateurs, caristes
651A	Conducteurs d’engin lourd de levage
644A	Conducteurs de véhicule de ramassage des ordures ménagères
624E	Ouvriers qualifiés de contrôle et d’essais en mécanique
641A	Conducteurs routiers et grands routiers
637D	Ouvriers qualifiés divers de type artisanal
636C	Boulangers, pâtissiers (sauf activité industrielle)
651B	Conducteurs d’engin lourd de manoeuvre
674A	Ouvriers de production non qualifiés : chimie, pharmacie, plasturgie
674B	Ouvriers de production non qualifiés de la transformation des viandes
674C	Autres ouvriers de production non qualifiés : industrie agroalimentaire
685A	Ouvriers non qualifiés divers de type artisanal
684B	Ouvriers non qualifiés de l’assainissement et du traitement des déchets
684A	Nettoyeurs
682A	Métalliers, serruriers, réparateurs en mécanique non qualifiés
681B	Ouvriers non qualifiés du second oeuvre du bâtiment
681A	Ouvriers non qualifiés du gros oeuvre du bâtiment
676E	Ouvriers non qualifiés divers de type industriel
676D	Agents non qualifiés des services d’exploitation des transports
676C	Ouvriers du tri, de l’emballage, de l’expédition, non qualifiés
676A	Manutentionnaires non qualifiés
675C	Ouvriers de production non qualifiés de l’imprimerie, presse, édition
675B	Ouvriers de production non qualifiés du travail du bois et de l’ameublement
675A	Ouvriers de production non qualifiés du textile et de la confection, de la tannerie-mégisserie et du travail du cuir
674E	Ouvriers de production non qualifiés : industrie lourde du bois, fabrication des papiers et cartons
674D	Ouvriers de production non qualifiés : métallurgie, production verrière, céramique, matériaux de construction
635A	Tailleurs et couturières qualifiés, ouvriers qualifiés du travail des étoffes (sauf fabrication de vêtements), ouvriers qualifiés de type artisanal du travail du cuir
634D	Mécaniciens qualifiés de maintenance, entretien : équipements non industriels
628G	Ouvriers qualifiés divers de type industriel
634B	Métalliers, serruriers qualifiés
634C	Mécaniciens qualifiés en maintenance, entretien, réparation : automobile
628A	Mécaniciens qualifiés de maintenance, entretien : équipements industriels

Continued on next page

Table 1: Occupations in the BRIM Community “Blue Collar Workers" (Continued)

PCS-ESE	Occupation Name in French
627F	Ouvriers de la composition et de l'impression, ouvriers qualifiés de la brochure, de la reliure et du façonnage du papier-carton
627D	Ouvriers qualifiés de scierie, de la menuiserie industrielle et de l'ameublement
627C	Ouvriers qualifiés du travail industriel du cuir
627B	Ouvriers qualifiés de la coupe des vêtements et de l'habillement, autres opérateurs de confection qualifiés
627A	Opérateurs qualifiés du textile et de la mégisserie
626C	Opérateurs et ouvriers qualifiés des industries lourdes du bois et de la fabrication du papier-carton
628C	Régleurs qualifiés d'équipements de fabrication (travail des métaux, mécanique)
626B	Autres opérateurs et ouvriers qualifiés : métallurgie, production verrière, matériaux de construction
625H	Ouvriers qualifiés des autres industries (eau, gaz, énergie, chauffage)
625G	Autres ouvriers de production qualifiés ne travaillant pas sur machine : industrie agroalimentaire (hors transformation des viandes)
625F	Autres opérateurs travaillant sur installations ou machines : industrie agroalimentaire (hors transformation des viandes)
625D	Opérateurs de la transformation des viandes
625C	Autres opérateurs et ouvriers qualifiés de la chimie (y.c. pharmacie) et de la plasturgie
625B	Ouvriers qualifiés et agents qualifiés de laboratoire : agroalimentaire, chimie, biologie, pharmacie
625A	Pilotes d'installation lourde des industries de transformation : agroalimentaire, chimie, plasturgie, énergie
624G	Autres mécaniciens ou ajusteurs qualifiés (ou spécialité non reconnue)
626A	Pilotes d'installation lourde des industries de transformation : métallurgie, production verrière, matériaux de construction
628D	Régleurs qualifiés d'équipements de fabrication (hors travail des métaux et mécanique)
628B	Electromécaniciens, électriciens qualifiés d'entretien : équipements industriels
632A	Maçons qualifiés
632C	Charpentiers en bois qualifiés
634A	Carrossiers d'automobiles qualifiés
633D	Electriciens, électroniciens qualifiés en maintenance, entretien : équipements non industriels
632D	Menuisiers qualifiés du bâtiment
633B	Dépanneurs qualifiés en radiotélévision, électroménager, matériel électronique (salariés)

Continued on next page

Table 1: Occupations in the BRIM Community “Blue Collar Workers" (Continued)

PCS-ESE	Occupation Name in French
632E	Couvreurs qualifiés
632F	Plombiers et chauffagistes qualifiés
633A	Electriciens qualifiés de type artisanal (y.c. bâtiment)
632K	Ouvriers qualifiés d'entretien général des bâtiments
628E	Ouvriers qualifiés de l'assainissement et du traitement des déchets
632G	Peintres et ouvriers qualifiés de pose de revêtements sur supports verticaux
632H	Soliers moquetteurs et ouvriers qualifiés de pose de revêtements souples sur supports horizontaux
632J	Monteurs qualifiés en agencement, isolation

Table 2: Occupations in the BRIM Community “Engineers & Executives"

PCS-ESE	Occupation Name in French
542A	Secrétaires
387B	Ingénieurs et cadres de la logistique, du planning et de l'ordonnancement
387A	Ingénieurs et cadres des achats et approvisionnements industriels
386E	Ingénieurs et cadres de fabrication des autres industries (imprimerie, matériaux souples, ameublement et bois)
386D	Ingénieurs et cadres de la production et de la distribution d'énergie, eau
386C	Ingénieurs et cadres d'étude, recherche et développement des autres industries (imprimerie, matériaux souples, ameublement et bois)
386B	Ingénieurs et cadres d'étude, recherche et développement de la distribution d'énergie, eau
385C	Ingénieurs et cadres technico-commerciaux des industries de transformations (biens intermédiaires)
385B	Ingénieurs et cadres de fabrication des industries de transformation (agroalimentaire, chimie, métallurgie, matériaux lourds)
385A	Ingénieurs et cadres d'étude, recherche et développement des industries de transformation (agroalimentaire, chimie, métallurgie, matériaux lourds)
384C	Ingénieurs et cadres technico-commerciaux en matériel mécanique professionnel
384B	Ingénieurs et cadres de fabrication en mécanique et travail des métaux
387C	Ingénieurs et cadres des méthodes de production
384A	Ingénieurs et cadres d'étude, recherche et développement en mécanique et travail des métaux
383B	Ingénieurs et cadres de fabrication en matériel électrique, électronique

Continued on next page

Table 2: Occupations in the BRIM Community "Engineers & Executives" (Continued)

PCS-ESE	Occupation Name in French
383A	Ingénieurs et cadres d'étude, recherche et développement en électricité, électronique
382D	Ingénieurs et cadres technico-commerciaux en bâtiment, travaux publics
382C	Ingénieurs, cadres de chantier et conducteurs de travaux (cadres) du bâtiment et des travaux publics
382B	Architectes salariés
461C	Secrétaires de niveau supérieur (non cadres, hors secrétaires de direction)
376G	Cadres de l'immobilier
376F	Cadres des services techniques des organismes de sécurité sociale et assimilés
376E	Cadres des services techniques des assurances
376D	Chefs d'établissements et responsables de l'exploitation bancaire
376C	Cadres commerciaux de la banque
383C	Ingénieurs et cadres technico-commerciaux en matériel électrique ou électronique professionnel
376B	Cadres des opérations bancaires
387D	Ingénieurs et cadres du contrôle-qualité
387F	Ingénieurs et cadres techniques de l'environnement
461D	Maîtrise et techniciens des services financiers ou comptables
461E	Maîtrise et techniciens administratifs des services juridiques ou du personnel
461F	Maîtrise et techniciens administratifs des autres services administratifs
462C	Acheteurs non classés cadres, aides-acheteurs
462E	Autres professions intermédiaires commerciales (sauf techniciens des forces de vente)
463B	Techniciens commerciaux et technico-commerciaux, représentants en biens d'équipement, en biens intermédiaires, commerce interindustriel (hors informatique)
463C	Techniciens commerciaux et technico-commerciaux, représentants en biens de consommation auprès d'entreprises
463D	Techniciens commerciaux et technico-commerciaux, représentants en services auprès d'entreprises ou de professionnels (hors banque, assurance, informatique)
463E	Techniciens commerciaux et technico-commerciaux, représentants auprès de particuliers (hors banque, assurance, informatique)
464A	Assistants de la publicité, des relations publiques
465A	Concepteurs et assistants techniques des arts graphiques, de la mode et de la décoration salariés
387E	Ingénieurs et cadres de la maintenance, de l'entretien et des travaux neufs
467A	Chargés de clientèle bancaire
467C	Professions intermédiaires techniques et commerciales des assurances

Continued on next page

Table 2: Occupations in the BRIM Community “Engineers & Executives" (Continued)

PCS-ESE	Occupation Name in French
543B	Employés qualifiés des services comptables ou financiers
467D	Professions intermédiaires techniques des organismes de sécurité sociale
472A	Dessinateurs en bâtiment, travaux publics
472B	Géomètres, topographes
473A	Dessinateurs en électricité, électromécanique et électronique
388E	Ingénieurs et cadres spécialistes des télécommunications
388D	Ingénieurs et cadres technico-commerciaux en informatique et télécommunications
388C	Chefs de projets informatiques, responsables informatiques
388B	Ingénieurs et cadres d’administration, maintenance, support et services aux utilisateurs en informatique
388A	Ingénieurs et cadres d’étude, recherche et développement en informatique
467B	Techniciens des opérations bancaires
376A	Cadres des marchés financiers
382A	Ingénieurs et cadres d’étude du bâtiment et des travaux publics
375A	Cadres de la publicité
354A	Artistes plasticiens
478D	Techniciens des télécommunications et de l’informatique des réseaux
353A	Directeurs de journaux, administrateurs de presse, directeurs d’éditions (littéraire, musicale, audiovisuelle et multimédia)
352A	Journalistes (y c. rédacteurs en chef)
375B	Cadres des relations publiques et de la communication
555A	Vendeurs par correspondance, télévendeurs
545D	Employés des services techniques des organismes de sécurité sociale et assimilés
545C	Employés des services techniques des assurances
545B	Employés des services commerciaux de la banque
545A	Employés administratifs des services techniques de la banque
544A	Employés et opérateurs d’exploitation en informatique
543H	Employés administratifs non qualifiés
543G	Employés administratifs qualifiés des autres services des entreprises
543F	Employés qualifiés des services commerciaux des entreprises (hors vente)
543E	Employés qualifiés des services du personnel et des services juridiques
231A	Chefs de grande entreprise de 500 salariés et plus
543C	Employés non qualifiés des services comptables ou financiers
478C	Techniciens d’installation, de maintenance, support et services aux utilisateurs en informatique
478B	Techniciens de production, d’exploitation en informatique

Continued on next page

Table 2: Occupations in the BRIM Community “Engineers & Executives" (Continued)

PCS-ESE	Occupation Name in French
461B	Secrétaires de direction, assistants de direction (non cadres)
374D	Cadres commerciaux des petites et moyennes entreprises (hors commerce de détail)
372A	Cadres chargés d'études économiques, financières, commerciales
372B	Cadres de l'organisation ou du contrôle des services administratifs et financiers
374C	Cadres commerciaux des grandes entreprises (hors commerce de détail)
372C	Cadres spécialistes des ressources humaines et du recrutement
374B	Chefs de produits, acheteurs du commerce et autres cadres de la mercatique
372E	Juristes
373B	Cadres des autres services administratifs des grandes entreprises
373A	Cadres des services financiers ou comptables des grandes entreprises
373D	Cadres des autres services administratifs des petites et moyennes entreprises
373C	Cadres des services financiers ou comptables des petites et moyennes entreprises
478A	Techniciens d'étude et de développement en informatique

Table 3: Occupations in the BRIM Community “Public, Health & Education"

PCS-ESE	Occupation Name in French
372F	Cadres de la documentation, de l'archivage (hors fonction publique)
525A	Agents de service des établissements primaires
524C	Agents administratifs des collectivités locales
524B	Agents administratifs de l'Etat et assimilés (sauf Poste, France Télécom)
523C	Adjoints administratifs des collectivités locales
523D	Adjoints administratifs des hôpitaux publics
523B	Adjoints administratifs de l'Etat et assimilés (sauf Poste, France Télécom)
422B	Professeurs de lycée professionnel
422A	Professeurs d'enseignement général des collèges
421B	Professeurs des écoles
372D	Cadres spécialistes de la formation
422C	Maîtres auxiliaires et professeurs contractuels de l'enseignement secondaire
351A	Bibliothécaires, archivistes, conservateurs et autres cadres du patrimoine
342B	Professeurs et maîtres de conférences
342G	Ingénieurs d'étude et de recherche de la recherche publique
526D	Aides médico-psychologiques

Continued on next page

Table 3: Occupations in the BRIM Community “Public, Health & Education" (Continued)

PCS-ESE	Occupation Name in French
526A	Aides-soignants
525D	Agents de service hospitaliers
525C	Agents de service de la fonction publique (sauf écoles, hôpitaux)
311D	Psychologues, psychanalystes, psychothérapeutes (non médecins)
331A	Personnels de direction de la fonction publique (Etat, collectivités locales, hôpitaux)
332A	Ingénieurs de l'Etat (y.c. ingénieurs militaires) et assimilés
332B	Ingénieurs des collectivités locales et des hôpitaux
333A	Magistrats*
525B	Agents de service des autres établissements d'enseignement
333E	Autres personnels administratifs de catégorie A de l'Etat (hors Enseignement, Patrimoine, Impôts, Trésor, Douanes)
333F	Personnels administratifs de catégorie A des collectivités locales et hôpitaux publics (hors Enseignement, Patrimoine)
341A	Professeurs agrégés et certifiés de l'enseignement secondaire
422E	Surveillants et aides-éducateurs des établissements d'enseignement
342C	Professeurs agrégés et certifiés en fonction dans l'enseignement supérieur
342D	Personnel enseignant temporaire de l'enseignement supérieur
342F	Directeurs et chargés de recherche de la recherche publique
343A	Psychologues spécialistes de l'orientation scolaire et professionnelle
423B	Formateurs et animateurs de formation continue
563A	Assistantes maternelles, gardiennes d'enfants, familles d'accueil
425A	Sous-bibliothécaires, cadres intermédiaires du patrimoine
424A	Moniteurs et éducateurs sportifs, sportifs professionnels
434G	Educateurs de jeunes enfants
435A	Directeurs de centres socioculturels et de loisirs
435B	Animateurs socioculturels et de loisirs
441A	Clergé séculier
441B	Clergé régulier
451C	Contrôleurs des Impôts, du Trésor, des Douanes et assimilés*
434F	Educateurs techniques spécialisés, moniteurs d'atelier
451E	Autres personnels administratifs de catégorie B de l'Etat (hors Enseignement, Patrimoine, Impôts, Trésor, Douanes)
451H	Professions intermédiaires administratives des hôpitaux
472D	Techniciens des travaux publics de l'Etat et des collectivités locales
479A	Techniciens des laboratoires de recherche publique ou de l'enseignement
621F	Ouvriers qualifiés des travaux publics (salariés de l'Etat et des collectivités locales)

Continued on next page

Table 3: Occupations in the BRIM Community “Public, Health & Education" (Continued)

PCS-ESE	Occupation Name in French
564B	Employés des services divers
563C	Employés de maison et personnels de ménage chez des particuliers
563B	Aides à domicile, aides ménagères, travailleuses familiales
451G	Professions intermédiaires administratives des collectivités locales
434E	Moniteurs éducateurs
422D	Conseillers principaux d'éducation
431F	Infirmiers en soins généraux, salariés
431A	Cadres infirmiers et assimilés
431B	Infirmiers psychiatriques
431C	Puéricultrices
431D	Infirmiers spécialisés (autres qu'infirmiers psychiatriques et puéricultrices)
628F	Agents qualifiés de laboratoire (sauf chimie, santé)
434C	Conseillers en économie sociale familiale
434D	Educateurs spécialisés
432B	Masseurs-kinésithérapeutes rééducateurs, salariés
432D	Autres spécialistes de la rééducation, salariés
434B	Assistants de service social
434A	Cadres de l'intervention socio-éducative
433A	Techniciens médicaux

Table 4: Occupations in the BRIM Community “Sales & Retail"

PCS-ESE	Occupation Name in French
433D	Préparateurs en pharmacie
468A	Maîtrise de restauration : salle et service
462D	Animateurs commerciaux des magasins de vente, marchandiseurs (non cadres)
462B	Maîtrise de l'exploitation des magasins de vente
462A	Chefs de petites surfaces de vente
468B	Maîtrise de l'hébergement : hall et étages
541D	Standardistes, téléphonistes
541C	Agents d'accueil non qualifiés
488A	Maîtrise de restauration : cuisine/production
636A	Bouchers (sauf industrie de la viande)
636B	Charcutiers (sauf industrie de la viande)
377A	Cadres de l'hôtellerie et de la restauration

Continued on next page

Table 4: Occupations in the BRIM Community “Sales & Retail" (Continued)

PCS-ESE	Occupation Name in French
636D	Cuisiniers et commis de cuisine
374A	Cadres de l’exploitation des magasins de vente du commerce de détail
233D	Chefs d’entreprise de services, de 10 à 49 salariés
233C	Chefs d’entreprise commerciale, de 10 à 49 salariés
233B	Chefs d’entreprise de l’industrie ou des transports, de 10 à 49 salariés
233A	Chefs d’entreprise du bâtiment et des travaux publics, de 10 à 49 salariés
232A	Chefs de moyenne entreprise, de 50 à 499 salariés
220X	Commerçants et assimilés, salariés de leur entreprise
433B	Opticiens lunetiers et audioprothésistes salariés
488B	Maîtrise de restauration : gestion d’établissement
627E	Ouvriers de la photogravure et des laboratoires photographiques et cinématographiques
210X	Artisans salariés de leur entreprise
554C	Vendeurs en droguerie, bazar, quincaillerie, bricolage
554B	Vendeurs en ameublement, décor, équipement du foyer
554A	Vendeurs en alimentation
554E	Vendeurs en habillement et articles de sport
554F	Vendeurs en produits de beauté, de luxe (hors biens culturels) et optique
554G	Vendeurs de biens culturels (livres, disques, multimédia, objets d’art)
553C	Autres vendeurs non spécialisés
554H	Vendeurs de tabac, presse et articles divers
556A	Vendeurs en gros de biens d’équipement, biens intermédiaires
561C	Serveurs, commis de restaurant, garçons non qualifiés
561D	Aides de cuisine, apprentis de cuisine et employés polyvalents de la restauration
561B	Serveurs, commis de restaurant, garçons qualifiés
562A	Manucures, esthéticiens
551A	Employés de libre service du commerce et magasiniers
552A	Caissiers de magasin
541B	Agents d’accueil qualifiés, hôtesses d’accueil et d’information
561E	Employés de l’hôtellerie : réception et hall
561F	Employés d’étage et employés polyvalents de l’hôtellerie
562B	Coiffeurs salariés

Table 5: Occupations in the BRIM Community “Transport & Logistic"

PCS-ESE	Occupation Name in French
642A	Conducteurs de taxi
642B	Conducteurs de voiture particulière
546A	Contrôleurs des transports (personnels roulants)
546B	Agents des services commerciaux des transports de voyageurs et du tourisme
546C	Employés administratifs d'exploitation des transports de marchandises
546D	Hôtesses de l'air et stewards
546E	Autres agents et hôtesses d'accompagnement (transports, tourisme)
654B	Conducteurs qualifiés d'engins de transport guidés (sauf remontées mécaniques)
656B	Matelots de la marine marchande
656C	Capitaines et matelots timoniers de la navigation fluviale
643A	Conducteurs livreurs, coursiers
641B	Conducteurs de véhicule routier de transport en commun
692A	Marins-pêcheurs et ouvriers de l'aquaculture
389C	Officiers et cadres navigants techniques de la marine marchande
389B	Officiers et cadres navigants techniques et commerciaux de l'aviation civile
389A	Ingénieurs et cadres techniques de l'exploitation des transports
451A	Professions intermédiaires de la Poste
466A	Responsables commerciaux et administratifs des transports de voyageurs et du tourisme (non cadres)
521A	Employés de la Poste
333C	Cadres de la Poste*
466C	Responsables d'exploitation des transports de voyageurs et de marchandises (non cadres)
526E	Ambulanciers salariés
480B	Maîtres d'équipage de la marine marchande et de la pêche
466B	Responsables commerciaux et administratifs des transports de marchandises (non cadres)

Table 6: Occupations in the BRIM Community “Agriculture"

PCS-ESE	Occupation Name in French
480A	Contremaîtres et agents d'encadrement (non cadres) en agriculture, sylviculture
691E	Ouvriers agricoles sans spécialisation particulière
691A	Conducteurs d'engin agricole ou forestier
631A	Jardiniers

Continued on next page

Table 6: Occupations in the BRIM Community “Agriculture" (Continued)

PCS-ESE	Occupation Name in French
381B	Ingénieurs et cadres d’étude et développement de l’agriculture, la pêche, les eaux et forêts
691B	Ouvriers de l’élevage
691C	Ouvriers du maraîchage ou de l’horticulture
691D	Ouvriers de la viticulture ou de l’arboriculture fruitière
381C	Ingénieurs et cadres de production et d’exploitation de l’agriculture, la pêche, les eaux et forêts
471A	Techniciens d’étude et de conseil en agriculture, eaux et forêt
471B	Techniciens d’exploitation et de contrôle de la production en agriculture, eaux et forêt

Table 7: Occupations in the BRIM Community “Artist and Creative Work"

PCS-ESE	Occupation Name in French
465B	Assistants techniques de la réalisation des spectacles vivants et audiovisuels salariés
354G	Professeurs d’art (hors établissements scolaires)
354E	Artistes de la danse
354C	Artistes dramatiques
637C	Ouvriers et techniciens des spectacles vivants et audiovisuels
353B	Directeurs, responsables de programmation et de production de l’audiovisuel et des spectacles
354F	Artistes du cirque et des spectacles divers
353C	Cadres artistiques et technico-artistiques de la réalisation de l’audiovisuel et des spectacles
354B	Artistes de la musique et du chant

Table 8: Occupations in the BRIM Community “Medical Service"

PCS-ESE	Occupation Name in French
311C	Chirurgiens dentistes
433C	Autres spécialistes de l’appareillage médical salariés
344D	Pharmaciens salariés
344C	Internes en médecine, odontologie et pharmacie
344A	Médecins hospitaliers sans activité libérale

Continued on next page

Table 8: Occupations in the BRIM Community “Medical Service" (Continued)

PCS-ESE	Occupation Name in French
526B	Assistants dentaires, médicaux et vétérinaires, aides de techniciens médicaux
344B	Médecins salariés non hospitaliers

B.5. Taxonomy Clusters

Table 1: Occupations in the cluster “Hub"

PCS-ESE	Occupation Name in French	T_i (log)	A_i (log)
389B	Officiers et cadres navigants techniques et commerciaux de l’aviation civile	3.36	-2.45
627E	Ouvriers de la photogravure et des laboratoires photographiques et cinématographiques	2.97	-1.79
431F	Infirmiers en soins généraux, salariés	2.86	0.39
623E	Soudeurs manuels	2.59	-0.29
465A	Concepteurs et assistants techniques des arts graphiques, de la mode et de la décoration salariés	2.39	-0.34
636A	Bouchers (sauf industrie de la viande)	1.86	-0.65
554G	Vendeurs de biens culturels (livres, disques, multimédia, objets d’art)	1.85	-0.7
381C	Ingénieurs et cadres de production et d’exploitation de l’agriculture, la pêche, les eaux et forêts	1.83	-1.38

Table 2: Occupations in the cluster “Diffuser"

PCS-ESE	Occupation Name in French	T_i (log)	A_i (log)
656B	Matelots de la marine marchande	4.22	-4.31
480B	Maîtres d’équipage de la marine marchande et de la pêche	4.15	-4.82
389C	Officiers et cadres navigants techniques de la marine marchande	4.14	-3.33
656C	Capitaines et matelots timoniers de la navigation fluviale	3.99	-4.35
441B	Clergé régulier	2.05	-2.63
546D	Hôtesses de l’air et stewards	1.86	-3.67
441A	Clergé séculier	1.85	-2.55

Table 3: Occupations in the cluster “Channel"

PCS-ESE	Occupation Name in French	T_i (log)	A_i (log)
692A	Marins-pêcheurs et ouvriers de l’aquaculture	1.69	-4.89
465C	Photographes	0.86	-4.05
623D	Opérateurs qualifiés sur machine de soudage	0.4	-3.75
354A	Artistes plasticiens	0.26	-3.74
636B	Charcutiers (sauf industrie de la viande)	0.19	-3.26
480A	Contremaîtres et agents d’encadrement (non cadres) en agriculture, sylviculture	0.18	-2.8

Continued on next page

Table 3: Occupations in the cluster “Channel" (Continued)

PCS-ESE	Occupation Name in French	T_i (log)	A_i (log)
431B	Infirmiers psychiatriques	0.06	-4.36
554H	Vendeurs de tabac, presse et articles divers	-0.08	-3.18

Table 4: Occupations in the cluster “Condenser"

PCS-ESE	Occupation Name in French	T_i (log)	A_i (log)
471B	Techniciens d’exploitation et de contrôle de la production en agriculture, eaux et forêt	1.6	-2
471A	Techniciens d’étude et de conseil en agriculture, eaux et forêt	1.58	-1.44
635A	Tailleurs et couturières qualifiés, ouvriers qualifiés du travail des étoffes (sauf fabrication de vêtements), ouvriers qualifiés de type artisanal du travail du cuir	1.4	-1.53
674E	Ouvriers de production non qualifiés : industrie lourde du bois, fabrication des papiers et cartons	1.39	-2.04
627F	Ouvriers de la composition et de l’impression, ouvriers qualifiés de la brochure, de la reliure et du façonnage du papier-carton	1.38	-1.02
675A	Ouvriers de production non qualifiés du textile et de la confection, de la tannerie-mégisserie et du travail du cuir	1.36	-1.37
627B	Ouvriers qualifiés de la coupe des vêtements et de l’habillement, autres opérateurs de confection qualifiés	1.3	-1.59
627C	Ouvriers qualifiés du travail industriel du cuir	1.27	-1.87
626C	Opérateurs et ouvriers qualifiés des industries lourdes du bois et de la fabrication du papier-carton	1.12	-2.12
691E	Ouvriers agricoles sans spécialisation particulière	1.08	-0.65
691B	Ouvriers de l’élevage	1	-1.5
627A	Opérateurs qualifiés du textile et de la mégisserie	0.98	-1.86
485B	Agents de maîtrise en fabrication des autres industries (imprimerie, matériaux souples, ameublement et bois)	0.92	-1.12
675C	Ouvriers de production non qualifiés de l’imprimerie, presse, édition	0.89	-1.79
476A	Assistants techniques, techniciens de l’imprimerie et de l’édition	0.86	-1.57
628F	Agents qualifiés de laboratoire (sauf chimie, santé)	0.82	-0.81
691A	Conducteurs d’engin agricole ou forestier	0.82	-1.75
691D	Ouvriers de la viticulture ou de l’arboriculture fruitière	0.81	-1.25
691C	Ouvriers du maraîchage ou de l’horticulture	0.8	-1.16

Continued on next page

Table 4: Occupations in the cluster “Condenser" (Continued)

PCS-ESE	**Occupation Name in French**	T_i **(log)**	A_i **(log)**
642A	Conducteurs de taxi	0.77	-1.01
354C	Artistes dramatiques	0.75	-0.52
333E	Autres personnels administratifs de catégorie A de l'Etat (hors Enseignement, Patrimoine, Impôts, Trésor, Douanes)	0.73	0.57
344A	Médecins hospitaliers sans activité libérale	0.71	-0.04
627D	Ouvriers qualifiés de scierie, de la menuiserie industrielle et de l'ameublement	0.7	-0.65
451A	Professions intermédiaires de la Poste	0.68	-0.3
354E	Artistes de la danse	0.68	-1.22
432D	Autres spécialistes de la rééducation, salariés	0.67	-0.66
344C	Internes en médecine, odontologie et pharmacie	0.67	-0.63
622E	Autres monteurs câbleurs en électronique	0.66	-1.62
621C	Conducteurs qualifiés d'engins de chantiers du bâtiment et des travaux publics	0.65	-0.19
632G	Peintres et ouvriers qualifiés de pose de revêtements sur supports verticaux	0.65	-0.21
623A	Chaudronniers-tôliers industriels, opérateurs qualifiés du travail en forge, conducteurs qualifiés d'équipement de formage, traceurs qualifiés	0.63	-0.38
641B	Conducteurs de véhicule routier de transport en commun	0.61	-0.49
434E	Moniteurs éducateurs	0.61	-0.18
381B	Ingénieurs et cadres d'étude et développement de l'agriculture, la pêche, les eaux et forêts	0.59	-2.34
354B	Artistes de la musique et du chant	0.58	-0.8
431A	Cadres infirmiers et assimilés	0.58	-0.08
434D	Educateurs spécialisés	0.57	0.2
651B	Conducteurs d'engin lourd de manoeuvre	0.57	-0.9
554F	Vendeurs en produits de beauté, de luxe (hors biens culturels) et optique	0.56	0.14
472A	Dessinateurs en bâtiment, travaux publics	0.55	-0.19
623B	Tuyauteurs industriels qualifiés	0.55	-1.24
632A	Maçons qualifiés	0.54	0.32
624C	Monteurs qualifiés d'ensembles mécaniques travaillant en moyenne ou en grande série	0.54	-0.29
628C	Régleurs qualifiés d'équipements de fabrication (travail des métaux, mécanique)	0.54	-0.83
354F	Artistes du cirque et des spectacles divers	0.54	-1.14
622C	Monteurs câbleurs qualifiés en électricité	0.54	-0.29
432B	Masseurs-kinésithérapeutes rééducateurs, salariés	0.53	-1.28

Continued on next page

Table 4: Occupations in the cluster “Condenser" (Continued)

PCS-ESE	Occupation Name in French	T_i (log)	A_i (log)
626B	Autres opérateurs et ouvriers qualifiés : métallurgie, production verrière, matériaux de construction	0.53	-0.08
643A	Conducteurs livreurs, coursiers	0.52	0.51
634C	Mécaniciens qualifiés en maintenance, entretien, réparation : automobile	0.51	-0.01
633B	Dépanneurs qualifiés en radiotélévision, électroménager, matériel électronique (salariés)	0.5	-1.8
465B	Assistants techniques de la réalisation des spectacles vivants et audiovisuels salariés	0.49	-0.08
425A	Sous-bibliothécaires, cadres intermédiaires du patrimoine	0.49	-0.69
386E	Ingénieurs et cadres de fabrication des autres industries (imprimerie, matériaux souples, ameublement et bois)	0.49	-0.72
684B	Ouvriers non qualifiés de l’assainissement et du traitement des déchets	0.49	-0.61
451E	Autres personnels administratifs de catégorie B de l’Etat (hors Enseignement, Patrimoine, Impôts, Trésor, Douanes)	0.49	-0.08
546B	Agents des services commerciaux des transports de voyageurs et du tourisme	0.48	-0.05
628D	Régleurs qualifiés d’équipements de fabrication (hors travail des métaux et mécanique)	0.48	-1.19
621G	Mineurs de fond qualifiés et autres ouvriers qualifiés des industries d’extraction (carrières, pétrole, gaz...)	0.48	-1.61
673C	Ouvriers non qualifiés de montage, contrôle en mécanique et travail des métaux	0.48	-0.08
672A	Ouvriers non qualifiés de l’électricité et de l’électronique	0.48	-0.63
642B	Conducteurs de voiture particulière	0.47	-1.03
353A	Directeurs de journaux, administrateurs de presse, directeurs d’éditions (littéraire, musicale, audiovisuelle et multimédia)	0.47	-1.03
523B	Adjoints administratifs de l’Etat et assimilés (sauf Poste, France Télécom)	0.47	0.05
633A	Electriciens qualifiés de type artisanal (y.c. bâtiment)	0.46	-0.04
628A	Mécaniciens qualifiés de maintenance, entretien : équipements industriels	0.45	-0.08
633D	Electriciens, électroniciens qualifiés en maintenance, entretien : équipements non industriels	0.45	-1.16
623G	Opérateurs qualifiés d’usinage des métaux sur autres machines (sauf moulistes)	0.45	-0.37
311C	Chirurgiens dentistes	0.44	-1.85
353C	Cadres artistiques et technico-artistiques de la réalisation de l’audiovisuel et des spectacles	0.44	-0.24

Continued on next page

Table 4: Occupations in the cluster “Condenser" (Continued)

PCS-ESE	**Occupation Name in French**	T_i **(log)**	A_i **(log)**
625D	Opérateurs de la transformation des viandes	0.43	-0.23
628E	Ouvriers qualifiés de l'assainissement et du traitement des déchets	0.43	-0.9
632C	Charpentiers en bois qualifiés	0.43	-0.91
476B	Techniciens de l'industrie des matériaux souples, de l'ameublement et du bois	0.42	-1.84
682A	Métalliers, serruriers, réparateurs en mécanique non qualifiés	0.42	-0.63
673A	Ouvriers de production non qualifiés travaillant par enlèvement de métal	0.4	-0.66
654B	Conducteurs qualifiés d'engins de transport guidés (sauf remontées mécaniques)	0.4	-1.24
526E	Ambulanciers salariés	0.4	-1.07
634D	Mécaniciens qualifiés de maintenance, entretien : équipements non industriels	0.4	-0.8
422A	Professeurs d'enseignement général des collèges	0.4	-0.26
634A	Carrossiers d'automobiles qualifiés	0.4	-1.01
625C	Autres opérateurs et ouvriers qualifiés de la chimie (y.c. pharmacie) et de la plasturgie	0.39	-0.16
624D	Monteurs qualifiés en structures métalliques	0.39	-0.83
632D	Menuisiers qualifiés du bâtiment	0.39	-0.35
624G	Autres mécaniciens ou ajusteurs qualifiés (ou spécialité non reconnue)	0.39	-1.35
466C	Responsables d'exploitation des transports de voyageurs et de marchandises (non cadres)	0.38	0.06
342B	Professeurs et maîtres de conférences	0.38	-0.2
342D	Personnel enseignant temporaire de l'enseignement supérieur	0.38	-0.3
623F	Opérateurs qualifiés d'usinage des métaux travaillant à l'unité ou en petite série, moulistes qualifiés	0.38	-0.43
422E	Surveillants et aides-éducateurs des établissements d'enseignement	0.37	0.34
477D	Techniciens de l'environnement et du traitement des pollutions	0.36	-0.79
353B	Directeurs, responsables de programmation et de production de l'audiovisuel et des spectacles	0.36	-0.92
434B	Assistants de service social	0.36	-0.48
624E	Ouvriers qualifiés de contrôle et d'essais en mécanique	0.36	-1.15
644A	Conducteurs de véhicule de ramassage des ordures ménagères	0.35	-1.02
634B	Métalliers, serruriers qualifiés	0.35	-0.74

Continued on next page

Table 4: Occupations in the cluster “Condenser" (Continued)

PCS-ESE	Occupation Name in French	T_i (log)	A_i (log)
625A	Pilotes d'installation lourde des industries de transformation : agroalimentaire, chimie, plasturgie, énergie	0.35	-0.75
625F	Autres opérateurs travaillant sur installations ou machines : industrie agroalimentaire (hors transformation des viandes)	0.35	-0.29
354G	Professeurs d'art (hors établissements scolaires)	0.35	-1.02
624B	Monteurs, metteurs au point très qualifiés d'ensembles mécaniques travaillant à l'unité ou en petite série	0.34	-0.95
351A	Bibliothécaires, archivistes, conservateurs et autres cadres du patrimoine	0.34	-0.82
433A	Techniciens médicaux	0.34	-0.31
474A	Dessinateurs en construction mécanique et travail des métaux	0.34	-0.69
641A	Conducteurs routiers et grands routiers	0.34	0.33
485A	Agents de maîtrise et techniciens en production et distribution d'énergie, eau, chauffage	0.33	-0.06
674C	Autres ouvriers de production non qualifiés : industrie agroalimentaire	0.32	0.46
624F	Ouvriers qualifiés des traitements thermiques et de surface sur métaux	0.32	-1.23
632J	Monteurs qualifiés en agencement, isolation	0.32	-0.24
431C	Puéricultrices	0.32	-1.26
675B	Ouvriers de production non qualifiés du travail du bois et de l'ameublement	0.31	-1.19
341A	Professeurs agrégés et certifiés de l'enseignement secondaire	0.3	0.22
621E	Autres ouvriers qualifiés des travaux publics	0.3	-0.27
632E	Couvreurs qualifiés	0.29	-0.89
343A	Psychologues spécialistes de l'orientation scolaire et professionnelle	0.29	-0.45
474C	Techniciens de fabrication et de contrôle-qualité en construction mécanique et travail des métaux	0.28	0.21
673B	Ouvriers de production non qualifiés travaillant par formage de métal	0.28	-1.64
546A	Contrôleurs des transports (personnels roulants)	0.28	-1.23
631A	Jardiniers	0.27	-0.72
486D	Agents de maîtrise en maintenance, installation en mécanique	0.27	-0.83
344D	Pharmaciens salariés	0.26	-0.53
342G	Ingénieurs d'étude et de recherche de la recherche publique	0.26	-0.13
484B	Agents de maîtrise en fabrication : métallurgie, matériaux lourds et autres industries de transformation	0.26	-1.11

Continued on next page

Table 4: Occupations in the cluster “Condenser" (Continued)

PCS-ESE	Occupation Name in French	T_i (log)	A_i (log)
628B	Electromécaniciens, électriciens qualifiés d'entretien : équipements industriels	0.25	-0.5
473A	Dessinateurs en électricité, électromécanique et électronique	0.24	-1.18
681A	Ouvriers non qualifiés du gros oeuvre du bâtiment	0.24	0.13
526B	Assistants dentaires, médicaux et vétérinaires, aides de techniciens médicaux	0.24	-0.21
475B	Techniciens de production et de contrôle-qualité des industries de transformation	0.24	0.06
232A	Chefs de moyenne entreprise, de 50 à 499 salariés	0.24	0.32
475A	Techniciens de recherche-développement et des méthodes de production des industries de transformation	0.23	-0.23
352A	Journalistes (y c. rédacteurs en chef)	0.23	-0.78
472B	Géomètres, topographes	0.23	-1.23
481B	Chefs de chantier (non cadres)	0.23	-0.01
423B	Formateurs et animateurs de formation continue	0.23	0.4
473B	Techniciens de recherche-développement et des méthodes de fabrication en électricité, électromécanique et électronique	0.23	-0.48
632F	Plombiers et chauffagistes qualifiés	0.22	-0.29
486B	Agents de maîtrise en maintenance, installation en électricité et électronique	0.21	-1
524C	Agents administratifs des collectivités locales	0.2	0.74
674A	Ouvriers de production non qualifiés : chimie, pharmacie, plasturgie	0.2	-0.33
451C	Contrôleurs des Impôts, du Trésor, des Douanes et assimilés*	0.2	-1.22
434A	Cadres de l'intervention socio-éducative	0.19	-0.05
433D	Préparateurs en pharmacie	0.19	-0.78
311D	Psychologues, psychanalystes, psychothérapeutes (non médecins)	0.19	-0.45
637C	Ouvriers et techniciens des spectacles vivants et audiovisuels	0.19	-0.62
621D	Ouvriers des travaux publics en installations électriques et de télécommunications	0.19	-0.85
477B	Techniciens d'installation et de maintenance des équipements industriels (électriques, électromécaniques, mécaniques, hors informatique)	0.18	0.38
479B	Experts salariés de niveau technicien, techniciens divers	0.18	0.2
422C	Maîtres auxiliaires et professeurs contractuels de l'enseignement secondaire	0.18	-0.29
466B	Responsables commerciaux et administratifs des transports de marchandises (non cadres)	0.18	-0.92

Continued on next page

Table 4: Occupations in the cluster “Condenser" (Continued)

PCS-ESE	**Occupation Name in French**	T_i **(log)**	A_i **(log)**
562B	Coiffeurs salariés	0.17	-0.37
626A	Pilotes d'installation lourde des industries de transformation : métallurgie, production verrière, matériaux de construction	0.17	-1.02
632H	Soliers moquetteurs et ouvriers qualifiés de pose de revêtements souples sur supports horizontaux	0.16	-2.02
389A	Ingénieurs et cadres techniques de l'exploitation des transports	0.16	0.02
625H	Ouvriers qualifiés des autres industries (eau, gaz, énergie, chauffage)	0.16	-0.86
484A	Agents de maîtrise en fabrication : agroalimentaire, chimie, plasturgie, pharmacie.	0.16	-0.41
473C	Techniciens de fabrication et de contrôle-qualité en électricité, électromécanique et électronique	0.16	-0.33
685A	Ouvriers non qualifiés divers de type artisanal	0.16	-0.66
372F	Cadres de la documentation, de l'archivage (hors fonction publique)	0.15	-1.21
621A	Chefs d'équipe du gros oeuvre et des travaux publics	0.15	-0.36
681B	Ouvriers non qualifiés du second oeuvre du bâtiment	0.15	0.03
434G	Educateurs de jeunes enfants	0.13	-1.13
561E	Employés de l'hôtellerie : réception et hall	0.12	-0.12
342C	Professeurs agrégés et certifiés en fonction dans l'enseignement supérieur	0.12	-1.32
637D	Ouvriers qualifiés divers de type artisanal	0.11	-1.18
541B	Agents d'accueil qualifiés, hôtesses d'accueil et d'information	0.11	0.24
622A	Opérateurs qualifiés sur machines automatiques en production électrique ou électronique	0.11	-1.1
655A	Autres agents et ouvriers qualifiés (sédentaires) des services d'exploitation des transports	0.11	-0.3
674D	Ouvriers de production non qualifiés : métallurgie, production verrière, céramique, matériaux de construction	0.11	-0.5
676D	Agents non qualifiés des services d'exploitation des transports	0.11	-0.59
651A	Conducteurs d'engin lourd de levage	0.1	-1.07
474B	Techniciens de recherche-développement et des méthodes de fabrication en construction mécanique et travail des métaux	0.1	-0.5
625G	Autres ouvriers de production qualifiés ne travaillant pas sur machine : industrie agroalimentaire (hors transformation des viandes)	0.09	-1.2

Continued on next page

Table 4: Occupations in the cluster “Condenser" (Continued)

PCS-ESE	**Occupation Name in French**	T_i **(log)**	A_i **(log)**
676E	Ouvriers non qualifiés divers de type industriel	0.09	0.46
421B	Professeurs des écoles	0.08	-0.63
477C	Techniciens d'installation et de maintenance des équipements non industriels (hors informatique et télécommunications)	0.08	-0.07
562A	Manucures, esthéticiens	0.08	-0.7
435B	Animateurs socioculturels et de loisirs	0.08	0.34
546C	Employés administratifs d'exploitation des transports de marchandises	0.06	-0.48
472C	Métreurs et techniciens divers du bâtiment et des travaux publics	0.06	0.13
384A	Ingénieurs et cadres d'étude, recherche et développement en mécanique et travail des métaux	0.06	0.1
344B	Médecins salariés non hospitaliers	0.06	-0.61
433C	Autres spécialistes de l'appareillage médical salariés	0.06	-1.31
524B	Agents administratifs de l'Etat et assimilés (sauf Poste, France Télécom)	0.05	0.36
468A	Maîtrise de restauration : salle et service	0.05	-0.08
451H	Professions intermédiaires administratives des hôpitaux	0.05	0.03
478D	Techniciens des télécommunications et de l'informatique des réseaux	0.04	0.15
382A	Ingénieurs et cadres d'étude du bâtiment et des travaux publics	0.04	0.18
382B	Architectes salariés	0.04	-0.83
220X	Commerçants et assimilés, salariés de leur entreprise	0.03	0.31
621F	Ouvriers qualifiés des travaux publics (salariés de l'Etat et des collectivités locales)	0.03	0.68
382C	Ingénieurs, cadres de chantier et conducteurs de travaux (cadres) du bâtiment et des travaux publics	0.03	0.2
233D	Chefs d'entreprise de services, de 10 à 49 salariés	0.03	0.12
384B	Ingénieurs et cadres de fabrication en mécanique et travail des métaux	0.02	0.26
671C	Ouvriers non qualifiés des travaux publics et du travail du béton	0.02	-0.3
628G	Ouvriers qualifiés divers de type industriel	0.02	0.07
434F	Educateurs techniques spécialisés, moniteurs d'atelier	0.02	-1.47
487A	Responsables d'entrepôt, de magasinage	0.02	-0.25
386B	Ingénieurs et cadres d'étude, recherche et développement de la distribution d'énergie, eau	0.02	-0.19

Continued on next page

Table 4: Occupations in the cluster “Condenser" (Continued)

PCS-ESE	Occupation Name in French	T_i (log)	A_i (log)
486E	Agents de maîtrise en entretien général, installation, travaux neufs (hors mécanique, électromécanique, électronique)	0.01	-1.25
431D	Infirmiers spécialisés (autres qu'infirmiers psychiatriques et puéricultrices)	0.01	-1.25
374A	Cadres de l'exploitation des magasins de vente du commerce de détail	0.01	0.46
674B	Ouvriers de production non qualifiés de la transformation des viandes	0.01	-0.7
621B	Ouvriers qualifiés du travail du béton	0.004	-0.5
488A	Maîtrise de restauration : cuisine/production	0.002	-0.17
210X	Artisans salariés de leur entreprise	0.002	-0.09
422B	Professeurs de lycée professionnel	-0.002	-0.48
377A	Cadres de l'hôtellerie et de la restauration	-0.005	-0.004
383A	Ingénieurs et cadres d'étude, recherche et développement en électricité, électronique	-0.01	0.21
383C	Ingénieurs et cadres technico-commerciaux en matériel électrique ou électronique professionnel	-0.01	-0.63
477A	Techniciens de la logistique, du planning et de l'ordonnancement	-0.01	0.01
468B	Maîtrise de l'hébergement : hall et étages	-0.01	-0.92
488B	Maîtrise de restauration : gestion d'établissement	-0.01	-0.61
479A	Techniciens des laboratoires de recherche publique ou de l'enseignement	-0.01	-0.73
625B	Ouvriers qualifiés et agents qualifiés de laboratoire : agroalimentaire, chimie, biologie, pharmacie	-0.01	-0.53
481A	Conducteurs de travaux (non cadres)	-0.02	-0.28
546E	Autres agents et hôtesses d'accompagnement (transports, tourisme)	-0.02	-1.05
467D	Professions intermédiaires techniques des organismes de sécurité sociale	-0.02	-0.25
434C	Conseillers en économie sociale familiale	-0.03	-1.2
387F	Ingénieurs et cadres techniques de l'environnement	-0.03	-0.72
462B	Maîtrise de l'exploitation des magasins de vente	-0.03	0.26
385B	Ingénieurs et cadres de fabrication des industries de transformation (agroalimentaire, chimie, métallurgie, matériaux lourds)	-0.03	-0.07
387E	Ingénieurs et cadres de la maintenance, de l'entretien et des travaux neufs	-0.03	0
541C	Agents d'accueil non qualifiés	-0.04	-0.78

Continued on next page

Table 4: Occupations in the cluster “Condenser" (Continued)

PCS-ESE	Occupation Name in French	T_i (log)	A_i (log)
383B	Ingénieurs et cadres de fabrication en matériel électrique, électronique	-0.04	-0.9
233A	Chefs d'entreprise du bâtiment et des travaux publics, de 10 à 49 salariés	-0.04	-0.49
545D	Employés des services techniques des organismes de sécurité sociale et assimilés	-0.05	0.02
332B	Ingénieurs des collectivités locales et des hôpitaux	-0.06	-0.07
386D	Ingénieurs et cadres de la production et de la distribution d'énergie, eau	-0.06	-0.57
385A	Ingénieurs et cadres d'étude, recherche et développement des industries de transformation (agroalimentaire, chimie, métallurgie, matériaux lourds)	-0.07	0.16
553C	Autres vendeurs non spécialisés	-0.07	0.24
342F	Directeurs et chargés de recherche de la recherche publique	-0.07	-1.05
554B	Vendeurs en ameublement, décor, équipement du foyer	-0.07	-0.04
463E	Techniciens commerciaux et technico-commerciaux, représentants auprès de particuliers (hors banque, assurance, informatique)	-0.07	0.22
375A	Cadres de la publicité	-0.08	-0.29
472D	Techniciens des travaux publics de l'Etat et des collectivités locales	-0.08	-0.42
376B	Cadres des opérations bancaires	-0.08	0.04
487B	Responsables du tri, de l'emballage, de l'expédition et autres responsables de la manutention	-0.08	-0.99
525D	Agents de service hospitaliers	-0.08	0.5
466A	Responsables commerciaux et administratifs des transports de voyageurs et du tourisme (non cadres)	-0.08	-0.39
387D	Ingénieurs et cadres du contrôle-qualité	-0.09	-0
385C	Ingénieurs et cadres technico-commerciaux des industries de transformations (biens intermédiaires)	-0.09	-0.13
386C	Ingénieurs et cadres d'étude, recherche et développement des autres industries (imprimerie, matériaux souples, ameublement et bois)	-0.09	-0.87
526A	Aides-soignants	-0.09	0.11
433B	Opticiens lunetiers et audioprothésistes salariés	-0.09	-1.29
233C	Chefs d'entreprise commerciale, de 10 à 49 salariés	-0.09	-0.36
464A	Assistants de la publicité, des relations publiques	-0.09	-0.64
467A	Chargés de clientèle bancaire	-0.1	0.43
382D	Ingénieurs et cadres technico-commerciaux en bâtiment, travaux publics	-0.11	-0.27

Continued on next page

Table 4: Occupations in the cluster “Condenser" (Continued)

PCS-ESE	Occupation Name in French	T_i (log)	A_i (log)
387A	Ingénieurs et cadres des achats et approvisionnements industriels	-0.11	-0.19
422D	Conseillers principaux d'éducation	-0.12	-1.84
376A	Cadres des marchés financiers	-0.12	-0.53
525C	Agents de service de la fonction publique (sauf écoles, hôpitaux)	-0.13	-1.68
545C	Employés des services techniques des assurances	-0.13	0.29
684A	Nettoyeurs	-0.14	0.67
233B	Chefs d'entreprise de l'industrie ou des transports, de 10 à 49 salariés	-0.14	-0.62
435A	Directeurs de centres socioculturels et de loisirs	-0.14	-1.13
384C	Ingénieurs et cadres technico-commerciaux en matériel mécanique professionnel	-0.14	-0.66
521A	Employés de la Poste	-0.14	-1.42
478C	Techniciens d'installation, de maintenance, support et services aux utilisateurs en informatique	-0.14	0.07
376C	Cadres commerciaux de la banque	-0.14	0.24
333F	Personnels administratifs de catégorie A des collectivités locales et hôpitaux publics (hors Enseignement, Patrimoine)	-0.14	0.31
333C	Cadres de la Poste*	-0.14	-0.88
541D	Standardistes, téléphonistes	-0.15	-0.35
332A	Ingénieurs de l'Etat (y.c. ingénieurs militaires) et assimilés	-0.15	-0.64
372E	Juristes	-0.15	0.03
463C	Techniciens commerciaux et technico-commerciaux, représentants en biens de consommation auprès d'entreprises	-0.16	-0.11
561F	Employés d'étage et employés polyvalents de l'hôtellerie	-0.16	-0.14
636D	Cuisiniers et commis de cuisine	-0.16	0.33
387C	Ingénieurs et cadres des méthodes de production	-0.17	-0.09
556A	Vendeurs en gros de biens d'équipement, biens intermédiaires	-0.17	0.04
632K	Ouvriers qualifiés d'entretien général des bâtiments	-0.17	-0.17
333A	Magistrats*	-0.18	-1.34
451G	Professions intermédiaires administratives des collectivités locales	-0.18	-0.04
467B	Techniciens des opérations bancaires	-0.18	-0.09
563B	Aides à domicile, aides ménagères, travailleuses familiales	-0.19	0.55
555A	Vendeurs par correspondance, télévendeurs	-0.19	-0.49
523D	Adjoints administratifs des hôpitaux publics	-0.19	-0.68
526D	Aides médico-psychologiques	-0.2	-0.25

Continued on next page

Table 4: Occupations in the cluster “Condenser" (Continued)

PCS-ESE	**Occupation Name in French**	T_i **(log)**	A_i **(log)**
231A	Chefs de grande entreprise de 500 salariés et plus	-0.21	-1.25
545A	Employés administratifs des services techniques de la banque	-0.21	-0.18
333B	Inspecteurs et autres personnels de catégorie A des Impôts, du Trésor et des Douanes	-0.21	-0.24
463D	Techniciens commerciaux et technico-commerciaux, représentants en services auprès d'entreprises ou de professionnels (hors banque, assurance, informatique)	-0.22	-0.03
387B	Ingénieurs et cadres de la logistique, du planning et de l'ordonnancement	-0.23	-0
374D	Cadres commerciaux des petites et moyennes entreprises (hors commerce de détail)	-0.24	0.94
462A	Chefs de petites surfaces de vente	-0.24	-0.4
463B	Techniciens commerciaux et technico-commerciaux, représentants en biens d'équipement, en biens intermédiaires, commerce interindustriel (hors informatique)	-0.24	-0.03
424A	Moniteurs et éducateurs sportifs, sportifs professionnels	-0.25	-0.5
375B	Cadres des relations publiques et de la communication	-0.25	-0.07
554A	Vendeurs en alimentation	-0.25	0.31
478A	Techniciens d'étude et de développement en informatique	-0.25	0.26
554E	Vendeurs en habillement et articles de sport	-0.25	0.29
563A	Assistantes maternelles, gardiennes d'enfants, familles d'accueil	-0.26	-0.15
462C	Acheteurs non classés cadres, aides-acheteurs	-0.26	-0.4
462D	Animateurs commerciaux des magasins de vente, marchandiseurs (non cadres)	-0.26	-0.7
525A	Agents de service des établissements primaires	-0.26	-0.48
478B	Techniciens de production, d'exploitation en informatique	-0.26	-0.23
461E	Maîtrise et techniciens administratifs des services juridiques ou du personnel	-0.29	0.34
467C	Professions intermédiaires techniques et commerciales des assurances	-0.29	-0.01
545B	Employés des services commerciaux de la banque	-0.29	0.01
374B	Chefs de produits, acheteurs du commerce et autres cadres de la mercatique	-0.29	0.33
376D	Chefs d'établissements et responsables de l'exploitation bancaire	-0.29	-0.02
542A	Secrétaires	-0.29	1.03
652A	Ouvriers qualifiés de la manutention, conducteurs de chariots élévateurs, caristes	-0.29	0.43

Continued on next page

Table 4: Occupations in the cluster “Condenser" (Continued)

PCS-ESE	**Occupation Name in French**	T_i **(log)**	A_i **(log)**
676C	Ouvriers du tri, de l'emballage, de l'expédition, non qualifiés	-0.3	0.73
554C	Vendeurs en droguerie, bazar, quincaillerie, bricolage	-0.3	-0.41
376E	Cadres des services techniques des assurances	-0.3	0.12
543B	Employés qualifiés des services comptables ou financiers	-0.3	0.87
551A	Employés de libre service du commerce et magasiniers	-0.31	0.75
376F	Cadres des services techniques des organismes de sécurité sociale et assimilés	-0.32	-0.48
561D	Aides de cuisine, apprentis de cuisine et employés polyvalents de la restauration	-0.32	0.57
373B	Cadres des autres services administratifs des grandes entreprises	-0.32	0.73
561B	Serveurs, commis de restaurant, garçons qualifiés	-0.32	0.31
561C	Serveurs, commis de restaurant, garçons non qualifiés	-0.33	0.49
461F	Maîtrise et techniciens administratifs des autres services administratifs	-0.35	0.65
374C	Cadres commerciaux des grandes entreprises (hors commerce de détail)	-0.35	0.62
544A	Employés et opérateurs d'exploitation en informatique	-0.35	-0.23
543F	Employés qualifiés des services commerciaux des entreprises (hors vente)	-0.36	0.74
373D	Cadres des autres services administratifs des petites et moyennes entreprises	-0.38	0.91
331A	Personnels de direction de la fonction publique (Etat, collectivités locales, hôpitaux)	-0.38	0.04
676A	Manutentionnaires non qualifiés	-0.39	0.5
563C	Employés de maison et personnels de ménage chez des particuliers	-0.4	0.03
653A	Magasiniers qualifiés	-0.4	0.43
462E	Autres professions intermédiaires commerciales (sauf techniciens des forces de vente)	-0.4	0.58
636C	Boulangers, pâtissiers (sauf activité industrielle)	-0.41	-0.55
552A	Caissiers de magasin	-0.42	0.15
372A	Cadres chargés d'études économiques, financières, commerciales	-0.42	0.69
372C	Cadres spécialistes des ressources humaines et du recrutement	-0.43	0.39
525B	Agents de service des autres établissements d'enseignement	-0.44	-0.22
543H	Employés administratifs non qualifiés	-0.45	0.73
543E	Employés qualifiés des services du personnel et des services juridiques	-0.46	0.03

Continued on next page

Table 4: Occupations in the cluster “Condenser" (Continued)

PCS-ESE	**Occupation Name in French**	T_i **(log)**	A_i **(log)**
372B	Cadres de l'organisation ou du contrôle des services administratifs et financiers	-0.46	0.83
372D	Cadres spécialistes de la formation	-0.47	-0.26
388A	Ingénieurs et cadres d'étude, recherche et développement en informatique	-0.47	1.06
376G	Cadres de l'immobilier	-0.47	-0.22
534A	Agents civils de sécurité et de surveillance	-0.48	-0.36
388E	Ingénieurs et cadres spécialistes des télécommunications	-0.49	-0.11
461D	Maîtrise et techniciens des services financiers ou comptables	-0.51	0.44
523C	Adjoints administratifs des collectivités locales	-0.52	0.01
373C	Cadres des services financiers ou comptables des petites et moyennes entreprises	-0.53	0.64
388B	Ingénieurs et cadres d'administration, maintenance, support et services aux utilisateurs en informatique	-0.53	0.33
388C	Chefs de projets informatiques, responsables informatiques	-0.54	0.84
543G	Employés administratifs qualifiés des autres services des entreprises	-0.54	1.04
564B	Employés des services divers	-0.59	0.48
461B	Secrétaires de direction, assistants de direction (non cadres)	-0.62	0.15
543C	Employés non qualifiés des services comptables ou financiers	-0.63	-0.46
373A	Cadres des services financiers ou comptables des grandes entreprises	-0.64	0.08
461C	Secrétaires de niveau supérieur (non cadres, hors secrétaires de direction)	-0.64	-0.26
388D	Ingénieurs et cadres technico-commerciaux en informatique et télécommunications	-0.64	-0.09

B.6. Inter- vs. Intra-Community Transitions

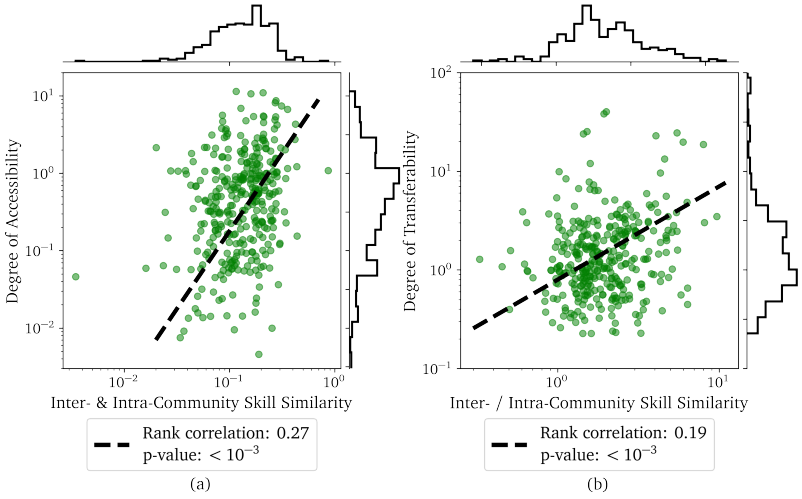


Figure B.3.: Correlation of Accessibility and Transferability Metrics with Inter- vs. Intra-Community Transitions. (a) Scatter plot depicting the relationship between the accessibility of an occupation and weighted inter- vs. intra-community transitions, showing a significant positive rank correlation of 0.27. (b) Scatter plot depicting the relationship between the transferability of an occupation and weighted inter- vs. intra-community transitions, revealing a significant positive rank correlation of 0.13. The color of the points corresponds to the respective accessibility and transferability taxonomy described in Section 6.3. A dashed line serves as a visual guide and represents the linear fit of the points.

B.7. Age Effect on Accessibility and Transferability

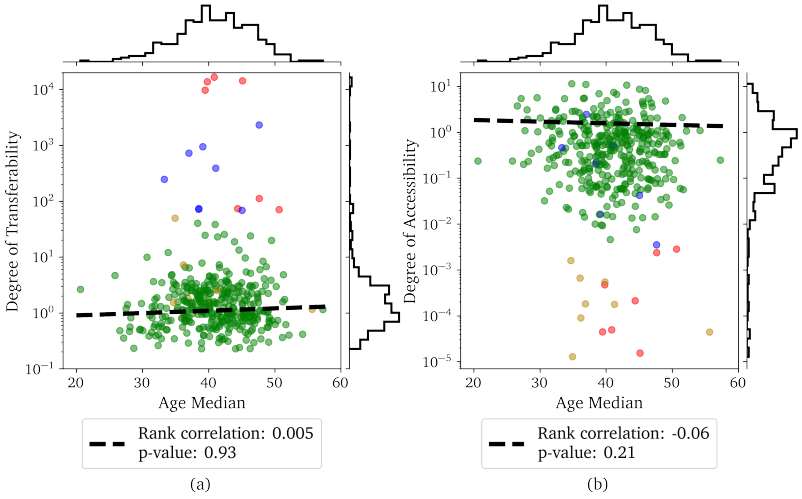


Figure B.4.: Age Effect on Accessibility and Transferability. Scatter plot depicts the relationship between age median and (a) the degree of transferability and (b) the degree of accessibility for each occupation. Color coding corresponds to the introduced occupation taxonomy. A dashed line serves as a visual guide and represents the linear fit of the points. There is no significant rank correlation observed between age and transferability nor between age and accessibility (legend of (a,b)).

Figures B.4 (a) and (b), show the relationship between the median age of workers in an occupation and its (a) degree of transferability and (b) accessibility. There is no significant rank correlation between the median age and either transferability or accessibility. This indicates that our metrics for accessibility and transferability are note correlated to age-related effects.

To investigate the impact of age on our findings, we construct a labor flow network separately for workers below the age of 30 who are typically still in the learning or training phase for specific occupations and for workers above 60 who might experience seniority effects Eurostat (2018) [266].

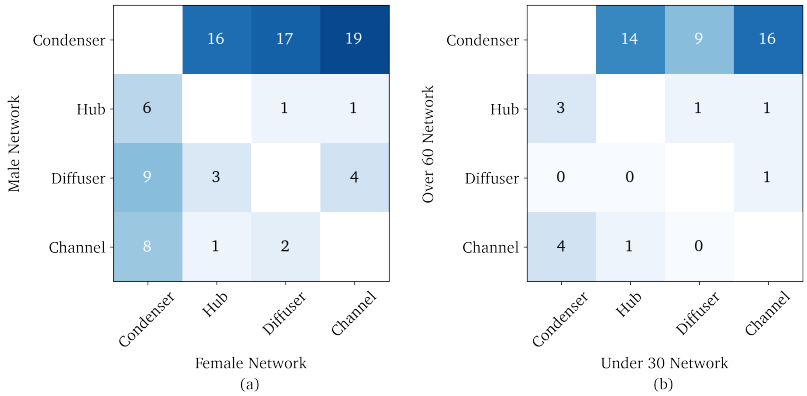


Figure B.5.: Confusion Matrix for Gender and Age. Panel (a) shows the confusion matrix for taxonomy clustering in the male and female networks, excluding diagonal elements (occupations assigned to the same cluster in both networks). A significant number of occupations in the male Condenser cluster are spread across Hubs, Diffusers, and Channels in the female network, indicating a broader dispersion in accessibility and transferability for females. Panel (b) displays the confusion matrix for clustering in age groups below 30 and above 60, excluding diagonal elements. Many occupations in the Condenser cluster for the 60+ age group are distributed across Hubs, Diffusers, and Channels for the under-30 group, reflecting a wider dispersion in accessibility and transferability for younger individuals.

When analyzing transitions separately for workers under 30 and those over 60, we can discover on a single occupational level the presence of entry positions and those offering more opportunities for older workers. Figure B.5 (b), shows the confusion matrix illustrating the classification of the same occupations across different age groups[1]. Fifty occupations are classified differently between the age groups. The majority of shifts in classification occur from the "Condenser" category in the over-60 network to other categories in the under-30 network. Suggesting a more dispersed spectrum of accessibility and transferability for the under-30 age group, which indicates a higher prevalence of occupations that are less accessible but possess a greater degree of transferability.

1: We exclude diagonal elements, representing occupations classified the same in both age groups.

If an occupation in the network of workers below 30 is in the Hub (Condenser) category and switches to a Diffuser position (Channel) in the network above 60, it suggests higher accessibility to younger workers, indicating potential entry-level roles. Examples include Unskilled employees of accounting or financial services (543C) or Employed eyewear opticians and audioprosthetists (433B).

Conversely, if transferability of an occupation increases with age, leading to a shift from Condenser (Channel) to Hub (Diffuser), it suggests that opportunities for specific occupations tend to expand with age. Examples of such jobs include Qualified employees of accounting or financial services (543B) and Sellers of beauty, luxury and optical products (554F) (see Figure B.4). For the first occupation, one might argue that specific skills necessitate time for acquisition, whereas in the latter occupation, acquiring skills like selling products to individuals and cultivating a network of trusted customers could create new opportunities for occupational transitions.

B.8. Policy Evaluation with Matrix Perturbation

The occupational transition matrix $\mathcal{P}$ is a right stochastic matrix where each entry

$$\mathcal{P}_{i,j} = \mathbf{Pr}(X_i = i | X_j = j) \tag{B.7}$$

corresponds to the probability of transitioning to occupation j, given an occupation i, with $\sum_j \mathcal{P}_{i,j} = 1$. Then there exists one eigenvalue $\lambda_1 = 1$ corresponding to an eigenvector $\vec{v}_1 = \vec{\pi}$ that is the stationary distribution $\vec{\pi}$ for a markov process of $N \to \infty$ steps.

Let $\mathcal{P}$ be a symmetric matrix with eigenvectors $\vec{v}_i$ and λ_i, then

$$\mathcal{P}\vec{v}_i = \lambda_i \vec{v}_i \tag{B.8}$$

Now let us consider a perturbation ϵV, then

$$(\mathcal{P} + \epsilon V)\vec{v}_2^\star = \lambda_2^\star \vec{v}_2^\star \tag{B.9}$$

We can write $\vec{v}_2^\star$ and $\lambda_2^\star$ as

$$\vec{v}_2^\star = \vec{v}_2 + \epsilon\vec{v} + ... \tag{B.10}$$
$$\lambda_2^\star = \lambda_2 + \epsilon\lambda + ... \tag{B.11}$$

with this

$$(\mathcal{P} + \epsilon V)(\vec{v}_2 + \epsilon\vec{v}) = (\lambda_2 + \epsilon\lambda)(\vec{v}_2 + \epsilon\vec{v}) \tag{B.12}$$

If we separate the equation for each order of ϵ, neglect higher orders of ϵ and use the symmetry of $\mathcal{P}$, such that $\vec{v}_2^T\mathcal{P}\vec{v} = \vec{v}_2^T\lambda_2\vec{v}$, then

$$\lambda = \frac{\vec{v}_2^T V \vec{v}_2}{\vec{v}_2^T \vec{v}_2} \tag{B.13}$$

$$\lambda_2^\star = \lambda_2 + \epsilon\frac{\vec{v}_2^T V \vec{v}_2}{\vec{v}_2^T \vec{v}_2} \tag{B.14}$$

$$\Delta\lambda = \lambda_2^\star - \lambda_2 = \epsilon\frac{\vec{v}_2^T V \vec{v}_2}{\vec{v}_2^T \vec{v}_2} \tag{B.15}$$

with $\vec{v}_i^T\vec{v}_i = 1$. Knowing the eigenvectors $\vec{v}_i$ of the unperturbed occupational transition matrix $\mathcal{P}$, we can calculate the change of eigenvalue $\Delta\lambda$ w.r.t. a small perturbation of a rate of a single link. We are looking for a policy V where $\Delta\lambda_2 < 0$ as with this, the convergence of the perturbed matrix is faster since the spectral gap of the perturbed matrix gets larger. Let's say we denote the target occupation with k and increase the flow between k and a destination occupation l^+, simultaneously we decrease the flow from target occupation k to source occupation l^-. Then the perturbation V is

$$V = \begin{pmatrix} 0 & \dots & \dots & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ -1 & \dots & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & \dots & \dots & 0 \end{pmatrix} \tag{B.16}$$

or (B.17)

$$(V)_{i,j} = \begin{cases} -1, & i = l^-, j = k \\ 1, & i = l^+, j = k \\ 0, & \text{else} \end{cases} \tag{B.18}$$

with this the product in eq. B.15 gets

$$\vec{v}_2^T V \vec{v}_2 = (\vec{v}_2^T)_k \cdot \sum_i V_{i,k} (\vec{v}_2)_i \tag{B.19}$$

$$= (\vec{v}_2^T)_k \cdot ((\vec{v}_2)_{l^+} - (\vec{v}_2)_{l^-}) \tag{B.20}$$

Here we can derive four conditions for the sign of $\Delta\lambda$

$$\Delta\lambda \begin{cases} > 0, & \text{if } (\vec{v}_2)_{l^+} > (\vec{v}_2)_{l^-} \wedge (\vec{v}_2)_k > 0 \\ > 0, & \text{if } (\vec{v}_2)_{l^+} < (\vec{v}_2)_{l^-} \wedge (\vec{v}_2)_k < 0 \\ < 0, & \text{if } (\vec{v}_2)_{l^+} > (\vec{v}_2)_{l^-} \wedge (\vec{v}_2)_k < 0 \\ < 0, & \text{if } (\vec{v}_2)_{l^+} < (\vec{v}_2)_{l^-} \wedge (\vec{v}_2)_k > 0 \end{cases} \tag{B.21}$$

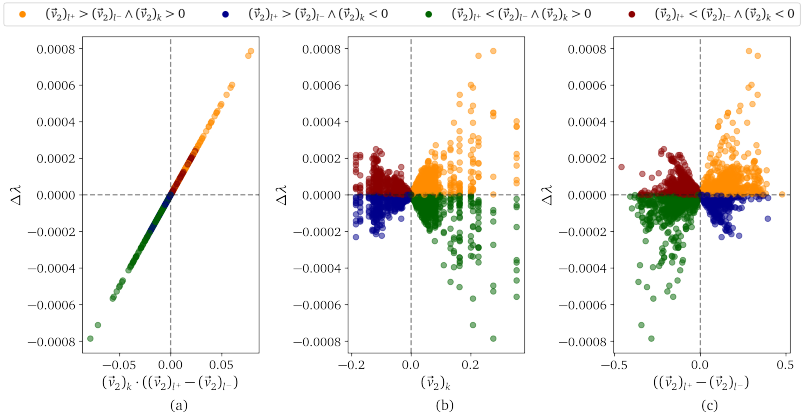


Figure B.6.: Matrix Perturbation of the French Labor Market. (a) Scatter plot showing eq. B.20, demonstrating a strong positive correlation that supports the validity of the assumptions made for first-order matrix perturbation. The color coding of points corresponds to the sign cases of $\Delta\lambda$ (see eq. B.21). (b) Scatter plot showing the relationship between the eigenvalue difference and the value of the eigenvector for the target occupation, $(\vec{v}_2)_k$. (c) Scatter plot showing the eigenvalue difference against the difference in eigenvector values between the destination and source occupations, $((\vec{v}_2)_{l^+} - (\vec{v}_2)_{l^-})$.

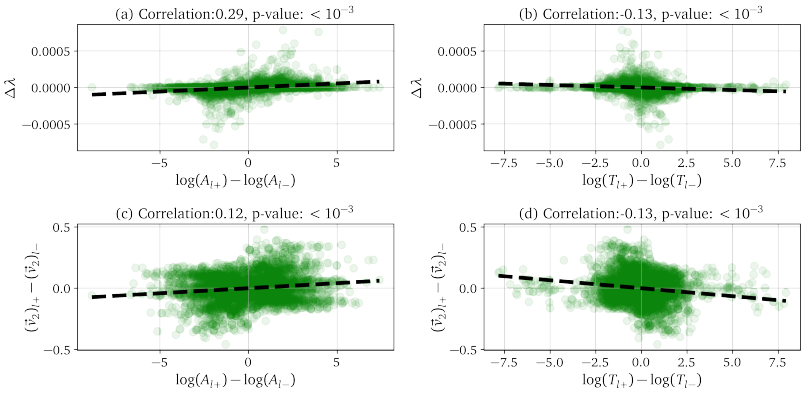


Figure B.7.: Correlation of Accessibility and Transferability Metrics with Changes in Eigenvalue (a) Scatter plot showing the log difference in accessibility metrics between destination and source occupations against the change in the second eigenvalue. (b) Scatter plot showing the log difference in transferability metrics between destination and source occupations against the change in the second eigenvalue. (c) Scatter plot showing the log difference in accessibility metrics between destination and source occupations against the difference in the corresponding values of the second eigenvector. (d) Scatter plot showing the log difference in transferability metrics between destination and source occupations against the difference in the corresponding values of the second eigenvector.

B.9. Task-based Cluster

Table 1: Occupations in the cluster “Non-Routine Cognitive"

ROME	Occupation in French	Occupation in English
A1301	Conseil et assistance technique en agriculture et environnement naturel	Advice and technical assistance in agriculture and natural environment
A1302	Contrôle et diagnostic technique en agriculture	Control and technical diagnosis in agriculture
A1303	Ingénierie en agriculture et environnement naturel	Engineering in agriculture and natural environment
A1406	Encadrement équipage de la pêche	Fishing crew supervision
A1413	Fermentation de boissons alcoolisées	Fermentation of alcoholic beverages
A1504	Santé animale	Animal health
B1101	Création en arts plastiques	Creation in plastic arts
B1301	Décoration d’espaces de vente et d’exposition	Decoration of sales and exhibition spaces
B1801	Réalisation d’articles de chapellerie	Production of hats
B1805	Stylisme	Styling
C1101	Conception - développement produits d’assurances	Insurance product development
C1102	Conseil clientèle en assurances	Insurance customer advice
C1103	Courtage en assurances	Insurance brokerage
C1104	Direction d’exploitation en assurances	Insurance operating direction
C1105	Études actuarielles en assurances	Acturable insurance studies
C1106	Expertise risques en assurances	Insurance risks expertise
C1107	Indemnisations en assurances	Insurance compensation
C1108	Management de groupe et de service en assurances	Group management and insurance service
C1110	Souscription d’assurances	Insurance subscription
C1202	Analyse de crédits et risques bancaires	Credit analysis and banking risks
C1203	Relation clients banque/finance	Relation clients banque/finance
C1204	Conception et expertise produits bancaires et financiers	Design and expertise banking and financial products
C1205	Conseiller / conseillère en gestion de patrimoine	Advisor to Personal Heritage Management
C1206	Gestion de clientèle bancaire	Banking customer management
C1207	Management en exploitation bancaire	Management in banking
C1208	Courtier / courtière en banque	Banking broker
C1301	Front office marchés financiers	Front Office Financial Markets
C1302	Gestion back et middle-office marchés financiers	Back and Middle-Office Management Financial Markets
C1303	Gestion de portefeuilles sur les marchés financiers	Portfolio management on the financial markets
C1501	Gérance immobilière	Real estate management
C1502	Gestion locative immobilière	Real estate rental management

Continued on next page

Table 1: Occupations in the cluster “Non-Routine Cognitive" (Continued)

ROME	Occupation in French	Occupation in English
C1503	Management de projet immobilier	Real estate project management
C1504	Conseiller / conseillère immobilier	Real estate advisor
C1505	Responsable d'agence immobilière	Real estate agency manager employment employment
D1107	Vente en gros de produits frais	Wholesale of fresh products
D1201	Achat vente d'objets d'art, anciens ou d'occasion	Purchase Sale of works of art, old or used
D1204	Location de véhicules ou de matériel de loisirs	Rental of vehicles or leisure equipment
D1210	Vente en animalerie	Sales in pet stores
D1211	Vente en articles de sport et loisirs	Sales of sports and leisure articles
D1212	Vente en décoration et équipement du foyer	Sales of decoration and home equipment
D1213	Vente en gros de matériel et équipement	Wholesale of equipment and equipment
D1214	Vente en habillement et accessoires de la personne	Personal clothing and accessories sale
D1301	Management de magasin de détail	Retail store management
D1402	Relation commerciale grands comptes et entreprises	Large accounts and companies commercial relationship
D1403	Relation commerciale auprès de particuliers	Commercial relationship with individuals
D1404	Relation commerciale en vente de véhicules	Commercial relationship for sale of vehicles
D1405	Conseil en information médicale	Medical Information Advice
D1406	Management en force de vente	Sales force management
D1407	Technico-commercial / technico-commerciale	Technico-commercial executive employment
D1410	Commercial / commerciale	Salesperson
D1501	Animateur / animatrice des ventes	Framework sales manager
D1502	Management/gestion de rayon produits alimentaires	Management of food products
D1503	Management/gestion de rayon produits non alimentaires	Management of non-food products
D1504	Direction de magasin de grande distribution	Directorate of large retailers
D1506	Marchandisage	Merchandising
D1508	Encadrement du personnel de caisses	Supervision of cashiers
D1509	Responsable grande distribution	Large distribution manager
D1510	Chef / cheffe de secteur commercial	Head of sales area
E1101	Animation de site multimédia	Multimedia site animation

Continued on next page

Table 1: Occupations in the cluster “Non-Routine Cognitive" (Continued)

ROME	Occupation in French	Occupation in English
E1102	Écriture d'ouvrages, de livres	Writing works, books
E1103	Chargé / chargée de communication	Communication manager
E1104	Conception de contenus multimédias	Multimedia content design
E1105	Coordination d'édition	Edition coordination
E1106	Journalisme et information média	Media journalism and information
E1107	Chef / cheffe de projet événementiel	Event project manager
E1108	Traduction, interprétariat	Translation, interpretation
E1109	Directeur / directrice de la communication	Communication director
E1201	Photographie	Photography
E1203	Production en laboratoire photographique	Photographic laboratory production
E1205	Réalisation de contenus multimédias	Multimedia content production
E1303	Encadrement des industries graphiques	Supervision of graphic industries
E1305	Préparation et correction en édition et presse	Preparation and correction in edition and press
E1306	Prépresse	Prepress
E1308	Intervention technique en industrie graphique	Technical intervention in graphic industry
E1401	Développement et promotion publicitaire	Development and advertising promotion
E1402	Élaboration de plan média	Media plan development
F1101	Architecture du BTP et du paysage	Connection of construction and landscape
F1102	Architecte d'intérieur	Interior designer
F1103	Contrôle et diagnostic technique du bâtiment	Building technical control and diagnosis
F1104	Dessin BTP et paysage	Drawing in construction and landscaping
F1105	Études géologiques	Geological studies
F1106	Ingénierie et études du BTP	Engineering and studies in construction
F1107	Géomètre-topographe	Surveyor-topographer
F1108	Métré de la construction	Construction Management
F1109	Mesures topographiques	Topographic measures
F1110	Dessinateur / dessinatrice enveloppe du bâtiment	Building envelope draftsman
F1111	Ingénieur / ingénieure génie civil	Civil engineer
F1112	Ingénieur / ingénieure calcul et structure	Calculation and structure engineer

Continued on next page

Table 1: Occupations in the cluster “Non-Routine Cognitive" (Continued)

ROME	Occupation in French	Occupation in English
F1201	Conducteur / conductrice de travaux	Construction supervisor
F1202	Chef / cheffe de chantier	Site manager
F1203	Direction et ingénierie d'exploitation de gisements et de carrières	Management and engineering of deposits and quarries
F1204	Qualité sécurité environnement et protection santé du BTP	Environmental safety and health protection of construction
F1205	Responsable de travaux BTP	Public works manager
G1101	Accueil touristique	Tourist reception
G1102	Promotion du tourisme local	Promotion of local tourism
G1201	Accompagnement de voyages, d'activités culturelles ou sportives	Support for travel, cultural or sports activities
G1202	Animation d'activités culturelles ou ludiques	Animation of cultural or fun activities
G1203	Animation de loisirs auprès d'enfants ou d'adolescents	Leisure activities with children or adolescents
G1204	Éducateur sportif / éducatrice sportive	Sports educator
G1206	Personnel technique des jeux	Technical Games Staff
G1208	Entraîneur / entraîneuse de sportifs de haut niveau	Coach of professional athletes
G1209	Animateur de loisirs sportifs / animatrice de loisirs sportifs	Sports leisure facilitator regulated employment
G1301	Conception de produits touristiques	Design of tourist products
G1302	Optimisation de produits touristiques	Optimization of tourism products
G1303	Vente de voyages	Travel agency
G1401	Assistance de direction d'hôtel-restaurant	Hotel/Restaurant assistance
G1402	Management d'hôtel-restaurant	Hotel Management
G1403	Gestion de structure de loisirs ou d'hébergement touristique	Leisure structure or tourist accommodation management
G1404	Management d'établissement de restauration collective	Management of catering establishments
G1503	Management du personnel d'étage	Floor staff management
G1601	Chef / cheffe de cuisine	Chef / Kitchen chef
G1701	Conciergerie en hôtellerie	Hotel concierge
G1703	Réception en hôtellerie	Hotel reception
G1802	Management du service en restauration	Catering service management
G1804	Sommellerie	Sommellerie
H1101	Assistance et support technique client	Customer assistance and technical support

Continued on next page

Table 1: Occupations in the cluster “Non-Routine Cognitive" (Continued)

ROME	Occupation in French	Occupation in English
H1102	Management et ingénierie d'affaires	Management and business engineering
H1201	Expertise technique couleur en industrie	Technical color expertise in industry
H1202	Conception et dessin de produits électriques et électroniques	Design and drawing of electrical and electronic products
H1203	Conception et dessin produits mécaniques	Mechanical product design and drawing
H1204	Design industriel	Design industrial
H1205	Études - modèles en industrie des matériaux souples	Research in the soft materials industry
H1206	Management et ingénierie études, recherche et développement industriel	Management and engineering studies, research and industrial development
H1207	Rédaction technique	Technical writing
H1208	Intervention technique en études et conception en automatisme	Technical intervention in studies and design in automation
H1209	Intervention technique en études et développement électronique	Technical intervention in studies and electronic development
H1210	Intervention technique en études, recherche et développement	Technical intervention in studies, research and development
H1301	Inspection de conformité	Compliance inspection
H1302	Management et ingénierie hygiène sécurité environnement	Management and engineering Hygiene Safety Environment
H1303	Intervention technique en hygiène sécurité environnement	Technical intervention in hygiene Safety environment
H1401	Management et ingénierie gestion industrielle et logistique	Management and engineering Industrial and Logistics Management
H1402	Management et ingénierie méthodes et industrialisation	Management and engineering methods and industrialization
H1403	Intervention technique en gestion industrielle et logistique	Technical intervention in industrial and logistics management
H1404	Intervention technique en méthodes et industrialisation	Technical intervention in methods and industrialization
H1501	Direction de laboratoire d'analyse industrielle	Industrial analysis laboratory department
H1502	Management et ingénierie qualité industrielle	Industrial quality management and engineering
H1503	Intervention technique en laboratoire d'analyse industrielle	Technical intervention in the industrial analysis laboratory

Continued on next page

Table 1: Occupations in the cluster “Non-Routine Cognitive" (Continued)

ROME	Occupation in French	Occupation in English
H1504	Intervention technique en contrôle essai qualité en électricité et électronique	Technical intervention in quality test in electricity and electronics
H1505	Intervention technique en formulation et analyse sensorielle	Technical intervention in formulation and sensory analysis
H1506	Intervention technique qualité en mécanique et travail des métaux	Technical intervention quality in mechanics and metal work
H2204	Encadrement des industries de l'ameublement et du bois	Supervision of furniture and wood industries
H2209	Intervention technique en ameublement et bois	Technical intervention in furnishings and woods
H2301	Conduite d'équipement de production chimique ou pharmaceutique	Driving of chemical or pharmaceutical production equipment
H2412	Patronnage - gradation	Pattern making - grading
H2501	Encadrement de production de matériel électrique et électronique	Management of production of electrical and electronic equipment
H2502	Management et ingénierie de production	Management and production engineering
H2503	Pilotage d'unité élémentaire de production mécanique ou de travail des métaux	Management of elementary unit for mechanical production or metal work
H2504	Encadrement d'équipe en industrie de transformation	Team supervision in processing industry
H2505	Encadrement d'équipe ou d'atelier en matériaux souples	Team supervision or workshop in flexible materials
H2701	Pilotage d'installation énergétique et pétrochimique	Energy and petrochemical installation
H2804	Pilotage de centrale à béton prêt à l'emploi, ciment, enrobés et granulats	Piloting of ready -to -use concrete center, cement, coated and aggregates
H2805	Pilotage d'installation de production verrière	Verrière production installation installation
H2906	Conduite d'installation automatisée ou robotisée de fabrication mécanique	Automated or robotic installation of mechanical manufacturing
I1101	Direction et ingénierie en entretien infrastructure	Direction and engineering in Infrastructure
I1102	Management et ingénierie de maintenance industrielle	Management and industrial maintenance engineering
I1103	Supervision d'entretien et gestion de véhicules	Maintenance supervision and vehicle management
I1302	Installation et maintenance d'automatismes	Installation and maintenance of automation

Continued on next page

Table 1: Occupations in the cluster “Non-Routine Cognitive" (Continued)

ROME	Occupation in French	Occupation in English
I1305	Installation et maintenance électronique	Electronic installation and maintenance
I1401	Maintenance informatique et bureautique	IT and office automation maintenance
J1101	Médecine de prévention	Preventive medicine
J1102	Médecine généraliste et spécialisée	General and specialized medicine
J1103	Médecine dentaire	Dental medicine
J1104	Suivi de la grossesse et de l’accouchement	Follow -up of pregnancy and childbirth
J1201	Biologie médicale	Medical biology
J1202	Pharmacie	Pharmacy
J1302	Analyses médicales	Medical analysis
J1306	Imagerie médicale	Medical imaging
J1307	Préparation en pharmacie	Pharmacy preparation
J1401	Audioprothèses	Hearing aids
J1402	Diététique	Dietetics
J1403	Ergothérapie	Occupational therapy
J1405	Optique - lunetterie	Optics - eyewear
J1406	Orthophonie	Speech therapy
J1407	Orthoptique	Orthoptic
J1408	Ostéopathie et chiropraxie	Osteopathy and chiropraxis
J1410	Prothèses dentaires	Dental prostheses
J1411	Prothèses et orthèses	Prostheses and orthotics
J1412	Rééducation en psychomotricité	Rehabilitation in psychomotricity
J1502	Coordination de services médicaux ou paramédicaux	Coordination of medical or paramedical services
J1503	Soins infirmiers spécialisés en anesthésie	Nursing specialized in anesthesia
J1504	Soins infirmiers spécialisés en bloc opératoire	Nursing specialized in operating room
J1505	Soins infirmiers spécialisés en prévention	Nursing specialized in prevention
J1506	Soins infirmiers généralistes	General nursing
J1507	Soins infirmiers spécialisés en puériculture	Nursing specialized in childcare
K1101	Accompagnement et médiation familiale	Support and family mediation
K1102	Aide aux bénéficiaires d’une mesure de protection juridique	Assistance to beneficiaries of a legal protection measure
K1103	Développement personnel et bien-être de la personne	Personal development and well-being of the person

Continued on next page

Table 1: Occupations in the cluster “Non-Routine Cognitive" (Continued)

ROME	Occupation in French	Occupation in English
K1104	Psychologie	Psychology
K1201	Action sociale	Social work
K1202	Éducation de jeunes enfants	Youth education
K1203	Encadrement technique en insertion professionnelle	Technical supervision in professional integration
K1204	Médiation sociale et facilitation de la vie en société	Social mediation and facilitation of life in society
K1205	Information sociale	Social Information Service
K1206	Intervention socioculturelle	Socio -cultural intervention
K1207	Intervention socioéducative	Socio -educational intervention
K1301	Accompagnement médicosocial	Medico -social support
K1303	Assistance auprès d'enfants	Children's assistance
K1305	Intervention sociale et familiale	Social and family intervention
K1401	Conception et pilotage de la politique des pouvoirs publics	Design and management of the public authorities
K1402	Conseil en santé publique	Public health advice
K1403	Management de structure de santé, sociale ou pénitentiaire	Health, social or penitentiary structure management
K1404	Mise en oeuvre et pilotage de la politique des pouvoirs publics	Implementation and management of public authorities policy
K1405	Représentation de l'état sur le territoire national ou international	Representation of the State on the national or international territory
K1502	Contrôle et inspection des affaires sociales	Control and inspection of social affairs
K1503	Contrôle et inspection des impôts	Tax control and inspection
K1504	Contrôle et inspection du trésor public	Treasury control and inspection
K1505	Protection des consommateurs et contrôle des échanges commerciaux	Consumer protection and trade control
K1601	Gestion de l'information et de la documentation	Information and documentation management
K1602	Gestion de patrimoine culturel	Cultural heritage management
K1702	Direction de la sécurité civile et des secours	Directorate of Civil Security and Rescue
K1703	Direction opérationnelle de la défense	Operational Department of Defense
K1704	Management de la sécurité publique	Public security management
K1706	Sécurité publique	Public security
K1707	Surveillance municipale	Municipal surveillance
K1801	Conseil en emploi et insertion socio-professionnelle	Employment advice and socio -professional integration
K1802	Développement local	Local development
K1901	Aide et médiation judiciaire	Judicial aid and mediation

Continued on next page

Table 1: Occupations in the cluster “Non-Routine Cognitive" (Continued)

ROME	Occupation in French	Occupation in English
K1902	Collaboration juridique	Legal collaboration
K1903	Défense et conseil juridique	Defense and legal advice
K1904	Magistrature	Magistracy
K2101	Conseil en formation	Training Advice
K2102	Coordination pédagogique	Educational coordination
K2103	Direction d'établissement et d'enseignement	Establishment and education department
K2104	Éducation et surveillance au sein d'établissements d'enseignement	Education and surveillance within educational establishments
K2105	Enseignement artistique	Artistic education
K2106	Enseignement des écoles	School teaching
K2107	Enseignement général du second degré	General second degree education
K2108	Enseignement supérieur	Higher education
K2109	Enseignement technique et professionnel	Technical and vocational education
K2110	Formation en conduite de véhicules	Training in vehicle driving
K2111	Formation professionnelle	Vocational training
K2112	Orientation scolaire et professionnelle	School and professional orientation
K2203	Management et inspection en propreté de locaux	Management and cleaning inspection of premises
K2302	Management et inspection en environnement urbain	Management and inspection in an urban environment
K2306	Supervision d'exploitation éco-industrielle	Eco-industrial operating supervision
K2401	Recherche en sciences de l'homme et de la société	Research in human and society sciences
K2402	Recherche en sciences de l'univers, de la matière et du vivant	Research in universe, matter and living sciences
K2502	Management de sécurité privée	Private security management
K2602	Conseil en services funéraires	Funeral Services Advice
L1101	Animation musicale et scénique	Musical and scenic animation
L1102	Mannequinat et pose artistique	Modeling and artistic pose
L1103	Présentation de spectacles ou d'émissions	Presentation of shows or programs
L1201	Danse	Dance
L1202	Musique et chant	Music and song
L1203	Art dramatique	Drama
L1301	Mise en scène de spectacles vivants	Direction of living shows
L1302	Production et administration spectacle, cinéma et audiovisuel	Production and administration show, cinema and audiovisual
L1303	Promotion d'artistes et de spectacles	Promotion of artists and shows

Continued on next page

Table 1: Occupations in the cluster “Non-Routine Cognitive" (Continued)

ROME	Occupation in French	Occupation in English
L1304	Réalisation cinématographique et audiovisuelle	Cinematographic and audiovisual realization
L1501	Coiffure et maquillage spectacle	Hairstyle and makeup
L1502	Costume et habillage spectacle	Costume and dressing
L1503	Décor et accessoires spectacle	Decor and Accessories
L1504	Éclairage spectacle	Lighting show
L1505	Image cinématographique et télévisuelle	Cinematographic and television image
L1507	Montage audiovisuel et post-production	Audiovisual assembly and post-production
L1508	Prise de son et sonorisation	Sound and sound recording
L1509	Régie générale	General management
L1510	Films d'animation et effets spéciaux	Animated films and special effects
M1101	Achats	Purchases
M1102	Direction des achats	Purchasing management
M1201	Analyse et ingénierie financière	Financial analysis and engineering
M1202	Audit et contrôle comptables et financiers	Accounting and financial audit and control
M1204	Contrôle de gestion	Management control
M1205	Direction administrative et financière	Administrative and Financial Department
M1206	Management de groupe ou de service comptable	Group or accounting service management
M1207	Trésorerie et financement	Treasury and financing
M1301	Direction de grande entreprise ou d'établissement public	Directorate of Grande Entreprise or Public Establishment
M1302	Direction de petite ou moyenne entreprise	Small or medium business management
M1401	Conduite d'enquêtes	Conducting surveys
M1402	Conseil en organisation et management d'entreprise	Business organization and management advice
M1403	Études et prospectives socio-économiques	Socio-economic studies and prospective
M1404	Management et gestion d'enquêtes	Management and management of surveys
M1502	Développement des ressources humaines	Human Resources Development
M1503	Management des ressources humaines	Human Resources Management
M1701	Administration des ventes	Sales administration
M1702	Analyse de tendance	Trend analysis

Continued on next page

Table 1: Occupations in the cluster “Non-Routine Cognitive" (Continued)

ROME	Occupation in French	Occupation in English
M1703	Management et gestion de produit	Management and product management
M1704	Management relation clientèle	Customer relations management
M1705	Marketing	Marketing
M1706	Promotion des ventes	Sales promotion
M1707	Stratégie commerciale	Business strategy
M1801	Administration de systèmes d'information	Information systems administration
M1802	Expertise et support en systèmes d'information	Expertise and support in information systems
M1803	Direction des systèmes d'information	Information Systems Department
M1804	Études et développement de réseaux de télécoms	Studies and development of telecoms networks
M1805	Études et développement informatique	Studies and IT development
M1806	Conseil et maîtrise d'ouvrage en systèmes d'information	Advice and project management in information systems
M1807	Exploitation de systèmes de communication et de commandement	Exploitation of communication and command systems
M1808	Information géographique	Geographic information
M1809	Information météorologique	Meteorological information
M1810	Production et exploitation de systèmes d'information	Production and operation of information systems
N1201	Affrètement transport	Transport charter
N1202	Gestion des opérations de circulation internationale des marchandises	Management of international traffic operations
N1301	Conception et organisation de la chaîne logistique	Design and organization of the logistics chain
N1302	Direction de site logistique	Logistics site management
N1303	Intervention technique d'exploitation logistique	Technical logistics operating intervention
N2101	Navigation commerciale aérienne	Air commercial navigation
N2102	Pilotage et navigation technique aérienne	Aerial management and technical navigation
N2201	Personnel d'escale aéroportuaire	Airport stopover staff
N2202	Contrôle de la navigation aérienne	Air navigation control
N2204	Préparation des vols	Flight preparation
N2205	Direction d'escale et exploitation aéroportuaire	Stopover management and airport operation
N3101	Encadrement de la navigation maritime	Maritime navigation supervision

Continued on next page

Table 1: Occupations in the cluster “Non-Routine Cognitive" (Continued)

ROME	Occupation in French	Occupation in English
N3201	Exploitation des opérations portuaires et du transport maritime	Exploitation of port operations and maritime transport
N3202	Exploitation du transport fluvial	Operating river transport
N4201	Direction d'exploitation des transports routiers de marchandises	Directorate to operate goods from goods
N4202	Direction d'exploitation des transports routiers de personnes	Directorate of operating transport of people
N4203	Intervention technique d'exploitation des transports routiers de marchandises	Technical intervention for the operating transport of goods of goods
N4204	Intervention technique d'exploitation des transports routiers de personnes	Technical intervention to operate road transport of people
N4301	Conduite sur rails	Rail driving
N4401	Circulation du réseau ferré	Rail network circulation
A1304	Conseiller / conseillère technique agricole	Agricultural technical advisor
A1420	Chef / cheffe de culture responsable d'unité de production agricole	Head of culture responsible for agricultural production unit
E1206	Ux - ui designer	UX - UI designer
G1210	Enseignant / enseignante d'équitation	Horseback Riding Teacher
J1105	Médecin coordonnateur	Coordinating physician
K1307	Animateur / animatrice petite enfance	Early childhood animator
M1405	Data scientist	Data Scientist
M1811	Data engineer	Data Engineer
M1812	Responsable de la sécurité des systèmes d'information (rssi)	Responsible for information systems security
M1813	Intégrateur / intégratrice logiciels métiers	Business software integrator

Table 2: Occupations in the cluster “Non-Routine Manual"

ROME	Occupation in French	Occupation in English
A1101	Conduite d'engins agricoles et forestiers	Conducting agricultural and forestry machinery
A1201	Bûcheronnage et élagage	Logging and pruning
A1202	Entretien des espaces naturels	Maintenance of natural spaces
A1203	Aménagement et entretien des espaces verts	Development and maintenance of green spaces
A1204	Protection du patrimoine naturel	Natural heritage protection
A1205	Sylviculture	Forestry

Continued on next page

Table 2: Occupations in the cluster “Non-Routine Manual" (Continued)

ROME	Occupation in French	Occupation in English
A1401	Aide agricole de production fruitière ou viticole	Fruit or wine production agricultural assistance
A1402	Aide agricole de production légumière ou végétale	Agricultural vegetable or vegetable production assistance
A1403	Aide d'élevage agricole et aquacole	Agricultural and aquaculture farming aid
A1404	Aquaculture	Aquaculture
A1405	Arboriculture et viticulture	Arboriculture and viticulture
A1407	Élevage bovin ou équin	Cattle or horse breeding
A1408	Élevage d'animaux sauvages ou de compagnie	Breeding of wild or companion animals
A1409	Élevage de lapins et volailles	Rabbits and poultry farming
A1410	Élevage ovin ou caprin	Sheep farming or goat
A1411	Élevage porcin	Porcine farm
A1412	Fabrication et affinage de fromages	Manufacturing and refining cheeses
A1414	Horticulture et maraîchage	Horticulture and market gardening
A1415	Équipage de la pêche	Fishing crew
A1416	Polyculture, élevage	Polyculture, breeding
A1417	Saliculture	Salt production
A1501	Aide aux soins animaux	Animal care assistance
A1502	Podologie animale	Animal podiatry
A1503	Toilettage des animaux	Animal grooming
B1201	Réalisation d'objets décoratifs et utilitaires en céramique et matériaux de synthèse	Realization of decorative and utility objects in ceramic and synthetic materials
B1302	Décoration d'objets d'art et artisanaux	Decoration of art and crafts
B1303	Gravure - ciselure	Engraving
B1401	Réalisation d'objets en lianes, fibres et brins végétaux	Realization of objects in lianas, fibers and plant strands
B1402	Reliure et restauration de livres et archives	Binding and restoration of books and archives
B1501	Fabrication et réparation d'instruments de musique	Manufacturing and repair of music instruments
B1601	Métallerie d'art	Art metalwork
B1602	Réalisation d'objets artistiques et fonctionnels en verre	Realization of artistic and functional glass objects
B1603	Réalisation d'ouvrages en bijouterie, joaillerie et orfèvrerie	Realization of works in jewelry, jewelry and goldsmithery
B1604	Réparation - montage en systèmes horlogers	Repair - Mounting in watchmaking systems

Continued on next page

Table 2: Occupations in the cluster “Non-Routine Manual" (Continued)

ROME	Occupation in French	Occupation in English
B1701	Conservation et reconstitution d'espèces animales	Conservation and reconstruction of animal species
B1802	Réalisation d'articles en cuir et matériaux souples (hors vêtement)	Creation of leather items and flexible materials (excluding clothing)
B1803	Réalisation de vêtements sur mesure ou en petite série	Creation of custom clothes or in small series
B1804	Réalisation d'ouvrages d'art en fils	Creation of art structures
B1806	Tapisserie - décoration en ameublement	Tapestry - Decoration in furnishings
D1101	Boucherie	Butchery
D1102	Boulangerie - viennoiserie	Bakery
D1103	Charcuterie - traiteur	Butchery
D1104	Pâtisserie, confiserie, chocolaterie et glacerie	Pastry, confectionery, chocolate factory and ice cream
D1105	Poissonnerie	Fishmonger
D1106	Vente en alimentation	Food sale
D1202	Coiffure	Coiffure
D1203	Hydrothérapie	Hydrotherapy
D1206	Réparation d'articles en cuir et matériaux souples	Repair of leather and flexible materials
D1207	Retouches en habillement	Alterations in clothing
D1208	Soins esthétiques et corporels	Aesthetic and bodily care
D1209	Vente de végétaux	Sale of plants
E1202	Production en laboratoire cinématographique	Cinematographic laboratory production
E1204	Projection cinéma	Cinema projection
F1301	Conduite de grue	Crane driving
F1302	Conduite d'engins de terrassement et de carrière	Driving of earthworks and career
F1401	Extraction liquide et gazeuse	Liquid and gaseous extraction
F1402	Extraction solide	Solid extraction
F1501	Montage de structures et de charpentes bois	Mounting wooden structures and frames
F1502	Montage de structures métalliques	Mounting metal structures
F1503	Réalisation - installation d'ossatures bois	Installation of wood frames
F1601	Application et décoration en plâtre, stuc et staff	Application and decoration in plaster, stucco and staff
F1602	Électricité bâtiment	Building electricity
F1603	Installation d'équipements sanitaires et thermiques	Installation of sanitary and thermal equipment

Continued on next page

Table 2: Occupations in the cluster “Non-Routine Manual" (Continued)

ROME	Occupation in French	Occupation in English
F1604	Montage d'agencements	Fitting of layouts
F1605	Montage de réseaux électriques et télécoms	Mounting electrical and telecoms networks
F1606	Peinture en bâtiment	Building painting
F1607	Menuisier / menuisière aluminium	Aluminum carpenter
F1608	Pose de revêtements rigides	Laying of rigid coatings
F1609	Pose de revêtements souples	Installation of flexible coverings
F1610	Pose et restauration de couvertures	Installation and restoration of blankets
F1611	Réalisation et restauration de façades	Creation and restoration of facades
F1612	Taille et décoration de pierres	Decoration
F1613	Travaux d'étanchéité et d'isolation	Sealing and insulation work
F1614	Pose de fermetures menuisées	Laying of carpeated closures
F1615	Poseur / poseuse de cloisons démontables et mobiles	Installer of demountable and mobile partitions
F1616	Poseur / poseuse de menuiseries extérieures	Carpentry
F1617	Poseur / poseuse de véranda	Veranda installer
F1618	Poseur / poseuse de façade vitrée	Facade installer
F1701	Construction en béton	Concrete worker
F1702	Construction de routes et voies	Construction of roads and tracks
F1703	Maçonnerie	Masonry
F1704	Préparation du gros oeuvre et des travaux publics	Preparation of structural work and public works
F1705	Pose de canalisations	Pipe laying
F1706	Préfabrication en béton industriel	Industrial concrete prefabrication
G1205	Personnel d'attractions ou de structures de loisirs	Personal attractions or leisure structures
G1207	Animation d'activités sportives	Animation of sports activities
G1502	Personnel polyvalent d'hôtellerie	Versatile hotel staff
G1602	Personnel de cuisine	Cooking staff
G1702	Personnel du hall	Hall staff
G1801	Café, bar brasserie	Coffee, Bar owner
G1803	Service en restauration	Catering service
H2205	Première transformation de bois d'oeuvre	First transformation of sawn timber
H2206	Réalisation de menuiserie bois et tonnellerie	Wood carpentry and cooperage
H2207	Réalisation de meubles en bois	Wooden furniture making
H2208	Réalisation d'ouvrages décoratifs en bois	Decorative wooden works

Continued on next page

Table 2: Occupations in the cluster “Non-Routine Manual" (Continued)

ROME	Occupation in French	Occupation in English
H2401	Assemblage - montage d'articles en cuirs, peaux	Assembly of articles in leathers, skins
H2407	Conduite de machine de transformation et de finition des cuirs et peaux	Transformation and finishing machine pipe and skins
H2408	Conduite de machine d'impression textile	Textile printing machine driving
H2409	Coupe cuir, textile et matériaux souples	Leather cut, textiles and flexible materials
H2411	Montage de prototype cuir et matériaux souples	Leather prototype assembly and flexible materials
H2414	Préparation et finition d'articles en cuir et matériaux souples	Preparation and finish of leather items and flexible materials
H2415	Contrôle en industrie du cuir et du textile	Control in leather and textile industry
H2601	Bobinage électrique	Electrical winding
H2602	Câblage électrique et électromécanique	Electric and electromechanical wiring
H2803	Façonnage et émaillage en industrie céramique	Shaping and enameling in ceramic industry
H2901	Ajustement et montage de fabrication	Manufacturing adjustment and assembly
H2902	Chaudronnerie - tôlerie	Boilermaking - sheet metal
H2903	Conduite d'équipement d'usinage	Machining equipment driving
H2904	Conduite d'équipement de déformation des métaux	Metal deformation equipment conduct
H2907	Conduite d'installation de production des métaux	Metal production installation conduct
H2908	Modelage de matériaux non métalliques	Non-metallic material modeling
H2910	Moulage sable	Sand molding
H2911	Réalisation de structures métalliques	Creation of metal structures
H2912	Réglage d'équipement de production industrielle	Industrial production equipment adjustment
H2913	Soudage manuel	Manual welding
H2914	Réalisation et montage en tuyauterie	Piping realization and assembly
H3203	Fabrication de pièces en matériaux composites	Manufacture of parts in composite materials
H3401	Conduite de traitement d'abrasion de surface	Surface abrasion treatment conduct
H3403	Conduite de traitement thermique	Heat treatment conduct
H3404	Peinture industrielle	Industrial paint

Continued on next page

Table 2: Occupations in the cluster “Non-Routine Manual" (Continued)

ROME	Occupation in French	Occupation in English
I1201	Entretien d'affichage et mobilier urbain	Display maintenance and urban furniture
I1202	Entretien et surveillance du tracé routier	Maintenance and monitoring of the road route
I1203	Maintenance des bâtiments et des locaux	Maintenance of buildings and premises
I1301	Installation et maintenance d'ascenseurs	Installation and maintenance of elevators
I1303	Installation et maintenance de distributeurs automatiques	Installation and maintenance of automatic distributors
I1304	Installation et maintenance d'équipements industriels et d'exploitation	Installation and maintenance of industrial and operating equipment
I1306	Installation et maintenance en froid, conditionnement d'air	Cold installation and maintenance, air conditioning
I1307	Installation et maintenance télécoms et courants faibles	Telecom installation and maintenance and low currents
I1308	Maintenance d'installation de chauffage	Heating installation maintenance
I1309	Maintenance électrique	Electrical maintenance
I1310	Maintenance mécanique industrielle	Industrial mechanical maintenance
I1402	Réparation de biens électrodomestiques et multimédia	Repair of electrodode and multimedia goods
I1501	Intervention en grande hauteur	Great height intervention
I1502	Intervention en milieu subaquatique	Underwater intervention
I1503	Intervention en milieux et produits nocifs	Intervention in harmful environments and products
I1601	Installation et maintenance en nautisme	Installation and maintenance in boating
I1602	Maintenance d'aéronefs	Aircraft maintenance
I1603	Maintenance d'engins de chantier, levage, manutention et de machines agricoles	Maintenance of construction machinery, lifting, handling and agricultural machines
I1604	Mécanique automobile et entretien de véhicules	Automotive mechanics and vehicle maintenance
I1605	Mécanique de marine	Naval mechanics
I1606	Réparation de carrosserie	Body repair
I1607	Réparation de cycles, motocycles et motoculteurs de loisirs	Repair of cycles, motorcycles and leisure motorcycles
J1301	Personnel polyvalent des services hospitaliers	Multipurpose hospital staff
J1303	Assistance médico-technique	Medical-technical assistance

Continued on next page

Table 2: Occupations in the cluster “Non-Routine Manual" (Continued)

ROME	Occupation in French	Occupation in English
J1304	Aide en puériculture	Childcare assistance
J1305	Conduite de véhicules sanitaires	Driving of health vehicles
J1404	Kinésithérapie	Physiotherapy
J1409	Pédicurie et podologie	Pedicuria and podology
J1501	Soins d'hygiène, de confort du patient	Hygiene, patient comfort care
K1302	Assistance auprès d'adultes	Assistance to adults
K1304	Services domestiques	Domestic services
K1701	Personnel de la défense	Defense staff
K1705	Sécurité civile et secours	Civil security and help
K2202	Lavage de vitres	Window washing
K2301	Distribution et assainissement d'eau	Distribution and sanitation of water
K2303	Nettoyage des espaces urbains	Cleaning urban spaces
K2304	Revalorisation de produits industriels	Revaluation of industrial products
K2305	Salubrité et traitement de nuisibles	Healthy and treatment of harmful
K2501	Gardiennage de locaux	Guarding of premises
K2503	Sécurité et surveillance privées	Private security and surveillance
K2601	Conduite d'opérations funéraires	Funeral operations conduct
K2603	Thanatopraxie	Thanatopraxy
L1204	Arts du cirque et arts visuels	Circus and visual arts
L1401	Sportif professionnel / sportive professionnelle	Professional sportsman
L1506	Machinerie spectacle	Show machinery
N1101	Conduite d'engins de déplacement des charges	Carrying of load displacement of loads
N1102	Déménagement	Moving
N1104	Manoeuvre et conduite d'engins lourds de manutention	Maneuver and driving of heavy handling machines
N2203	Exploitation des pistes aéroportuaires	Operating airport tracks
N3102	Équipage de la navigation maritime	Maritime navigation crew
N3103	Navigation fluviale	Navigation fluviale
N3203	Manutention portuaire	Port handling
N4101	Conduite de transport de marchandises sur longue distance	Long -distance freight transportation driving
N4102	Conduite de transport de particuliers	Transportation of individuals
N4103	Conduite de transport en commun sur route	Road public transport driving
N4104	Courses et livraisons express	Express races and deliveries
N4105	Conduite et livraison par tournées sur courte distance	Driving and delivery on short distance tours
N4302	Contrôle des transports en commun	Public transportation control

Continued on next page

Table 2: Occupations in the cluster “Non-Routine Manual" (Continued)

ROME	Occupation in French	Occupation in English
N4402	Exploitation et manoeuvre des remontées mécaniques	Operation and handling of ski lifts
N4403	Manoeuvre du réseau ferré	Rail network handling
A1102	Conducteur / conductrice d'engins d'exploitation forestière	Driver of Forest operating devices
A1418	Viticulteur / viticultrice	Winegrower
A1419	Ouvrier agricole polyvalent / ouvrière agricole polyvalente	Versatile agricultural worker
A1421	Palefrenier soigneur / palefrenière soigneuse	Groom
F1619	Couvreur / couvreuse	Roofer
F1707	Maçon / maçonne du paysage	Landscape mason
F1708	Installateur-poseur / installatrice-poseuse de piscines préfabriquées	Installer of prefabricated swimming pools
F1709	Démolisseur / démolisseuse	Demolisher
G1606	Cuisinier / cuisinière de collectivité	Community cook
J1308	Brancardier / brancardière	Stretcher bearer
K1306	Auxiliaire de vie sociale (avs)	Social life assistant
K1308	Agent territorial spécialisé / agente territoriale spécialisée des écoles maternelles (atsem)	Specialized territorial agent for nursery schools

Table 3: Occupations in the cluster “Routine Cognitive"

ROME	Occupation in French	Occupation in English
C1109	Rédaction et gestion en assurances	Drafting and insurance management
C1201	Accueil et services bancaires	Home and banking services
C1401	Gestion en banque et assurance	Bank management and insurance
D1401	Assistant commercial / assistante commerciale	Sales assistant
D1408	Téléconseil et télévente	Television and televenture
D1409	Gestion administrative des ventes	Administrative sales management
K1501	Application des règles financières publiques	Application of public financial rules
M1203	Comptabilité	Accounting
M1501	Assistanat en ressources humaines	Human resources assistantship
M1601	Accueil et renseignements	Home and information
M1602	Opérations administratives	Administrative operations
M1604	Assistanat de direction	Management assistant

Continued on next page

Table 3: Occupations in the cluster “Routine Cognitive" (Continued)

ROME	Occupation in French	Occupation in English
M1605	Assistanat technique et administratif	Technical and administrative assistantship
M1606	Saisie de données	Data entry
M1607	Secrétariat	Secretariat
M1608	Secrétariat comptable	Accounting secretariat
M1609	Secrétariat et assistanat médical ou médico-social	Secretariat and medical or medico-social assistant

Table 4: Occupations in the cluster “Routine Manual"

ROME	Occupation in French	Occupation in English
D1205	Nettoyage d’articles textiles ou cuirs	Cleaning of textiles or leathers
D1505	Personnel de caisse	Cashier
D1507	Mise en rayon libre-service	Shelf stocking self-service
E1301	Conduite de machines d’impression	Printing machine operation
E1302	Conduite de machines de façonnage routage	Routing machine operation
E1304	Façonnage et routage	Shaping and routing
E1307	Reprographie	Reprography
G1501	Personnel d’étage	Floor staff
G1603	Personnel polyvalent en restauration	Multipurpose staff in catering
G1604	Fabrication de crêpes ou pizzas	Manufacture of pancakes or pizzas
G1605	Plonge en restauration	Diving in catering
H2101	Abattage et découpe des viandes	Molding and cutting of meats
H2102	Conduite d’équipement de production alimentaire	Operation of food production equipment
H2201	Assemblage d’ouvrages en bois	Assembly of wooden structures
H2202	Conduite d’équipement de fabrication de l’ameublement et du bois	Building equipment for furnishing and wood manufacturing
H2203	Conduite d’installation de production de panneaux bois	Installation of wood panel production
H2402	Assemblage - montage de vêtements et produits textiles	Assembly of clothing and textile products
H2403	Conduite de machine de fabrication de produits textiles	Operation of machine for the manufacture of textile products
H2404	Conduite de machine de production et transformation des fils	Production machine and transformation of wires
H2405	Conduite de machine de textiles non-tissés	Operation of a nonwoven textile machine

Continued on next page

Table 4: Occupations in the cluster “Routine Manual" (Continued)

ROME	Occupation in French	Occupation in English
H2406	Conduite de machine de traitement textile	Textile processing machine pipe
H2410	Mise en forme, repassage et finitions en industrie textile	Formatting, ironing and finishes in textile industry
H2413	Préparation de fils, montage de métiers textiles	Preparation of wires, assembly of textile trades
H2603	Conduite d'installation automatisée de production électrique, électronique et microélectronique	Electric, electronic and microelectronic automated installation of power production
H2604	Montage de produits électriques et électroniques	Electrical and electronic products
H2605	Montage et câblage électronique	Electronic assembly and wiring
H2801	Conduite d'équipement de transformation du verre	Glass transformation equipment driving
H2802	Conduite d'installation de production de matériaux de construction	Installation of construction of building materials
H2905	Conduite d'équipement de formage et découpage des matériaux	Forming equipment and cutting of materials
H2909	Montage-assemblage mécanique	Mechanical assembly
H3101	Conduite d'équipement de fabrication de papier ou de carton	Building paper or cardboard manufacturing equipment
H3102	Conduite d'installation de pâte à papier	Operation of a paper pulp installation
H3201	Conduite d'équipement de formage des plastiques et caoutchoucs	Operation of plastic and rubber forming equipment
H3202	Réglage d'équipement de formage des plastiques et caoutchoucs	Plastal forming and rubbers forming equipment adjustment
H3301	Conduite d'équipement de conditionnement	Condition of packaging equipment
H3302	Opérations manuelles d'assemblage, tri ou emballage	Manual assembly operations, sorting or packaging
H3303	Préparation de matières et produits industriels (broyage, mélange, ...)	Preparation of industrial materials and products (grinding, mixture, ...)
H3402	Conduite de traitement par dépôt de surface	Surface deposition treatment operation
K2201	Blanchisserie industrielle	Industrial laundry
K2204	Nettoyage de locaux	Cleaning of premises
M1603	Distribution de documents	Document distribution
N1103	Magasinage et préparation de commandes	Shopping and preparation of orders
N1105	Manutention manuelle de charges	Manual handling of loads

Continued on next page

Table 4: Occupations in the cluster “Routine Manual" (Continued)

ROME	Occupation in French	Occupation in English
G1607	Employé / employée de restauration collective	Collective catering employee
H2103	Opérateur / opératrice de transformation des viandes	Immersion transformation operator
J1309	Agent / agente de stérilisation de service hospitalier	Hospital service agent

ECOLE DOCTORALE

Titre : Modélisation économique par les systèmes complexes : de la simulation multi-agents aux goulets d'étranglement du marché du travail

Mots clés : Modèles d'agents macroéconomiques, Dynamiques de l'inflation, Économie de la complexité

Résumé : En mêlant modélisation multi-agents, science des réseaux et approches des systèmes complexes, ce travail apporte de nouvelles perspectives face aux défis économiques. Les modèles multi-agents (ABM) offrent un cadre d'analyse novateur pour simuler des dynamiques macroéconomiques, s'intéressant aux phénomènes émergents tels que les cycles économiques, l'inflation et les crises. En étendant le modèle Mark-0, cette thèse introduit un mécanisme d'ancrage des anticipations pour analyser l'inflation post-COVID et l'efficacité des politiques économiques qui suivirent. Les résultats montrent que le succès de la politique monétaire repose davantage sur la gestion des anticipations que sur ses effets économiques directs, la spirale prix-salaires apparaissant comme un moteur clé. Les arbitrages nécessaires entre le contrôle de l'inflation et le chômage soulignent l'importance des mesures budgétaires pour assurer la stabilité économique. Afin d'améliorer la robustesse des ABM, une approche inspirée de la physique statistique est introduite pour explorer des espaces de paramètres de haute dimension. En exploitant le cadre de la sloppiness, l'étude identifie un petit ensemble de paramètres rigides qui gouvernent le comportement du modèle, tout en minimisant l'importance des paramètres moins influents, améliorant ainsi son interprétabilité et son efficacité. La deuxième partie se concentre sur les marchés du travail en tant que systèmes complexes adaptatifs. En représentant les professions comme des nœuds dans un réseau de mobilité professionnelle, l'analyse met en évidence des déséquilibres structurels où certaines professions jouent le rôle de hubs centraux tandis que d'autres restent isolées. De nouveaux indicateurs—l'accessibilité et la transférabilité—quantifient la mobilité des travailleurs et identifient les professions condensatrices comme des goulets d'étranglement avec une forte accessibilité mais une faible transférabilité. L'intégration des données de compétences révèle que les professions à forte transférabilité facilitent la mobilité entre clusters d'emplois, tandis que celles à forte accessibilité servent de points d'entrée dans de nouveaux secteurs. Ces résultats offrent des perspectives nouvelles pour améliorer les politiques publiques dans ce domaine. Plus particulièrement, cela permettrait de réduire les inadéquations travail-compétences, d'optimiser les programmes de reconversion et d'améliorer la fluidité du marché du travail face aux transformations économiques et technologiques.

Title : A Complex Systems Perspective on Economic Modelling : From Macroeconomic Agent-Based Models to Labor Market Bottlenecks

Keywords : Agent-based Modelling, Inflation Dynamics, Complexity Economics

Abstract : By integrating agent-based modelling, network science, and complex systems approaches, this work offers new perspectives on economic challenges. Agent-based models (ABMs) provide a bottom-up framework to simulate macroeconomic dynamics, capturing emergent phenomena such as business cycles, inflation, and crises. Extending the Mark-0 ABM, this research incorporates an expectation-anchoring mechanism to analyze post-COVID inflation and policy effectiveness. The findings highlight that monetary policy success depends more on managing expectations than on direct economic effects, with the wage-price spiral emerging as a key driver. Trade-offs between inflation control and unemployment underscore the importance of fiscal measures to sustain economic stability. To enhance the robustness of ABMs, a statistical physics-inspired approach is introduced to explore high-dimensional parameter spaces. Leveraging the sloppiness framework, the study identifies a small set of stiff parameters that govern model behavior while de-emphasizing less influential ones, improving interpretability and efficiency. The second part shifts focus to labor markets as complex adaptive systems. Viewing occupations as nodes in a labor mobility network reveals structural imbalances, where some occupations act as central hubs while others remain isolated. New metrics—accessibility and transferability—quantify worker mobility, identifying condenser occupations as bottlenecks with high accessibility but low transferability. Integrating skill similarity data further reveals that high-transferability occupations facilitate movement across job clusters, while high-accessibility roles serve as entry points into new sectors. These insights inform policies on skill mismatches, retraining programs, and labor market fluidity in the face of economic and technological shifts.

Institut Polytechnique de Paris
91120 Palaiseau, France