Institut d'études politiques de Paris
ÉCOLE DOCTORALE DE SCIENCES PO
Programme doctoral en économie
Département d'Économie
Doctorat en sciences économiques

Search, Matching and Heterogeneity

Julien Pascal

Thesis supervised by
Jean-Marc Robin, Professor of Economics, IEP de Paris and
University College London
Defended on 30 June, 2020

Jury:
M. Fabrice COLLARD, Directeur de recherche, CNRS-TSE (*reviewer*)
M. Olivier C. PIERRARD, Economiste senior, Banque centrale du Luxembourg
M. Xavier RAGOT, Directeur de recherche, CNRS-OFCE
M. Jean-Marc ROBIN, Professeur des universités, IEP de Paris
Mme Thepthida SOPRASEUTH, Professeur des universités, Université de Cergy-Pontoise
M. Etienne WASMER, Professor, New York University Abu Dhabi (*reviewer*)

Acknowledgements

I would like to dedicate this thesis to my parents, Marie-Hélène and Gilles, for their unconditional love and support. I also have a special thought to my grand-father, Roland, for showing me the value of being a lifelong learner.

My most grateful thanks goes to my supervisor Jean-Marc Robin for the wisest words of advice and for his invaluable guidance. At the beginning of my thesis, we had a discussion during which he compared academic research to classical music. To reach his outstanding level as a cellist, even Yo-Yo Ma had to first master the solfeggio, learn simple melodies, before playing increasingly sophisticated musical pieces. When in doubt, I tried to remember this analogy. There is no shame in first having to learn the solfeggio.

I would like to thank Florian Oswald for his constant support throughout my thesis. His expertise on Urban and Computational Economics were pivotal for my research. I am a strong believer in the human-machine synergy, and his knowledge of computers helped me to harness the computational power of the machines. As explained by Judd (1998), while there is no discernible upward trend in human cognitive capacities, improvements in algorithms, software and semiconductor technologies have improved our computing abilities by several orders of magnitude since they were invented. As our computing capabilities keep improving and more data is generated by users and connected devices, it is probably safe to assume that interconnectedness between economists and external computing devices will keep increasing in the future.

I also would like to warmly thank Xavier Ragot for guiding me through the vast and rich Macroeconomics literature. The literature on heterogeneous agent models can be intimidating at first sight. Having the possibility to discuss my research with a world-class expert on heterogeneous agent models was essential for my intellectual progression.

I also would like to thank Jeremy Lise for having me at the University of Minnesota and helping me to sharpen my understanding of state-of-the-art search-and-matching models.

I am especially grateful to Wouter Den Haan for having me at the London School of Economics. Professor Den Haan kindly encouraged me to study wages and how they behave along the business cycle. This suggestion encouraged me to write the first chapter of this

thesis.

I also feel lucky to have had the opportunity to interact with many brilliant PhD students at Sciences Po. I would like to thank them for the interactions we had along the way. While a PhD is often pictured as an individual experience, I really see it as a team effort. This team was composed of people not afraid of climbing the steep stairs leading to "les combles" on a daily basis, the open-space reserved for Econ PhD students at Sciences Po. Here are some, though I am afraid not all, of the people that were part of these interactions: Tyler Abbot, Victor Augias, Vladimir Avetian, Montserrat Botey, Thomas Bourany, Olivier Cassagneau-Francis, Sophie Cetre, Guillaume Chapelle, Pauline Corblet, Pierre Cotterlaz-Carraz, Cyril Couaillier, Florin Cucu, Nicolò Dalvit, Samuel Delpeuch, Pierre Deschamps, Edgard Dewitte, Jean-Benoit Eymeoud, Federico Danilo Filetti, Etienne Fize, Arthur Guillouzouic-le-Corff, Dàniel Gyetvai, Dorian Henricot, Jean-Louis Keene, Gustave Kenedi, Ségal Le Guern Herry, Clémence Lobut, José Rodrigo Lopez-Kolkovsky, Mario Luca, Alaïs Martin-Baillon, Daniel Martins de Almedia Barreto, Clément Mazet-Sonilhac, Julia Mink, Elisa Mougin, Marco Palladino, Ludovic Panon, Asem Patel, Stefan Pauly, Victor Saint-Jean, Jérôme Sansonetti, Jan Sonntag, Charles Louis-Sidois, Camille Urvoy, Quentin Vandeweyer, Lucas Vernet, Paul Vertier, Pierre Villedieu, Max Viskanic, Paulus Wagner and Zydney Wong.

I would like to thank the people at Sciences Po and at the LIEPP who made my PhD experience a smooth experience. This includes, but is not limited to, Guillaume Sarrat de Tramezaigues, Sandrine Le Goff, Pilar Calvo, Melissa Mundell, Claudine Lamaze, Samira Jebli, Latifa Lousao and Juliette Seban.

I could not have completed this endeavor without the constant encouragement of my partner Christina. I would like to thank her for being such a terrific human being. A PhD is an emotional roller coaster. When I hit the bottom, Christina was there to say to the right words to put me back on track. She also helped in polishing my arguments by being a tremendous listener and thinker. She proof-read most of this thesis and detected a unfathomable number of typos.

I would like to acknowledge the help of an army of anonymous, who contributed to this thesis directly or indirectly. I am infinitely grateful to the open source community, loosely defined as the collection of individuals who release, maintain and help to improve free and open source software on the Internet. By releasing high-quality open-source software, they are pushing the world towards the good optimal equilibrium of fruitful collaboration. This thesis was written on a Linux laptop. When complex number crunching was not feasible on my local machine, I used a cluster operated on a Linux server. Thank you Linus Torvalds, and by extension anyone who contributed to this operating system, for creating Linux. This thesis has been written in Latex, created by Leslie Lamport. For computational intensive

tasks, I used Julia (Bezanson et al. (2017)). For statistical analysis, I mainly used R, created by Ross Ihaka and Robert Gentleman. For web scraping and machine-learning analysis, I used Python, created by Guido van Rossum. When in doubt with a technical question, most often than not the answer was already posted on Stack Overflow by an anonymous. All the programming languages mentioned above are made stronger by individuals taking on their free time to write libraries/packages. Little credit is usually given to academics who invest time in writing packages that may be used by other scholars or practitioners. While the social impact of releasing a professional-level package library is certainly high, I am not aware of any University counting the release of a high-level package as a requirement towards reaching tenure. The quality of a contribution could be measured by the number of downloads and/or GitHub/GitLab "stars".

On a side note, I wonder if anyone has attempted to estimate the market value of the open source community. In this doctoral dissertation, I focused on "traditional" market activities: the labor and housing markets. However, I have the feeling that I only saw the tip of an iceberg. If I had to start a PhD again, I would probably attempt the quantification exercise I just mentioned.

Finally, as I am finishing this thesis, the world is in the middle of the Covid-19 outbreak. My thoughts go to the medical professionals currently fighting the war against the invisible enemy.

Foreword

> If I have seen further it is by standing on the shoulders of Giants
>
> *Isaac Newton*

Based on my small-sample experience of trying to explain to an audience of non-specialists what a PhD in Economics is, I came to realize that in the popular mind, the role of an economist is often confounded with the activity of an accountant, a banker or a trader. To clarify the object of this doctoral thesis, I feel compelled to give a definition of what Economics — as a scientific discipline — is. Let me use the definition given by Robbins (2007):

> *"Economics is a science which studies human behavior as a relationship between ends and scarce means which have alternative uses."*

This definition has the advantage of emphasizing that the object of Economics is to analyze how human beings interact. While an economic model may lack many dimensions that make humans human, and that the use of equations might obscure the final objective, the goal of Economists is always to study people and societies.

Having defined what Economics is, I also would like to state the research agenda that I have tried to follow in this doctoral dissertation. Sargent (2015), during a speech delivered at the PASC15 conference, expressed a clear research agenda that guides my own research:

> *"Some people accuse Economists to have Physics envy. And I plead guilty, happily.*
>
> *Because I think the goal is to make Economics more like Physics, in the sense that we want to take models seriously, and take them to the data.*

There is an interaction between the successes and failures of models in explaining the data, and the revisions of the models. That is the business we inspire to be in.

We also have Physics envy in the sense that we are ruthless in importing Mathematics from every place we can get it. We are limited only by our ability. The purpose is to interpret historical data in a way that distinguishes cause from coincidence. [...]

The reason we want to do that is that we want to evaluate the consequences of alternative government policies."

I sincerely hope that the three research papers contained in this thesis contribute to the ambitious program elicited by Thomas Sargent.

Contents

Abstract **17**

Sommaire **18**

Introduction **21**

1 Labor Income Shocks along the Business Cycle **33**

1 Data . . . 37
1.1 Wage dynamics . . . 37
1.2 Labor Income Shocks . . . 41
2 Model . . . 46
2.1 Workers, Firms and Timing . . . 46
2.2 Wage setting and employment flows . . . 48
3 Solving for the wages . . . 50
3.1 The value of a job to workers . . . 51
3.2 Dimension reduction . . . 52
3.3 Evolution of wages . . . 54
3.4 Distribution of (starting) wages . . . 55
4 Parametrization and estimation . . . 56
4.1 Parametrization . . . 56
4.2 Estimated values and model fit . . . 57
4.3 Identification . . . 63
5 Labor income shocks and sorting . . . 64
6 Optimal unemployment insurance . . . 68
7 Conclusion . . . 72
A Numerical details . . . 77
B Algorithm to solve and simulate the model . . . 77
C Accuracy of approximations . . . 77

C.1 Forecasting rules accuracy . . . 78
C.2 Accuracy of approximating $q_t(y)$ using a parametric function . . . 81
C.3 Accuracy of approximating $q_t(y)$ using a histogram . . . 82
D The BKM Algorithm . . . 84
E Flow equations for the joint distribution of wage and employment . . . 85
F Identification . . . 87
G Inference . . . 88
H Extension: Human Capital Accumulation . . . 89

2 Spatial Equilibrium and Transportation Costs 93
1 A Brief History of the Reform . . . 95
2 Data and Empirical Facts . . . 99
3 Sample Selection . . . 104
4 Empirical Specifications and Results . . . 106
4.1 Empirical Specifications . . . 107
4.2 Impact on employment . . . 108
4.3 Discussion on internal migrations . . . 113
4.4 Impact on public transport usage . . . 113
5 Model . . . 114
5.1 Setting and value functions . . . 115
5.2 Flow equations . . . 118
5.3 Equilibrium . . . 119
5.4 Employment effects of the FTZ reform . . . 119
5.5 Algorithm to find an equilibrium . . . 120
5.6 Parametrization and calibration . . . 120
5.7 FTZ Counterfactual . . . 123
6 Conclusion . . . 125
A Data . . . 129
B Pole Emploi Categories . . . 133
C Robustness . . . 134
C.1 Placebo tests . . . 134
C.2 Main regression and categories C - E . . . 138
D Existence and uniqueness of a solution to $U(l)$ . . . 141
D.1 Monotonicity . . . 141
D.2 Discounting . . . 142

3 Rental Housing Market and Directed Search **143**
1 Key facts on the French rental housing market and data collection 146
1.1 The rental housing market in France and in the Paris metropolitan area 146
1.2 Data collection 147
2 Stylized facts 148
2.1 Descriptive statistics 149
2.2 Hedonic pricing model 153
2.3 Residual rent and search behavior 158
2.4 Pricing dynamics 161
3 Model 164
3.1 Setting 165
3.2 The value of searching for tenants 166
3.3 The value of renting for landlords 168
3.4 Competitive search equilibrium 169
3.5 Advertised rent decreases: efficiency and welfare considerations 170
4 Conclusion 171
A Data collection 175
B Maps 177
C Rent control legislation in the Paris metropolitan area 181

Conclusion **183**

Résumé en Français **2**

List of Figures

1.1 Weekly earnings, hourly wage and hours worked . . . 44
1.2 Timing of events . . . 48
1.3 Impulse response function after a productivity shock . . . 60
1.4 Value of net production at the match level $s(x, y, 1)$ . . . 61
1.5 Displaced workers and income losses . . . 66
1.6 Change in labor income risk over the cycle and sorting . . . 68
1.7 Optimal UI . . . 71
1.8 Accuracy of forecasting rules . . . 79
1.9 Percentage error when approximating $q_t(y)$ . . . 82
1.10 Forecasting rules for $\boldsymbol{q_{t+1}}$ . . . 83

2.1 Impact of the reform on travel pass cots . . . 97
2.2 Number of subscribers by travel pass and fare area in 2015 . . . 98
2.3 Average change in commuting costs for travel pass users . . . 98
2.4 Fare areas at the IRIS level . . . 101
2.5 Square meter price and population density . . . 102
2.6 Square meter price and population density . . . 103
2.7 Unemployment rate and distance to the CBD . . . 122
2.8 Model's primitives and distance to the city center . . . 123
2.9 Simulation of the FTZ reform: impact on unemployment . . . 124
2.10 Simulation of the FTZ reform: impact on migration decisions . . . 125
2.11 Public transport networks fare areas in the region Ile-de-France . . . 129
2.12 Standardized square meter price index, 2013 - Q4 . . . 130
2.13 Minimization of the objective function in (2.2) . . . 131
2.14 Municipalities included in the sample . . . 132
2.15 Number of workers registered to Pole Emploi by fare area . . . 133

3.1 Trends in rents and household disposable income in France: 2000 - 2017 . . . 147
3.2 Ads characteristics at the municipality level . . . 152

3.3 Number of photos per ads and distribution of aesthetic scores 155
3.4 Selected sample of photos from the top 1% in terms of aesthetic score 156
3.5 Selected sample of photos from the bottom 1% in terms of aesthetic score . . 157
3.6 Deviation from predicted rent and number of contacts 160
3.7 Advertised rent reductions and number of days before decline 163
3.8 Example of a typical ad on the website 175
3.9 Example of an ad in which the landlord was contacted 32 times 176
3.10 Population in the Paris metropolitan area in 2017 177
3.11 Number of ads 177
3.12 Median rent per m^2 178
3.13 Median surface of advertised accommodation (m^2) 178
3.14 Mean number of rooms 179
3.15 Mean number of bedrooms 179
3.16 Percentage of accommodation rented furnished 180
3.17 Percentage of ads managed by a real estate agency 180

List of Tables

1.1 Elasticity of wage and starting wage deciles in the US: 1984 - 2017 40
1.2 Volatility of wage and starting wage deciles in the US: 1984 - 2017 41
1.3 Regression log wages and weekly hours worked 43
1.4 Probability of changes in hourly wage, weekly earnings and hours along the business cycle 45
1.5 Difference in the probability of changes in hourly wage, weekly earnings and hours 45
1.6 Estimated Parameters 59
1.7 Empirical and Simulated Employment Moments. US (1951-2012) 62
1.8 Elasticity of wage and starting wage deciles 63
1.9 Variations in the probability of labor income changes over the cycle 67
1.10 Variations in the probability of labor income changes over the cycle with the optimal UI 70
1.11 Forecasting rule accuracy for λ_t 80
1.12 Forecasting rule accuracy for a_t 80
1.13 Forecasting rule accuracy for b_t 81
1.14 Jacobian of the function mapping parameter values to simulated moments $f : \boldsymbol{p} \to \boldsymbol{m}$ 87

2.1 Selected city characteristics included in the objective function in (2.2) 106
2.2 Number of workers registered to Pole Emploi 110
2.3 Number of workers registered to Pole Emploi in category A 111
2.4 Number of workers registered to Pole Emploi in category B 112
2.5 Impact of the reform on public transport usage 114
2.6 Calibrated parameters 122
2.7 Placebo test (all categories) 134
2.8 Placebo test category A 135
2.9 Placebo test category B 136

2.10 Number of workers registered to Pole Emploi in category C 138
2.11 Number of workers registered to Pole Emploi in category D 139
2.12 Number of workers registered to Pole Emploi in category E 140

3.1 Dwelling characteristics 150
3.2 Dwelling characteristics by municipality 151
3.3 Hedonic regression model 158
3.4 Truncated regression 161
3.5 Welch t-tests 164

Abstract

This thesis consists of three chapters that study frictional markets. The first chapter asks the question of what are the sources of labor income shocks, with a special focus on the scarring effects of recessions. I develop and estimate a dynamic frictional model of the labor market with heterogeneous workers and firms. The economic contribution of the first chapter is to show that sorting — the degree of complementarity between firms and workers — is a key component of idiosyncratic labor income risk. My technical contribution is to show that, while the determination of wage is *a priori* complex in a dynamic search model with heterogeneity, an efficient and robust algorithm exists.

The second chapter explores to what extent a localized drop in commuting costs may lead to an increase in local employment. This chapter makes use of a discontinuity introduced by a French reform in September 2015 in the Paris metropolitan area. This reform decreased the cost of public transportation in some cities, but not all. I carefully construct control and treatment groups by selecting cities that are geographically close to each other, to limit the impact of unobservable factors on employment. However, the control and treatment groups experienced different changes in commuting costs. I find that cities that enjoyed a decrease in commuting costs experienced an increase in local employment. The effects are concentrated on long-term unemployed workers. To rationalize these findings, I build and calibrate a search-and-matching model with a spatial dimension, to underline the mechanisms at play.

While the first two chapters analyze the labor market, the last chapter focuses on another key frictional market: the housing market. This chapter fills a gap in the empirical literature on the housing market: while several studies exist on the real estate market, little is known on the rental market because there are no comprehensive datasets recording rental agreements. To circumvent this issue, I collected data on rental ads in the Paris metropolitan area using web scraping techniques for a period of three months. One novelty compared to the existing literature is that I also collected data on tenants' search behavior. I show that the rental housing market is well described by a directed search model. However, a non-negligible proportion of landlords use a two-step pricing approach when setting the rent. The landlords rent-setting mechanism raises interesting welfare and modeling questions.

Résumé

Cette thèse est constituée de trois chapitres. Le dénominateur commun de ces chapitres est l'analyse des marchés frictionnels, c'est à dire les marchés pour lesquels le processus de tâtonnement walrasien ne permet pas à l'offre et à la demande de s'équilibrer instantanément.

Le premier chapitre s'intéresse aux origines des fluctuations du revenu du travail au cours du cycle économique. En particulier, ce premier chapitre revisite la question de la persistance des chocs de revenu du travail, appelée *l'effet scarification des récessions* dans la littérature anglo-saxonne. Je revisite la question des fluctuations du revenu du travail au cours du cycle en développant et en estimant un modèle d'appariement du marché du travail, avec une incertitude sur le niveau de la productivité agrégée des travailleurs. Je fais l'hypothèse que les travailleurs et les entreprises possèdent différents niveaux de productivité, ce qui me permet de mettre en évidence que la qualité de l'appariement entre travailleurs et entreprises au cours du cycle est un facteur de premier ordre dans la fluctuation des revenus du travail. Une seconde contribution du premier chapitre, à un niveau plus technique, est de montrer qu'il existe un algorithme robuste et fiable pour résoudre la détermination des salaires dans un modèle d'appariement dynamique, avec travailleurs et entreprises hétérogènes se rencontrant de manière aléatoire. Cet algorithme, appliqué au marché du travail, offre des enseignements plus généraux sur les marchés frictionnels à agents hétérogènes. Pour pouvoir résoudre ces modèles compliqués, il est crucial de pouvoir séparer le problème de faisabilité de l'appariement (l'emploi est-il réalisable ?) du problème de la répartition de la rente générée par l'appariement (quel salaire sera offert au travailleur ?). Cependant, le problème de répartition de la rente peut de son côté dépendre du problème de faisabilité et de l'état du marché du travail en général.

Le second chapitre analyse le lien causal entre une baisse du coût des transports en commun et les dynamiques d'emploi local. Mon argumentation se base sur la discontinuité créée par la réforme de la tarification du forfait Navigo en septembre 2015. Cette réforme a égalisé le coût des transports en commun en Île-de-France — pour la majorité des titres hebdomadaires, mensuels et annuels — qui auparavant différaient en fonction des zones tarifaires traversées par l'utilisateur. Ce chapitre se base sur une observation factuelle et une

idée simple. Premièrement, les zones tarifaires des transports en commun en Île-de-France forment des disques concentriques, ayant pour centre Paris. Ces zones tarifaires ont été déterminées bien avant la réforme de 2015. Deuxièmement, bien que les villes appartenant à différentes zones tarifaires diffèrent à bien des égards, si l'on considère deux villes proches l'une de l'autre, les dynamiques locales d'emploi non observables par l'économètre ont de grandes chances d'être similaires. Je construis donc mon groupe de contrôle et mon groupe traité en choisissant des villes de part et d'autre d'une frontière tarifaire, qui a cessé d'être pertinente à partir de septembre 2015. Puis j'estime l'impact causal d'une baisse du coûts des transports en commun sur les dynamiques d'emploi local en utilisant une méthode des doubles différences. La méthode des doubles différences met en évidence que la baisse du coût des transports en commun pour les villes bénéficiant de la réforme a généré une dynamique d'emploi positive. La baisse du chômage identifiée est particulièrement robuste pour les chômeurs de longue durée. Pour expliquer ces résultats empiriques, je développe un modèle d'appariement avec une dimension spatiale.

Alors que les deux premiers chapitres se concentrent sur le marché du travail, le troisième chapitre se focalise sur un autre marché frictionnel d'importance capitale : le marché immobilier. Ce chapitre est motivé par l'observation que peu de données existent sur le marché locatif, malgré son importance pour les individus — 39.3% des Français étaient locataires en 2015. Alors que les transactions immobilières sont enregistrées par les notaires, il n'existe pas de base de données centralisant l'ensemble des contrats locatifs. L'analyse de ce chapitre repose sur une base de données collectées via des méthodes de *web scraping* — qui pourrait se traduire comme *collecte automatique des données du web*. La collecte a été réalisée pour le marché locatif de la région parisienne, pendant une période de trois mois. Une des nouveautés de ce chapitre par rapport à la littérature existante est liée au fait que j'observe non seulement les annonces locatives, mais aussi le comportement des chercheurs de logement. Combinant l'information sur les logements et les locataires potentiels, je montre que le marché locatif de la région parisienne est bien décrit par un modèle d'appariement directionnel. Le modèle d'appariement directionnel repose notamment sur l'hypothèse que les acteurs peuvent choisir le sous-marché dans lequel ils cherchent un partenaire, ce qui le différencie des modèles à recherche aléatoire analysés dans les deux premiers chapitres Cependant, contrairement aux hypothèses classiques du modèle d'appariement directionnel, je trouve qu'une proportion non négligeable des propriétaires utilisent une stratégie de prix dynamique descendante. Cette stratégie soulève à la fois des questions d'efficience et de nouveaux challenges théoriques.

Introduction

This thesis consists of three chapters that analyze frictional markets. While the abstraction of a Walrasian market — in which the price adjusts instantaneously to equilibrate demand and supply — is a good description of reality in many instances, several key markets do not fit this framework. One leading example is the labor market, in which both workers and firms exert significant efforts to search for each other. This process is well exemplified by the yearly "rookie market" in Economics, during which freshly graduated doctors in Economics spend several months interacting with potential employers to possibly receive one, or for the most promising candidates, several job offers. The process is also extremely time-consuming and costly for employers. They need to process a vast amount of information on each candidate. They commit a substantial amount of time to leading interviews and are required to listen to lengthy job market talks in order to determine the research potential of a given candidate.

What is true for young PhD graduates also holds for the rest of the population. For instance, it is standard for young graduates to have to go through 4 to 8 interviews before getting an offer from a competitive consulting company. It could take months in between the time a resume has been submitted and the time an offer is made. Time-consuming and costly for workers, the hiring process is also costly for firms. Some studies have estimated the average cost of hiring a worker to be between 10 and 17 weeks of wage payments (Blatter et al. (2012)). In many instances, such elevated friction costs have to be taken into consideration when answering questions related to the labor market.

In this thesis, I make an attempt to extend our knowledge on frictional markets, both theoretically and empirically. In the first chapter, I revisit the question of the cost of a job loss and the issue of the scarring effects of recessions — the fact that losing a job during a recession comes with an extra penalty, compared to losing a job during an expansion. In the second chapter of this thesis, in which the focus is more empirical, I analyze how a change in the cost of public transportation might change local employment dynamics. In the third chapter, I analyze another crucial and under-studied frictional market: the rental housing market. Based on a new sample of user-generated content, I show new empirical facts for

the rental market that raise efficiency issues and offer new modeling challenges.

Chapter 1: Labor Income Shocks along the Business Cycle

The first chapter of this thesis revisits the question of the cost of a job loss and the question of the scarring effects of recessions — the fact that losing a job during a recession comes with an extra penalty, compared to losing a job during an expansion.

Standard search-and-matching models have a hard time generating realistic costs of a losing a job. A simple back-of-the-envelope calculation illustrates why this is the case. Using the average daily earning of American workers and the average number of days to fill a vacancy, it is easy to estimate that it costs American firms on average approximately 1000 USD to hire a worker (Hall and Rogerson (2011)). Assuming a symmetric Nash bargaining, the worker's loss from a layoff that occurs immediately after the hire is about 1000 USD. However, empirical analyses indicate that the present value of job losses is in between 1 and 3 years of pre-displacement earnings. How can we reconcile the fact that the cost of destroying a new job is so inconsequential with the fact that a job loss after several years of tenure amounts to several thousands of dollars? Robert Hall summarizes the challenges posed by this observation:

> "The master model faces an important challenge in explaining how workers move from having, on average, only a roughly $1,000 stake in a brand-new job to having around $100,000 at stake after more than 3 years of tenure. The implied gradient of accumulation of the worker's share of job-specific capital is remarkably steep, and thus a real challenge to empirical model builders." Hall and Rogerson (2011)

In addition to the fact that the cost of a job loss is hard to rationalize in a standard search-and-matching model, another challenge that empirical model builders face is that the cost of a job loss *varies along the business cycle*. The fact that workers losing a job during recessions incur an extra penalty has been coined the *scarring effects of recessions*. The scarring effects of recessions has been identified in several empirical works, in particular in Jacobson et al. (1993), Stevens (1997) and Davis and Till (2011). In the latter study, the authors find that workers displaced during recessions experience losses (in present value terms) that are 30 to 40% higher than the losses of similar workers being displaced during expansions. It has also been established that most of losses are due to a drop in wages, while the number of hours worked recovers after one year or two.

To explain why the impact of job losses is so persistent over time and why workers losing their jobs in recessions incur an additional penalty, the literature has proposed two explanations. The first explanation is based on the theory of human capital accumulation. As workers keep working at the same firm, they accumulate job-specific human capital. When losing their jobs, workers cannot transfer their previously accumulated job-specific human capital to their new job. As a result, workers are less productive in their new job and are thus less paid. This human-capital mechanism is discussed for instance in Ljungqvist and Sargent (1998).

While this theory is appealing for its simplicity, empirical evidence suggests that the job-specific human capital explanation cannot entirely explain business cycle variations of the cost of a job loss. In particular, the same magnitude and persistence of labor income losses has been identified in the group of freshly graduated students (Kahn (2010)). Ceteris paribus, students graduating in recessions experience persistent income losses relative to similar students graduating during expansions. What is troubling for the human capital story is that both groups of students have zero job-specific human capital. The group of new college graduates, for whom the story of human capital cannot apply, invites us to wonder whether there is something specific about the labor market, not about workers, during recessions that is scarring workers.

An alternative theory that would explain the cost of a job loss is a signaled-based theory. Workers differ in many ways and many differences are not observed by the Econometrician. Some factors are not even observed by employers (for instance, how hard-working an employee is hard to determine) and they need to make educated guesses about the types of workers they are hiring based on observations from the past and the information collected during interviews. An employer might interpret a job loss as a sign that this person is a "poor" worker. As a result, employers may be reluctant to hire them, or may hire them at a discounted rate. The signal-based theory has been investigated for instance in Gibbons and Katz (1991).

While the signal theory is also appealing and certainly explains parts of why it is costly to lose a job, the signal theory has difficulties in explaining business cycle variations in the cost of a job loss. In particular, the probability of losing a job during a recession is higher than the probability of losing a job during an expansion. In a recession, the general state of the economy leads to many bankruptcies and many "good" workers end up unemployed (Nakamura (2008)). Why do firms not penalize workers who have lost their jobs during an expansion *more* than workers who have lost their jobs during a recession?

The first chapter aims to shed light on the cost of a job loss using a *sorting-based* explanation, which complements the signal and human capital theories. The sorting-based

explanation is quite intuitive. Workers not only climb-up a wage ladder as their tenure increases, they also climb-up a "firm ladder". By firm ladder, I mean that workers may start their career in a job that is not the perfect match for their skills, and that by transitioning from one job to another, the degree of complementarity between worker skills and firm characteristics improves. When they lose their jobs, workers fall from both the wage ladder *and* the firm ladder. Because unemployed workers fall in two dimensions, the model naturally generates large and persistent wage losses. The model has also a natural channel to generate business cycle fluctuations in labor income losses. In the model, wage increases are generated by the occurrences of outside options: a worker may meet with an alternative employer and decide to change jobs, or the worker may use the alternative employer as a threat to receive a promotion. In a recession, the expected profits of posting a vacancy decreases and the likelihood of meeting with an alternative employer collapses. As a result, the mechanism that generates both intra-firm wage increases (the threat of leaving the company for another firm) and the mechanism that reallocates workers from bad to better positions is less strong during recessions. Workers losing their jobs during recessions and finding a job in the recovery phase stay poorly matched for longer and their wage growth is sluggish.

My sorting-based explanation relies on an estimated model of the labor market with different types of workers and firms, to allow for differences in complementarities. The employment side is identical to the model of Lise and Robin (2017). My contribution is to show how to solve for the wage dynamics in a model with heterogeneous workers and firms, aggregate uncertainty on productivity and random search of agents. The algorithm I describe is based on the observation that (i) as a first step, the employment feasibility problem can be solved independently of wages; (ii) as a second step, wages do depend on a high-dimensional state variable, but dimension reduction techniques using the information in step (i) make the wage-allocation problem solvable. My belief is that the technical take-aways from the first chapter are more general: in a frictional model with random search, heterogeneity and aggregate uncertainty, tractability is obtained by separating the feasibility problem (is the job feasible?) from the allocation problem (what is the wage?). Even though the allocation problem might depend in a fairly complicated way on the state of the economy, the model is still tractable.

A more exhaustive first chapter would help us in quantifying the different contributions of the signal, human capital and sorting channels to the cost of a job loss. While I was not able to achieve this task in this thesis, I show how the model with sorting can also accommodate learning-by-doing job-specific human capital accumulation. Solving and estimating the model with human capital accumulation is definitely a priority for my future research projects.

Chapter 2: Commuting Costs and Local Employment Dynamics

The second chapter analyzes how local employment dynamics are influenced by the cost of public transport. This chapter is also an analysis of the labor market, but the focus is more empirical relative to the first chapter. Instead of considering heterogeneity across time, this chapter abstracts from business cycle fluctuations and studies how heterogeneity across space in a metropolitan area impacts labor market outcomes.

The motivation for this chapter is based on the observation that several countries have initiated multi-billion projects to increase the usage of public transportation. For instance, the Grand Paris Express is a 42 billion USD investment project aiming at adding 200 km of new track in the Paris metropolitan area. Large scale infrastructure projects are also under way in the United Kingdom (London Crossrail 1), in Spain (Barcelona metro line 9), in the US (New York Second Avenue Subway). In addition to extending public transport networks, some governments are passing reforms to change the cost of public transport. For instance, public transport in Luxembourg became free of charge as of March 2020.

These projects and reforms are motivated by several goals, including the willingness to curb traffic, decrease pollution levels and reduce C02 emissions. One argument that has been put forth is that a decrease in the cost of public transport might boost employment. For instance, the institution managing the construction of the Grand Paris Express estimates that 100 000 jobs will be created by the new transport system, once the project is in place.[1]. The main goal of this chapter is to make an empirical contribution to the literature that analyzes how commuting costs may change local employment dynamics. My contribution is based on a discontinuity in the Parisian metropolitan area created by a French reform in 2015.

In September 2015, a unique travel pass for public transportation in the Parisian metropolitan area was created, called "Forfait Toutes Zones" (FTZ). Before the FTZ reform, users had to pay commuting costs that depended on the fare areas (also called "zones") crossed during the trip. Fare areas are ring-shaped regions, having for a center the city of Paris *intra-muros*, which constitutes zone 1. Zone 2 is constituted of municipalities that have access to metro stations, excluding Paris *intra-muros*. Zone 3 is constituted of municipalities from the close Parisian suburbs. Zones 4 and 5 are annulus situated further away from the center of Paris.

The FTZ reform created a discontinuity in the cost of commuting using public transport. For instance, users of the yearly Navigo pass zones 1-3 — likely to be people living in the close suburbs and commuting to Paris — experienced a 14.3 € monthly decrease

[1]Source: `https://www.societedugrandparis.fr/info/economie-et-emploi-201`

in their commuting costs, while users of the yearly Navigo pass zones 1-2 experienced a 3.5 € monthly increase. My identification strategy relies on the assumption that two municipalities geographically close to each other, but belonging to different fare areas, share similar employment dynamics (after conditioning for differences in observables). I use a difference-in-difference strategy with a control group constituted of municipalities from zone 2 and a treatment group constituted of municipalities from zone 3. I make sure that the municipalities from the treatment and control groups are located close to each other.

Related work includes papers using instrumental variable strategies. For instance, Mayer and Trevien (2017) document that the arrival of the Regional Express Rail (RER) in the Paris metropolitan area caused a 8.8% increase in employment for the municipalities connected to the network between 1975 and 1990. Garcia-López et al. (2017) show that improvements in the Parisian transit system led to the emergence of employment sub-centers in suburban municipalities that had a rail station. Duranton and Turner (2012) show that a 10% increase in a provision of buses caused the *population* to increase by 0.8% in the US. Instead of using an instrumental variable approach, I use a *spatial regression discontinuity* (Neumark and Simpson (2015)). The methodology used in this paper is very close to the one used by Chapelle et al. (2018), Einiö and Overman (2016), Kline and Moretti (2013) and Hilber et al. (2019).

My empirical results are consistent with the above mentioned studies, in the sense that I find an increase in the access to public transportation — here, a decrease in the use cost — boosts local employment dynamics. I find the reform generated a 2% *decrease* in the number of unemployed workers registered to Pole Emploi — the French unemployment agency — for municipalities in zone 3, relative to municipalities in zone 2. The positive employment effects are concentrated on long-term unemployed workers.

In this chapter, I claim that we can think of transport infrastructure projects and public transport cost reforms as local employment policies. The central idea is that commuting to work is costly. If commuting to work is too expensive, a worker would simply prefer to stay unemployed (Zenou (2000)). The FTZ reform can be seen as a *spatially-located employment premium*, hence potentially boosting employment for municipalities benefiting from the reform. One way to think about this mechanism is that the FTZ reform decreased the reservation wage of some workers. A different but related idea is that the FTZ could also boost employment by changing the search behavior of unemployed workers. Indeed, it is likely that the reform expanded the job-search area of unemployed workers living in the Parisian suburbs. I formalize these ideas by solving and calibrating a search-and-matching model with a spatial dimension, which borrows elements from Brueckner et al. (1999), Wasmer and Zenou (2002) and the discrete choice literature to take into consideration that residents may

want to relocate after the reform. I use the model to understand the employment effects of the FTZ reform.

Chapter 3: The Rental Housing Market and Directed Search

The third chapter analyzes another key frictional market: the housing market. The housing sector is essential for the rest of the economy because real estate constitutes the largest asset class. The Great Recession was triggered by the bursting of a housing bubble in the United States (Farmer (2012)). The housing market has also been linked to many social and economic issues. It has been shown that the lack of affordable housing in many countries has distorted labor market outcomes (Brown and Matsa (2019)) and depressed productivity growth (Herkenhoff et al. (2018)). It has also been established that the state of the housing market impacts how people marry (Wei et al. (2017)) and could fuel support for populism (Adler and Ansell (2020)).

This chapter is motivated by the fact that while there is a growing literature analyzing the real estate market, the rental housing market has been understudied. I believe that the main reason for this knowledge gap is the lack of comprehensive datasets on rental agreements. For instance in France, while real estate transactions have to be recorded by notaries, it is not the case for rental agreements.

This chapter helps to fill the knowledge gap on the rental housing market by constructing and utilizing a dataset of online ads for the Parisian market. Hence, it is related to the growing number of papers using web scraping and user-generated datasets to shed new light on previously unobservable parts of the Economy (Edelman (2012)).

This chapter makes two contributions to a growing literature on the rental housing market. Firstly, information on both sides of the market was collected at a detailed level. In particular, I observe the number of contacts received by landlords for each ad. Housing search behavior is generally unobserved. As predicted by a standard directed search model (for instance, the competitive search model of Moen (1997)), I show that accommodation that is cheaper than expected — controlling for observable characteristics — attracts more searchers. Landlords discounting their apartment are rewarded by a higher arrival rate of tenants. By the same token, tenants searching for a discounted apartment have to "pay" a higher price by lowering their probability to find such a good deal.

A second contribution is to show that a non-negligible proportion of landlords uses a descending auction-like rent setting strategy. Approximately 7% of landlords set a rent that

is higher than what is predicted by a hedonic regression model, before lowering the advertised rent to a level in line with the predictions based on observable characteristics. To the best of my knowledge, this empirical fact is new in the literature. I discuss what this empirical observation could mean in terms of welfare and efficiency for the rental housing market.

References

Adler, D. and Ansell, B. (2020). Housing and populism. *West European Politics*, 43(2):344–365.

Bezanson, J., Edelman, A., Karpinski, S., and Shah, V. B. (2017). Julia: A fresh approach to numerical computing. *SIAM review*, 59(1):65–98.

Blatter, M., Muehlemann, S., and Schenker, S. (2012). The costs of hiring skilled workers. *European Economic Review*, 56(1):20–35.

Brown, J. and Matsa, D. A. (2019). Locked in by leverage: Job search during the housing crisis. *Journal of Financial Economics*.

Brueckner, J. K., Thisse, J.-F., and Zenou, Y. (1999). Why is central paris rich and downtown detroit poor?: An amenity-based theory. *European economic review*, 43(1):91–107.

Chapelle, G., Vignolles, B., and Wolf, C. (2018). Impact of a housing tax credit on local housing markets: Evidence from france. *Annals of Economics and Statistics/Annales d'Économie et de Statistique*, (130):101–131.

Davis, S. J. and Till, M. (2011). Recessions and the cost of job loss. Technical report, Working Paper. National Bureau of Economic Research.

Duranton, G. and Turner, M. A. (2012). Urban growth and transportation. *Review of Economic Studies*, 79(4):1407–1440.

Edelman, B. (2012). Using internet data for economic research. *Journal of Economic Perspectives*, 26(2):189–206.

Einiö, E. and Overman, H. G. (2016). The (displacement) effects of spatially targeted enterprise initiatives: Evidence from uk legi.

Farmer, R. E. (2012). The stock market crash of 2008 caused the great recession: Theory and evidence. *Journal of Economic Dynamics and Control*, 36(5):693–707.

Garcia-López, M.-À., Hémet, C., and Viladecans-Marsal, E. (2017). Next train to the polycentric city: The effect of railroads on subcenter formation. *Regional Science and Urban Economics*, 67:50–63.

Gibbons, R. and Katz, L. F. (1991). Layoffs and lemons. *Journal of labor Economics*, 9(4):351–380.

Hall, R. and Rogerson, R. (2011). Recessions and the cost of job loss: Comments and discussion. Technical report, Working Paper. National Bureau of Economic Research.

Herkenhoff, K. F., Ohanian, L. E., and Prescott, E. C. (2018). Tarnishing the golden and empire states: Land-use restrictions and the us economic slowdown. *Journal of Monetary Economics*, 93:89–109.

Hilber, C. A., Carozzi, F., and Xiaolun, Y. (2019). The economic impacts of help to buy.

Jacobson, L. S., LaLonde, R. J., and Sullivan, D. G. (1993). Earnings losses of displaced workers. *The American economic review*, pages 685–709.

Judd, K. L. (1998). *Numerical methods in economics*. MIT press.

Kahn, L. B. (2010). The long-term labor market consequences of graduating from college in a bad economy. *Labour economics*, 17(2):303–316.

Kline, P. and Moretti, E. (2013). Local economic development, agglomeration economies, and the big push: 100 years of evidence from the tennessee valley authority. *The Quarterly Journal of Economics*, 129(1):275–331.

Lise, J. and Robin, J.-M. (2017). The macrodynamics of sorting between workers and firms. *American Economic Review*, 107(4):1104–35.

Ljungqvist, L. and Sargent, T. J. (1998). The european unemployment dilemma. *Journal of political Economy*, 106(3):514–550.

Mayer, T. and Trevien, C. (2017). The impact of urban public transportation evidence from the paris region. *Journal of Urban Economics*, 102:1–21.

Moen, E. R. (1997). Competitive search equilibrium. *Journal of political Economy*, 105(2):385–411.

Nakamura, E. (2008). Layoffs and lemons over the business cycle. *Economics Letters*, 99(1):55–58.

Neumark, D. and Simpson, H. (2015). Place-based policies. In *Handbook of regional and urban economics*, volume 5, pages 1197–1287. Elsevier.

Robbins, L. (2007). *An essay on the nature and significance of economic science*. Ludwig von Mises Institute.

Sargent, T. (2015). Computational challenges in macroeconomics.

Stevens, A. H. (1997). Persistent effects of job displacement: The importance of multiple job losses. *Journal of Labor Economics*, 15(1, Part 1):165–188.

Wasmer, E. and Zenou, Y. (2002). Does city structure affect job search and welfare? *Journal of urban Economics*, 51(3):515–541.

Wei, S.-J., Zhang, X., and Liu, Y. (2017). Home ownership as status competition: Some theory and evidence. *Journal of Development Economics*, 127:169–186.

Zenou, Y. (2000). Urban unemployment, agglomeration and transportation policies. *Journal of Public Economics*, 77(1):97–133.

Chapter 1

Labor Income Shocks along the Business Cycle

Abstract

This paper analyzes the determinants of labor income shocks along the business cycle. My main finding is that sorting between firms and workers is a key component of idiosyncratic risk. Labor income shocks are analyzed through the lenses of a dynamic search-and-matching model, which I estimate using US data. Because of search frictions and mismatches between firms and workers, the *laissez-faire* equilibrium is not necessarily optimal. My results underline that the government can tame business cycle fluctuations by designing a simple unemployment policy improving sorting between firms and workers.

Keywords: Sorting, Labor Income Risk, Business Cycle
JEL Classification: E32, J31

Introduction

This paper has a triple objective. The first one is to deepen our understanding on the sources of labor income shocks. That is, the unpredictable part of labor income changes. It is now well established that fluctuations in earnings at the individual level is an order of magnitude bigger than fluctuations at the macro level (Parker and Vissing-Jorgensen (2009)). The assumption of normality of labor income shocks has been attacked by several recent publications, in particular in Guvenen et al. (2014) and Guvenen et al. (2015). Recessions are periods marked by intense negative labor income shocks, underlined by a spike in left-skewness in the distribution of labor income changes. The aim of this paper is to unpack the black-box of the complex labor income process and to analyze its determinants. Why should we care? What does left-skewness mean for an average worker? In practical terms, it means that some categories of workers are badly hurt by recessions, with persistent consequences. The scarring effects of recessions are now well identified (see Ouyang (2009)). The persis-

tence of labor income shocks can be in part explained by search models with human capital depreciation, which creates a negative feedback loop on aggregate variables, as in Walentin and Westermark (2018). If negative feedback loops are involved, preventing bad shocks from happening or helping workers to recover from them is probably a welfare enhancing policy.

This paper sheds light on a previously ignored component of idiosyncratic income risk: sorting between workers and firms. By sorting, I mean the extent to which the market allocates the right workers to the right jobs, where "right" is captured by complementarities in the production function. Why is sorting an important mechanism for the labor income process? When considering a labor market with search frictions and random search, the pairing between firms and workers is not necessarily optimal. An inefficient match in turn translates into lower wages as long as the match persists. Being fired not only has a direct consequence on someone's labor income, it also has dynamic consequences. To go back to her/his previous income level, a newly unemployed worker has to climb up the *intra-firm wage ladder* and the *inter-firm ladder*. That is, a worker has to spend some time searching on the labor market before finding a firm that is the right match, and conversely. Numerical simulations show that the inter-firm ladder is far from being negligible. Long-tails in the distribution of labor income shocks hinges on the economy featuring heterogeneous firms, hence an inter-firm ladder. To the best of my knowledge, empirical work on this component of risk is rather scarce. One notable exception is Huckfeldt et al. (2016), who shows that earnings costs of job loss are concentrated among workers who find reemployment in lower-paying occupations. Using CPS and PSID data, the author estimates that the initial earnings losses of workers losing their job and subsequently switching occupations are *four times larger* than losses for workers finding a job in their previous occupation. Persistence of the initial wage loss is only observed for occupation switchers. These empirical facts can be consistently explained by the existence of an inter-firm ladder combined with some degree of random search on the labor market.

Related literature on sorting includes the seminal contribution of Abowd et al. (1999), decomposing real total annual compensation per worker into an employee, an employer and a residual effect. Bonhomme et al. (2019) introduce a framework that can accommodate interactions between worker and firm attributes. In a variance decomposition exercise, Song et al. (2018) show that two-thirds of the rise in the dispersion of log earnings between 1978 and 2013 can be attributed to a rise in the dispersion of average earnings *between* firms. In the search-and-matching literature, Lise and Robin (2017) study how sorting patterns are altered along the business cycle. To study labor income shocks across the cycle, my strategy is to use and extend their model. My contribution is twofold. Firstly, I extend the model of Lise and Robin (2017) by solving for the wages, which were left implicit in

their contribution. Secondly, I estimate the model using the simulated method of moments (SMM) and wage moments. In particular, I focus on starting wage moments, which are particularly well-defined within the model and have strong identification power.

A second objective of this paper explores new techniques to solve and estimate dynamic search-and-matching models with heterogeneity. Solving labor models with both search frictions and heterogeneous agents is a notoriously difficult task. If workers do not have access to full information on the state of the labor market, which includes the number of vacancies posted by each firm and their associated wage, a commonly held view is that such frameworks cannot be solved using standard numerical techniques. To avoid these complications, the literature on dynamic search-and-matching models has focused on *block-recursive* equilibrium, following the seminal contribution of Menzio and Shi (2010). In short, a block-recursive model is one in which value functions and market tightness are independent from the distribution of employment across worker types. Such knife-edge conditions are met when search is *directed*. That is, (i) when firms make public the wage associated to each vacancy they post (ii) workers have full information over wages and the types of vacancy posted. Armed with full knowledge of the labor market conditions, workers direct their search efforts towards a specific sub-market. A free entry condition for firms pins down the number of vacancies created. While particularly clever and numerically efficient, block-recursive models are constrained efficient as a by-product of the modeling tricks involved (see Schaal (2017)). Thus, in a block recursive model, the *laissez-faire* equilibrium is necessarily optimal. When the goal of a paper is to explain a mechanism through the eyes of a model, constrained efficiency is mostly harmless. However, if the objective is to understand how a government may or may not improve the market outcome, it seems more appropriate to come up with a new concept of equilibrium that does not rule out inefficiencies in the first place. This is the route I explore in this paper. The strategy I use to solve a non block-recursive search model can be seen as a variant of the Krusell and Smith (1998) algorithm. Agents are endowed with a simple forecasting rule that needs to be estimated via Monte-Carlo. The particularity in my setting is that the time series needed to estimate the forecasting rule do not depend on the value functions to be calculated. This property, specific to the model under scrutiny, leads me to design an algorithm that is both rapid and robust to solve the model. Because the model is half way between a fully non-block recursive model and a block recursive one, I propose to name it as *semi-block recursive*.

To a lesser extent, my work is related to the burgeoning literature on how to solve and estimate models with both aggregate uncertainty and heterogeneity. Following the seminal contribution of Krusell and Smith (1998), several techniques have been developed (see Reiter (2009), Algan et al. (2014) and Winberry (2018)). While some of these techniques

are global and the other ones use linearization around the non-stochastic steady-state, the common denominator of the above mentioned methods is that they rely on the recursive representation of a multi-stage decision process. An interesting line of research has recently used the sequence representation of the dynamic choice problem (see Le Grand et al. (2017), Boppart et al. (2018) and Auclert et al. (2019)). In this paper, I stick to the more commonly used recursive form, but I note that the sequence form is particularly well-suited in my setting. My contribution is not to develop a new method *per se*, but to realize that there is a space in between models that are fully non-recursive and models that are fully recursive. Within that thin space, exiting methods can easily be applied.

A third objective of this paper is to analyze the potential gains from designing an optimal unemployment insurance (UI). Having defined a concept of equilibrium in which the *laissez-faire* equilibrium is not optimal by design, I explore simple unemployment policies that have the potential of being welfare-improving. The design of an optimal UI and the extent to which it can stabilize the cycle has received a comprehensive treatment in the macroeconomic literature. In a framework with heterogeneous agents and aggregate uncertainty, Ragot and Le Grand (2019) solve for the optimal Ramsey problem. The optimal replacement rate is pro-cyclical and stabilizes aggregate demand. In the present paper welfare gains are realized by improving the improving the sorting between firms and workers, boosting the value of production. In a similar setting, Lise et al. (2016) show how an optimal replacement rate might improve the market equilibrium and transfer utility across groups of workers. They find that the optimal unemployment scheme can deliver a welfare improvement of 1.4%, concentrated on low-skill workers. I contribute to this literature by analyzing a similar unemployment insurance scheme in a dynamic setting. I find that the optimal unemployment insurance scheme generates a 0.25% increase in welfare at the steady-state. While the gains are rather modest at the steady-state, the policy is successful in stabilizing labor income shocks over the business cycle by approximately 2%. These gains are achieved by transferring income from high-skilled to low-skilled workers and by a stabilization of the inter-firm channel. The mechanism is quite intuitive: by making unemployment workers better off, especially low-skilled workers, they become more selective when choosing a job. The congestion effects of low-skilled workers are mitigated and high-skilled workers end up in better matches. As high-skilled workers are less likely to lose their job when the economy enters a recession, fluctuations in labor income are less severe.

Section 1 presents key facts for the distribution of wages along the business cycle in the US, as well as some important stylized facts for the process of labor income shocks. Section 2 presents a frictional model of the labor market, with random search and heterogeneous workers and firms, that can rationalize the data. Section 3 explains why solving for the

wages in a dynamic random search model with heterogeneity is challenging. An algorithm solving these challenges in a fast and robust way is then presented. Section **??** presents the parametrization of the model and examines how well the estimated model can replicate keys facts of labor market in the US. Section 5 discusses how the model can replicate the labor income shock process in the US, and why sorting appears to be key to generate realistic business cycle shocks in labor income. Section 6 presents the results of analyzing an optimal unemployment insurance scheme. The last section concludes.

1. Data

In this section, I calculate key statistics for the wage distribution along the business cycle in the US. This section also analyzes the drivers of labor income shocks.

With regard to statistics on the wage distribution, I focus on the elasticity and the standard deviation of wages for both the entire workforce and for the *new hires* (workers who just exited unemployment). Putting the spotlight on the wage of newly employed workers is important for two reasons. Firstly, the elasticity of starting wages has been shown to be key in solving the Shimer's puzzle. The Shimer's puzzle is the observation that standard search models predict fluctuations in the unemployment rate of an order of magnitude smaller than what is observed in empirical time series (see Shimer (2005) and Pissarides (2009)). Secondly, I focus on the wage of new hires for practical reasons linked to my modeling strategy. In the model developed below, the distribution of wages for new hires is well-behaved and has strong identification power, while the distribution of wages for the entire workforce is much more complicated. I first describe the methodology used to calculate starting wage moments, which closely follows Haefke et al. (2013), before analyzing the determinants of labor income shocks in the US.

1.1. Wage dynamics

To calculate statistics on the starting wage distribution in the US, I use the CPS Merged Outgoing Rotation Groups (CPS MORG), which contain both employment and wage variables for the period 1979 until nowadays.[1] More specifically, I use the Center for Economic and Policy Research ORG extracts[2], which contain time-consistent CPS MORG variables and wage variables corrected for top-coding in declared wages. To measure real hourly wage,

[1] Respondents are asked to report their hourly wage if they are hourly workers or to report their weekly earning if they are paid by the hour.

[2] `http://ceprdata.org/cps-uniform-data-extracts/cps-outgoing-rotation-group/cps-org-programs/`

I use the CEPR MORG variable *rw*, which excludes overtime, tips and commissions for hourly workers; but includes overtime, tips and commissions for non-hourly workers. The dataset is trimmed to excludes observations where real 1989 hourly wage is smaller than \$0.50 or bigger than \$200.

While the CPS was designed to offer repeated cross-section views of the US population, it also has a longitudinal dimension. Every household that enters the CPS is interviewed each month for 4 months, then ignored for 8 months, then interviewed again for 4 more months. Usual weekly hours/earning questions are asked only at households in their 4th and 8th interviews.[3] Hence, by comparing the same individual's hours/earnings in the 4th and 8th interviews, one can calculate the evolution of hourly wage over a year period. One can also determine which workers transitioned from unemployment to employment in between two interviews.

The process is complicated by the fact that keeping track of individuals in the CPS MORG is not straightforward, as the unit of reference is a housing unit. About 60,000 housing units are designated for data collection each month. Each house is assigned a household identifier (HHID) and an individual within the household is assigned an individual line number (LINENO). Individuals not changing location in between two interviews are in theory uniquely identified by the pair (HHID-LINENO). When a new household moves in, a "household counter" variable (HHNUM) is incremented by 1. To control for households moving in, I build an individual identifier as the combination of (HHID-LINENO-HHNUM) and immutable characteristics (gender and ethnicity). As an additional safety check, I compare the age of each potential individual between two observations. If the age difference is less than or equal to 2, I validate the match. Otherwise, the match is discarded. By choosing an age difference of 2 instead of 1, I allow for some degree of coding error. Using the (HHID-LINENO-HHNUM-gender-ethnicity) identifier, I can keep track of changes in the labor force status for individuals (employed, unemployed or not in the labor force). I calculate deciles at the yearly frequency for the starting wage distribution (workers currently employed and previously unemployed or out of the labor force) and for the entire distribution of wages.

Key statistics used to estimate the model in the subsequent period are the volatility and elasticity of (real) wage deciles with respect to changes in labor productivity. The response of wage deciles to productivity is measured by the coefficient of a regression of the log real wage deciles on log real labor productivity. As in Haefke et al. (2013), I estimate the regression in first differences to avoid spurious correlation if wages and productivity are integrated:

$$\Delta \log(w_{d,t}) = \alpha_d + \eta_j \Delta \log(y_t) + \varepsilon_{d,t} \tag{1.1}$$

[3]See `https://www.nber.org/data/morg.html`

where $w_{d,t}$ denotes the dth wage decile at time t and y_t is a measurement of labor market productivity at time t.[4] I estimate equation (1.1) on a sample restricted to men and women in between 25 and 60 years old working in the private sector, with series aggregated at the yearly frequency. While data is available starting in 1979, I restrict the sample to a period starting with the Great Moderation in 1984, as in Haefke et al. (2013). The period 1979-1983 is marked by large volatility in macro variables and by a substantial drop in the real minimum wage, which pushes downward the estimates of Table 1.1. Results are presented in Table 1.1. Consistent with the empirical literature (see Pissarides (2009)), I find that wages for new hires are much more sensitive to variations in labor productivity than in the series for aggregate wages. My point estimates are slightly less than reported by Haefke et al. (2013)[5], but of the same order of magnitude. Wage rigidity in ongoing contracts suggests is at odds with the assumption of a continuously-renegotiated Nash bargaining to determine wages, in the tradition of Mortensen and Pissarides (1994). If wages are the solution of a continuously-renegotiated Nash sharing rule, each change in aggregate productivity leads to variations in workers' wages, which is contradicted by the data.

To estimate the volatility of each wage decile, I first detrend series using a linear trend or a HP filter. Results are presented Table 1.2. Two facts are worth noting. Firstly, wage deciles for new hires are much more volatile than the overall population. Depending on whether detrending is done using a linear trend or an HP filter, the volatility of the median wage for new hires is between 40% and 100% higher than the general population. Secondly, when considering the sample of all workers, lower percentiles tend to be more volatile compared to the top of the wage distribution. This pattern is easily explained by the fact that the lower percentiles of the wage distribution are predominantly impacted by the wage of new hires, which are more volatile than the wage of workers in ongoing contracts.

[4]To measure labor productivity, I use real output per hour of all persons in the non-farm (OPHNFB)

[5]When the authors do not control for education (as it is the case in this paper), they find an aggregate wage elasticity of 0.16 for all workers and 0.54 for new hires for the period 1984 - 2006. When the authors control for difference in observables characteristics, they find and elasticity of 0.24 for all workers and 0.79 for new hires.

Table 1.1: Elasticity of wage and starting wage deciles in the US: 1984 - 2017

	New hires		All workers	
	η_j	p-value	η_j	p-value
P10	0.44	0.07	0.20	0.37
P20	0.32	0.27	0.13	0.45
P30	0.34	0.21	0.18	0.29
P40	0.31	0.32	0.18	0.24
P50	0.35	0.21	0.12	0.41
P60	0.42	0.09	0.30	0.06
P70	0.37	0.17	0.26	0.04
P80	0.45	0.20	0.28	0.02
P90	0.65	0.08	0.46	0.00

Notes: This table shows p-values and point estimates for the regression (1.1), measuring the sensitivity of wage deciles to changes in aggregate labor productivity. To measure productivity, I use real output per hour of all persons in the non-farm sector (OPHNFB). To measure hourly wage, I use the series 'rw' from the CEPR CPS ORG Extract, which converts hourly pay to constant 2018 dollars using the CPI-U-RS and corrects for top-coding. For calculations involving starting wages, the years 1986, 1995 and 1996 are excluded from the sample. For years 1986 and 1995-1996, I find only a limited number of workers transiting from unemployment to employment relative to other years (less than 1250 workers). Digits were rounded to the nearest hundredth.

Sources: CEPR CPS ORG Extract (`http://ceprdata.org/cps-uniform-data-extracts/cps-outgoing-rotation-group/cps-org-data/`) and U.S. Bureau of Labor Statistics retrieved from FRED, Federal Reserve Bank of St. Louis `https://fred.stlouisfed.org/series/OPHNFB`

Table 1.2: Volatility of wage and starting wage deciles in the US: 1984 - 2017

	New Hires		All workers		Volatility Ratio New hires/All workers	
	Linear trend	HP-filter	Linear trend	HP-filter	Linear trend	HP-filter
P10	0.041	0.019	0.036	0.017	1.14	1.14
P20	0.044	0.021	0.038	0.011	1.17	1.98
P30	0.051	0.019	0.029	0.012	1.77	1.61
P40	0.049	0.024	0.028	0.010	1.75	2.27
P50	0.039	0.020	0.028	0.010	1.39	1.98
P60	0.043	0.018	0.026	0.012	1.65	1.53
P70	0.041	0.020	0.022	0.010	1.85	2.00
P80	0.047	0.025	0.022	0.009	2.11	2.81
P90	0.045	0.026	0.024	0.010	1.82	2.59

Notes: This table the standard deviation of deciles of the log real wage distribution for new hires and all workers. The log of wage deciles were detrended using a linear trend or using a HP-filter with a smoothing parameter equal to 6.5. For calculations involving starting wages, the years 1986, 1995 and 1996 are excluded from the sample. For these years, I am able to find only a limited number of workers transiting from unemployment to employment relative to other years (less than 1250 workers).
Sources: CEPR CPS ORG Extract (`http://ceprdata.org/cps-uniform-data-extracts/cps-outgoing-rotation-group/cps-org-data/`).

1.2. Labor Income Shocks

Labor income shocks are defined as the unpredictable part of labor income changes. Let $w_{i,t}$ denote the real hourly wage of individual i at time t. I first project $w_{i,t}$ on a set of observable characteristics:

$$\log(w_{i,t}) = \boldsymbol{x_{i,t}}'\boldsymbol{\beta_w} + \varepsilon_{i,t} \tag{1.2}$$

where $x_{i,t}$ includes a constant, age and age square, marital status, education level and a linear time trend to capture long-run dynamics impacting real wage. Results are presented in the first column of Table 1.3. Log hourly wage is an increasing and concave function of age; a higher education level is associated with a higher hourly wage; holding other factors constant, women earn 28% less than men. By construction, the residual $e_{i,t} \equiv \log(w_{i,t}) - \boldsymbol{x_{i,t}}'\hat{\boldsymbol{\beta}}_{\boldsymbol{w}}$ is

orthogonal to the set of observable variables included in the right hand side of equation (1.2). I construct an hourly wage index orthogonal to observable characteristics as follows:

$$\log(\tilde{w}_{i,t}) = \log(w_{i,t}) + \left(\bar{\boldsymbol{x}}' - \boldsymbol{x_{i,t}}'\right)\hat{\boldsymbol{\beta}}_{\boldsymbol{w}} \tag{1.3}$$

where $\bar{\boldsymbol{x}}'$ denotes the average observable characteristics across individuals and periods. By construction $\Delta \log(\tilde{w}_{i,t}) \equiv \log(\tilde{w}_{i,t}) - \log(\tilde{w}_{i,t-1})$ measures the (log) difference in hourly wage that cannot be explained by observable factors. Note that by taking the difference for the same individual i, unobservable individual fixed effects that may have explained parts of labor income changes are removed. I apply the same methodology for weekly hours worked, by first fitting a linear model and then removing predictable changes in hours. Point estimates for the linear model are presented in the second column of Table 1.3. Let $\Delta\tilde{h}_{i,t}$ denote the unpredictable part of changes in weekly hours worked, which is calculated according to the following formula:

$$\Delta\tilde{h}_{i,t} = \left(h_{i,t} - h_{i,t-1}\right) - \left(\boldsymbol{x_{i,t-1}}' - \boldsymbol{x_{i,t}}'\right)\hat{\boldsymbol{\beta}}_{\boldsymbol{h}} \tag{1.4}$$

I generate a measurement of weekly labor income orthogonal to changes in observable factors, denoted by $\tilde{y}_{i,t}$, using $\tilde{w}_{i,t}$ and $\tilde{h}_{i,t}$:

$$\tilde{y}_{i,t} = \tilde{h}_{i,t} \times \tilde{w}_{i,t} \tag{1.5}$$

The resulting $\tilde{y}_{i,t}$, as well as $\tilde{h}_{i,t}$ and $\tilde{w}_{i,t}$ and are reported in Figure 1.4. Visual inspection of Figure 1.4 indicates that recessions years are marked by a contemporaneous drop in mean hours worked. The early 1980s and 1990s recessions were characterized by a contemporaneous drop in the mean real hourly wage, but the early 2000s recession and the Great Recession of 2008-2009 were inflexion points, with a decrease in the mean real hourly wage following with a lag. To complement this picture at the aggregate level, I calculate the probability that a worker experiences certain events. At the individual level, the probability of losing more than 0.5 times the standard deviation in real weekly earnings (approximately \$579) jumps by 1.24 percentage points in recession (see tables 1.4 and 1.5). The drop in real weekly earnings is caused by a decrease in hourly wage and a decrease in hours worked. Large negative changes in hours worked are more frequent in recession[6] and the probability of a large increase in hours worked declines.

[6] A 2σ decrease in hours worked corresponds to a drop in weekly hours worked of approximately 19 hours.

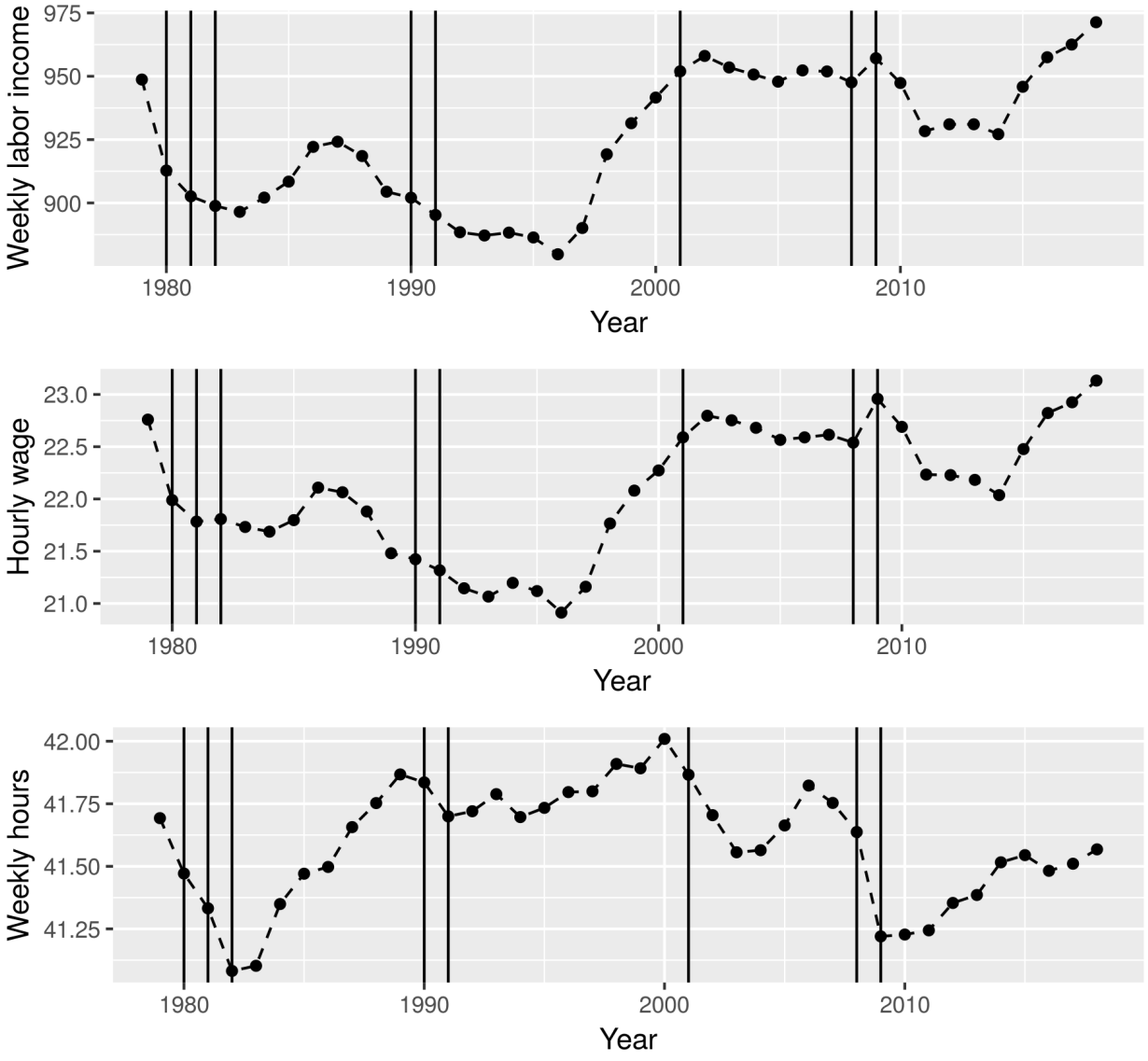
Table 1.3: Regression log wages and weekly hours worked

	Dependent variable: log hourly wage $\log(w_{i,t})$	*Dependent variable:* weekly hours worked $h_{i,t}$
age	0.044^{***}	0.280^{***}
	(0.0002)	(0.004)
age^2	-0.0004^{***}	-0.003^{***}
	(0.00000)	(0.00005)
married	0.080^{***}	-0.438^{***}
	(0.0005)	(0.009)
trend	0.001^{***}	-0.004^{***}
	(0.00002)	(0.0004)
HS	0.242^{***}	0.776^{***}
	(0.001)	(0.015)
some college	0.392^{***}	1.101^{***}
	(0.001)	(0.016)
college	0.691^{***}	2.495^{***}
	(0.001)	(0.017)
advanced	0.868^{***}	4.315^{***}
	(0.001)	(0.019)
woman	-0.280^{***}	-5.398^{***}
	(0.0004)	(0.008)
constant	1.601^{***}	35.462^{***}
	(0.004)	(0.077)
Observations	5,309,050	4,735,952
R^2	0.281	0.096
Adjusted R^2	0.281	0.096
Residual Std. Error	0.487	9.087
F Statistic	$230{,}227.100^{***}$	$55{,}871.350^{***}$
Note:		*p<0.1; **p<0.05; ***p<0.01

Notes: This table shows the results of estimating a linear regression model for the log of hourly real wage (column 1) and a linear regression model for the number of weekly hours worked (column 2).

Sources: CEPR CPS ORG Extract.

Fig. 1.1. Weekly earnings, hourly wage and hours worked



Notes: This figure shows the year-specific mean values for the residualized weekly hours worked $\tilde{h}_{i,t}$, the residualized real hourly wage $\tilde{w}_{i,t}$, and the implied weekly labor income $\tilde{y}_{i,t} = \tilde{h}_{i,t} \times \tilde{w}_{i,t}$. Vertical lines represent NBER recessions.

Table 1.4: Probability of changes in hourly wage, weekly earnings and hours along the business cycle

	$\tilde{w}_{i,t}$	$\tilde{w}_{i,t}$	$\tilde{y}_{i,t}$	$\tilde{y}_{i,t}$	$\tilde{h}_{i,t}$	$\tilde{h}_{i,t}$
Recession	No	Yes	No	Yes	No	Yes
$\Pr(\Delta x_{it} < -2\sigma)$	2.36%	2.37%	2.63%	2.67%	3.81%	4.00%
$\Pr(\Delta x_{it} < -\sigma)$	9.20%	9.64%	9.77%	10.45%	10.82%	10.94%
$\Pr(\Delta x_{it} < -0.5\sigma)$	19.47%	20.26%	19.99%	21.23%	18.14%	18.78%
$\Pr(\Delta x_{it} > 0.5\sigma)$	20.56%	20.37%	21.24%	21.04%	18.84%	17.72%
$\Pr(\Delta x_{it} > \sigma)$	9.68%	9.72%	10.36%	10.34%	11.37%	10.40%
$\Pr(\Delta x_{it} > 2\sigma)$	2.54%	2.42%	2.85%	2.66%	4.04%	3.90%

Notes: This table displays the probability that a yearly difference in x_{it} is above or below certain threshold σ, which denotes the standard deviation of Δx_{it}. Probabilities are calculated conditional on the state of the economy (recession or expansion). $\tilde{w}_{i,t}$ denotes the hourly wage of worker i at time t, net of net of predictable factors (age, education, etc). $\tilde{h}_{i,t}$ denotes weekly hours worked net of predictable factors. $\tilde{y}_{i,t}$ denotes the weekly labor income of individual i at time t.

Table 1.5: Difference in the probability of changes in hourly wage, weekly earnings and hours

	$\tilde{w}_{i,t}$	$\tilde{y}_{i,t}$	$\tilde{h}_{i,t}$
$\Delta \Pr(\Delta x_{it} < -2\sigma)$	0.01%	0.04%	0.19%
$\Delta \Pr(\Delta x_{it} < -\sigma)$	0.44%	0.68%	0.12%
$\Delta \Pr(\Delta x_{it} < -0.5\sigma)$	0.78%	1.24%	0.64%
$\Delta \Pr(\Delta x_{it} > 0.5\sigma)$	-0.19%	-0.20%	-1.12%
$\Delta \Pr(\Delta x_{it} > \sigma)$	0.03%	-0.02%	-0.97%
$\Delta \Pr(\Delta x_{it} > 2\sigma)$	0.13%	-0.19%	-0.14%

Notes: This table displays the difference (recession minus expansion) in the probability that a yearly difference in x_{it} is above or below certain threshold. The threshold value σ denotes the standard deviation of Δx_{it}. $\tilde{w}_{i,t}$ denotes the hourly wage of worker i at time t, net of net of predictable factors (age, education, etc). $\tilde{h}_{i,t}$ denotes weekly hours worked net of predictable factors. $\tilde{y}_{i,t}$ denotes the weekly labor income of individual i at time t.

2. Model

The empirical section underlined that (i) wages of newly hired workers are correlated with productivity, while wages in ongoing contracts are quite rigid; (ii) short-term downside risks in labor income are mainly driven by losses in hours. In this section, I develop a model that delivers both features and that allows me to decompose labor income risk between a part that is driven by workers' characteristics and another that depends on firms.

To model the employment side, I choose the framework of Lise and Robin (2017), which comes with a natural notion of sorting. Because in their model search is assumed to be random, workers may not necessarily meet with their optimal firm type, as measured by complementarities in the production function. Workers may accept "good-enough" jobs, knowing that future job opportunities might show up in the future. An alternative framework featuring two-sided heterogeneity and aggregate uncertainty is the model of Schaal (2017), built upon the directed search model of Menzio and Shi (2010). However, it is hard to define sorting in the latter framework because the directed search assumption generates a constrained efficient market outcome. Specifically, conditionally on the state of the economy, workers are always in the best match they can achieve.

My contribution in terms of modeling is to develop an efficient technique to solve for the wages. The key insight is that (i) the employment problem is independent from determination of wages; (ii) while the wage process depends on the employment process, this dependency is rather mild. That is, while the wage problem is not recursive (the state variable is infinite dimensional), a dimension reduction in the spirit of Krusell and Smith (1998) can be used. The wage problem is particularly well-behaved because the time series needed for the dimension-reduction step are independent from the value function for wages. In practical terms, the algorithm I describe is a one-step Krusell and Smith (1998) algorithm, with no need for an outer loop to estimate the parameter values of the forecasting rules. Besides being numerically efficient, the algorithm I describe has nice theoretical properties. Conditional on a forecasting rule to be estimated, which can be high-dimensional and arbitrarily complex, the value function characterizing the evolution of wages exists and is unique.

2.1. Workers, Firms and Timing

The economy is populated by a continuum of risk-neutral and infinitely-lived workers with mass 1. A worker can be either employed or unemployed. Workers differ in a skill parameter x distributed according to a density $\ell(x)$, which is time-invariant. There is a continuum of risk-neutral firms, differing by a productivity parameter y uniformly distributed on $[0, 1]$. Firms do not need capital to operate and can only hire one worker at a time. Firms can

freely enter the market and do so until the value of an unfilled vacancy is zero. Firms advertise positions through job placement agencies. The cost of posting v vacancies is given by a strictly increasing and convex cost function $c(.)$. In equilibrium, the marginal cost of creating a vacancy is equal to the expected return of doing so. Aggregate uncertainty stems from an aggregate productivity parameter z that follows an $AR(1)$ process:

$$z_{t+1} = \rho_z z_t + \sigma_z \varepsilon_{t+1} \tag{1.6}$$

with ε_{t+1} an i.i.d. Gaussian random variable with zero mean and unit variance. The value of home production for a worker of type x is given by $b(x, z_t)$ and the value of output for a worker of x working with a firm of type y is denoted by $p(x, y, z_t)$.

Let $h_t(x, y)$ denote the measure of matches (jobs) between workers of type x and firms of type y at the beginning of period t. I make the simplifying assumption that workers are either employed or unemployed. As a result, the following accounting identity holds:

$$\ell(x) = u_t(x) + \int_0^1 h_t(x, y)dy \tag{1.7}$$

where $u_t(x)$ denotes the measure of workers of type x unemployed at the beginning of period t.

The timing of the model is as follows: at the beginning of period t, the aggregate productivity variable changes from z_{t-1} to z_t (see Figure 1.2). Right after the change in productivity, workers may lose their job for exogenous or endogenous reasons. Job search, matching and wage setting happen in a sub-period, denoted by $t+$. The measure of matches surviving job destruction is denoted by $h_{t+}(x, y)$ and the measure of unemployed workers in the sub-period $t+$ is denoted by $u_{t+}(x)$.

During the sub-period $t+$, unemployed workers search for a job with an intensity normalized to 1, while employed workers search for an alternative job with an intensity equal to s. The resulting aggregate search effort, denoted by L_t, is a linear aggregation of individual effort:[7]

$$L_t = \int_0^1 u_{t+}dx + s\int_0^1 \int_0^1 h_{t+}(x, y)dxdy \tag{1.8}$$

The free entry condition on the firm's side implies that the marginal cost of posting one vacancy is equal to the expected value of a job opening:

[7]One strong assumption in equation (1.8) is that workers do not choose their search effort. During recessions, workers may want to search more to offset the decrease in the number of vacancies posted. As a result, assuming a constant search effort probably amplifies labor income shocks in this economy.

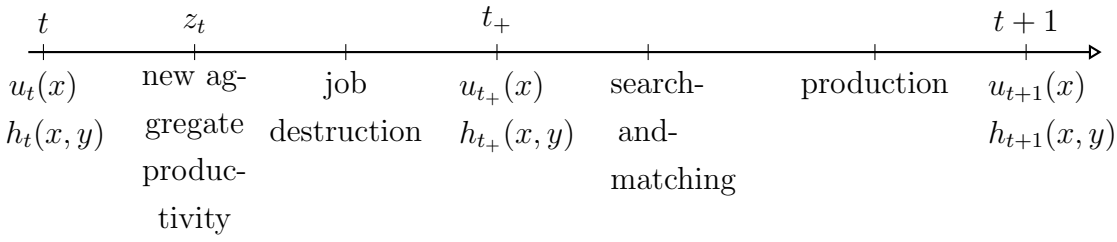
$$c'(v_t(y)) = q_t J_t(y) \tag{1.9}$$

where $J_t(y)$ denotes the expected value of a contact by a vacancy of type y and q_t is the probability (per recruiting effort) that a firm contacts a worker. Because the cost function is assumed to be increasing and convex, $c'(.)$ can be inverted $v_t(y) = (c')^{-1}\big(q_t J_t(y)\big)$. The total number of vacancies in period t is obtained by integrating over firm types $V_t = \int_0^1 v_t(y)dy$. The total number of meetings at time t, denoted by M_t, is the result of both the aggregate search effort of workers and the number of vacancies posted by firms. A matching function $M(.)$ is used to model the meeting of both sides of the labor market:

$$M_t = M(L_t, V_t) \tag{1.10}$$

The probability for a worker to meet a vacancy is the ratio of the number of meeting to the aggregate search effort $\lambda_t = \frac{Mt}{L_t}$. The probability that a firm contacts any searching work q_t is the ratio of the number of meeting to the total number of vacancies $q_t = \frac{Mt}{V_t}$.

Fig. 1.2. Timing of events



2.2. Wage setting and employment flows

Wages are determined according to the sequential auction framework, as in Robin (2011) or Postel-Vinay and Turon (2010). Unemployed workers have zero bargaining power and receive their reservation wage. Employed workers search for alternative employers. When they meet another firm, workers reveal the meeting to their current employer. A Bertrand competition between the incumbent and the poaching firm is triggered, which results in either a wage increase and/or a job-to-job transition.

In this environment, two properties are absolutely essential. Firstly, the value of unemployment to a worker of type x when the aggregate productivity level is z_t, denoted by $U(x, z_t)$, is independent from the distribution of matches $h_t(x, y)$. Secondly, the joint surplus of a match, denoted by $S(x, y, z)$, does not depend on $h_t(x, y)$ either. The functions $U(.)$

and $S(.)$ are solution of the following functional equations:

$$U(x, z_t) = b(x, z_t) + \frac{1}{1+r}\,\mathbb{E}_{z_{t+1}|z_t}\Big[U(x, z_{t+1})\Big] \tag{1.11}$$

$$S(x, y, z_t) = p(x, y, z_t) - b(x, z_t) + \frac{1-\delta}{1+r}\,\mathbb{E}_{z_{t+1}|z_t}\Big[\max\big(0, S(x, y, z_{t+1})\big)\Big] \tag{1.12}$$

where r is the interest rate. The expectation operator is taken with respect to next period's aggregate productivity level only. Equations (2.7) and (1.12) can be trivially solved by value function iteration.[8]

Two points are worth emphasizing. Firstly, the joint surplus of a match $S(x, y, z_t)$ does not depend on the job meeting rate λ_t. Independence from the job meeting rate implies the joint surplus of a match does not depends on the distribution of matches across skill and firm productivity types $h_t(x, y)$, an infinite dimensional object. Independence of $S(x, y, z_t)$ from the job meeting rate hinges on unemployed workers having zero bargaining power. Secondly, the joint surplus of a match does not depend on wages. The fact that $S(x, y, z_t)$ does not depend on wages rests on the assumption that workers and firms are risk neutral. Within a match, the wage is an instrument to decide the split of the joint surplus between workers and firms. Because utility is linear, the allocation of the surplus between the two parties does not modify the surplus itself. Independence of the joint surplus of a match from $h_t(x, y)$ and $h_t(x, y, w)$ are both essential when developing an algorithm to solve efficiently the wage process.

Knowledge of the joint surplus of a match $S(x, y, z_t)$ is sufficient to determine both job feasibility and job-to-job movements. Using this fact, the employment side of the model can be closed. The assumption of zero bargaining power for unemployed workers and the sequential auction hypothesis yield the following expression for the the expected value of a contact:

$$\begin{aligned} J_t(y) = &\int_0^1 \frac{u_{t+}(x)}{Lt} \max\big(0, S(x, y, z_t)\big)dx \\ &+ s\int_0^1\int_0^1 \frac{h_{t+}(x, y')}{L_t} \max\Big(0, S(x, y, z_t) - S(x, y', z_t)\Big)dxdy \end{aligned} \tag{1.13}$$

The first line in equation (1.13) is the expected value of hiring from the pool of unemployed workers. When a firm hires an unemployed worker, it offers the reservation wage

[8]See Lise and Robin (2017) for a derivation of these value functions.

and captures the entire surplus. The second line is the expected value of poaching workers from less productive firms. When a firm y poaches from another firm y', it has to offer the maximum wage the firm y' is willing to offer. After a successful poaching, the firm y enjoys a surplus equal to the difference between the total value of the match and what it has to pay to attract the worker.

The measure of $x-y$ matches in the sub-period $t+$, surviving exogenous and endogenous job destruction is given by:

$$h_{t+}(x,y) = (1-\delta)\mathbb{1}\{S(x,y,z_t) \geq 0\}h_t(x,y) \tag{1.14}$$

The measure of matches at the end of period t, denoted by $h_{t+1}(x,y)$, writes:

$$\begin{aligned} h_{t+1}(x,y) = h_{t+}(x,y)\Big[1 - s\int_0^1 \lambda_t \frac{v_t(y')}{V_t}\mathbb{1}\{S(x,y',z_t) > S(x,y,z_t)\}dy'\Big] \\ + s\int_0^1 h_{t+}(x,y')\lambda_t \frac{v_t(y)}{Vt}\mathbb{1}\{S(x,y,z_t) > S(x,y',z_t)\}dy' \\ + u_{t+}(x)\lambda_t \frac{v_t(y)}{Vt}\mathbb{1}\{S(x,y,z_t) \geq 0\} \end{aligned} \tag{1.15}$$

The first line in equation (1.15) takes into account the measure of workers moving to more productive poachers. The second line takes into account the measure of workers poached from less productive firms. The third line is the inflow of workers hired from unemployment.

3. Solving for the wages

This section expresses the Bellman function for the value a job to workers. While the relevant state variable is infinite-dimensional, this section shows that an efficient and robust algorithm exists. The flow equation for the distribution of starting wages is then derived.

To preview the difficulties that lie ahead, let me explain why the relevant state variable is infinite-dimensional. Let $W_t(x,y)$ denote the value of a job to a worker x working with a firm of type y at time t. In the class of models with search frictions and aggregate uncertainty, $W_t(x,y)$ generally depends on the next period's job meeting rate λ_{t+1}. If the job meeting rate is high, an employed worker is more likely to receive a promotion or to change job. The job meeting in the future thus impacts the reservation wage of workers today. The job meeting rate is itself a function of the current distribution of matches $h_t(x,y)$. Indeed, $h_t(x,y)$ affects both firms' expected value of posting vacancies (equation (1.13)) and workers' search effort (equation (1.8)). If no additional assumptions are made, the relevant aggregate state variable

for the determination of wages contains *the joint distribution of matches* $h_t(x, y)$.

A convenient assumption is to posit that a contract is an agreement to receive a given share of the match surplus $S(x, y, z_t)$. Given that the surplus does not depend on the distribution of matches $h(x, y)$, this property is inherited by $W_t(x, y)$. This path is followed by Lise et al. (2017). Yet, this assumption implies that the wage changes every time z does, even when workers or firms have no credible threat to quit or change job. However, the empirical analysis underlined that wages in ongoing contracts are rigid, which suggests that an alternative route could be considered. One alternative is to define a contract as an agreement to a constant wage w, which can be re-bargained by mutual consent only. This is the path I explore in this paper. I resolve the difficulty of having $h_t(x, y)$ in the aggregate state variable by realizing that the dependency of $W_t(x, y)$ on $h_t(x, y)$ is rather mild. As a result, dimension reduction tools from the literature on heterogeneous agents can easily be applied.

3.1. The value of a job to workers

The value of a job to a worker of type x, when working with a company of type y offering a wage w, when the aggregate state variable is $\Gamma_t \equiv \big(z_t, h_t(x, y)\big)$ writes:

$$\begin{aligned} W(x, y, w, \Gamma_t) = w + \frac{1}{1+r}\, \mathbb{E}_{\Gamma_{t+1}|\Gamma_t} \Big[\big(\delta + (1-\delta)\mathbb{1}(S(x, y, z_{t+1}) < 0)\big) U(x, z_{t+1}) \\ (1-\delta)\mathbb{1}(S(x, y, z_{t+1}) \geq 0) \Big(s\lambda(\Gamma_{t+1}) \int_0^1 \frac{v_t(y', \Gamma_{t+1})}{V(\Gamma_{t+1})} I(x, y, y', z_{t+1}) dy' \\ + (1 - s\lambda(\Gamma_{t+1})) R(x, y, w, \Gamma_{t+1}) \Big) \Big] \end{aligned} \tag{1.16}$$

with

$$I(x, y, y', z_t) = \begin{cases} S(x, y, z_t) & \text{if } S(x, y', z_t) > S(x, y, z_t) \geq 0 \\ S(x, y', z_t) & \text{if } S(x, y, z_t) > S(x, y', z_t) > 0 \\ 0 & \text{else} \end{cases} \tag{1.17}$$

and

$$R(x, y, w, \Gamma_t) = \begin{cases} W(x, y, \phi^0(x, y, \Gamma_t), \Gamma_t) & \text{if } \Delta(x, y, w, \Gamma_t) < 0 \leq S(x, y, z_t) \\ W(x, y, \phi^1(x, y, \Gamma_t), \Gamma_t) & \text{if } \Delta(x, y, w, \Gamma_t) > S(x, y, z_t) \geq 0 \\ W(x, y, w, \Gamma_t) & \text{else} \end{cases} \tag{1.18}$$

The flow value in equation (1.16) is simply the wage received by the worker. The continuation value in equation (1.16) contains three components. The first line takes into account the probability that an unemployed worker loses her job, which could happen for exogenous or endogenous reasons. If the worker loses her job next period, she receives the value of unemployment $U(x, z_{t+1})$. If the worker stays employed, two cases can occur. Either the worker meets with another firm, which happens with probability $s\lambda(\Gamma_{t+1})$ and is taken into account in the second line of equation (1.16). Or no meeting occurs, which happens with probability $1 - s\lambda(\Gamma_{t+1})$, but re-bargaining is still possible if one of the two parties has a credible threat to break the match. Intra-firm re-bargaining is taken into consideration in the third line of equation (1.16).

The function $I(.)$ captures the change in wages that may result from a worker meeting with an alternative employer. If the poaching firm, characterized by a productivity parameter y', is a better match for a worker of type x, the worker changes job and gets the full match surplus from her *previous* employer. If the poaching firm is not a better match for a worker of type x, the worker reveals the meeting to her current employer to trigger a wage increase. The current employer, who wants to retain her employee, makes a counter-offer that matches the best offer that firm y' can make. At this new wage, the worker has no incentive to move to firm y' and she stays with firm y. If the poaching firm is a not a credible employer, the meeting is not revealed and has no impact on the current match.

The function $R(.)$ takes into account *intra-firm* re-bargaining. That is, the renegotiations that may still occur within the same company, even though a worker did not meet with an alternative employer. If a worker has a credible threat to leave because market conditions are such that the current wage is too low, the wage is re-bargained up to $\phi^0(x, y, \Gamma_t)$. This happens when the worker's surplus $\Delta(x, y, w, \Gamma_t) \equiv W(x, y, w, \Gamma_t) - U(x, z_t)$ is negative. When market conditions are such that a firm's surplus $\Pi(x, y, w, \Gamma_t)$ is negative, the wage is re-bargained down to $\phi^1(x, y, \Gamma_t)$. The firm has a credible threat to break the match when the worker's surplus is greater than the joint surplus of a match $S(x, y, z_t) \equiv \Delta(x, y, w, \Gamma_t) + \Pi(x, y, w, \Gamma_t)$.

3.2. Dimension reduction

An important feature of equation (1.16) is that the endogenous distribution of matches $h_t(x, y)$ only matters through two objects: (i) the next period's job meeting rate $\lambda(\Gamma_{t+1})$ (ii) the next period's endogenous distribution of vacancies $q(y, \Gamma_{t+1}) \equiv v(y, \Gamma_{t+1})/V(\Gamma_{t+1})$. This fact implies that to make a decision today, firms and workers only need to forecast three quantities: the next period's aggregate productivity level z_{t+1} (which is trivial given

the assumption (1.6) that aggregate productivity follows an $AR(1)$ process) and the duo $(\lambda(\Gamma_{t+1}), q(y, \Gamma_{t+1}))$. As previously explained, the job meeting rate impacts the likelihood of experiencing job-to-job transitions and promotions, so it makes sense that the job meeting rate is an element of the state variable. The distribution of vacancies across types $q(y, \Gamma_{t+1})$ is also relevant for workers, because the types of firms currently posting vacancies influence the profitability of meetings.

To reduce the dimension of the relevant state variable to a finite dimensional object, let us follow a strategy similar to the one employed by Krusell and Smith (1998). Let us postulate that agents use a simple parametric forecasting rule to predict next period's job meeting rate:

$$\lambda_{t+1} = f_\lambda(\Omega|\boldsymbol{\theta_\lambda}) \tag{1.19}$$

with $\Omega \equiv (z_{t+1}, \lambda_t) \cup \Phi$, where Φ contains relevant variables known in the current period.[9] Agents also use a simple parametric rule to keep track of the endogenous distribution of vacancies $q(y, \Gamma_{t+1})$. As in Algan et al. (2008) or Winberry (2018), my strategy is to use a parametric function $\tilde{q}(y|\boldsymbol{q_{t+1}})$ to approximate $q(y, \Gamma_{t+1})$, where $\boldsymbol{q_{t+1}}$ is a finite-dimensional vector. In practice, I use a Beta density, which performs extremely well.[10] Agents are endowed with an additional forecasting rules to keep track of the shape parameters of the Beta density $\boldsymbol{q_{t+1}} = (a_{t+1}, b_{t+1})$:

$$\boldsymbol{q_{t+1}} = f_q(\Omega|\boldsymbol{\theta_q}) \tag{1.20}$$

Conditional on the forecasting rules $\boldsymbol{\theta} \equiv (\boldsymbol{\theta_\lambda}, \boldsymbol{\theta_q})$, the value of a y job to a worker of type x with wage w can be written as:

$$\begin{aligned} W(x, y, w, z_t, \lambda_t|\boldsymbol{\theta}) = w + \frac{1}{1+r}\mathbb{E}_{\Gamma_{t+1}|\Gamma_t}\Big[\big(\delta + (1-\delta)\mathbb{1}(S(x, y, z_{t+1}) < 0)\big)U(x, z_{t+1}) \\ (1-\delta)\mathbb{1}(S(x, y, z_{t+1}) \geq 0)\Big(s\lambda_{t+1}\int_0^1 \tilde{q}(y|\boldsymbol{q_{t+1}})I(x, y, y', z_{t+1})dy' \\ + (1 - s\lambda_{t+1})R(x, y, w, \lambda_{t+1}, \tilde{q}(y|\boldsymbol{q_{t+1}}))\Big)\Big] \end{aligned} \tag{1.21}$$

The parameter values for the forecasting rules $\boldsymbol{\theta}$ can be estimated by Monte-Carlo by simulating an economy for a long period of time. Importantly, unlike in the model of Krusell and Smith (1998), the Monte-Carlo step and the calculation of the value function step are

[9] For instance, Φ may contain z_t, the square of λ_t and an interaction term between λ_t and z_t.

[10] Using a histogram to approximate $q(y, \Gamma_{t+1})$, as in Reiter (2009), is an attractive alternative. Agents would forecast the value of $q(y, \Gamma_{t+1})$ on a deterministic grid with N elements: $\boldsymbol{q_{t+1}} = \big(q_{1;t+1}(y), q_{2;t+1}(y), ..., q_{N;t+1}(y)\big) = f_q(\Omega_t|\Theta_q)\big)$. If the economy features N discrete firm types instead of a continuum, this approach amounts to assuming that agents forecast the number of vacancies posted by each firm type. I discuss this alternative in section C.3 of the Appendix.

independent from each others. The time series needed to estimate $\boldsymbol{\theta}$, can be simulated without any reference to the value of a job $W(.)$. In a setting with heterogeneity and aggregate uncertainty, one generally has to find a fixed point for the value functions *and* the forecasting rules. In the present framework, because the forecasting rules and the value functions are orthogonal to each others, $W(.)$ has to be calculated only *once*. Besides the computational edge of the present setting, I see the main advantage of the current setting as a theoretical one. Conditional on the forecasting rules $\boldsymbol{\theta}$, the Bellman operator implicitly defined in equation (1.21) is a contraction. Hence, $W(.)$ exists and is unique.

In practice, I build Ω so that it contains $(z_{t+1}, z_t, \lambda_t)$, their square and interactions terms. I use the LASSO to determine which variables are important to forecast next period's job meeting rate. A simple forecasting rule with only first order terms emerges from this procedure:

$$\lambda_t = -0.141 + 0.675\lambda_{t-1} + 0.194 z_t$$

This result is reminiscent of the *approximate aggregation* finding, prevalent in macroeconomic models with heterogeneous agents. While a priori the state variable is infinite dimensional, once individual policy rules are aggregated, a simple rule emerges from complexity. Using the LASSO generates forecasting rules that are robust to over-fitting, as measured by the out-of-the-sample maximum absolute percentage error. A series of accuracy tests, reported in section C.1 the Appendix, underlines that the forecasting rule used by agents is extremely accurate.

3.3. Evolution of wages

It has been shown that one may solve for the function $U(.)$ and $S(.)$ in a first step, while $W(.)$ may be obtained in a second step. This subsection gives the formal expressions for the wages implied by the assumptions that unemployed workers have zero bargaining power and that workers and firms interact through the *sequential auction* framework.

Let us introduce the notation $\hat{\Gamma}_t \equiv (z_t, \lambda_t | \hat{\boldsymbol{\theta}})$, denoting the approximate aggregate state variable (conditional on the forecasting rule $\hat{\boldsymbol{\theta}}$). At every period t, each firm y offers three types of wages for each worker of type x. A starting wage $\phi^0(x, y, \hat{\Gamma}_t)$ is offered to worker moving out of unemployment, or when the worker's surplus gets negative (and the match is still feasible). The starting wage $\phi^0(x, y, \hat{\Gamma}_t)$ is implicitly defined by

$$W(x, y, \phi^0(x, y, \hat{\Gamma}_t), \hat{\Gamma}_t) = U(x, z_t) \tag{1.22}$$

When the firm's surplus is negative (and the match is still feasible), the wage is re-bargained down to $\phi^1(x, y, \hat{\Gamma}_t)$. This wage is implicitly defined by

$$W(x, y, \phi^1(x, y, \hat{\Gamma}_t), \hat{\Gamma}_t) = S(x, y, z_t) \tag{1.23}$$

When an employed worker contacts another firm, the resulting wage, denoted by $\phi^2(x, y, y', \hat{\Gamma}_t)$, depends on the incumbent and the poaching firms:

$$\phi^2(x, y, y', \hat{\Gamma}_t) = \begin{cases} \phi^1(x, y', \hat{\Gamma}_t) & \text{if } S(x, y, z_t) > S(x, y', z_t) > 0 \\ \phi^1(x, y, \hat{\Gamma}_t) & \text{if } S(x, y', z_t) > S(x, y, z_t) \geq 0 \end{cases} \tag{1.24}$$

If a worker x, currently working with firm y, meets with y' (a less productive match compared to firm y), firm y responds by offering the maximum wage firm y' can offer: $\phi^1(x, y', \hat{\Gamma}_t)$. If firm y' is a better match for worker x, firm y' offers a wage such that firm y's maximum offer is not enough to retain worker x.

While the wages defined by equation (1.22), (1.23) are implicitly defined, they can trivially be obtained numerically using a one-dimensional root-finding algorithm, such as the Brent's method.

3.4. Distribution of (starting) wages

The flow equation for the joint evolution of matches and wages, denoted by $h_t(x, y, w)$, is complicated since it involves its past values (see section E of the Appendix). It is also cumbersome to approximate $h_t(x, y, w)$ since the wage dimension is inherently continuous. One could use the method developed in Young (2010) to approximate $h_t(x, y, w)$, or use a panel with a sufficiently high number of agents.

When estimating the model, I use a much simpler endogenous object: the distribution of *starting wages* for workers exiting unemployment, denoted by $h_{t,0}(x, y) \equiv h_t(x, y, \phi^0(x, y, \hat{\Gamma}_t))$. Each period, the equation for $h_{t,0}(x, y)$ solves:

$$h_{t,0}(x, y) = u_{t+}(x)\lambda_t \frac{v(y, \hat{\Gamma}_t)}{V(\hat{\Gamma}_t)} \mathbb{1}\{S(x, y, z_t) \geq 0\} \tag{1.25}$$

Contrary to the joint distribution of matches and wages, $h_{t,0}(x, y)$ is only bi-dimensional and is memory-less, which makes it an appealing object for estimation purposes.

The model predicts interesting wage dynamics. On the one hand, when the aggregate productivity parameter is high, workers expect it to stay high in the future. Higher future prospects decrease today's reservation wage, putting a downward pressure on the starting wage ("expectation effect"). The expectation effect is particularly strong in the present setting because workers have zero bargaining power. On the other hand, today's output

value $p(x, y, z)$ goes up, increasing the value of workers from the firms' perspective, making them more willing to pay high wages ("output value effect").

These two effects are combined with the dynamics of sorting along the cycle, which may alter the dynamics of aggregate wages through composition effects. A "cleansing" effect implies that only the better matched workers stay employed during a downturn. Simultaneously, firms may find it harder to employ good workers in recessions, driving them to post more low-quality jobs (the "sullying effect", see Barlevy (2002)).

4. Parametrization and estimation

This section presents the parametrization of the model before discussing to what extent the model can fit US data. This section also contains a heuristic discussion of identification.

4.1. Parametrization

I choose a parametrization that nests the one used in Lise and Robin (2017).[11] More specifically, I use a home production of the form:

$$b(x, z) = (b_0 + b_1 z) \times \bar{b}(x) \tag{1.26}$$

where $\bar{b}(x) \equiv 0.7 \times p(x, y^*(x), 1)$ and $y^*(x) = \arg\max_y p(x, y, 1)$. The function $\bar{b}(x)$ captures a fixed proportion of output when the worker is in his/her optimal match, at the neutral productivity state.

It turns out that a key ingredient to jointly match wage and employment moments is to allow for home production $b(.)$ to be a function of the aggregate state variable z_t. A tendency of $b(x, z_t)$ to be increasing with z would limit the "expectation effect", mentioned above, which pushes down starting wages in good times. Why would home production depend on z in the first place? One leading reason is that home production contains unemployment benefits, which are generally equal to a fraction of past labor income. Empirically, labor income tends to increase in booms, causing unemployment benefits to rise.

For the aggregate productivity level, I make the assumption that z_t follows an $AR(1)$ process of the form:

$$z_{t+1} = \rho z_t + \sigma\sqrt{1 - \rho^2}\varepsilon_{t+1} \tag{1.27}$$

[11] I do this for both theoretical and practical reasons. From a theoretical point of view, it is reassuring that the parametrization can simplify to one that was proven to successfully replicate a vast array of employment moments. From a practical perspective, it is convenient that my parametrization nests an existing one, as it helps me in formulating a meaningful set of starting values.

Worker's ability types x are distributed according to a Beta distribution with shape parameters β_1 and β_2. Output at the match level is given by a polynomial of the form:

$$p(x, y, z) = z(p_1 + p_2 x + p_3 y + p_4 x^2 + p_5 y^2 + p_6 xy) \tag{1.28}$$

Such a specification allows for rich interactions between worker and firm types, but is quite restrictive with regard to the aggregate productivity variable, which enters the production function multiplicatively.

I assume the following simple form for the cost of posting v vacancies:

$$c(v) = c_0 \frac{v^{1+c_1}}{1+c_1} \tag{1.29}$$

For the matching function, the number of meetings is given by a standard Cobb-Douglas function, combining search efforts and the total number of vacancies:

$$M_t = \alpha L_t^{\omega} V_t^{1-\omega} \tag{1.30}$$

4.2. Estimated values and model fit

In the present setting, a likelihood-based estimation is unfeasible because the likelihood function is not tractable. Instead, I estimate the model using the Simulated Method of Moments (SMM) with US data. I include employment-related moments and moments from the starting wage distribution. In practice, I include the elasticity of the first 9th deciles of the starting wage distribution in the list of moments to be matched (see Tables 1.8 and 1.7). In total, I use 35 moments to estimate 16 parameters. Practical details regarding the estimation procedure are listed in the section G of the Appendix .

Estimated parameters are presented in Table 1.6. The parameter b_1 is found to be positive. Indexation of unemployment benefits on past labor income would imply this feature. Because home production is slightly pro-cyclical, unemployed workers are more picky during expansions. Having more selective workers mitigates the "expectation effect" on the worker's side, potentially pushing down the reservation wage.

Table 1.7 underlines that the model does a great job at reproducing employment and production moments for the US between 1951 and 2012. Correlations of the unemployment rate with important economic variables, such as the market tightness, transition rates and GDP have the same signs as their empirical counterparts. Table1.8 shows that the model can replicate fairly well moments for the wage of new hires. The model captures well the fact that high wage deciles are more correlated with labor productivity than low wage deciles. In

empirical time series, wages for all workers are in between two and three times less volatile than wages of new hires. The model qualitatively replicates this feature. The sluggishness of simulated wages comes from the assumption that wages are constant, unless a credible threat to an ongoing job arises. However, wages in empirical series are even more sluggish. Adding a renegotiation cost when an intra-firm renegotiation occurs would lead to more sluggishness, without significantly altering the model.

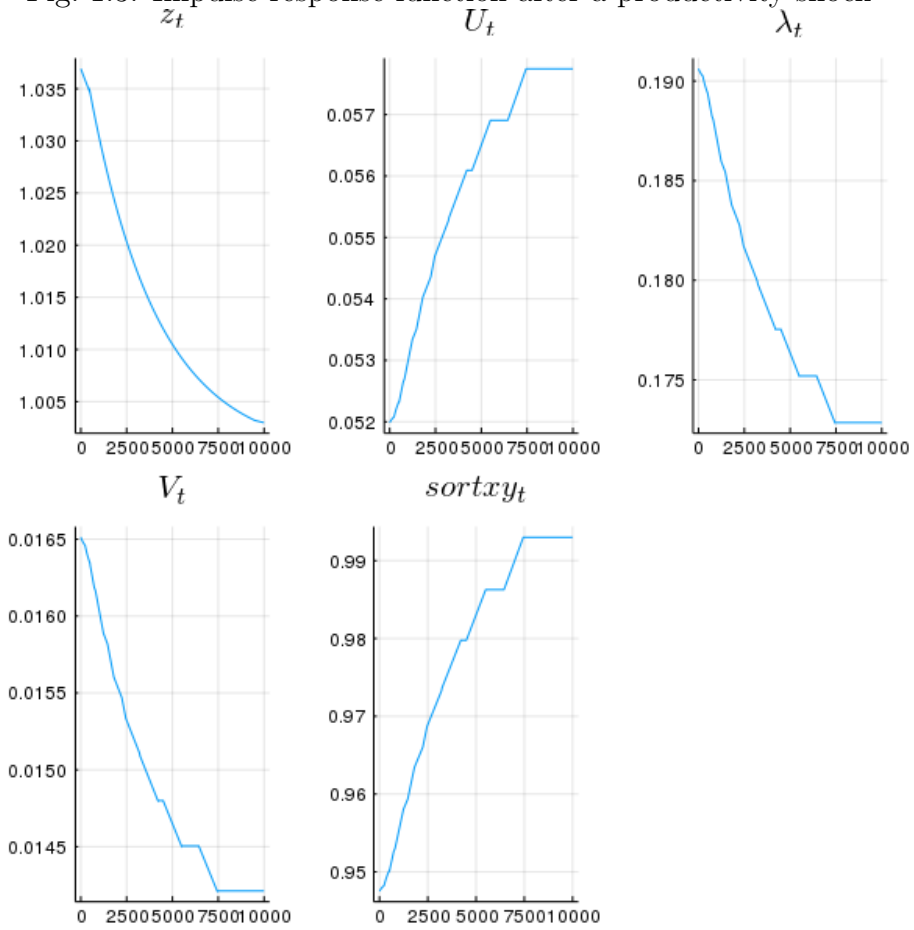
The impulse response function indicates that the model features a "cleansing effect": the quality of sorting between firms and workers declines by approximately 5% after a one-standard deviation productivity shock (Figure 1.3). The parameter estimates imply that the labor market features associative matching, with high productivity firms preferring to match with high productivity workers (see Figure 1.4). The value for the parameter s suggests that unemployed workers are searching for a job with an intensity that is 16 times bigger than the intensity of already employed workers.

Table 1.6: Estimated Parameters

Parameter	Value	Estimated	Description
r	0.05	No	annual interest rate
α	0.501	Yes	matching function parameter
ω	0.5	No	matching function parameter
s	0.061	Yes	search intensity parameter
c_0	0.405	Yes	vacancy posting cost parameter
c_1	0.030	Yes	vacancy posting cost parameter
δ	0.012	Yes	exogenous separation rate
ϕ	0.083	Yes	productivity shock parameter
ρ	0.999	No	productivity shock parameter
β_1	13.490	Yes	worker heterogeneity parameter
β_2	16.735	Yes	worker heterogeneity parameter
p_1	0.053	Yes	value added parameter
p_2	2.162	Yes	value added parameter
p_3	-0.157	Yes	value added parameter
p_4	8.818	Yes	value added parameter
p_5	-1.880	Yes	value added parameter
p_6	7.126	Yes	value added parameter
b_0	0.478	Yes	home production parameter
b_1	0.814	Yes	home production parameter

Notes: Parameter values were estimated using the Simulated Method of Moments. Parameter values were rounded to the nearest thousandths. I take the value of parameter ρ, characterizing the persistence of TFP (at the weekly frequency) from Lise and Robin (2017).

Fig. 1.3. Impulse response function after a productivity shock

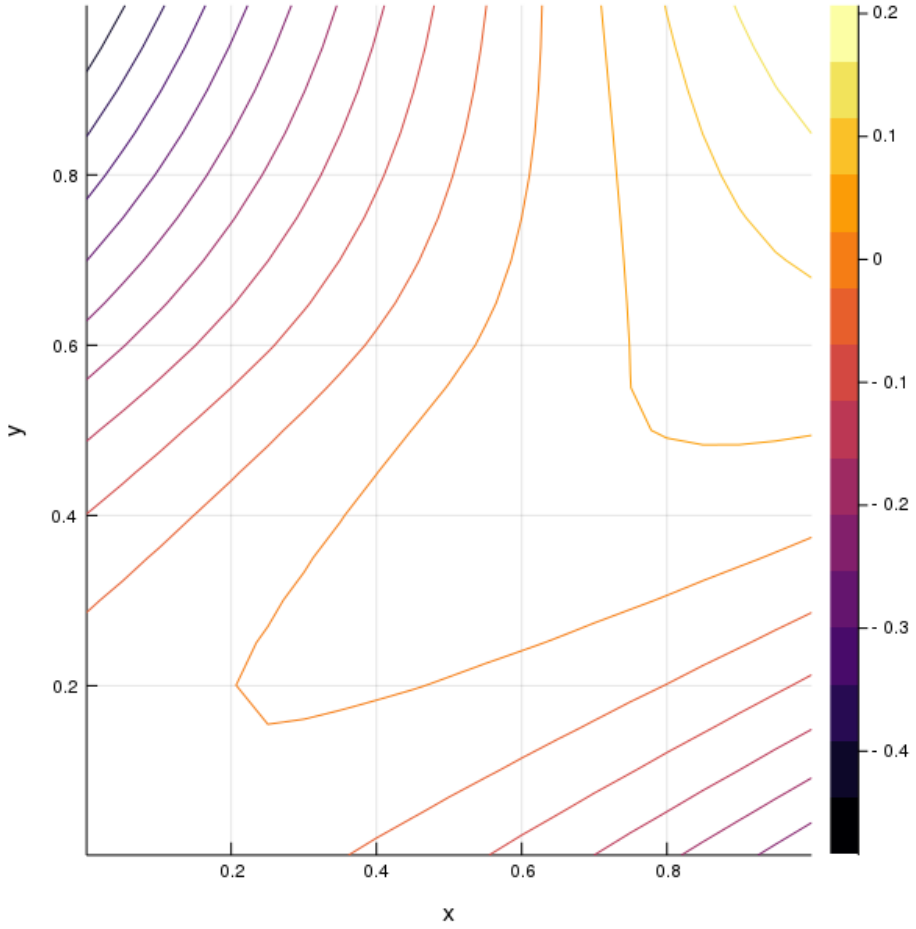


Notes: This figure shows the response of the unemployment rate U_t, the job meeting rate λ_t, the number of vacancies V_t, and sorting after a one-standard deviation positive TFP shock. I quantify sorting using the formula

$$sortxy_t = \frac{1}{C} \exp\Big(- \int_0^1 \int_0^1 h_t(x,y)|y - y^*(x,z_t)|dxdy \Big)$$

where $y^*(x,z)$ is the firm type that maximizes the joint surplus of a match for a worker of type x when the productivity parameter is equal to z_t. C is a normalizing constant chosen such that the best observed sorting value is equal to 1. A value of $sortxy_t$ below 1 indicates a sub-optimal pairing between firms and workers, in the sense that production could be improved by reallocating workers to firms with productivity types closer to $y^*(x,z_t)$.

Fig. 1.4. Value of net production at the match level $s(x, y, 1)$



Notes: This figure shows a contour plot for the value of net production at the match level $s(x, y, 1) = p(x, y, 1) - b(x, 1)$ when the aggregate state variable z_t is at its neutral state ($z_t = 1$).

Table 1.7: Empirical and Simulated Employment Moments. US (1951-2012)

	Data	Model		Data	Model
$E[U]$	0.058	0.059	$std[U_{27+}]$	0.478	0.150
$E[U_{5+}]$	0.035	0.028	$std[UE]$	0.127	0.089
$E[U_{15+}]$	0.018	0.010	$std[EU]$	0.100	0.058
$E[U_{27+}]$	0.010	0.006	$std[EE]$	0.095	0.116
$E[UE]$	0.421	0.527	$std[V/U]$	0.381	0.333
$E[EU]$	0.025	0.032	$corr[U,V]$	-0.846	-0.985
$E[EE]$	0.025	0.015	$corr[U,GDP]$	-0.860	-0.983
$E[V/U]$	0.634	0.251	$corr[V,GDP]$	0.721	0.994
$std[V]$	0.206	0.188	$corr[UE,GDP]$	0.878	0.957
$std[GDP]$	0.033	0.046	$corr[EU,GDP]$	-0.716	-0.970
$std[U]$	0.191	0.146	$corr[UE,EE]$	0.695	0.947
$std[U_{5+}]$	0.281	0.236	$corr[EE,GDP]$	0.495	0.947
$std[U_{15+}]$	0.395	0.257	$autocorr[GDP]$	0.932	0.987

Notes: Data columns are from Lise and Robin (2017) for the period 1951-2012, with the exception of the moments involving job-to-job transitions, which is for the period 1994-2011. $E[U]$ is the average quarterly unemployment rate. $E[U_{5+}]$, $E[U_{15+}]$ and $E[U_{27+}]$ are the average quarterly unemployment rates for more than 5, 15 and 27 weeks respectively. $E[UE]$, $E[EU]$ and $E[EE]$ are the average quarterly job-finding, job-losing and job-to-job transition rates. V denotes the number of vacancies. $std[x]$ denotes the standard deviation of the variable x. $corr[x,y]$ denotes the correlation between variables x and y. $autocorr[x]$ denotes the auto-correlation of variables x.

Table 1.8: Elasticity of wage and starting wage deciles

	New hires		All workers	
	Data	Model	Data	Model
P10	0.44	0.45	0.20	0.40
P20	0.32	0.38	0.13	0.32
P30	0.34	0.35	0.18	0.33
P40	0.31	0.34	0.18	0.30
P50	0.35	0.35	0.12	0.28
P60	0.42	0.40	0.30	0.31
P70	0.37	0.42	0.26	0.32
P80	0.45	0.68	0.28	0.39
P90	0.65	0.77	0.46	0.55

Notes: This table displays the elasticity of starting wage deciles with respect to changes in aggregate labor productivity. The first column is based on observations from CEPR CPS ORG dataset. The second column is based on simulated data. The fourth column is based on simulated data generated using a sample of 1000 agents during 6000 periods (weeks), discarding the first 1000 periods.

4.3. Identification

I provide a heuristic justification of my identification strategy and I rely on numerical tools to support my intuitions. First of all, exit from unemployment and job-to-job mobility are key in determining the value of search efficiency α and the relative search intensity between unemployed and unemployed worker, denoted by s. Holding other parameters constant, a higher job-finding rate is associated with a higher search efficiency and more job-to-job transitions indicate a higher relative search intensity of employed workers. The steady-state value (or its long-run average) of the unemployment rate is informative on the exogenous job-destruction rate δ. The unemployment rates by duration are informative on the distribution of types within the economy. For instance, more long-term unemployment indicates a distribution of worker types tilted towards low types. The elasticity of starting wages captures to what extent the "cleansing" effect (survival of better workers) and the "sullying" effect (firms posting low quality jobs) dominate during recessions. Thus, deciles of the starting wage distribution provide valuable information on the matching function, the vacancy cost function, the distribution of types and parameters for home production $b(.)$ and market production $p(.)$. Table 1.14, reporting values for the Jacobian of the function

mapping parameters to simulated moments $f : \boldsymbol{p} \to \boldsymbol{m}$, largely confirms these intuitions. In addition to the statistics discussed above, market tightness is found to be key in disciplining the model. Because the market tightness contains information on both sides of the market, it is not surprising that it has a strong identifying power. Interestingly, the partial derivatives of f with respect to the deciles of the wage distribution are in between one and two orders of magnitude larger the partial derivatives involving employment moments. The information contained in (starting) wages is substantial. More complex parametrizations could potentially be estimated using the information contained in the variation of wages.

5. Labor income shocks and sorting

Having estimated the model using US data, it is now possible to use the model to draw new insights about labor income shocks along the business cycle. What are the main mechanisms behind labor income shocks? In particular, what is the contribution of the inter-firm channel to variations over the business cycle?

To quantify the importance of the inter-firm channel on labor income shocks, I first measure the persistence of labor income losses for displaced workers. If follow the literature on displaced workers (see for instance Stevens (1997)) and run the following regression on simulated data:

$$y_{it} = \sum_{k=0}^{10} \delta_k D_{it}^k + \varepsilon_{it} \tag{1.31}$$

The variable D_{it}^k is an indicator variable equal to 1 if worker i was a displaced worker k periods ago. The coefficients δ_k measures the current effect of job displacement on y_{it}. I classify a worker as displaced if the worker experienced unemployment in year k for at least 10 weeks and if the worker was poorly matched upon finding a new job. I consider a worker to be poorly matched if the (absolute value) of the distance between the current firm y and the optimal match y^* is bigger than a threshold value [12]. The left panel of Figure 1.5 displays the recovery of yearly earnings, the number of weeks worked and the hourly wage for displaced workers relative to non-displaced workers using simulated data. Displaced workers experience on average a 37% drop in yearly labor income the year of displacement, mainly driven by a loss in weeks worked. As the number of weeks worked quickly recovers, the milder initial loss in terms of hourly wage takes much longer to recover. Long-lasting labor earnings losses are driven by mismatches after a job loss. As sorting improves over time, the

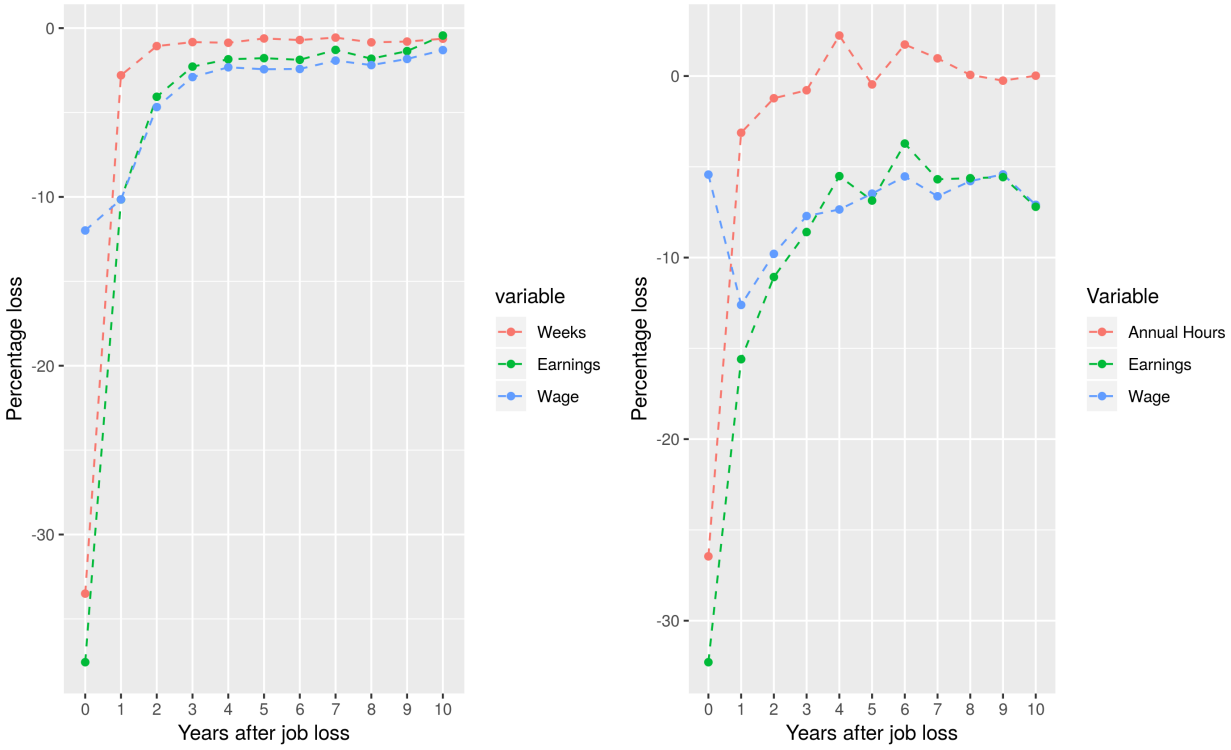
[12] As a threshold value, I use 2 times the standard deviation of the distances (absolute value difference) between the current firm types and the optimal firm type y^*: $|y - y^*|$)

initial loss in hourly wage slowly vanishes. Because search for better firm types is random, improvements in sorting takes time to materialize. These results are consistent with the empirical patterns reported by the literature on job displacement. The right panel of Figure 1.5, based on empirical work from Huckfeldt et al. (2016), also shows that the initial impact of a job loss is mainly driven by a loss in annual hours worked. However, as predicted by the model, the persistence in labor earning losses is to be attributed to an enduring loss in wages. While the model predicts that displaced workers eventually recover from a job loss, the right panel of Figure 1.5 indicates that empirically there is a 5% permanent drop in wages.

The permanent drop in wages can be explained by a loss in human capital accumulation for displaced workers, which is not included in the present model. In section H of the Appendix, I explain how the model could be augmented to include human capital accumulation, using the learning-by-doing approach. I also show that if workers accumulate human capital in a learning-by-doing fashion, the determination of job feasibility is still independent from the determination of wages. Hence, the algorithm discussed in section 3 may still be used.

To further understand how labor income shocks and sorting are related, I run an experiment in which I keep only one type of firm in the economy (the median firm). I simulate both economies forward, keeping the series of aggregate shocks constant. Results are presented in Table 1.9. When the inter-firm channel is nonexistent, business cycle fluctuations in yearly labor income are reduced by approximately 12% compared to the baseline model with firm heterogeneity. The inter-firm channel is particularly important for the *tails* of labor income changes, as visible on Figure 1.6. When the inter-firm channel is shut down, the *change* over the business cycle in the probability of experiencing more than a 50% drop in labor income over a year is reduced by approximately 19%. The change over the business cycle in the likelihood of experiencing more than a 50% increase in labor income over a year is reduced by 97.5%.

Why does sorting matter for changes in labor income over the cycle? In the present model, changes in labor income are either caused by a change in employment status (employed or unemployed) or by a change in wages. Changes in wages are mainly driven by the inter-firm channel. That is, the reallocation of workers from bad matches to better matches. The inter-firm channel depends on the number of meetings occurring each period. In a recession, the number of on-the-job meeting $s\lambda_t$ plummets causing the inter-firm channel to dry up. As a result, in the aftermath of a bad productivity shock, workers stay mismatched longer, generating labor income losses. In an economy with only one type of firm, this dimension vanishes and the model generates more modest fluctuations in labor along the business.



(a) Model

(b) Data

Fig. 1.5. Displaced workers and income losses

Notes: The left panel shows the effect of job displacement on yearly labor income, the annual number of weeks worked, and the hourly wage relative to workers who have not been displaced within the past ten years. Calculations are based on simulated data. The right panel is from Huckfeldt et al. (2016), based on data from the Panel Study of Income Dynamics from 1968 to 1997. It shows the (relative) effect of job displacement on annual hours worked, yearly labor income, and hourly wage.

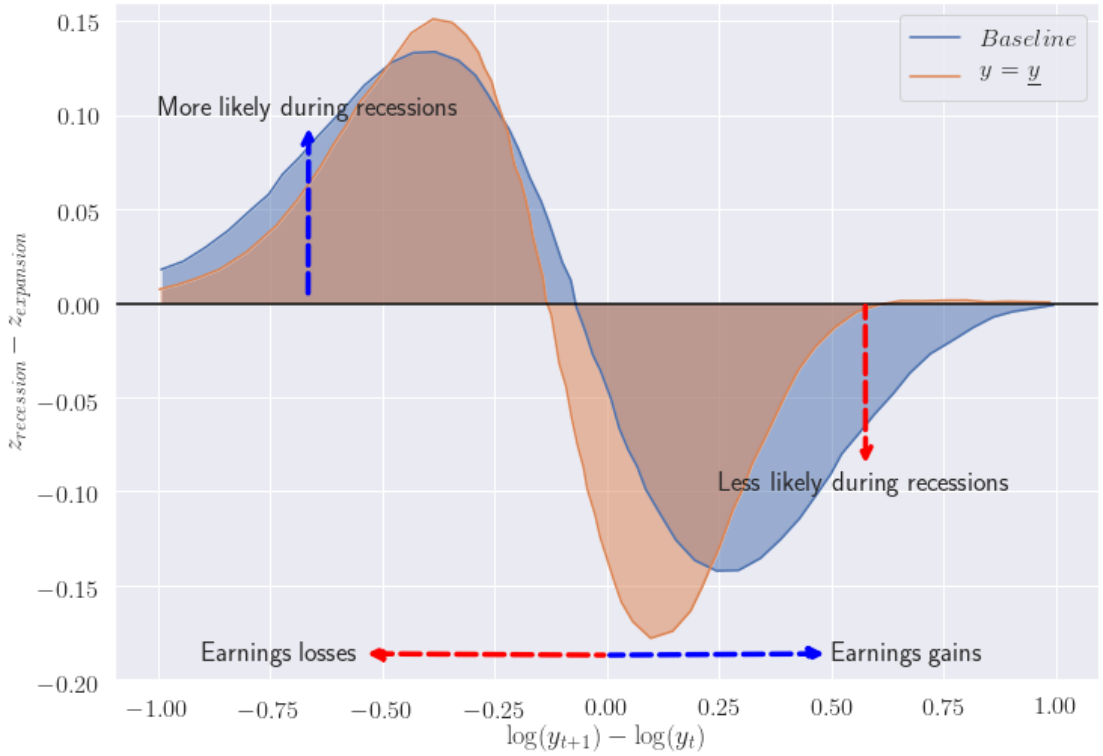
Table 1.9: Variations in the probability of labor income changes over the cycle

	Changes in probabilities over the cycle		
	Data	Baseline	$y = \underline{y}$
A	0.164	0.149	0.133
$\Delta \Pr(\Delta Income_{it} < 50\%)$	1.3%	3.1%	2.5%
$\Delta \Pr(\Delta Income_{it} > 50\%)$	-1.5%	-1.4%	-0.035%
$\Delta \Pr(\Delta Income_{it} < 25\%)$	2.9%	6.3%	5.5%
$\Delta \Pr(\Delta Income_{it} > 25\%)$	-3.0%	-4.4%	-1.8%

Notes: The first column is based on data from Guvenen et al. (2014) (US data, 1995–96 versus 2008–9.). The second and third columns are based on simulated data. The variable A is defined as the area between the line representing the density gap and the x-axis. the pdf gap is defined as the pdf for $(\log(y_{t+1}) - \log(y_t))$ in recession minus the pdf for $(\log(y_{t+1}) - \log(y_t))$ in expansion. If the business cycle did not generate changes in labor income, the density gap would be null and A would be equal to 0 (see Figure 1.6).

Other rows measure the *change* in the probability of a certain event relative to yearly labor income. For instance, the row $\Delta \Pr(\Delta Income < 50\%)$ measures the *change* in the probability that a worker loses more than 50% of her labor income over a year (recession minus expansion). The second column is based on simulated data from the model with the estimated parameters. The third column is based on simulated data with only one type of firm (the median firm).

Fig. 1.6. Change in labor income risk over the cycle and sorting



Notes: This figure shows the density gap, defined as the pdf for $(\log(y_{t+1}) - \log(y_t))$ in recession minus the pdf for $(\log(y_{t+1}) - \log(y_t))$ in expansion. If the business cycle did not generate changes in labor income, the density gap would be null. The blue line is the baseline model. The orange line is based on a simulation with only one type of firm (the median firm).

6. Optimal unemployment insurance

The analysis has established that sorting between firms and workers is central for idiosyncratic labor income risk, especially when considering extreme labor income changes. Can a government improve the market outcome using a simple unemployment benefit rule?

This question has already been explored by Lise et al. (2016) in a similar setting, but without aggregate uncertainty. The authors find that the optimal unemployment scheme delivers an improvement of 1.4%. One may wonder to what extent their findings extend to a model with aggregate risk. One may also wonder to what extent the government may attenuate fluctuations in labor income with a simple unemployment benefit scheme. Before analyzing the optimal unemployment benefit policy, one may wonder what are the

imperfections that would justify government intervention in the first place. In a model featuring matching and search frictions, there exist congestion externalities (Hosios (1990)). Some types of workers may be searching too much, especially in the present setting in which workers and firms are heterogeneous. Some high-type firms would probably post more vacancies if they were certain to find only high-type workers on the labor market. For this reason, it might be optimal that low-type workers search less, which could be incentives by providing a higher replacement rate. I solve for an optimal unemployment benefit equal to a fix proportion of the expected labor income at the steady-state:

$$b_{UI}(x) = b \int_0^1 p(x, y, 1)h(y|x)dy \tag{1.32}$$

The unemployment insurance is funded by a proportional tax on match output:

$$\int b_{UI}(x)u(x)dx = \tau \int p(x, y, 1)h(x, y)dxdy \tag{1.33}$$

The welfare criterion I use is simply the sum of market output (net of taxes), plus home production and unemployment insurance, minus the cost of creating vacancies:

$$\begin{aligned} W_{b_0} = &\int_0^1 \int_0^1 (1-\tau)h(x, y)p(x, y, 1)dxdy + \int_0^1 \big(b(x, 1) + b_{UI}(x)\big)u(x)dx + \\ &- \int_0^1 c(v(y))dy \end{aligned} \tag{1.34}$$

Equation (1.34) underlines the mechanism at play. Taxing output hurts employed workers (the first term) and improves the welfare of unemployed workers (the second term). Yet, the composition of matches is altered — the first term in (1.34) might change because of $h(x, y)$ — and firms may change the amount of vacancies they post, potentially generating savings. Results are presented in Figure 1.7.

The optimal value for b_0 is approximately 0.06, which is funded by a 1.3% tax on output at the match level. This is slightly higher than the 0.95% tax rate found in Lise et al. (2016). This tax on output is used to fund an unemployment insurance that represents approximately 19.2% of the aggregate value of home production. The unemployment insurance scheme rises the unemployment rate to approximately 7%. The welfare gains are driven by a 13.16% increase in sorting and a decrease in the cost of creating vacancies, which are enough to offset the distortions created by the tax on output. Table 1.10 underlines that UI stabilizes labor income by approximately 2%. This policy fosters employment for high type workers and firms. The tax on output provides a safety net for low-skilled workers more likely to

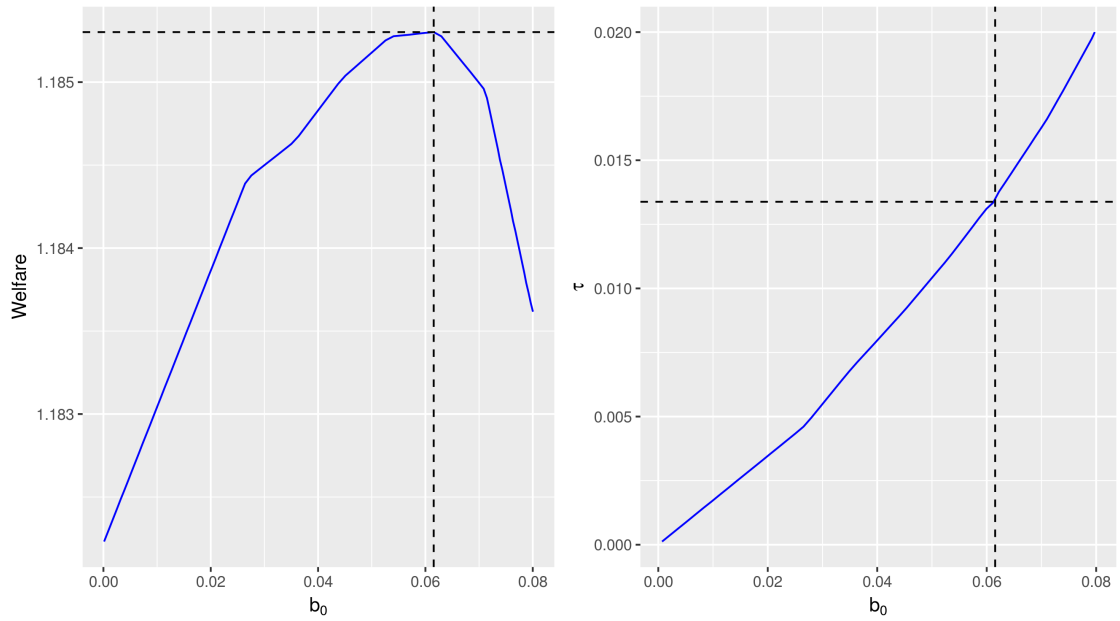
bear the burden of unemployment along the business cycle.

Table 1.10: Variations in the probability of labor income changes over the cycle with the optimal UI

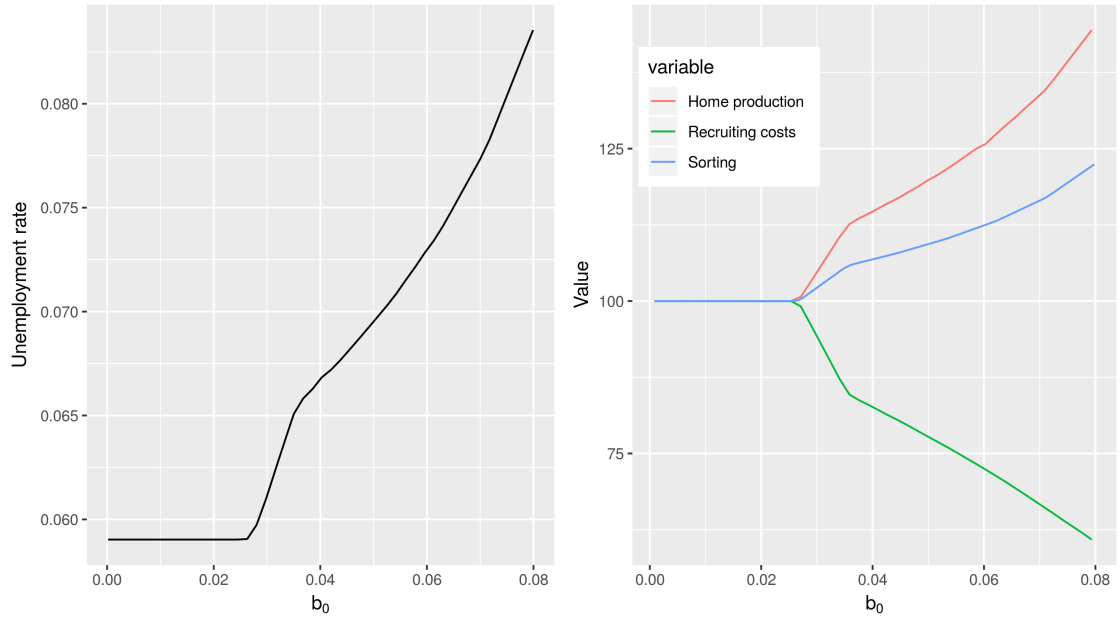
	Changes in probabilities over the cycle		
	Data	Baseline	UI
A	0.164	0.149	0.146
$\Delta P(\Delta Income_{it} < 50\%)$	1.3%	3.1%	3.1%
$\Delta P(\Delta Income_{it} > 50\%)$	-1.5%	-1.4%	-1.2%
$\Delta P(\Delta Income_{it} < 25\%)$	2.9%	6.3%	6.2%
$\Delta P(\Delta Income_{it} > 25\%)$	-3.0%	-4.4%	-4.2%

Notes: The first column is based on data from Guvenen et al. (2014) (US data, 1995-96 versus 2008-9.). The second and third columns are based on simulated data. The variable A is defined as the area between the line representing the density gap and the x-axis. the pdf gap is defined as the pdf for $(\log(y_{t+1}) - \log(y_t))$ in recession minus the pdf for $(\log(y_{t+1}) - \log(y_t))$ in expansion. If the business cycle did not generate changes in labor income, the density gap would be null and A would be equal to 0.

Other rows measure the *change* in the probability of a certain event relative to yearly labor income. For instance, the row $P(Income < 50\% decrease)$ measures the *increase* in the probability that a worker loses more than 50% of her labor income over a year when the economy is in recession. The second column is based on simulated data from the model with the estimated parameters. The third column is based on simulated data with only one type of firm (the median firm).



(a) Welfare, b_0 and τ



(b) b_0 and labor market outcome

Fig. 1.7. Optimal UI

Notes: The top-left quadrant shows the link between the replacement rate b_0 and welfare. Welfare is the sum of net market production, home production, the total value of UI, minus the cost of creating vacancies. The top right-quadrant shows the link between b_0 and the corresponding tax rate on production τ. The bottom-left quadrant shows the link between b_0 and the unemployment rate. The bottom-right quadrant shows the link between b_0, home production, recruiting costs and sorting quality (normalized to 100 when $b_0 = 0$).

7. Conclusion

This paper analyzes the determinants of labor income changes over the business cycle. The novelty in this analysis is to show that sorting between firms and workers matters when considering fluctuations in labor income. The mechanism is quite intuitive. Because of search frictions on the labor market, workers and firms are not necessarily well matched. The pairing between firms and workers is improved by the slow process of job-to-job transitions. In a recession, less vacancies are posted and the inter-firm channel dries up. Workers accumulate labor income losses by working with firms that are not optimal for them. While the primary driver of income losses in recession is unemployment, the sorting channel accounts for 12% of fluctuations in labor income. A simple policy can generate welfare gains in that context: unemployment insurance. By varying the replacement rate, the government alters incentives for different types of workers. In particular, low-skilled workers are more patient. High-skilled workers benefit from the resulting reduction in congestion effects. Improved sorting is more than enough to offset the distortion effects created by taxing output.

To arrive to this conclusion, I developed and estimated a dynamic search-and-matching model with heterogeneous firms and workers. While a priori not tractable, the model simplifies considerably by realizing that the state variable can be reduced to a finite dimensional vector, without losing much generality nor stability. The key to tractability lies in the fact that employment flows can be determined in a first step, independently from the wage allocation problem. While wages do dependent on the next period's job meeting rate and distribution of jobs across firm and worker types, as it is generally the case in dynamic search-and-matching models with heterogeneity, dimension reduction techniques can be used. Because the simulated series needed to perform the dimension reduction step are independent from the wages, the value function characterizing to the wage problem exists and is obtained immediately, without the need of an outer loop.

The clear separation of the employment problem from the determination of wages results from the combination of three assumptions: zero bargaining power for unemployed workers, determination of wages according to the sequential auction framework and linear utility. The existence of an other set of assumptions implying a similar separation between employment flows and wages is still an open question. In particular, relaxing the assumption of linear utility is important to analyze precautionary savings for this class of models. The extent to which the conclusions of this paper can be extended to a framework with risk-averse workers, having access to a risk-free asset, is currently being investigated.

References

Abowd, J. M., Kramarz, F., and Margolis, D. N. (1999). High wage workers and high wage firms. *Econometrica*, 67(2):251–333.

Algan, Y., Allais, O., and Den Haan, W. J. (2008). Solving heterogeneous-agent models with parameterized cross-sectional distributions. *Journal of Economic Dynamics and Control*, 32(3):875–908.

Algan, Y., Allais, O., Den Haan, W. J., and Rendahl, P. (2014). Solving and simulating models with heterogeneous agents and aggregate uncertainty. In *Handbook of Computational Economics*, volume 3, pages 277–324. Elsevier.

Auclert, A., Bardóczy, B., Rognlie, M., and Straub, L. (2019). Using the sequence-space jacobian to solve and estimate heterogeneous-agent models. Technical report, National Bureau of Economic Research.

Barlevy, G. (2002). The sullying effect of recessions. *The Review of Economic Studies*, 69(1):65–96.

Bezanson, J., Edelman, A., Karpinski, S., and Shah, V. B. (2017). Julia: A fresh approach to numerical computing. *SIAM review*, 59(1):65–98.

Bonhomme, S., Lamadon, T., and Manresa, E. (2019). A distributional framework for matched employer employee data. *Econometrica*, 87(3):699–739.

Boppart, T., Krusell, P., and Mitman, K. (2018). Exploiting mit shocks in heterogeneous-agent economies: the impulse response as a numerical derivative. *Journal of Economic Dynamics and Control*, 89:68–92.

Duffie, D. and Singleton, K. J. (1990). Simulated moments estimation of markov models of asset prices.

Gourieroux, C., Monfort, A., and Renault, E. (1993). Indirect inference. *Journal of applied econometrics*, 8(S1):S85–S118.

Guvenen, F., Karahan, F., Ozkan, S., and Song, J. (2015). What do data on millions of us workers reveal about life-cycle earnings risk? Technical report, National Bureau of Economic Research.

Guvenen, F., Ozkan, S., and Song, J. (2014). The nature of countercyclical income risk. *Journal of Political Economy*, 122(3):621–660.

Haefke, C., Sonntag, M., and Van Rens, T. (2013). Wage rigidity and job creation. *Journal of monetary economics*, 60(8):887–899.

Hansen, L. P. (1982). Large sample properties of generalized method of moments estimators. *Econometrica: Journal of the Econometric Society*, pages 1029–1054.

Hosios, A. J. (1990). On the efficiency of matching and related models of search and unemployment. *The Review of Economic Studies*, 57(2):279–298.

Huckfeldt, C. et al. (2016). Understanding the scarring effect of recessions. *Report, Economics Department*.

Krusell, P. and Smith, Jr, A. A. (1998). Income and wealth heterogeneity in the macroeconomy. *Journal of political Economy*, 106(5):867–896.

Le Grand, F., Ragot, X., et al. (2017). Optimal fiscal policy with heterogeneous agents and aggregate shocks. *Document de travail*.

Lise, J., Meghir, C., and Robin, J.-M. (2016). Matching, sorting and wages. *Review of Economic Dynamics*, 19:63–87.

Lise, J. and Robin, J.-M. (2017). The macrodynamics of sorting between workers and firms. *American Economic Review*, 107(4):1104–35.

Lise, J., Robin, J. M., Lentz, R., et al. (2017). The macrodynamics of the wage distribution. *2017 Meeting Papers, 1220, Society for Economic Dynamics*.

Ljungqvist, L. and Sargent, T. J. (1998). The european unemployment dilemma. *Journal of political Economy*, 106(3):514–550.

Mackowiak, B. and Wiederholt, M. (2009). Optimal sticky prices under rational inattention. *American Economic Review*, 99(3):769–803.

Menzio, G. and Shi, S. (2010). Block recursive equilibria for stochastic models of search on the job. *Journal of Economic Theory*, 145(4):1453–1494.

Mortensen, D. T. and Pissarides, C. A. (1994). Job creation and job destruction in the theory of unemployment. *The review of economic studies*, 61(3):397–415.

Ouyang, M. (2009). The scarring effect of recessions. *Journal of Monetary Economics*, 56(2):184–199.

Parker, J. A. and Vissing-Jorgensen, A. (2009). Who bears aggregate fluctuations and how? *American Economic Review*, 99(2):399–405.

Pissarides, C. A. (2009). The unemployment volatility puzzle: Is wage stickiness the answer? *Econometrica*, 77(5):1339–1369.

Postel-Vinay, F. and Turon, H. (2010). On-the-job search, productivity shocks, and the individual earnings process. *International Economic Review*, 51(3):599–629.

Ragot, X. and Le Grand, F. (2019). Managing inequality over the business cycles: Optimal policies with heterogeneous agents and aggregate shocks. 2019 Meeting Papers 1090, Society for Economic Dynamics.

Reiter, M. (2009). Solving heterogeneous-agent models by projection and perturbation. *Journal of Economic Dynamics and Control*, 33(3):649–665.

Robin, J.-M. (2011). On the dynamics of unemployment and wage distributions. *Econometrica*, 79(5):1327–1355.

Schaal, E. (2017). Uncertainty and unemployment. *Econometrica*, 85(6):1675–1721.

Shimer, R. (2005). The cyclical behavior of equilibrium unemployment and vacancies. *American economic review*, 95(1):25–49.

Song, J., Price, D. J., Guvenen, F., Bloom, N., and Von Wachter, T. (2018). Firming up inequality. *The Quarterly journal of economics*, 134(1):1–50.

Stevens, A. H. (1997). Persistent effects of job displacement: The importance of multiple job losses. *Journal of Labor Economics*, 15(1, Part 1):165–188.

Walentin, K. and Westermark, A. (2018). Learning on the job and the cost of business cycles. *Sveriges Rikbank Working Paper Series*.

Winberry, T. (2018). A method for solving and estimating heterogeneous agent macro models. *Quantitative Economics*, 9(3):1123–1151.

Young, E. R. (2010). Solving the incomplete markets model with aggregate uncertainty using the krusell–smith algorithm and non-stochastic simulations. *Journal of Economic Dynamics and Control*, 34(1):36–41.

A. Numerical details

When estimating the model, I choose a rather coarse grid for the x and y dimensions using 11 points. Given that these dimensions represent non-time varying discrete types, this is done without loss of generality. When solving for the functions $U(.)$, $S(.)$ and $W(.)$, I treat z as a continuous dimension and I use Gaussian quadrature nodes. To produce results comparable with Lise and Robin (2017), I use an evenly-spaced discrete grid for z when simulating time series, with 51 different values for z. Regarding the forecasting rule $f(\Omega|\boldsymbol{\theta})$, Ω contains a constant, $log(z_t)$, $log(z_{t-1})$, λ_t, λ_{t-1}, their square and interactions terms. I use the LASSO to estimate the forecasting rule, which has the advantage of automatically selecting the relevant variables needed to make accurate and reliable predictions. [13] I solve $W(x, y, z, \lambda|\boldsymbol{\theta})$ by value function iteration. To calculate moments on the full distribution of wages, I use a panel of 1000 agents. I simulate the panel for 6000 periods (weeks) and I drop the first 1000 observations to get rid of the potential impact of initial values. To minimize the SMM objective function, I use a multi-start approach. That is, I start several Nelder-Mead algorithms in parallel with different starting values. The global minimum is the minimum of the several minima for which convergence was reached. The code is implemented in Julia 0.6.4 (see Bezanson et al. (2017)).

B. Algorithm to solve and simulate the model

To solve the model, I proceed as follows:

1. Solve S(.) and U(.) by value function iteration, as in Lise and Robin (2017).
2. Simulate an economy for a long period of time and discard the first 10th observations. This step can be achieved independently of $W(.)$. This generates a synthetic sample containing $(z_t, \lambda_t)_{t=1}^T$ and any other variables of interest.
3. Find an estimate of the parametric forecasting rule, denoted by $\hat{\boldsymbol{\theta}}$.
4. Find a solution to $W(x, y, w, \hat{\Gamma}_t|\hat{\boldsymbol{\theta}})$ by value function iteration.
5. Using $W(x, y, w, \hat{\Gamma}_t|\hat{\boldsymbol{\theta}})$, $U(x, z_t)$ and $S(x, y, z_t)$, calculate the value of ϕ^0, ϕ^1 and ϕ^2 for every t of the synthetic data.

C. Accuracy of approximations

My computational strategy relies on two approximations: (i) the next period's job meeting rate and the parameters of $\tilde{q}(y|\boldsymbol{q_t})$ are well predicted by simple forecasting rules (ii) the

[13] For the LASSO, I use the package GLMNet.jl

endogenous distribution of vacancies across types $q(y)$ can reasonably be approximated by a parametric function. This section shows that both approximations are accurate. This section also shows that one may dispense with the parametric assumption for the distribution of vacancies and instead use a histogram.

C.1. Forecasting rules accuracy

I calculate how well forecasting rules predict the paths for λ_t and $\boldsymbol{q_t}$ within the sample that was used to estimate the forecasting rules (within-sample prediction power). I also simulate a new sample and I compare the actual realization of time series to the predictions implied by the forecasting rule (out-of-the-sample prediction power). The within-sample accuracy can be visualized in Figure 1.8. The estimated forecasting rules are quite precise, as underlined in Table 1.11. For instance, the median percentage error for the job meeting rate is only 0.13%. Is this number within a credible range? The literature on rational inattention has described reasons why firms may optimally commit small errors. The full optimization problem may too hard or too resource-consuming to solve. Firms may decide to use only a fraction of the full information set each period (Mackowiak and Wiederholt (2009))). The estimated forecasting rule for the job meeting rate is the following simple linear relationship:

$$\lambda_t = -0.141 + 0.675\lambda_{t-1} + 0.194 z_t$$

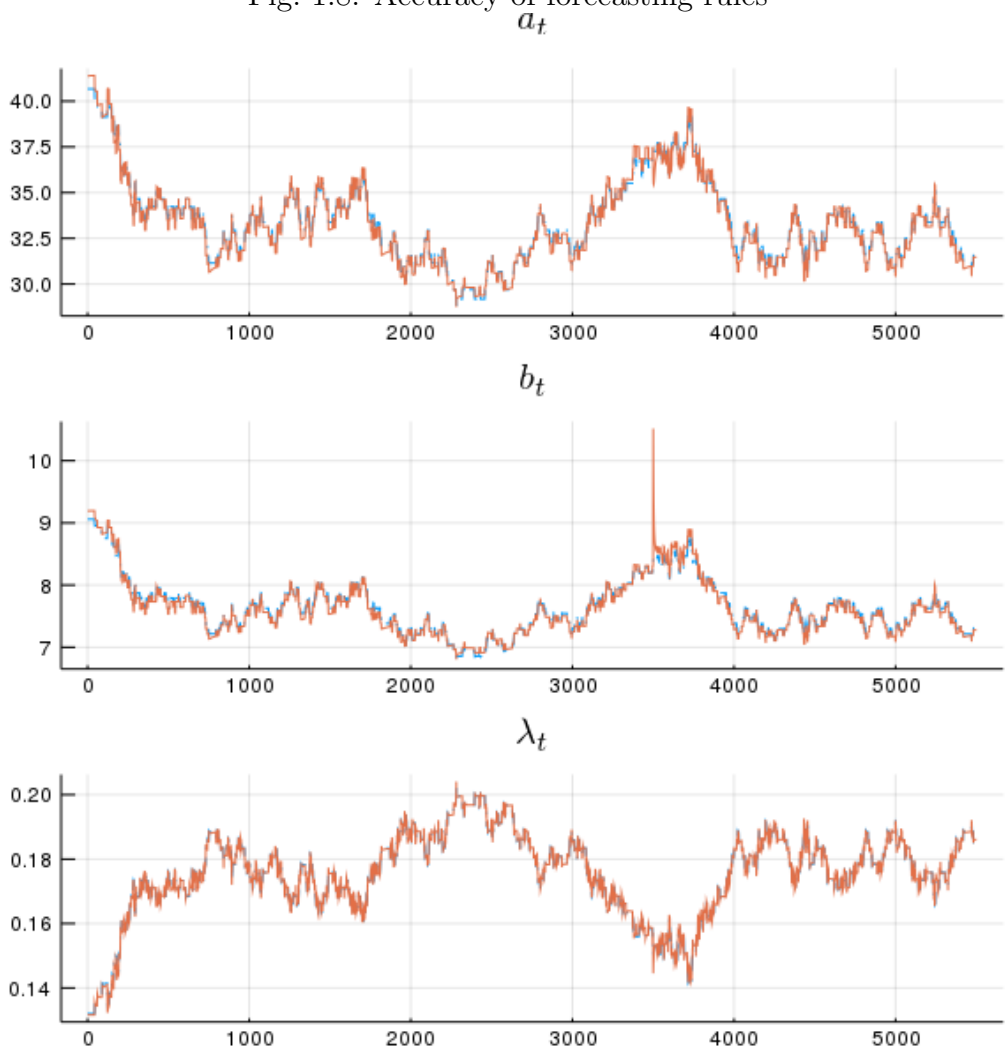
The forecasting rules for the shape parameter of the Beta density are given by:

$$a_t = 129.236 - 20.121\lambda_{t-1} - 90.718 z_t$$

$$b_t = 18.227 - 21.445\lambda_{t-1} - 6.704 z_t$$

While the list of potential predictors contains higher order terms and interaction terms, the LASSO selects only first order terms.

Fig. 1.8. Accuracy of forecasting rules



Notes: This figure shows the series for the job meeting rate λ_t and for the shape parameters of the Beta distribution a_t and b_t approximating the distribution of vacancies across firm types. The solid orange lines represent the actual realization of the series. The blue lines are the series implied by the forecasting rules.

Table 1.11: Forecasting rule accuracy for λ_t

	Within-sample	Out-of-the-sample
max abs % error	13.35	13.28
min abs % error	5.65e-6	3.59e-6
mean abs % error	0.27	0.22
median abs % error	0.13	0.08

Notes: This table shows the maximum, the minimum, the mean and the median absolute percentage error made when using the predicted job-finding rate $\hat{\lambda}$ instead of its actual value λ, calculated as $|\frac{\lambda-\hat{\lambda}}{\lambda}|$. The left column reports the prediction error in the sample used to calculate the forecasting rule; the right column shows the error made in a new sample, without re-estimating the forecasting rule.

Table 1.12: Forecasting rule accuracy for a_t

	Within sample	Out of the sample
max abs % error	35.09	35.10
min abs % error	3.70e-5	1.00e-5
mean abs % error	0.96	0.76
median abs % error	0.69	0.59

Notes: This table shows the maximum, the minimum, the mean and the median absolute percentage error made when using the predicted $\hat{a}$ instead of the actual a, calculated as $|\frac{a-\hat{a}}{a}|$. The left column reports the prediction error in the sample used to calculate the forecasting rule; the right column shows the error made in a new sample, without re-estimating the forecasting rule.

Table 1.13: Forecasting rule accuracy for b_t

	Within sample	Out of the sample
max abs % error	20.53	20.19
min abs % error	5.76e-6	2.61e-5
mean abs % error	0.87	0.75
median abs % error	0.65	0.56

Notes: This table shows the maximum, the minimum, the mean and the median absolute percentage error made when using the predicted $\hat{b}$ instead of the actual b, calculated as $|\frac{b-\hat{b}}{b}|$. The left column reports the prediction error in the sample used to calculate the forecasting rule; the right column shows the error made in a new sample, without re-estimating the forecasting rule.

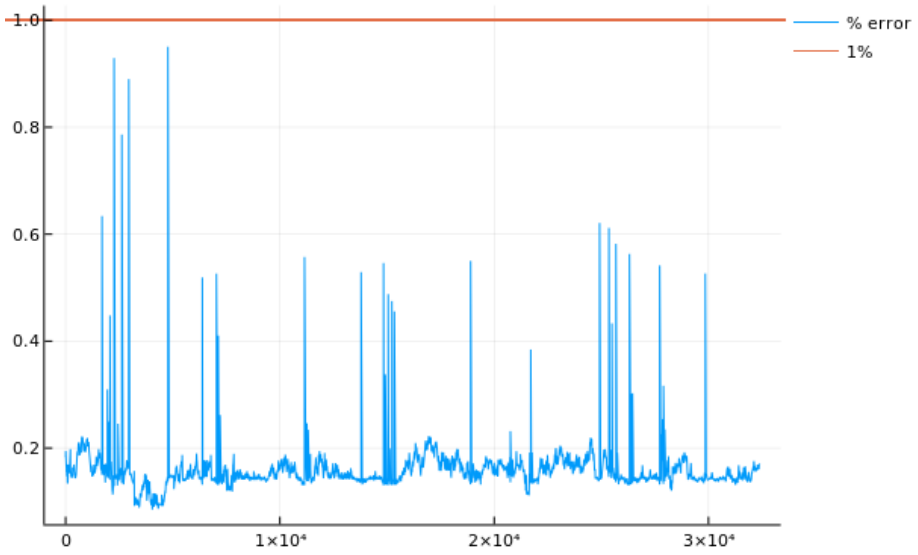
C.2. Accuracy of approximating $q_t(y)$ using a parametric function

For the results presented in the main section of the paper, I approximate the distribution of vacancies across types $q_t(y)$ using a parametric function $\tilde{q}(y|\boldsymbol{q_t})$. Because the support for y is $[0, 1]$ and $q(y)$ is uni-modal, I use a Beta density characterized by two parameters a_t and b_t. To quantify the error made when approximating the endogenous distribution of vacancies by a parametric counterpart, I use the following measurement:

$$e_t = \frac{\int_0^1 c\big(|\tilde{q}(y|(\hat{a}_t, \hat{b}_t)) - q(y, \Gamma_t)|\big)dy}{\int c\big(q(y, \Gamma_t)\big)dy} \tag{1.35}$$

where $\hat{a}_t$ and $\hat{b}_t$ are the shape parameters of the Beta density, calculated using the forecasting rule and $c(v)$ is the cost of posting v vacancies. The variable e_t measures the cost of misallocated vacancies (the numerator), relative to the total cost of vacancy posting (the denominator). If agents are perfectly rational, the realized distribution of vacancies and the perceived one are equal: $q(y, \Gamma_t) = \tilde{q}(y|(\hat{a}_t, \hat{b}_t))$. Hence, perfect rationality implies that the numerator is null each period.

Fig. 1.9. Percentage error when approximating $q_t(y)$



Notes: This graph shows e_t, which is a unit-less measurement of cost misallocation implied by using $\tilde{q}(y|(\hat{a_t}, \hat{b_t}))$ to approximate $q(y, \Gamma_t)$. The variable e_t is defined by

$$e_t = \frac{\int_0^1 c\big(|\tilde{q}(y|(\hat{a_t}, \hat{b_t})) - q(y, \Gamma_t)|\big)dy}{\int_0^1 c\big(q(y, \Gamma_t)\big)dy} \tag{1.36}$$

where $c(v)$ is the cost of posting v vacancies; $\hat{a_t}$ and $\hat{b_t}$ are the shape parameters of the Beta density implied by the forecasting rule. The maximum value for e_t is 0.950% and its median value is 0.151%.

C.3. Accuracy of approximating $q_t(y)$ using a histogram

This section shows that one may dispense with the parametric assumption on the distribution of vacancies across firm types $q_t(y)$. Instead, as in Reiter (2009), one may use a histogram to approximate the distribution of vacancies across firm types $q_t(y)$. If the economy features N discrete firm types instead of a continuum, this approach amounts to assuming that agents forecast the number of vacancies posted by each firm type:

$$\boldsymbol{q_{t+1}} = \big(q_{1;t+1}(y), q_{2;t+1}(y), ..., q_{N;t+1}(y)\big) = f_q(\Omega_t|\Theta_q)) \tag{1.37}$$

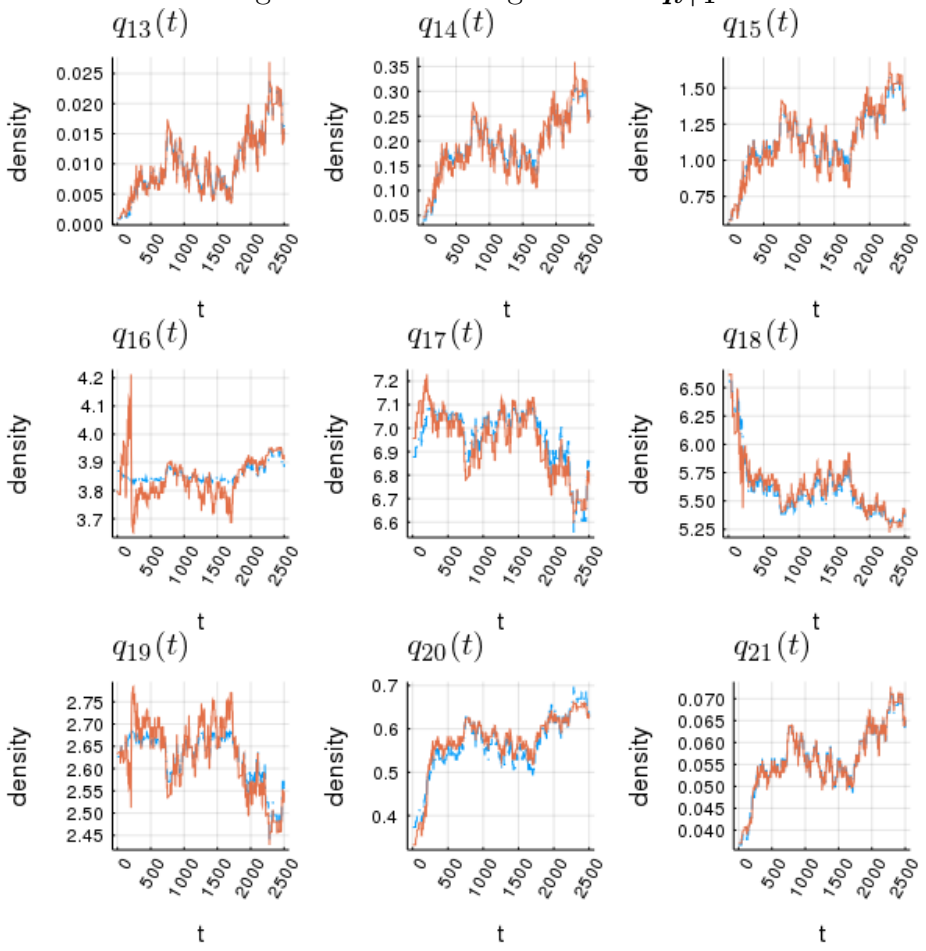
This approach is without loss of generality, because the y-dimension is already discretized when the model is numerically solved. Agents are endowed with a linear forecasting rule:

$$\boldsymbol{q_{t+1}} = \Theta_q \Omega_t \tag{1.38}$$

where $\boldsymbol{q_{t+1}}$ is a $N \times 1$ vector, Θ_q a $N \times k$ matrix containing the parameters for the forecasting rule, and Ω_t a $k \times 1$ matrix containing the information at time t relevant to predict $\boldsymbol{q_{t+1}}$. For instance, Ω_t is built to contain $(z_t, z_{t-1}, \lambda_{t-1})$, their squares and interaction terms. As illustrated in Figure 1.10, the estimated forecasting rule for (1.38) is successful in predicting the distribution of vacancies across firm types outside the steady-state.

It is important to note that the histogram approach described in this section is feasible because the vector-valued forecasting rule (1.38) has to be estimated only once. If an outer loop over that forecasting rule was needed, the convergence of the matrix Θ_q would be difficult to reach (and to verify) in practice.

Fig. 1.10. Forecasting rules for $\boldsymbol{q_{t+1}}$



Notes: This figure shows selected components of the distribution of vacancies across firm types $\boldsymbol{q_{t+1}} = \big(q_{1;t+1}(y), q_{2;t+1}(y), ..., q_{N;t+1}(y)\big)$. The orange lines are the actual realizations of the series and the blue lines are the values predicted by the forecasting rule $\boldsymbol{q_{t+1}} = \Theta_q \Omega_t$. Omitted components for $\boldsymbol{q_{t+1}}$ have a negligible value.

D. The BKM Algorithm

Throughout my exposition, I use the *recursive formulation* of a dynamic choice problem. However, I make the observation that the BKM algorithm designed by Boppart et al. (2018), which uses the *sequence formulation* of a dynamic choice problem, is particularly attractive in my setting.

The BKM algorithm uses the information contained in the perfect foresight path of the economy after a MIT shock. In general, a shooting algorithm has to be used to find this perfect foresight path, for which convergence properties are difficult to know a priori. In the present context, because the model is *semi-block recursive*[14], the algorithm of Boppart et al. (2018) can easily be used to find an approximation of the model. Semi-block recursivity ensures that the BKM algorithm is well-behaved, because the path for $\lambda_t, q_t(y)$ can be solved without reference to $W_t(x, y)$. In practice, one could proceed as follows:

1. Solve $S(.)$ and $U(.)$ by value function iteration
2. Starting from the steady-state with no aggregate uncertainty at $t = 0$, generate a one-standard-deviation aggregate shock at time $t = 1$, which goes back to its steady-state value ($z = 1$) at $t = 2$.
3. Solve for the transition path of for $t = 2, ..., T$.
4. Solve $W_t(x, y, \lambda_t, q_y(t))$ by backward induction, starting from:

$$W_t(x, y, \lambda_t, q_y(t)) = W_{SS}(x, y, \lambda_{SS}, q_{SS}(y))$$

The perfect foresight path is obtained in a single step. If the model were *not* semi-recursive, finding the perfect foresight transition path would be much more complicated. The algorithm would have to be modified, with the addition of an outer loop, with no guarantee of convergence. It would proceed as follows:

1. Assume a path for endogenous economic variables X_t, including $\{\lambda_t, q_y(t)\}$
2. Solve $S_t(.)$, $U_t(.)$ and $W_t(.)$ by backward induction, using the path previously assumed
3. Using $S_t(.)$, $U_t(.)$ and $W_t(.)$, simulate forward the path of economic variables, generating Y_t
4. If the distance between the paths X_t and Y_t is sufficiently small, stop
5. Otherwise, repeat the steps 1-4

The main advantage of using the BKM algorithm over the method presented in the

[14]The employment problem is independent from wages. The wage problem depends on employment only through the job meeting rate and the distribution of vacancies.

core of the text, is that the parametric assumption for the distribution of vacancies, or its approximation by an histogram, is not required anymore.

E. Flow equations for the joint distribution of wage and employment

Two methods are available to simulate the joint distribution of wages and employment status, denoted by $h_t(x, y, w)$. The first one is to simulate a panel with a sufficient number of agents. Then second one is to directly simulate the cross-sectional distribution $h_t(x, y, w)$. In this section, I derive the flow equations for the second approach.

Let Ω_t denote the support for wages at period t. Given our assumption on the wage process, Ω_t contains the starting wages and the promotions wages offers within that period:

$$\{\phi^0(x, y, \hat{\Gamma}_t), \phi^1(x, y, \hat{\Gamma}_t)\}_{(x,y)\in[0,1]^2} \subseteq \Omega_t$$

The set Ω_t also contains the wages inherited from past periods that were not altered by outside job offers or by intra-firm renegotiations. The flow equation for the distribution of starting wages $h_t(x, y, \phi^0(x, y, \hat{\Gamma}_t))$ solves:

$$\begin{aligned} h_{t+1}(x, y, \phi^0(x, y, \hat{\Gamma}_t)) = h_{t+}(x, y, \phi^0(x, y, \hat{\Gamma}_t)) \times \big(1 - \int_0^1 s\lambda_t \frac{v_t(y')}{V_t} \mathbb{1}\{S(x, y', z_t) \geq S(x, y, z_t)\} dy'\big) \\ + u_{t+}(x)\lambda_t \frac{v(y, \hat{\Gamma}_t)}{V(\hat{\Gamma}_t)} \mathbb{1}\{S(x, y, z_t) \geq 0\} \\ + \int_{w\in\Omega_t \setminus \phi^0(x,y,\hat{\Gamma}_t)} \mathbb{1}\{W(x, y, w, \hat{\Gamma}_t) - U(x, z) < 0\}(1 - s\lambda_t + s_{x,y}) h_{t+}(x, y, w) dw \end{aligned} \tag{1.39}$$

where $h_{t+}(x, y, \phi^0(x, y, \hat{\Gamma}_t)) \equiv (1 - \delta)\mathbb{1}\{S(x, y, z) \geq 0\} h_t(x, y, \phi^0(x, y, \hat{\Gamma}_t))$ denotes the measure of workers with wage $\phi^0(x, y, \hat{\Gamma}_t)$ after endogenous and exogenous job destruction. The measure of employed workers meeting with a firm, which is not a threat to the current match, is denoted by $s_{x,y} \equiv \int_0^1 s\lambda_t \frac{v_t(y')}{V_t} \mathbb{1}\{S(x, y, z_t) \geq 0 > S(x, y', z_t)\} dy'$. The first line in (1.39) takes into account the outflow of workers poached by more productive firms. The second line considers the inflow of hiring from the pool of unemployed workers. The third line takes into consideration intra-firm re-bargaining. The flow equation for the measure at the promotion wage $h_t(x, y, \phi^1(x, y, \hat{\Gamma}_t))$ solves:

$$h_{t+1}(x,y,\phi^1(x,y,\hat{\Gamma}_t)) = h_{t+}(x,y,\phi^1(x,y,\hat{\Gamma}_t)) \times \big(1 - \int_0^1 s\lambda_t \frac{v_t(y')}{V_t} \mathbb{1}\{S(x,y',z_t) \geq S(x,y,z_t)\} dy'\big)$$
$$+ \int_{w \in \Omega_t \setminus \phi^1(x,y,\hat{\Gamma}_t)} \mathbb{1}\{W(x,y,w,\hat{\Gamma}_t) - U(x,z) > S(x,y,z_t)\} \times$$
$$(1 - s\lambda_t + s_{x,y}) h_{t+}(x,y,w) dw \tag{1.40}$$

The first line in (1.40) takes into account the outflow of workers poached by more productive firms. The second and third lines take into consideration the measures of matches in which the firm had a credible threat to break the match. The expression for $h_{t+1}(x,y,\phi^1(x,y',\hat{\Gamma}_t))$ when $y' \neq y$ has to take into account workers poached by more productive firms, poaching from less productive firms and wage increases resulting from counter-offers:

$$h_{t+1}(x,y,\phi^1(x,y',\hat{\Gamma}_t)) = h_{t+}(x,y,\phi^1(x,y',\hat{\Gamma}_t)) \times \big(1 - \int_0^1 s\lambda_t \frac{v_t(y'')}{V_t} \mathbb{1}\{S(x,y'',z_t) \geq S(x,y,z_t)\} dy''\big)$$
$$+ \int_0^1 s\lambda_t h_{t+}(x,y') \frac{v_t(y)}{V_t} \mathbb{1}\{S(x,y,z_t) \geq S(x,y',z_t)\} dy$$
$$+ \int_{w \in \Omega_t \setminus \phi^1(x,y',\hat{\Gamma}_t)} \int_0^1 s\lambda_t h_{t+}(x,y,w) \mathbb{1}\{\phi^1(x,y',\hat{\Gamma}_t) \geq w\} \frac{v_t(y')}{V_t} \times$$
$$\mathbb{1}\{S(x,y,z_t) > S(x,y',z_t) \geq 0\} dy' dw \tag{1.41}$$

For wages that are *not* in the set of starting or promotion wages at time t, the flow equation takes into account workers (i) surviving both endogenous and exogenous job destruction (ii) workers with no on-the-job meeting or meeting (or choosing not to disclose unsuccessful ones) (iii) workers with no intra-firm re-bargaining:

$$h_{t+1}(x,y,w) = h_{t+}(x,y,w)(1 - s\lambda_t + s_{x,y}) \times \mathbb{1}\{0 \leq W(x,y,w,\hat{\Gamma}_t) - U(x,z) < S(x,y,z_t)\} \tag{1.42}$$

F. Identification

Table 1.14: Jacobian of the function mapping parameter values to simulated moments f : $\boldsymbol{p} \rightarrow \boldsymbol{m}$

Moment	α	δ	s	β_1	β_2	c_0	c_1	p_1	p_2	p_3	p_4	p_5	p_6	b_0	b_1	ϕ
U	-0.07	1.76	0.24	-0.05	0.00	0.19	0.00	-0.03	0.00	-0.01	0.00	0.00	0.00	0.05	0.05	0.00
U_{5+}	-0.02	0.24	0.06	-0.04	0.00	0.05	0.00	-0.01	0.00	0.00	0.00	0.00	0.00	0.01	0.01	0.00
U_{15+}	0.00	0.00	0.00	-0.03	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
U_{27+}	0.00	0.00	0.00	-0.03	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
$\frac{U}{V}$	4.78	-114.66	-17.38	4.07	-0.33	-31.70	3.43	4.37	0.71	1.07	0.15	0.08	0.05	-7.61	-7.66	1.07
$U2E$	-0.35	17.70	1.28	0.89	-0.03	1.00	-0.04	-0.15	-0.02	-0.03	0.00	0.00	0.00	0.23	0.24	-0.08
$J2J$	0.06	1.08	-0.04	0.02	0.00	-0.17	0.02	0.02	0.00	0.00	0.00	0.00	0.00	-0.04	-0.04	0.00
$E2U$	-0.05	1.59	0.19	-0.01	0.00	0.15	0.00	-0.02	0.00	-0.01	0.00	0.00	0.00	0.04	0.04	0.00
std_U	0.00	3.18	-0.43	0.17	-0.01	-0.01	0.16	-0.10	-0.01	-0.01	0.00	0.02	0.01	0.31	0.24	1.29
$std_V U$	0.00	3.62	-0.41	0.17	-0.01	0.01	-0.04	-0.11	-0.01	-0.01	0.00	0.02	0.01	0.65	0.49	2.97
std_{E2U}	0.29	-1.06	-1.53	0.11	-0.01	-0.84	0.19	0.02	0.01	0.02	0.00	0.02	0.01	0.14	0.06	1.46
std_{U2E}	0.25	-3.56	-0.95	-0.05	0.00	-0.70	0.03	0.10	0.02	0.02	0.00	0.00	0.00	-0.14	-0.15	0.18
std_V	-0.01	0.45	0.01	-0.01	0.00	0.02	-0.20	-0.01	0.00	-0.01	0.00	0.00	0.00	0.34	0.25	1.68
std_U	0.00	3.18	-0.43	0.17	-0.01	-0.01	0.16	-0.10	-0.01	-0.01	0.00	0.02	0.01	0.31	0.24	1.29
std_{U5+}	-0.46	8.75	1.35	0.19	-0.01	1.33	0.10	-0.29	-0.04	-0.05	-0.01	0.02	0.01	0.60	0.54	1.03
std_{U15+}	-0.04	0.52	0.14	0.00	0.00	0.12	0.00	-0.02	0.00	0.00	0.00	0.00	0.00	0.04	0.04	0.04
std_{U27+}	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
std_{GDP}	-0.01	0.28	0.02	0.00	0.00	0.02	-0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.01	0.47
std_{J2J}	-0.04	2.11	-0.51	0.01	0.00	0.12	0.03	-0.06	-0.01	-0.02	0.00	0.00	0.00	0.26	0.20	1.04
$corr(E2U, GDP)$	-0.02	-0.06	0.09	-0.01	0.00	0.04	-0.01	0.00	0.00	0.00	0.00	0.00	0.00	-0.01	-0.01	-0.27
$(corr(V, GDP)$	0.00	0.13	0.01	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.16
$corr(U2E, GDP)$	-4.88	72.69	18.10	1.11	0.03	13.96	-0.48	-2.06	-0.32	-0.47	-0.06	-0.01	-0.01	3.19	3.27	-1.75
$corr(U, GDP)$	0.00	-0.20	-0.01	0.00	0.00	-0.01	0.02	-0.01	0.00	0.00	0.00	0.00	0.00	0.01	0.01	-0.15
$corr(U, V)$	0.00	0.02	0.01	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
$corr(U2E, J2J)$	-5.12	76.49	18.88	1.16	0.03	14.64	-0.41	-2.22	-0.35	-0.56	-0.07	-0.03	-0.02	3.50	3.55	-1.14
$autocorr(VA)$	-0.01	0.21	0.01	0.00	0.00	0.02	-0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.01	0.34
$elasticity_P 10$	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
$elasticity_P 20$	-1.33	-82.78	-13.40	-0.08	-0.02	3.81	-0.80	0.34	0.05	0.88	-0.03	0.47	0.13	-3.68	-3.64	-0.38
$elasticity_P 30$	-29.27	-1815.16	-293.20	-4.28	0.17	83.70	-16.58	7.66	1.01	14.81	-0.66	8.44	2.25	-82.26	-80.81	-26.38
$elasticity_P40$	120.59	-457.84	1227.69	29.82	-3.16	-344.87	45.59	66.54	11.99	14.76	0.10	0.67	-0.10	-10.41	-8.35	-80.54
$elasticity_P 50$	61.92	47.30	635.25	14.84	-1.54	-177.09	23.58	30.80	5.60	6.63	0.19	0.57	0.10	6.09	6.97	-16.92
$elasticity_P 60$	231.49	172.99	2364.13	54.36	-5.51	-662.01	99.83	121.41	20.38	23.37	0.23	-2.10	-0.36	39.60	41.01	-40.17
$elasticity_P 70$	149.67	155.89	1530.98	35.27	-3.60	-428.03	63.61	77.49	12.98	12.10	0.27	-0.95	-0.17	26.00	26.90	-22.13
$elasticity_P 80$	50.52	85.40	517.90	11.95	-1.22	-144.49	21.52	24.26	4.16	1.17	0.11	-1.12	-0.28	9.65	10.31	-11.23
$elasticity_P 90$	-26.62	13.48	-288.15	-7.20	0.88	76.14	-5.84	-9.78	-2.11	-3.41	-0.08	-0.44	-0.35	-10.42	-9.92	-85.40

G. Inference

Because the likelihood function is untractable, I use the Simulated Method of Moments (SMM) to estimate parameter values (see for instance Duffie and Singleton (1990) and Gourieroux et al. (1993)). The SMM estimates are the one minimizing a weighted difference of simulated moments from their empirical counterparts:

$$\hat{\boldsymbol{\theta}}_{\boldsymbol{SMM}} = \arg\min_{\boldsymbol{\theta}} \big(\hat{\boldsymbol{m}} - \boldsymbol{m}^{\boldsymbol{S}}(\boldsymbol{\theta})\big) W \big(\hat{\boldsymbol{m}} - \boldsymbol{m}^{\boldsymbol{S}}(\boldsymbol{\theta})\big)' \tag{1.43}$$

where $\hat{\boldsymbol{m}}$ is a vector of empirical moments and $\boldsymbol{m}^{\boldsymbol{S}}(\boldsymbol{\theta})$ a vector of the same moments calculated using simulated data. More specifically, the l element of the vector $\hat{\boldsymbol{m}}$ is calculated as $m_l \equiv \frac{1}{T}\sum_{t=1}^{T} f^*_{l,t}$, where $f^*_{l;t} \equiv f_l(Y_t, Y_{t-1}, ..., Y_{t-k+1})$ with f_l a function mapping the finite l-history of state information $\{Y_t, Y_{t-1}, ..., Y_{t-k+1}\}$ to $\mathbb{R}$. I choose the weighting matrix W to be diagonal, with values representing subjective weights of the moments I deem more important to match. A more sophisticated approach would use the efficient weighting matrix $\hat{W}$, estimated in a first step (see Hansen (1982)).

In this paper, my goal is to match unconditional moments from time series. In this context, under some regularity conditions[15], the SMM estimate is asymptotically normally distributed with asymptotic variance $(1+\tau)$ times that of the GMM estimator, where $\tau = \frac{T}{S}$, with T the sample size and S the length of the simulated sample. As the size of the simulated sample increases relative to the actual sample size, the efficiency loss due to using simulated rather than empirical data becomes negligible. Note that $\hat{\theta}_{SMM}$ is a function of both τ and the weighting matrix W. The SMM estimator has the following formula for the asymptotic variance:

$$\sqrt{T}\big(\hat{\boldsymbol{\theta}}_{\boldsymbol{SMM}} - \boldsymbol{\theta_0}\big) \overset{A}{\sim} \mathcal{N}(0, (1+\tau)\Sigma_1^{-1}\Sigma_2\Sigma_1^{-1}) \tag{1.44}$$

with $\lim_{T\to\infty} \frac{T}{S(T)} = \tau$

$$\Sigma_0 = \sum_{j=-\infty}^{+\infty} \mathbb{E}\left([\boldsymbol{f_t^*} - \mathbb{E}(\boldsymbol{f_t^*})][\boldsymbol{f_{t-j}^*} - \mathbb{E}(\boldsymbol{f_{t-j}^*})]'\right)$$

$$\Sigma_1 = D'WD$$

$$\Sigma_2 = D'W\Sigma_0 WD$$

$$D = \mathbb{E}_0\left(\frac{\partial \boldsymbol{m}^{\boldsymbol{S}}(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}'}\right)$$

[15] see Gourieroux et al. (1993), page 31

In practice, I approximate D numerically using a finite difference scheme. To calculate Σ_0, I use a HAC estimator on simulated data rather than on empirical data. Given that the convergence rate of spectral estimators is low and that I control the length of the simulated sample, this potentially increases the accuracy of my estimate, as discussed in Duffie and Singleton (1990). The standard error for the estimate l is then calculated by taking the square root of the (l,l) element of the estimate for the asymptotic matrix, multiplied by $\frac{1}{T}$:

$$SE_l = \Big(\frac{1}{T}\Big((1+\tau)\Sigma_1^{-1}\Sigma_2\Sigma_1^{-1}\Big)_{ll}\Big)^{1/2}$$

H. Extension: Human Capital Accumulation

This section shows that the model can easily accommodate human capital accumulation. The key to tractability is to retain *semi-block recursivity*, that is, the independence of the employment problem from the wage dynamics. If one makes the following three assumptions: (i) unemployed workers have zero bargaining power (ii) wages are determined according to the sequential auction framework (iii) workers and firms are risk neutral; then adding the feature of learning-by-doing does not break semi-recursivity.

A common way to introduce learning-by-doing is to assume that, as in Ljungqvist and Sargent (1998), employed workers have a positive probability of skill appreciation while workers and that unemployed workers have a positive probability of skill depreciation. Instead of using x to denote a fixed ability dimension, let us use x to represent a human capital level. Let us assume that human capital x takes values on the interval $\big[\underline{x}, \bar{x}\big]$. Let use the notation $\pi_u(x'|x)$ to denote the probability that an unemployed worker experiences a skill depreciation from x to x' during a given period. The probability of a skill appreciation from x to x' for an employed worker is given by $\pi_e(x'|x)$.[16]

To prevent workers to accumulate an infinite amount of human capital as time passes, let us make the assumption that each worker (unemployed or employed) has a probability μ of dying each period. At the beginning of each period, there is a measure μ of newborns. Newborns start their career with the minimum value of human capital $\underline{x}$. Let us express the Bellman functions, before deriving the flow equations.

The value function for an unemployed worker with human capital x writes:

[16] In the model of Ljungqvist and Sargent (1998), workers also experience an extra lay-off human capital shock: during the period of the lay-off, their human capital evolves according to $\pi_l(x'|x)$. I do not add this feature to the model written in this section to keep the notations simple.

$$U(x,z) = b(x,z) + \frac{1-\mu}{1+r}\,\mathbb{E}_{((x',z')|(x,z))}\left((1-\lambda')U(x',z') + \lambda'\int_0^1 \max\left(U(x',z'), W(x',y,\phi_0,\Gamma)\right)\frac{v'(y)}{V'}dy\right) \tag{1.45}$$

where the prime notation is used to denote the value of a variable next period, and ϕ_0 denotes the wage offered in the first job. By virtue of the assumption of zero bargaining power, all firms choose a wage ϕ_0 equal to the reservation wage, so that $W(x',y,\phi_0,\Gamma) = U(x',z')$, and equation (1.45) simplifies to:

$$U(x,z) = b(x,z) + \frac{1-\mu}{1+r}\,\mathbb{E}_{((x',z')|(x,z))}\left(U(x',z')\right) \tag{1.46}$$

where the expectation is over the future human capital level and future aggregate state z. Equation (1.45) has two differences with equation (2.7): the discount factor takes into consideration that a worker may die next period with probability μ; the human capital level may change from x to x'.

The Bellman equation for the joint surplus of a match between a worker with human capital x and a firm of type y writes:

$$S(x,y,z) = p(x,y,z) - b(x,z) + \frac{(1-\mu)(1-\delta)}{1+r}\,\mathbb{E}_{((x',z')|(x,z))}\left(\max\left(S(x',y,z'),0\right)\right) \tag{1.47}$$

Because of the assumption that both human capital and the aggregate productivity level only depend on their past value, one may re-write equation (1.46) and (1.47) more explicitly:

$$U(x,z) = b(x,z) + \frac{1-\mu}{1+r}\int_{\underline{z}}^{\bar{z}}\int_{\underline{x}}^{\bar{x}} U(x',z')\pi_u(x'|x)\pi(z'|z)dx'dz' \tag{1.48}$$

$$S(x,y,z) = p(x,y,z) - b(x,z) + \frac{(1-\mu)(1-\delta)}{1+r}\int_{\underline{z}}^{\bar{z}}\int_{\underline{x}}^{\bar{x}} \max\left(S(x',y,z'),0\right)\pi_u(x'|x)\pi(z'|z)dx'dz' \tag{1.49}$$

Equations (1.46) and (1.47) are still contractions and may be solved by value function iteration or by the collocation method. The double integral (instead of a simple integral with respect to next period's aggregate productivity) may be approximated using standard numerical techniques.

The flow equation for the measure of unemployed in the period t_+ (after endogenous and exogenous job destruction) writes:

$$
\begin{aligned}
u_{t_+}(x) = \mu \mathbb{1}\{x = \underline{x}\} & \\
&+ (1-\mu)\Big(\int_{\underline{x}}^{\bar{x}} u_t(s)\pi_u(x|s)ds \\
&+ \int_0^1 \int_{\underline{x}}^{\bar{x}} \big(\mathbb{1}\{S(s,y,z) < 0\} + \delta \mathbb{1}\{S(s,y,z) \geq 0\} \big) h_t(s,y)\pi_e(x|s)dsdy \Big)
\end{aligned}
\tag{1.50}
$$

The first line takes into account the inflow of newborns in the economy. The second line takes into consideration the measure of unemployed workers, who may experience a skill depreciation. The third line takes into consideration the measure of employed workers experiencing job destruction, right after the realization of the aggregate shock.

The flow equation for the measure of employed workers writes:

$$
h_{t_+}(x,y) = (1-\delta)(1-\mu) \int_{\underline{x}}^{\bar{x}} \mathbb{1}\{S(s,y,z) \geq 0\} h_t(s,y)\pi_e(x|s)ds \tag{1.51}
$$

where equation (1.51) takes into account the measure of matches surviving both exogenous and endogenous job destruction, as well as the probability that a worker may die.

It is important to notice that despite the fact that human capital evolves the joint surplus of a match $S(x,y,z)$ is still a sufficient statistics that determines job feasibility, independently of wages. As human capital evolves the joint surplus of a match gets bigger or smaller, possibly altering job feasibility conditions for some workers. However, because the utility function is assumed to be linear, the wage is only a tool that dictates *how* the joint surplus is split between workers and firms, without changing the total value of a match.

Because $S(x,y,z)$ is still a sufficient statistics that determines job feasibility, the numerical algorithm described in this paper can still be used. One may solve for (1.46) and (1.47), simulate forward the flow equations (1.50) and (1.51), estimate a simple rule for next period's job-meeting rate and distribution of vacancies across types using simulated series, and then solve for the Bellman equations characterizing wages. This sequence of steps has be performed only *once*.

The setting and equations presented in this section share many similarities with the model of Walentin and Westermark (2018). Two differences are that in the model of Walentin and Westermark (2018), the distribution of firms is exogenous (each firm draws its type from a fixed distribution) and that workers have a strictly positive bargaining power. The first assumption is rather innocuous, in the sense it is does not impact whether or not employment and wage problems can be separated. However, because of the latter assumption, their model is not semi-block recursive and the authors have to rely on a full-fledged Krusell-

Smith approach, involving an outsider loop to check for the convergence of the forecasting rules used by agents, with no *a priori* guarantee of convergence.

What this section shows is that the *combination* of the three assumptions listed in the beginning of this section is essential to obtain semi-block recursivity, thus tractability. An interesting line of research would be to investigate whether or not there exists an alternative set of assumptions that yield semi-block recursivity. In particular, it would be interesting to know if there exists an alternative wage-setting mechanism that renders the employment problem independent from wages, once workers are assumed to have a positive bargaining power.

Chapter 2

Spatial Equilibrium and Transportation Costs

Abstract

I exploit a spatial discontinuity introduced by a French reform in September 2015 to measure the links between transportation costs and local employment dynamics. I find that the reform, which decreased the cost of public transportation in selected areas, led to a 2% drop in the number of unemployed workers registered in the local unemployment agency for the municipalities benefiting from the reform. The positive employment effect is concentrated on long-term unemployed workers. I build and calibrate a spatial search-and-matching model to underline the mechanisms at play.

Keywords: Commuting costs; Local employment
JEL Classification: C31, C21

Introduction

Several multi-billion infrastructure projects are currently under way throughout the world. To name a few, the California High-Speed Rail in the US, Crossrail 2 in the UK, the Barcelona metro line 9 and the Grand Paris Express are expected to cost 150 billions USD in total. These projects aim to reduce transport costs and encourage people to commute using public transport rather than private vehicles, in an effort to limit traffic congestion and to curb CO2 emissions. What will be the consequences of such projects in terms of local employment, business creation, household location decisions and house prices? Given the costs of these large-scale infrastructure projects, answers to these questions are essential from a central planner's perspective.

The main goal of this chapter is to provide new insights on the links between commuting costs and local employment dynamics, using a spatial discontinuity created by a French in September 2015 in the Parisian area. The idea for this discontinuity is simple. Prior to the

reform, owners of a public transport travel pass would pay different fares depending on the zones they crossed during their commute. Parisians and residents from the near suburbs would pay a given price for a travel pass allowing them to travel in the zones 1 and 2, including Paris *intra-muros* and the bordering municipalities. People living further away from the city center (in the zones 3 to 5) and commuting to Paris, where most jobs are located, had to pay a premium. In September 2015, the link between the zones crossed and the price of travel passes was removed, with the creation of a unique price scheme called "Forfait Toutes Zones" (FTZ). The FTZ new price scheme generated a substantial price discount for users of the travel passes in zones 3 to 5, while the price for users in zones 1 and 2 marginally increased. Hence, the removal of fare areas ("dézonage") created a discontinuity for cities located from different sides of a former price border. I exploit this discontinuity to quantify the link between commuting costs and local employment dynamics, by carefully selecting a sample of municipalities close to the former border.

Can we think of transport infrastructure projects and commuting cost reforms as local employment policies? The present chapter claims that, to some extent, they are local employment policies. The present chapter is related to the unemployment/inactivity trap literature, which sheds light on the structural barriers that unemployed workers face when searching for a job. In the French context, Anne and L'Horty (2009) show that the complex system of national and local social transfers creates a situation in which minimum wage workers having children are just better off not working. The present chapter focuses on a type of disincentive that is spatially located. Conditionally on looking for a job located within Paris *intra-muros*, where many jobs are located, an unemployed worker living in the outskirts of Paris used to face *more disincentives to work* than an unemployed worker living the zones 1 or 2. The FTZ reform reduced the disincentives to work in zones 3, 4 and 5, without altering incentives in zones 1 and 2. Individuals who would have previously rejected offers located far away from where they reside are now more keen to accept them. In short, the FTZ reform can be seen as an spatially-located employment premium. The employment impact of the FTZ reform is likely to be strong for workers at or close to the minimum wage, for whom the savings generated are more substantial relative to their monthly labor income.

Related work include Mayer and Trevien (2017) who using an instrumental variable (IV) strategy document that the arrival of the Regional Express Rail (RER) in the Paris metropolitan area caused a 8.8% increase in employment for the municipalities connected to the network in between 1975 and 1990. Using a similar methodology, Garcia-López et al. (2017) show that improvements in the Parisian transit system led to the emergence of employment sub-centers in suburbans municipalities that had a rail station. Duranton and Turner (2012) also relying on an IV strategy, show that a 10% increase in a provision

of buses causes the *population* to increase by 0.8% in the US.

Instead of using an IV, the current chapter relies on the *spatial regression discontinuity* literature (see Neumark and Simpson (2015) for a review). This literature has emphasised a trade-off between the necessity to compare geographical areas close to each others to control for unobservable characteristics and the threat to identification posed by spillovers between neighboring areas. The methodology used in this chapter is very close to the one used by Chapelle et al. (2018) who analyse the impact of a housing tax credit on local housing market outcomes using fine-resolution data. Other related empirical strategies can be found in Einiö and Overman (2016), Kline and Moretti (2013) and Hilber et al. (2019).

From a theoretical standpoint, the model used to underline the mechanisms at play relies on elements from Brueckner et al. (1999) and Wasmer and Zenou (2002). Workers want to reside close to the center of Paris to enjoy high amenities and the proximity to jobs. On the other hand, living close to the city-center entails paying a higher rent. Taking into consideration these factors, workers make an informed decision when choosing where to live. A decrease in commuting costs for workers living in the outskirts of Paris renders working more attractive relative to unemployment. Within the model, a decrease in commuting costs generates a drop in workers' reservation wage, which in turn generates an employment boost.

Section 1 presents the FTZ reform, entered into force in September 2015. Section 2 describes the data sources. Section 3 explains how the treatment and control groups are selected. Section 4 presents the estimation results. Section 5 explores a frictional model of the labor market that helps in explaining the employment effects identified in the data. The last section concludes.

1. A Brief History of the Reform

Commuting by public transport is widespread in the Parisian area. In Paris *intra-muros*, approximately 78% of workers opt for this solution. Many workers choose to buy one of the several public transport passes available, giving full access to the region Ile-de-France public transport network, which includes metros, buses, tramways, RER[1] and some trains.[2]. The most popular option is to buy a Navigo card, valid for a week, a month or a year. Students have access to the yearly equivalent of the Navigo card, called the ImaginR pass, which comes with a substantial student discount. Until September 2015, the price of the

[1]RER (Réseau Express Régional) are express train lines connecting Paris city centre to surrounding suburbs.

[2]On a typical month in 2016, there were more than 4 million people residing in the region Ile-de-France using one of the different public transport travel passes available and approximately 3 million people using single-ride tickets,[3] for a region with approximately 12 million inhabitants

Navigo and ImaginR passes depended on the fare areas ("zones") crossed during the travel. Typically, users living in Paris or in cities sharing a border with Paris (fare areas 1 or 2) and working in Paris would buy the Navigo pass zones 1-2. Users living in the close suburbs (fare area 3) and working in Paris would choose the Navigo pass zones 1-3, paying a premium for the extra distance traveled (see Figure 2.11 for the different fare areas in the region Ile-de-France).

In September 2015, the "dézonage" reform removed the link between the areas crossed and the travel pass price for the most popular travel passes.[4] A new travel pass was created, called "Forfait Toutes Zones" (FTZ), being sold at a unique price. The FTZ travel pass created a discontinuity in the cost of commuting using public transport. For instance, users of the yearly Navigo pass zones 1-3 experienced a 14.3 € monthly decrease in their commutng costs, while users of the yearly Navigo pass zones 1-2 experienced a 3.5 € monthly increase. Hence, the reform generated a discontinuity in commuting costs for people living on each side of the border separating the fare areas 2 and 3. Taking into account the number of users of each travel pass, I find that people commuting to the city center and living in the fare area 3 experienced on average a 15.3 € decrease in their monthly commuting costs[5] relative their neighbors living in the fare area 2 as a consequence of the reform (see Figure 2.3). This number is likely to be a *lower bound* on the real commuting costs discontinuity, as it has been reported that the reform encouraged users to buy weekly or monthly Navigo passes instead of pricier single-ride tickets, hence generating an additional cost-saving channel for people residing in the fare area 3.[6]

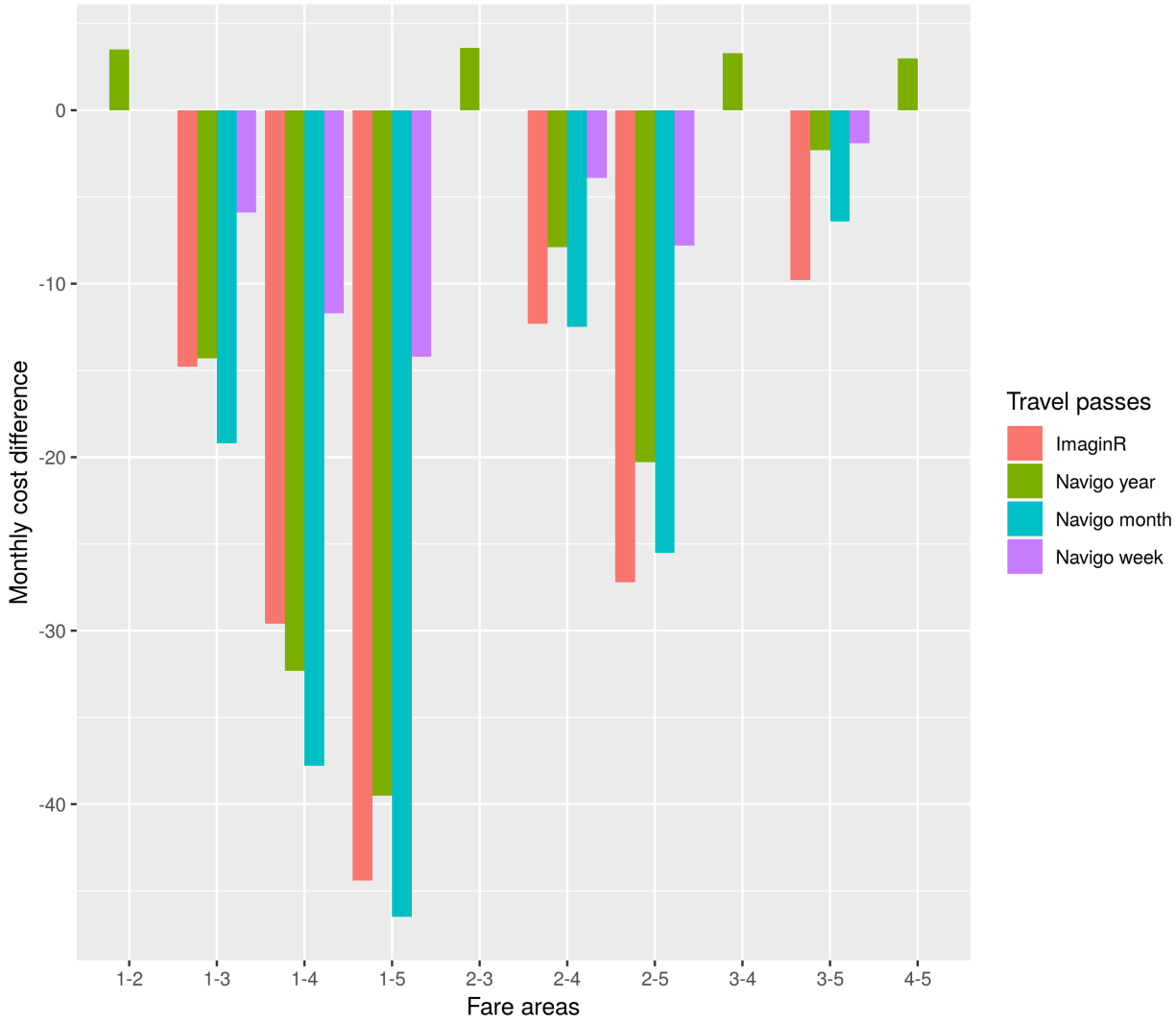
My empirical strategy relies on the fact that for cities close to the border separating the fare areas 2 and 3, the fact of falling on one side of the border can be considered as a random experiment. Hence, the "dézonage" reform generates a quasi-random variation in commuting costs for the group of "treated" (cities in the fare area 3 close to the fare area 2) relative to the group of "non-treated" (cities in the fare area 2, close to the fare area 3).

[4]The price of single-ride tickets is still dependent on the fare areas crossed during the travel.

[5]To give a point of comparison, a minimum wage worker earned a net monthly salary of 1136 €.

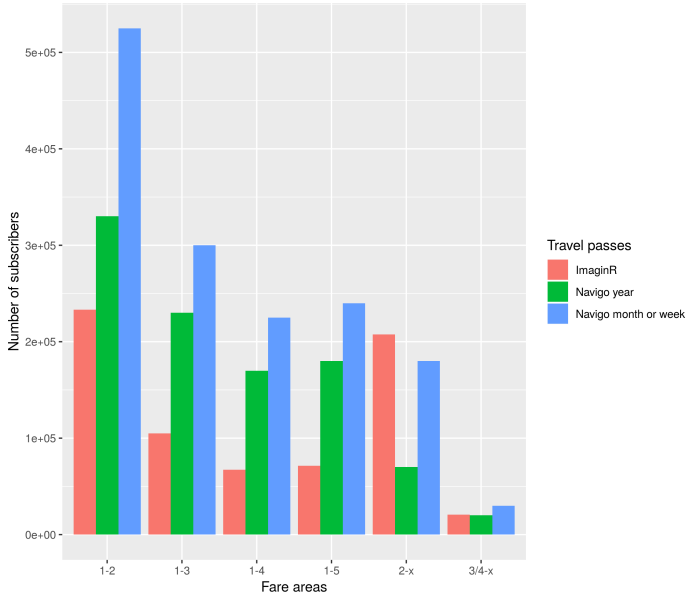
[6]"Sur la première année qui a suivi le lancement des Forfaits Toutes Zones, la structure des ventes des titres de transport [...] a significativement évolué : une nette progression des ventes de forfaits à usage illimité 'longs' [...] est allée de pair avec une diminution notable des ventes des tickets, billets et forfaits courts. Cela traduit notamment un effet de report vers les forfaits 'longs' dont le prix a baissé et qui sont, de ce fait, devenus plus aisément rentabilisables. Il en a résulté une baisse considérable des recettes" (OMNIL (2018))

Fig. 2.1. Impact of the reform on travel pass cots



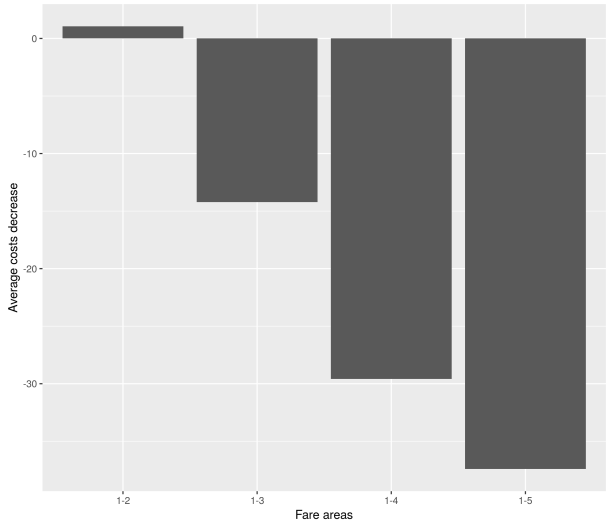
Notes: This figure shows the impact of the "dézonage" reform on the monthly cost of using a travel pass by fare area. Users of the yearly Navigo pass for fare areas 1 and 2 experienced a 3.5 € monthly increase, while users of the yearly Navigo pass for fare areas 1 and 3 saw their costs decrease by approximately 15 €. The absence of a bar indicates that the monthly cost was not impacted by the reform.

Fig. 2.2. Number of subscribers by travel pass and fare area in 2015



Notes: This figure shows the number of subscribers by travel pass and fare areas. $2-x$ and $3/4-x$ are suburbs-to-suburbs passes: $2-x$ is used to indicates passes valid for the fare areas $2-3$, $2-4$ and $2-5$; $3/4-x$ indicates passes valid for the fare areas $3-4$, $3-5$ and $4-5$.

Fig. 2.3. Average change in commuting costs for travel pass users



Notes: This figure shows the average cost impact (in euros) of the "dézonage" reform by fare area, for suburbs-to-center travel passes. The average cost impact is a weighted average, taking into account the number of users reported in figure 2.2.

2. Data and Empirical Facts

This section describes my data sources and discuss key empirical facts for the Parisian metropolitan area. For my empirical analysis, I use a combination of administrative data and observations collected on specific websites using web scraping techniques.

Municipalities characteristics and employment data

To measure the pre-reform characteristics of cities in region Ile-de-France, I use the database "Activité des résidents en 2015" from the INSEE, the national statistics bureau of France. This database, relying on observations from the French national census, includes characteristics on the population as of January 2015 at the IRIS level, which is a sub-city level containing approximately 2000 inhabitants. Typically, a city is composed of several IRIS. The "Activité des résidents" database contains detailed information on the population structure by age group, gender and employment status. For employed workers, the database offers insights on their sector of activity and their commuting habits. When needed, I aggregate this dataset to produce data at the city ("communes") level. To measure both pre-reform and post re-form employment dynamics, I use data from Pole Emploi, the French employment center. The dataset contains monthly observations on the number of unemployed workers registered to Pole Emploi at the city level for the period January 2014 - February 2019, for cities with more than 5000 inhabitants.[7]

It is important to keep in mind that the number of unemployed workers registered to Pole Emploi is not a perfect measurement of unemployment, as a worker can be registered to Pole Emploi without actively looking for a job, and unemployed workers may not register to Pole Emploi if they are not eligible. Senior workers are more likely to be enrolled to Pole Emploi while not actively looking for a job, while young workers are more likely to be unemployed, while not being registered to Pole Emploi.[8] An alternative would be to build a repeated cross-section dataset at the IRIS using several iterations of the database "Activité des résidents", which uses the ILO definition of unemployment. However, the latest year available is 2015.

Commuting time and fare ares

It is not straightforward to classify cities as belonging to the fare area 2 or 3, as cities may overlap the theoretical border separating them. To generate such a classification, I

[7] https://statistiques.pole-emploi.org/stmt/defm?fh=1&lk=0&pp=las&ss=

[8] https://www.insee.fr/fr/statistiques/fichier/2022025/insee-en-bref-chomage.pdf

gather data on the optimal route to the city center (Châtelet-les-Halles) at the IRIS level using the website Vianavigo[9]. For each IRIS, I draw an address at random, which gives me a starting point for the itinerary. I let Vianavigo find the optimal itinerary (the itinerary with shortest time) on a typical Monday morning to arrive at Châtelet-les-Halles at 9:00 a.m. The optimal itinerary may combine several public transport modes (bus, metro, train) and walking. Vianavigo gives me a classification of the optimal itinerary: I know what type of travel pass is required to perform the commute (fare areas 2 or 3). Results of this methodology are visible in Figure 2.4. Given that each city in my sample is composed of several IRIS, I calculate the city area as the median value across IRIS, which I round to the nearest integer.

Housing Price Data

To measure house prices at the IRIS level, I use the standardized price index as of the fourth quarter of 2013 provided by the Grand Paris notary agency. This standardized price index measures the value of a square meter, controlling for differences in the houses sold using a hedonic pricing model.

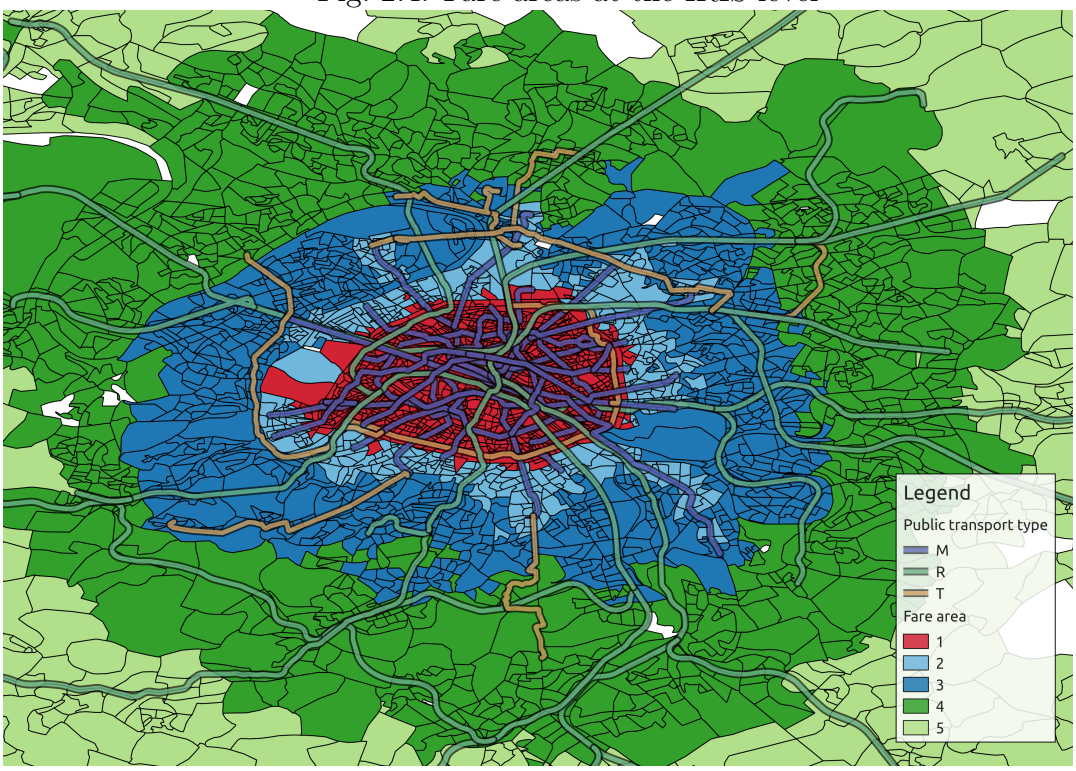
Empirical Facts

The Parisian region features several interesting characteristics. In particular, the region features three interesting gradients: (i) square meter price (ii) population density (iii) unemployment rate. Figure 2.5(b) shows that the price of a square meter decreases as the distance to the city center increases. Population density follows a similar patterns, with a density of approximately 20,000 inhabitants per km^2 in Paris (see Figure 2.5(a)) decreasing to less than 10,000 when distance to the center reaches 10 km. Job density closely follows population density: Chapelle et al. (2017) documents that 50% and 90% of jobs are situated less than 10 and 30 km away from Paris center respectively. Overall, the *monocentric assumption*, according to which jobs are located in a single area, seems to be appropriate for the Paris area. Interestingly, the unemployment rate first increases before reaching plateau at around 10 km from the city center, before starting to decrease again (see Figure 2.6(a)). Some wealthy workers choose to live in the city center to enjoy Paris amenities and the proximity to their jobs, at the expense of a higher rent. Unemployed workers will tend to settle in the *banlieues* (suburbs), where rent is more affordable and the city center is still quite accessible by public transport. As the distance to the city center increases again, the unemployment

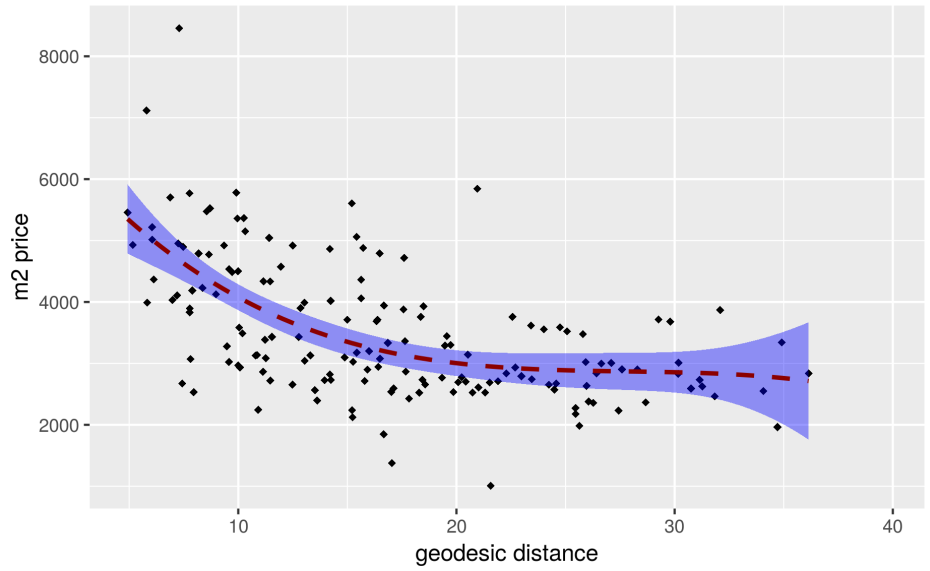
[9] `https://www.vianavigo.com/accueil`

rate decreases and workers are more likely to use their own means of transport rather than the public transport network (see Figure 2.6(a)).

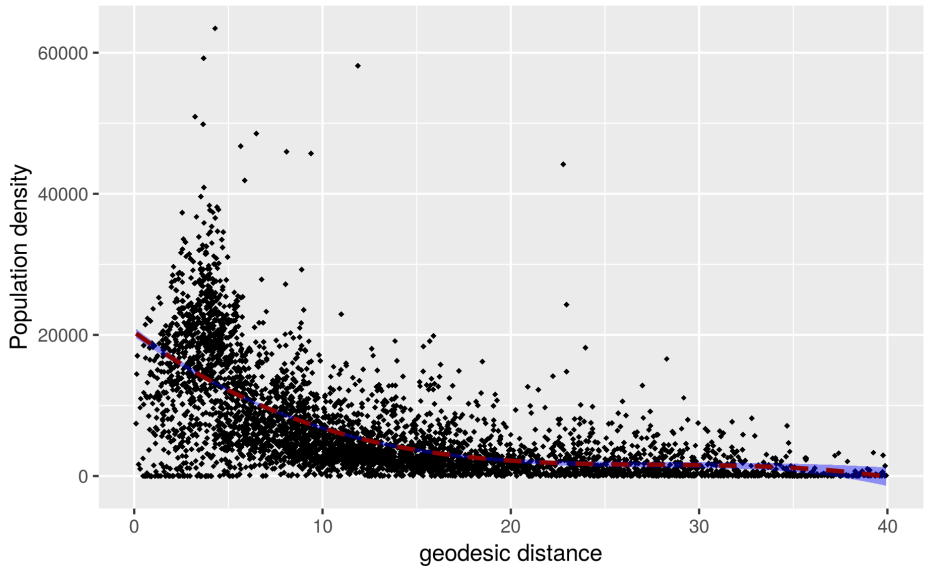
Fig. 2.4. Fare areas at the IRIS level



Notes: This figure shows the fare area associated to each IRIS. For each IRIS, a random address was drawn at random. The optimal itinerary from this address to the center of Paris (Châtelet-les-Halles) was determined using https://www.vianavigo.com/accueil. Using this optimal itinerary, it is straightforward to determine the fare areas crossed during the trop to Châtelet-les-Halles.
This figure also displays the metro lines (in purples), the RER lines (in green) and the tram lines (in orange).



(a) Standardized square meter price index and distance to the city center, 2013 - Q4



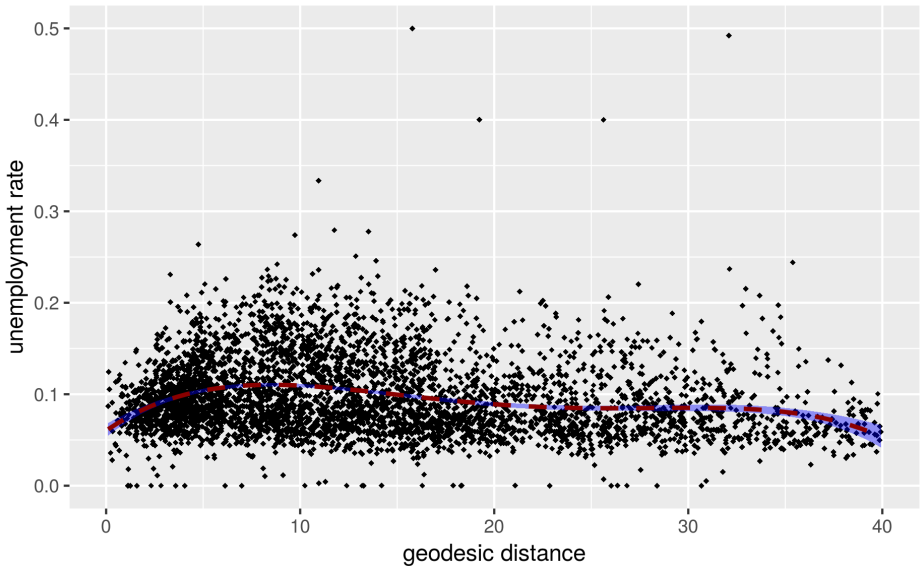
(b) Population density and distance to the city center, 2015

Fig. 2.5. Square meter price and population density

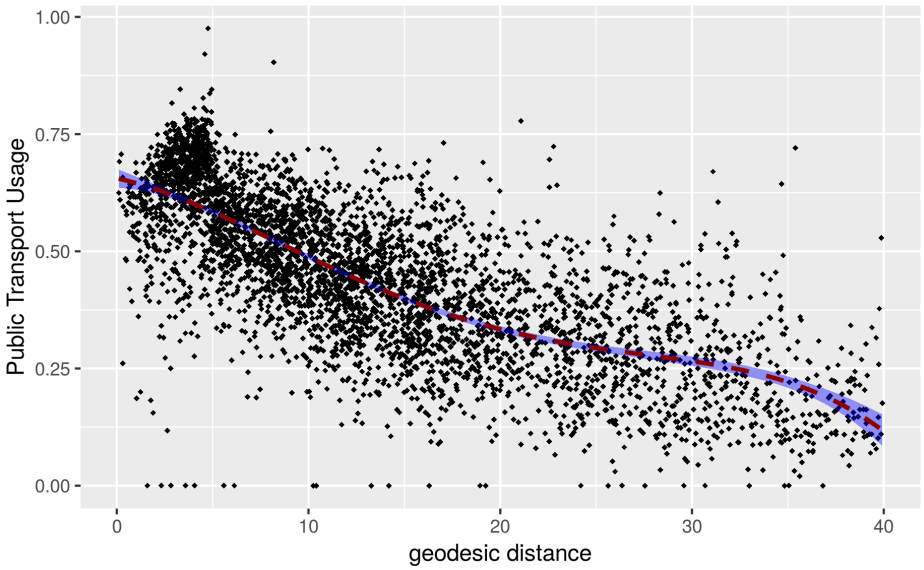
Sources: author's calculations based on data from the Grand Paris notary agency and INSEE, Activité des résidents.

Notes: Panel 2.5(a) shows the standardized square meter price index as a function of the geodesic distance to the center of Paris (Châtelet) as of 2013 - Q4. The dotted line is a fitted third-order polynomial and the shaded area represents a 95% confidence interval.

Panel 2.5(b) shows the number of inhabitant per km^2 a function of the geodesic distance to the center of Paris (Châtelet), as of 2015. The dotted line is a fitted fourth-order polynomial and the shaded area represents a 95% confidence interval.



(a) Unemployment rate and distance to the city center, 2015



(b) Public transport usage and distance to the city center, 2015

Fig. 2.6. Square meter price and population density

Sources: author's calculations based on data from the INSEE, Activité des résidents.
Notes: Panel 2.6(a) shows the unemployment rate (at the IRIS level) as a function of the geodesic distance to the center of Paris (Châtelet) as of 2015. The dotted line is a fitted fourth-order polynomial and the shaded area represents a 95% confidence interval.
Panel 2.6(b) shows the percentage of workers (at the IRIS level) using public transport to go to work as a function of the geodesic distance to the center of Paris (Châtelet), as of 2015. The dotted line is a fitted fourth-order polynomial and the shaded area represents a 95% confidence interval.

3. Sample Selection

In this section, I explain how I select my sample and how I construct my treatment and control groups. I build my sample with the double objective of maximizing (i) the proximity to the border separating the fare areas 2 and 3 (ii) the similarity between the treated and non-treated groups in terms of observable variables *before* the advent of the reform.

The rationale for (i) is that the proximity to the border ensures that cities from both groups are experiencing similar trends in terms local market conditions which I do not observe accurately for the pre and post-reform periods (rent, population dynamics, etc.). Also, the reform may have been motivated to favor certain geographical areas in the Paris metropolitan area. By considering cities close to each others, I nullify these potential political motives.

The reason for (ii) is that I do not have access to a wide set of observable variables post-reform. By maximizing similarity in terms of observable characteristics, I maximize the chances that the *common trend assumption* holds. That is, both groups were following a similar trend before the reform and the average deviation from the pre-treatment period is to be attributed to the "dézonage" reform only. It is also important to note that the main goal of this data selection step is not to build two groups that are perfectly identical, as permanent differences between cities will be taken into account using *entity fixed effects*. In practice, I include only cities in zones 2 and 3 meeting the following criterion:

$$\begin{cases} d_i \leq \bar{d}_{2,3} + \mu_{2,3} \times \sigma_{2,3} \\ d_i \geq \bar{d}_{2,3} + \mu_{2,3} \times \sigma_{2,3} \end{cases} \tag{2.1}$$

where d_i measures the geodesic distance to the city center; $\bar{d}_{2,3}$ denotes the mean distance to the city center for cities in zones 2 and 3; $\sigma_{2,3} \equiv \frac{1}{2}(\sigma_2 + \sigma_3)$ with σ_2 and σ_3 the standard deviation of the geodesic to the city center for the zones 2 and 3 respectively; $\mu_{2,3}$ denotes a coefficient to be determined empirically. The condition (2.1) defines a *ring* (or annulus) with origin the city center, defined to be the metro station Châtelet. The bigger $\mu_{2,3}$, the wider the ring and the more likely it is for one city to be retained in the sample. When $\mu_{2,3}$ gets sufficiently large, all the cities in the fare areas 2 and 3 are included in the sample.

Mathematically, I solve the following one-dimensional problem:

$$\mu_{2,3} \equiv \arg\min_{\mu} \sum_{k=1}^{K} \Big(\frac{\hat{x}_{2,k}(\mu) - \hat{x}_{3,k}(\mu)}{\hat{x}_{2,k}(\mu) + \hat{x}_{3,k}(\mu)} \Big)^2 \tag{2.2}$$

where $\hat{\boldsymbol{x}}_{\boldsymbol{2}}(\mu) \equiv (\hat{x}_{2,1}, \hat{x}_{2,2}, ..., \hat{x}_{2,K})$ and $\hat{\boldsymbol{x}}_{\boldsymbol{3}}(\mu) \equiv (\hat{x}_{3,1}, \hat{x}_{3,2}, ..., \hat{x}_{3,K})$ are vectors of mean characteristics for city in zone 2 and 3 respectively. Both vectors depend on the scalar μ because

changing its value alters the composition of the sample, thus the group-specific mean values. As characteristics of interest, I include a wide range of variables as of January 2015, informative on the structure of the active and non-active population by gender, age, occupation and commuting habits (going to work using public transport or by car). The results of the minimization is shown in Figure 2.13. The corresponding value for $\mu_{2,3}$ defines a ring that includes cities in the "petite couronne", an area that forms a ring around Paris (see Figure 2.14). Descriptive statistics for the treated and non-treated groups are reported in table 2.1. While cities located in the fare area 2 are slightly more populous than cities in the fare area 3 and are located closer to the city center, they are rather similar in terms of population characteristics, measured by age, gender, the percentage of the population having or looking for a job and the type of job (permanent or temporary position). In terms of commuting habits, workers in the fare area 2 are more likely to go to work using public transport rather than using a car compared to workers in the fare area 3.

Table 2.1: Selected city characteristics included in the objective function in (2.2)

	Fare area 2	Fare area 3
Population 15-64	34075	26512
Population 15-24/Population 15-64	17.92%	18.86%
Population 25-54/Population 15-64	66.58%	64.81%
Population 55-64/Population 15-64	15.50%	16.32%
Men 15-64/Population 15-64	48.51%	48.33%
Women 15-64/Population 15-64	51.49%	51.66%
Active Population 15-64/Population 15-64	77.33%	76.57%
Unemployed men 15-64/Active Population 15-64	6.96%	6.78%
Unemployed women 15-64/Active Population 15-64	7.06%	6.80%
Salaried workers 15-64/Active Population 15-64	76.83%	79.01%
Salaried permanent position/Active Population 15-64	66.29%	68.98%
Salaried temporary position/Active Population 15-64	0.84%	0.89%
Salaried subsidized job/Active Population 15-64	2.02%	2.06%
Working from home/Employed Workers	3.67%	3.14%
Walking to work/Employed Workers	9.37%	7.92%
Biking to work/Employed Workers	6.70%	5.77%
Going to work by car/Employed Workers	26.82%	34.93%
Going to work using public transport/Employed Workers	53.42%	48.19%
Latitude	48.865	48.845
Longitude	2.354	2.335
Commuting time to city center (mn)	31.9	38.5
Geodesic distance to city center	6.845	9.029

Notes: This table displays mean values for cities in the fare area 2 (left column) and fare area 3 (right column). Commuting time to city center is the theoretical time according to Vianavigo, using the optimal combination of public transport (bus, metro, train) to arrive at 9:00 a.m to the city center (Châtelet-les-Halles), on a usual Monday morning.

Sources: Vianavigo and INSEE, Activité des résidents en 2015.

4. Empirical Specifications and Results

In this section, I justify my empirical specification. I use a simple difference-in-difference specification, using the previously carefully selected control and treatment groups. I then

present the main empirical results, before discussing some factors that may pose a threat to identification.

4.1. Empirical Specifications

My baseline specification is the following:

$$y_{it} = \alpha_i + \gamma_t + \beta \times \delta_{i,t} + \boldsymbol{\eta x_{i,t}}' + \varepsilon_{i,t} \tag{2.3}$$

with $y_{i,t}$ the log of a dependent variable of interest; α_i a city-level fixed effect capturing permanent differences between cities; γ_t a time fixed effect capturing macro changes impacting all the cities in the sample over the period; $\delta_{i,t}$ and indicator variable equal to 1 if the city is in zone 3 *and* the time index t is such that the observation is after 1 September 2015; $\boldsymbol{x_{i,t}}$ a vector of control variables and $\varepsilon_{i,t}$ a city and time-specific $i.i.d$ error term. To quantify the dynamic impact of the reform, I estimate a year-dependent model, were the dummy $\delta_{i,t}$ is interacted with a year-specific dummy δ_y equal to 1 when the underlying year is equal to y:

$$y_{it} = \alpha_i + \gamma_t + \sum_{y=2015}^{2019} \beta_y \times \delta_{i,t} \times \delta_y + \boldsymbol{\eta x_{i,t}}' + \varepsilon_{i,t} \tag{2.4}$$

The parameter β_y measures the average treatment effect on year y. To test the robustness of results from regression (2.3) and (2.4), I estimate a continuous version of them. In the construction of my sample, I rounded the median fare area to the nearest integer (using observations at the IRIS level) to generate a clear cut between the treated and non-treated cities. Instead, one may want to use the fact that some cities overlap over the border separating the fare areas 2 and 3. One should expect the impact of the reform to be a function of the intensity of the "treatment" received. That is, the impact of the reform for a city with a median fare close to 3 should be higher than for cities with a median fare area close to 2. To test that hypothesis, I estimate the two following regressions:

$$y_{it} = \alpha_i + \gamma_t + \beta \times z_{i,t} + \boldsymbol{\eta x_{i,t}}' + \varepsilon_{i,t} \tag{2.5}$$

$$y_{it} = \alpha_i + \gamma_t + \sum_{y=2015}^{2019} \beta_y \times z_{i,t} \times \delta_y + \boldsymbol{\eta x_{i,t}}' + \varepsilon_{i,t} \tag{2.6}$$

where $z_{i,t}$ is an *intensity of treatment variable*, equal to the median fare area (minus 2) if t is such that the observation is after 1 September 2015; 0 otherwise. In the post treatment

period, $z_{i,t}$ is equal to (i) 1 for cities that are unambiguously located in the fare area 3, (ii) for cities that are unambiguously located in the fare area 2. The variable $z_{i,t}$ is in between 0 and 1 for cities that have neighborhoods from each side of the border.

The main identifying assumption for β to be an unbiased estimate for the treatment effect is that the *common trend assumption* holds: in absence of treatment, both the treated and untreated units would have evolved along the same path. I claim that the common trend assumption holds (conditional on my control variables) in my setting for two reasons. Firstly, my data selection procedure maximizes the similarity between the treated and untreated groups in terms of pre-reform observable characteristic. I also estimate models that are robust to pre-treatment differences in trends by interacting pre-treatment variables (as of January 2015) with a linear time trend and adding monthly dummy variables to capture the cyclical patterns of unemployment. The variables (in log) are interacted with a linear time trends; the size of the population in between 15 and 64; the number of employed workers in between 15 and 64; the number of them having a permanent contract and the number of them having a temporary contract. The rationale for the inclusion of these control variables is that the pre-treatment structure of the workforce would have caused the treated and the non-treated units to diverge, even in absence of the treatment. Secondly, by selecting municipalities that are close to each others, I control for the impact of potential unobservable characteristics on local employment dynamics.

4.2. Impact on employment

I first estimate the impact of the "dézonage" on employment. Results are presented in table 2.2. Column (1) indicates that the reform *decreased* the number of unemployed workers registered to Pole Emploi by approximately 2% workers for cities in the treatment group. The p-value is not below the usual 0.1 standard, which is in part due to the small size of the coefficient itself. To give an idea of the magnitude of the coefficient, Column (1) implies that on average 72 jobs were created per municipality in the treatment group.[10] Column (2) underlines that the treatment effect starts to materialize in 2016-2017. Given the frictional nature of unemployment, this result is expected: any reform impacting the labor market, having a direct impact on employment flows, takes a certain time to be visible on employment levels. Columns (3) and (4) indicate that using the "intensity of treatment" variable generates similar results in terms of magnitude and patterns.

The visual counterpart of the first column of table 2.2 is displayed in Figure 2.15, which shows the average number of people enrolled to their local branch of Pole Emploi for the

[10]On average, there were approximately 3606 workers registered to Pole Emploi in the treatment group for the period of observation.

treated (in blue) and non-treated cities (in red). Figure 2.15 indicates that the pre-trend assumption is slightly violated, with the treated units experiencing a slightly sharper increase in the number of workers registered to Pole Emploi. This visual inspection is confirmed by a Placebo test (See Table 2.7). This fact justifies the inclusion of control variables interacted with a linear time trend.

What types of workers benefited from the reform? Depending on their number of hours they worked last month and on their availability, workers registered to French unemployment agency (Pole Emploi) are assigned to 5 categories (see B for an exhaustive description of the categories). Unemployed workers having worked zero hours last month and actively searching for a job are in category A, while unemployed workers having worked less than 78 hours last month are in category B. Table 2.3 show results for the former and Table 2.4 for the latter. Column (1) of Table 2.3 suggests that the reform generated a 3.1% decrease in the number of category A workers. Column (2) indicates that the reform started to have employment effects in zone 3 in 2016. Table 2.4 indicates an overall decrease of category B workers for the treated units over the 2015 - 2019 period, but results are not statistically significant. Interestingly, Both columns (2) and (4) indicate the FTZ reform *increased* the number of category B workers in 2015. This results could be explained by the enrollment of new workers (previously not registered to Pole Emploi) and finding a temporary job and working less than 78 hours and/or by workers coming from other Pole Emploi categories. Results for categories C to E are presented in the Appendix (see Tables 2.10, 2.11 and 2.12). While no clear results are visible for categories C to D, Table 2.12 shows a spike in the number of workers registered in category E of approximately 12 % in 2016. Category E is composed of employed workers searching for an alternative job. One explanation is that the reform boosted the willingness of workers to look for jobs in other employment pools. To summarize, the reform mainly impacted category A workers, the category of workers with zero hours worked last month and mainly composed of long-term unemployed workers.[11]

[11]In 2011, 80% of French long-term unemployed workers were in caterogy A according to the Conseil d'Orientation pour l'Emploi: `http://www.coe.gouv.fr/COE_Chomage_de_longue_duree_Rapport-_version_finale-33ddd.pdf?file_url=IMG/pdf/COE_Chomage_de_longue_duree_Rapport-_version_finale-3.pdf`

Table 2.2: Number of workers registered to Pole Emploi

	Dependent variable: Log workers registered to Pole Emploi			
	(1)	(2)	(3)	(4)
β	-0.021		-0.021	
	$p = 0.152$		$p = 0.147$	
β_{2015}		0.006		0.004
		$p = 0.325$		$p = 0.513$
β_{2016}		-0.009		-0.013
		$p = 0.450$		$p = 0.299$
β_{2017}		-0.032^{*}		-0.038^{**}
		$p = 0.065$		$p = 0.033$
β_{2018}		-0.039^{*}		-0.044^{*}
		$p = 0.072$		$p = 0.067$
β_{2019}		-0.035		-0.047
		$p = 0.192$		$p = 0.108$

*p<0.1; **p<0.05; ***p<0.01

Notes: Each regression is based on 3008 observations. The panel contains 47 cities and 64 periods. Columns (1) and (2) correspond to regressions (2.3) and (2.4). Columns (3) and (4) correspond to regressions (2.5) and (2.6). For columns (1) and (2), the treatment (control) group includes cities in the fare area 3 (2), where the clear separation between the treated and non-treated is obtained by rounding the median fare area (measured at the IRIS level) to the nearest integer. Columns (3) and (4) measure the treatment effect by allowing for differences in the intensity of treatment. Cities entirely located in the fare area 2 are assigned an intensity of 0; cities entirely in the fare area 3 are assigned an intensity of 1; cities overlapping the border are assigned a value in $(0, 1)$.

Table 2.3: Number of workers registered to Pole Emploi in category A

	Dependent variable: Log workers registered to Pole Emploi in category A			
	(1)	(2)	(3)	(4)
β	-0.031^{*}		-0.033^{**}	
	$p = 0.078$		$p = 0.048$	
β_{2015}		-0.0001		-0.004
		$p = 0.991$		$p = 0.609$
β_{2016}		-0.027^{*}		-0.032^{**}
		$p = 0.078$		$p = 0.031$
β_{2017}		-0.041^{*}		-0.047^{**}
		$p = 0.052$		$p = 0.033$
β_{2018}		-0.038		-0.047
		$p = 0.132$		$p = 0.113$
β_{2019}		-0.045		-0.056
		$p = 0.140$		$p = 0.112$
Note:				*p<0.1; **p<0.05; ***p<0.01

Notes: Each regression is based on 3008 observations. The panel contains 47 cities and 64 periods. Columns (1) and (2) correspond to regressions (2.3) and (2.4). Columns (3) and (4) correspond to regressions (2.5) and (2.6). For columns (1) and (2), the treatment (control) group includes cities in the fare area 3 (2), where the clear separation between the treated and non-treated is obtained by rounding the median fare area (measured at the IRIS level) to the nearest integer. Columns (3) and (4) measure the treatment effect by allowing for differences in the intensity of treatment. Cities entirely located in the fare area 2 are assigned an intensity of 0; cities entirely in the fare area 3 are assigned an intensity of 1; cities overlapping the border are assigned a value in $(0, 1)$.

Table 2.4: Number of workers registered to Pole Emploi in category B

	Dependent variable: Log workers registered to Pole Emploi in category B			
	(1)	(2)	(3)	(4)
β	-0.017		-0.012	
	$p = 0.454$		$p = 0.541$	
β_{2015}		0.037^{**}		0.033^{**}
		$p = 0.044$		$p = 0.033$
β_{2016}		0.002		0.001
		$p = 0.915$		$p = 0.951$
β_{2017}		-0.027		-0.037
		$p = 0.307$		$p = 0.152$
β_{2018}		-0.060^{*}		-0.055^{*}
		$p = 0.082$		$p = 0.073$
β_{2019}		-0.059		-0.065^{*}
		$p = 0.113$		$p = 0.085$
Note:				*p<0.1; **p<0.05; ***p<0.01

Notes: Each regression is based on 3008 observations. The panel contains 47 cities and 64 periods. Columns (1) and (2) correspond to regressions (2.3) and (2.4). Columns (3) and (4) correspond to regressions (2.5) and (2.6). For columns (1) and (2), the treatment (control) group includes cities in the fare area 3 (2), where the clear separation between the treated and non-treated is obtained by rounding the median fare area (measured at the IRIS level) to the nearest integer. Columns (3) and (4) measure the treatment effect by allowing for differences in the intensity of treatment. Cities entirely located in the fare area 2 are assigned an intensity of 0; cities entirely in the fare area 3 are assigned an intensity of 1; cities overlapping the border are assigned a value in $(0, 1)$.

4.3. Discussion on internal migrations

In my sample, I am unable to track individuals. The results discussed above could be polluted by people migrating from zone 2 to 3, and conversely. In technical terms, the *single unit treatment value assumption* (SUTVA), according to which there should be no interference or spillovers between treated and untreated units, is likely to be violated. If unemployed workers from the fare area 3 were moving to the fare area 2, this would artificially increase the number of unemployed workers registered at Pole Emploi in zone 2 relative to zone 3, and create an upward bias in the estimates. But this case seems unlikely, as unemployed workers from the fare area 3 have more incentives to stay after the reform than before, as their commuting costs has decreased. The alternative scenario, in which unemployed workers from zone 2 move to zone 3 to enjoy a cheaper rent while keeping their commuting costs constant seems more plausible. Yet, this channel would push the number of unemployed in zone 3 upward and decrease the number of unemployed workers in zone 2. Thus, this would create a downward bias on my estimates (making it less negative). This intuition about the sign of the bias is supported by simulation of a simple structural model in the next section.

4.4. Impact on public transport usage

The "dézonage" reform led to a decrease in commuting costs for users living in the zones 2 - 5. The STIFF documents that this reform led to an increase in the number of people subscribing to annual, monthly and weekly travel passes. But did it lead to an increase in the number of travels via the public transport network? The answer to this question is essential, since a negative answer would cast serious doubt on the causal impact of the reform on employment. To the contrary, if usage of public transport in the treatment group increased relative to the control group, I interpret that as evidence that the cost differential boosted usage of public transport.

To answer this question, I estimate the specification (2.3) and (2.4) using a dataset with the yearly number of travellers passing trough each metro and RER stations. I restrict the sample to the stations in the zones 1, 2 and 3. I define the control group as stations in zones (1-2) and the treatment group as stations in zone 3. Results presented in Table 2.5. Column 1 shows that the reform led to a 5.1% increase in the number of travelers in zone 3 for the period 2016 - 2018. One word of caution is needed. The estimated value for β is likely to be positively biased because single-ride tickets are not taken into account. People in zone 3 previously commuting using single-ride tickets and switching to travel passes will raise the number of users in zone 3. Overall, Table 2.5 can be read as evidence of an increase

in public transport commuting in zone 3 and of a tendency to switch from single-ride tickets to cheaper travel passes.

Table 2.5: Impact of the reform on public transport usage

	Dependent variable: log traffic	
	(1)	(2)
β	0.051^{***}	
	$p = 0.00000$	
β_{2016}		0.036^{***}
		$p = 0.00001$
β_{2017}		0.053^{***}
		$p = 0.00002$
β_{2018}		0.065^{***}
		$p = 0.00000$
Note:	*p<0.1; **p<0.05; ***p<0.01	

Notes: Each regression is based on an unbalanced panel with data from 2013 to 2018: $n = 236, T = 5 - 6, N = 1415$. Columns (1) and (2) correspond to regressions (2.3) and (2.4). For columns (1) and (2), the treatment (control) group includes stations in the fare area 3 (1-2).

5. Model

The empirical analysis indicates the "dézonage" reform boosted employment for cities receiving the treatment, generating a 2% drop in the number of workers registered to Pole Emploi. To rationalize these findings, I develop a simple spatial search-and-matching model in which workers decide where to live, taking into considerations rent prices, commuting costs and the probability of finding a job. This is a partial equilibrium model: I do not

explicit the process of job creation from the firm's side and I abstract from the housing market. Developing a full-fledged general equilibrium model is outside the scope of the present chapter. The model features two channels through which commuting costs may affect unemployment: smaller commuting costs in some areas (i) decrease workers' reservation wage, hence boosts employment (ii) affect the attractiveness of these areas relative other cities, hence alter migrations patterns.

5.1. Setting and value functions

I model the Parisian region as a network of L nodes (cities/municipalities). Time is discrete and infinite. At each node is attributed a commuting time to the center c_l, a rent price r_l and a probability of finding a job λ_l. There is a mass 1 of ex-ante identical, risk-neutral and infinitely lived workers. Each worker, whether employed unemployed, consumes one unit of housing. When unemployed, they enjoy a flow of utility $b(l)$, which depends on location to capture differences in amenities between cities, in the spirit of Brueckner et al. (1999). When employed, they receive a wage w, which may depend on workers' location, and they produce an output of value py. The variable p is a random variable with cdf F and pdf f. One can think of p as measuring the dispersion of firms productivity in the economy, or as an idiosyncratic match-specific productivity shock (the "quality" of a match between similar workers and firms). The value of p is unknown to workers until meeting with firms, during which its value is fully revealed. I use the monocentric assumption according to which all jobs are located in the city center. Employed workers go to the CBD every weekday to work, while unemployed workers may go to the CBD to participate in interviews. The cost of going to the CBD every weekday while living in city L is denoted by c_l, while the cost of going to occasional interviews for unemployed workers is denoted by μc_l. As in Wasmer and Zenou (2002), I assume that unemployed workers have smaller commuting costs compared to employed workers residing in the same city: $\mu \in (0, 1)$. If employed workers go to the CBD 5 times a week and unemployed workers go to city center once a week, μ would be equal to $\frac{1}{5}$.

At the beginning of each period, unemployed workers search for a job. They meet with a firm with probability λ_l. Unemployed workers not meeting with firms have the possibility to stay in the same city, or to move to a new city l', in which case they incur a moving cost denoted by $m_{l,l'}$.[12] The ex-ante value of being unemployed to a worker is location l solves:

[12]The cost of staying put is $m_{l,l}$ is equal to 0.

$$U(l) = b(l) - \mu c_l - r_l + \frac{1}{1+r}\lambda_l E_p\big(\max\{W(l,p), U(l)\}\big) + \frac{1}{1+r}(1-\lambda_l)E_\epsilon \max_{l' \in 1,\ldots,L}\Big[\big(U(l') - m_{l,l'} + \epsilon\big)\Big] \tag{2.7}$$

where ϵ is a worker-specific $i.i.d$ shock capturing unobservable reasons for preferring a given location. I assume ϵ to be drawn from a type I extreme value distribution.[13] $W(l,p)$ denotes the value of being employed in city l with a firm of productivity type p. To simplify the model, let us make the assumption that workers have zero bargaining power, as in Postel-Vinay and Robin (2002). As a result, workers are offered a wage making them indifferent between being employed or unemployed: $U(l) = W(l,p)$. Equation (2.7) becomes:

$$U(l) = b(l) - \mu c_l - r_l + \frac{1}{1+r}\lambda_l U(l) + \frac{1}{1+r}(1-\lambda_l)E_\epsilon \max_{l' \in 1,\ldots,L}\Big[\big(U(l') - m_{l,l'} + \epsilon\big)\Big] \tag{2.8}$$

At the end of each period, unemployed workers solve the following discrete choice problem:

$$j \equiv \arg\max_{l \in 1,\ldots,L}\Big[\big(U(l') - m_{l,l'} + \epsilon\big)\Big] \tag{2.9}$$

Let us define the ex-ante choice-specific value function $V_l(j)$, denoting the net present value of payoffs conditional on taking action j, while in location l, before the exogenous taste shock is realized:

$$V_l(j) \equiv U(l') - m_{l,l'} \tag{2.10}$$

The probability for a worker in city l to move to city l', before the individual-specific idiosyncratic shock is ϵ is observed, is given by the multinomial logit formula:

$$P_l(l') = \frac{\exp\big(V_l(l')\big)}{\sum_{k=1}^{L}\exp\big(V_l(k)\big)} \tag{2.11}$$

Let us assume that employed workers cannot search for alternative jobs (no on-the-job search) or move to a new location. While these assumptions are rather strong, they greatly simplify the algebra. A match is destroyed exogenously at a rate δ, in which case newly unemployed workers receive the value $U(l)$. Employed workers stay employed with probability $1-\delta$, in which case they receive the value $W(l,p)$. The value of being employed in city l solves:

[13] With cdf $F(a) = \exp(\exp(-a))$.

$$W(l,p) = w(l,p) - c_l - r_l + \frac{\delta}{1+r}U(l) + \frac{1-\delta}{1+r}W(l,p) \tag{2.12}$$

Combining equations (2.8) and (2.12) and using the assumption that $U(l) = W(l,p)$ yields the following expression for wages, conditional on the job being feasible:

$$w(l) = b(l) + (1-\mu)c_l + \frac{1-\lambda_l}{1+r}E_\epsilon \max_{l' \in 1,\dots,L} \Big[U(l') - U(l) - m_{l,l'} + \epsilon\Big] \tag{2.13}$$

The intuition for equation (2.13) is simple: workers should be compensated (i) for the flow utility they enjoy while being unemployed $b(l)$ they lose by working (ii) the additional commuting cost to the CBD they incur when working (iii) the option value of changing location they lose when accepting a job. How do wages vary as workers move away from the city center? If one neglects expectations effects, equation (2.13) tells us that wages are decreasing in the distance to the city center, denoted by d, if amenities are decreasing *faster* than the rate at which commuting costs are increasing: $\frac{\partial b(l)}{\partial l}\frac{\partial l}{\partial d} < \mu \frac{\partial c_l}{\partial l}\frac{\partial l}{\partial d}$.

Remark that the maximum wage any firm can offer is the value of output py. If a firm were to offer a wage $w(l)$ greater than py, it would make negative profits. Any meeting for which $w(l) \leq py$ results in a new job being created. As a result, the wage offer function in city l for a firm with productivity p is:

$$w(l,p) = \begin{cases} w(l) & \text{if } w(l) \leq py \\ \varnothing & \text{if } w(l) > py \end{cases} \tag{2.14}$$

From equation (2.13), what is the impact of decreasing commuting costs on employment? A decrease in c_l leads to a lower reservation wage $w(l)$, thus one should expect a decrease in the unemployment rate, especially in the areas far away from the city center, as the set of feasible jobs expands. However, migration patterns are altered. Living further away from the city center is now more interesting from a worker' perspective. As a result some unemployed workers will move away from the city center. The final outcome depends on which of the two forces is more important.

Note that by virtue of the type I extreme value distributional assumption on ϵ, the expectation term in (2.8) admits the following closed-form solution

$$U(l) = b(l) - \mu c_l - r_l + \frac{\lambda_l}{1+r}U(l) + \frac{1-\lambda_l}{1+r}\big[\gamma + \log\big(\sum_{l'=1}^{L} \exp\{U(l') - m_{l,l'}\}\big)\big] \tag{2.15}$$

where γ is the Euler–Mascheroni constant.

5.2. Flow equations

Using equations (2.15) and (2.11), one can solve the model and find the probability that unemployed workers (not receiving an offer in the current period) relocate to each location. For workers meeting with firms, the job feasibility condition $w(l) \leq py$ tells us how many meetings are translated into jobs. Let us use $u_t(l)$ to denote the measure of unemployed workers residing in city l at time t. The flow equation for unemployment in city l has to take into account exogenous job destruction, job creation for unemployed workers in city l, unemployed workers from city l migrating to other cities and unemployed workers from other cities moving to city l:

$$\begin{aligned} u_{t+1}(l) = \delta\big(1-u_t(l)\big) &+ \lambda_l u_t(l)\big(P(w(l)>py) - P(w(l)\leq py)\big) \\ &- \big(1-\lambda_l\big)u_t(l)\big[\big(\sum_{k=1;k\neq l}^{L} P_l(k)\big) - P_l(l)\big] \\ &+ \sum_{k=1;k\neq l}^{L} \big(1-\lambda_k\big)P_k(l)u_t(k) \end{aligned} \tag{2.16}$$

To get a simpler expression for (2.16), let us make the assumption that firms' types p are normally distributed $p \sim \mathcal{N}(\mu_p, \sigma_p^2)$. Under this assumption, the term capturing the measure of unemployed workers finding a job in period t simplifies and the flow equation writes:

$$\begin{aligned} u_{t+1}(l) = \delta\big(1-u_t(l)\big) &+ \lambda_l u_t(l)\big(1 - 2\Phi\big(\frac{w(l)-\mu_p}{\sigma_p}\big)\big) \\ &- \big(1-\lambda_l\big)u_t(l)\big[\big(\sum_{k=1;k\neq l}^{L} P_l(k)\big) - P_l(l)\big] \\ &+ \sum_{k=1;k\neq l}^{L} \big(1-\lambda_k\big)P_k(l)u_t(k) \end{aligned} \tag{2.17}$$

Note that $f(l) = 1 - 2\Phi\big(\frac{w(l)-\mu_p}{\sigma_p}\big)$ is in between -1 and 1. If the reservation wage $w(l)$ in city l is so high that no match is feasible, $f(l)$ is equal to 1 and the measure of workers meeting with firms goes back to the pool of unemployed workers. On the other hand, if all matches are feasible $f(l)$ is equal to -1 and the measure of workers meeting with firms are joining the pool of employed workers.

5.3. *Equilibrium*

I define an equilibrium as a pair $(U, \boldsymbol{u})$ such that:

- U is a function satisfying equation (2.15)
- $\boldsymbol{u} \equiv (u(1), ..., u(l))$ is a joint distribution of employment and city location being such that each u_l is a fixed point of (2.16)

First note that there exists a unique solution to (2.15) (see section D). Consequently, flows between cities are uniquely determined by equation (2.11). As a result, equation (2.16) can be written as a matrix difference equation $\boldsymbol{u_{t+1}} = \mathbb{1}_L \delta + A\boldsymbol{u_t}$, where A is a deterministic matrix uniquely determined by the primitive parameters of the model. If an equilibrium exists, the steady state $\boldsymbol{u}$ solves $\boldsymbol{u} = (\boldsymbol{I} - A)^{-1}\mathbb{1}_L\delta$. Thus, existence and uniqueness of an equilibrium hinges on the invertibility of the matrix $B \equiv \boldsymbol{I} - A$. Without imposing further restrictions to the model, it is difficult to characterize conditions on parameters under which an equilibrium is reached.[14]

5.4. *Employment effects of the FTZ reform*

To gain insights on the model, let us focus on a non-degenerate "symmetric" equilibrium in which the outflows and inflows of workers from and to city L cancel out:

$$\big(1 - \lambda_l\big)u(l) \sum_{k=1;k\neq l}^{L} P_l(k) = \sum_{k=1;k\neq l}^{L} \big(1 - \lambda_k\big)P_k(l)u(k) \tag{2.18}$$

In this symmetric equilibrium (2.16) simplifies to:

$$u_{t+1}(l) = \delta\big(1 - u_t(l)\big) + \lambda_l u_t(l)\big(1 - 2\Phi\big(\frac{w(l) - \mu_p}{\sigma_p}\big)\big) + \big(1 - \lambda_l\big)u_t(l)P_l(l) \tag{2.19}$$

and the equilibrium is such that $\forall l, \in \{1, ..., L\}$:

$$u_l = \frac{\delta}{1 + \delta - \lambda_l\big(1 - 2\Phi\big(\frac{w(l)-\mu_p}{\sigma_p}\big)\big) - \big(1 - \lambda_l\big)P_l(l)} \tag{2.20}$$

Equation (2.20) illustrates the two channels through which a change in the commuting costs may impact the joint distribution of employment and location, modeled as a decrease in c_l for some cities. Firstly, there is a direct employment effect: the decrease in commuting costs pushes the wage $w(l)$ down, causing more jobs to be feasible, leading to a decrease in

[14] Numerically, convergence is reached with reasonable parameter values.

the unemployment rate for city l. Secondly, there is a reallocation effect: an asymmetric decrease in commuting costs c_l drives up the value of residing in city l relative to "non-treated" cities, leading to an increase in the probability of staying in city l $P_l(l)$. As a result, more unemployed workers decide to stay in city l rather than moving to an alternative city.

It is important to note the above analysis ignores the reaction of the land market: an increase in the attractiveness of city l should lead to an increase in rents for the same city, which may offset the desire of workers to relocate to city l. The discussion above also ignores the differences in timing: in the short run, the employment effect is likely to dominate over the reallocation effects if the moving costs are high and the proportion of movers is small.

Another channel is visible from equation (2.20). So far, I treated the job meeting rate in city l as exogenous. However, one could also view the FTZ reform as modifying the job meeting rate λ_l itself, if the unemployed workers of city l increase the size of the areas in which they search for a job after the reform. It is easy to see that if the job meeting rate λ_l increases, then the unemployment rate in city l decreases.

5.5. Algorithm to find an equilibrium

To find an equilibrium, I first solve for $U(l)$ and $P_l(l')$ $\forall l, l' \in \{1, ..., L\}^2$ using equations (2.8) and (2.11). Using equation (2.13), I determine which wages are feasible across space. I initialize an economy with an initial distribution of unemployed and employed workers across the cities $\{u_0(l)\}_{l=1,...,L}$, $\{e_0(l)\}_{l=1,...,L}$ with the constraint that the total mass integrates to 1:

$$\int u_0(l)dl + \int e_0(l)dl = 1$$

I then iterate forward the joint distribution of unemployment and city location using equation (2.16). I stop when the distance between two iterations reaches an arbitrary small value $d(u_{t+1}(l) - u_t(l)) < \varrho$

5.6. Parametrization and calibration

In this subsection, I present my parametrization of the model. I draw 100 cities at random in a square. I calculate the median distance to the city center, and I use the median to separate cities between two fare areas (zones 2 and 3). For the rent price associated to each city, I use a third order polynomial estimated from empirical data, using IRIS located in a distance between 0 and 20 km away from the city center (see Figure 2.5(b)). For the value of home production and leisure, I chose a simple linear function that depends on distance,

to capture that fact that amenities might be a decreasing function to the city-center in the Paris area

$$b(l) = b_0 + b_1 d(l) \tag{2.21}$$

For commuting costs, I choose a function that depends on the time travelled before reaching the CBD and on the fare area:

$$c(l) = c_0 + c_1 \times t(l) + c_2 \times f(l) \tag{2.22}$$

where t is the commuting time to the city center using public transport and f denotes an indicator variable equal to 1 if the city is in the fare area 3 and 0 otherwise. I follow the literature and assume that the job meeting rate is a decreasing function of distance:

$$\lambda(l) = \lambda_0 - \lambda_1 \times d(l) \tag{2.23}$$

In the age of internet and the widespread use of job-search websites, this assumption may be regarded with scepticism. But being close to the city center, where jobs are located, may have benefits through network effects (Bayer et al. (2008)). I set μ, which would imply that unemployed workers go on average to the CBD once a week, while employed workers go to the CBD 5 times a week, as in Wasmer and Zenou (2002). For the moving costs, I assume that there is a fixed cost associated with moving and an additional cost that is linked to the distance between the previous and the new location:

$$m_{l,l'} = m_0 \mathbb{1}\{l \neq l'\} + m_1 |l - l'| \tag{2.24}$$

For the yearly discount factor, I use $r = 0.05$. I calibrate the remaining 9 parameters in order to match the pattern of unemployment rate as a function of distance to the city center, as of 2015. I then simulate the "dézonage" reform by setting the parameter $c_2 = 0$.

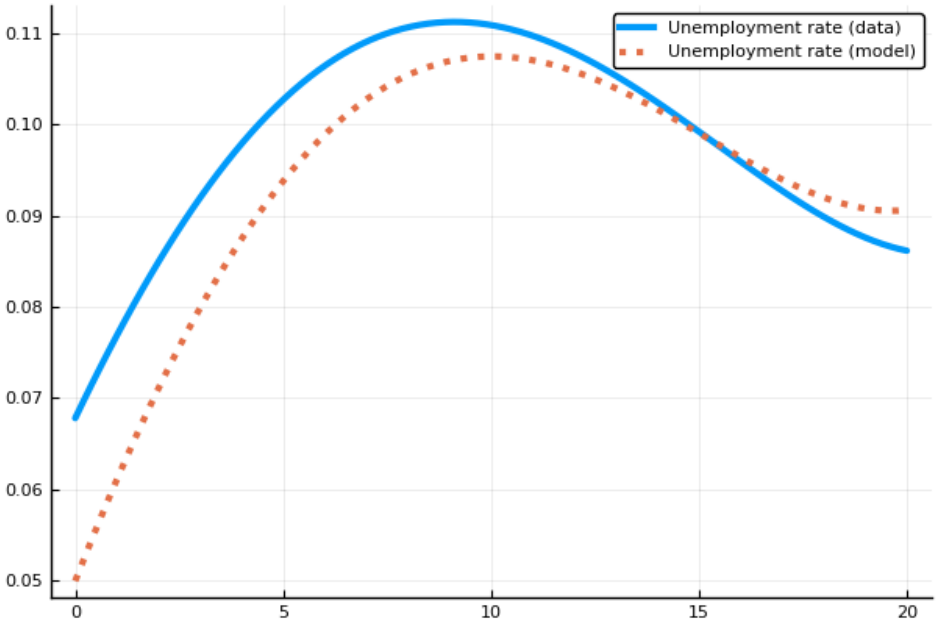
The model is able to reproduce the link between the local unemployment rate and the distance to the city center fairly well (see Figure 2.7). To generate this pattern, both the job-meeting rate and the level of amenity have to decrease quite substantially with distance to the center, which is compensated by lower rent prices (see Figure 2.8).

Table 2.6: Calibrated parameters

b_0	b_1	c_0	c_1	λ_0	λ_1	m_0	m_1	δ
0.6278	-0.2487	3.6684	0.1883	0.3999	0.0200	3.3616	1.6021	0.0952

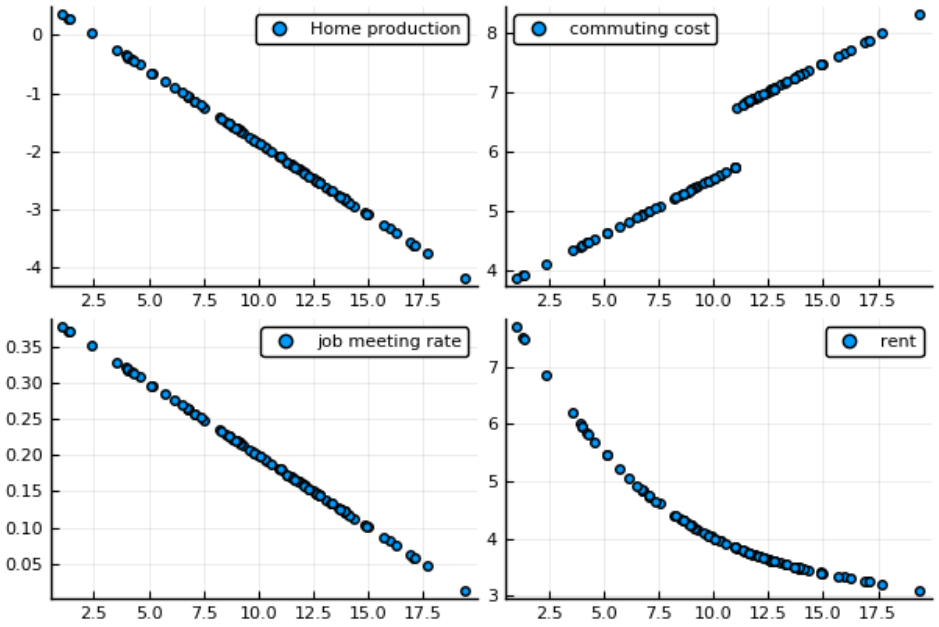
Notes: This table displays the calibrated parameter values. The parameters b_0 and b_1 are for the home production function. The parameters c_0 and c_1 capture the cost of commuting to the city center. The parameters λ_0 and λ_1 measure how the job meeting rate decreases with the distance to the city center. The parameters m_0 and m_1 measure how costly it is to move from one city to another. The parameter δ is the exogenous job destruction rate.

Fig. 2.7. Unemployment rate and distance to the CBD



Notes: This figure displays the unemployment rate (at the IRIS level) as a function of geodesic distance to Châtelet. The orange dotted line indicates is the model output, which is based on a simulation with 100 cities.

Fig. 2.8. Model's primitives and distance to the city center



Notes: This figure displays some of the model's primitives (the *y-axis*) as a function of distance to the city center (the *x-axis*). The top left panel displays how the value of home production changes with the distance to the city center. The top right panel displays the cost of commuting to the city center. The discontinuity at $d = 11$ km represents the border separating the zones 1-2 and 3 before the FTZ reform. The bottom left panel shows how the job meeting rate changes with the distance to the city center. The bottom right panel shows the rent gradient.

5.7. FTZ Counterfactual

To model the reform, I decrease the parameter from $c_2 = 1$ to $c_2 = 0$. This corresponds to approximately a 13% decrease in commuting costs for workers residing in zone 3.[15] I solve the model with the new set of parameter values and I simulate forward the distribution of unemployment, using as starting value the steady-state distribution of unemployment obtained with $c_2 = 1$.

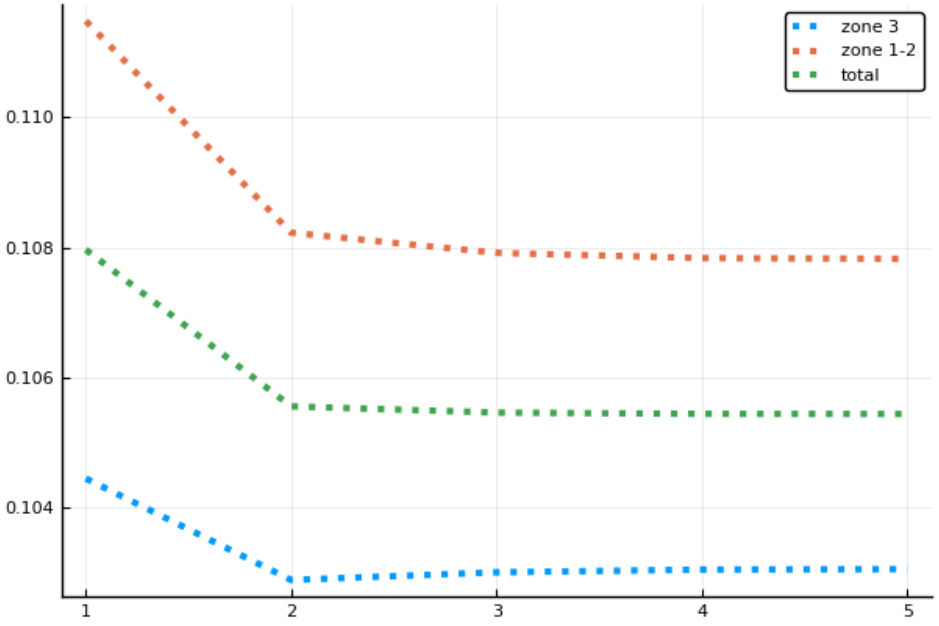
Results are presented in Figure 2.9. The reform generates a drop in the unemployment rate for cities in zones 3 *and* 2. Explaining the impact in zone 3, is straightforward. A decrease in the reservation wage for workers in zone 3 boosts employment. Why are cities in zone 2 also impacted? Lower transportation costs makes cities in the zone 3 more attractive

[15]The FTZ reform led to a $15.3/89.20 \approx 17\%$ decrease in commuting costs for workers previously owning a zone 1-3 annual Navigo pass.

for workers in zones 3 *and* 2. When having the opportunity to relocate to a new location, workers in zone 2 are more likely to choose a city in zone 3 after the reform (see Figure 2.10). As a result, the measure of unemployed workers in zone 2 decreases.

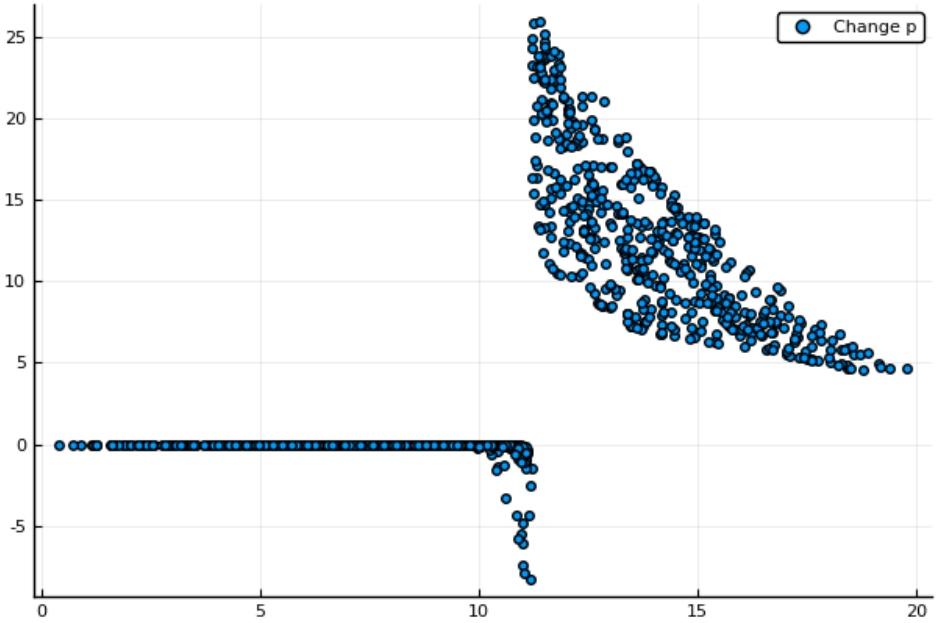
Hence, the model underlines that internal migration decisions are likely to create a downward bias in the empirical estimates. However, the severity of the bias is probably mitigated by the fact that, at the equilibrium, rent prices should react to the inflow of workers in zone 3. The feedback effect of rents is not taken into account in the current model.

Fig. 2.9. Simulation of the FTZ reform: impact on unemployment



Notes: This figure displays the evolution of the unemployment rate for the treated (zone 3) and non-treated groups (zone 1-2). The *y-axis* represents the unemployment rate for the treated and non treated, while the *x-axis* represents periods (years) after the implementation of the reform.

Fig. 2.10. Simulation of the FTZ reform: impact on migration decisions



Notes: This figure shows the percentage *change* in the probability of staying in the same city (post-reform probability minus the pre-reform probability), when facing the opportunity to change location (the *y-axis*), as a function of the geodesic distance to the city center (the *x-axis*). The discontinuity at $x = 11$ represents the border separating the zones 1-2 and 3.

This graph indicates that workers living in the zones 1-2 are more likely to relocate to zone 3 after the reform, while workers in zone 3 are more likely to stay in zone 3. Calculations are based on a simulation with 1000 cities.

6. Conclusion

The creation of the public transport travel pass "Forfait Toutes Zones" in the Paris metropolitan area provides a rich setup to measure the impact of a decrease of commuting costs on local employment dynamics. The cost of using public transport in the periphery decreased (zones 3-5), while it remained constant for Paris intra-muros and in the close suburbs (zones 1-2). My reduced form estimates, relying on the spatial discontinuity introduced by the reform between zones 2 and 3, indicate that the reform led to a 2% *decrease* in the number of unemployed workers registered in the local unemployment agency for the municipalities benefiting from the reform. The reform mainly impacted category A workers, more likely to be long-term unemployed workers.

I build and calibrate a simple spatial search-and-matching model of the labor market.

The model helps us in understanding how a localized drop in commuting costs might generate an increase in local employment. The main mechanism is that the a decrease in commuting costs makes employment more attractive relative to unemployment, hence decreasing the reservation wage. Unemployed workers might also have expanded the area in which they search for a job, creating an additional positive employment channel. More precise data and additional analyses would be needed to quantify the contribution of each channel to local employment dynamics.

Using the model, I also make the observation that using a repeated cross-section at the municipality level instead of a panel of workers is not a major threat to my identification strategy. The reason is that unemployed workers have more incentives to relocate to the suburbs of Paris after the reform, which should push up the unemployment number in these areas. However, if the FTZ reform did encourage people to move from Paris *intra-muros* to the suburbs, the housing market should have reacted to the extra demand generated by the reform. It is likely that the FTZ reform increased house prices and rents in the suburbs.

The links between the FTZ reform and house prices are currently being investigated and the model is being extended to take into consideration the feedback effects that higher rents may have on migration towards the suburbs.

References

Anne, D. and L'Horty, Y. (2009). Aides sociales locales, revenu de solidarité active (rsa) et gains du retour à l'emploi. *Économie et statistique*, 429(1):129–157.

Bayer, P., Ross, S. L., and Topa, G. (2008). Place of work and place of residence: Informal hiring networks and labor market outcomes. *Journal of Political Economy*, 116(6):1150–1196.

Brueckner, J. K., Thisse, J.-F., and Zenou, Y. (1999). Why is central paris rich and downtown detroit poor?: An amenity-based theory. *European economic review*, 43(1):91–107.

Chapelle, G., Vignolles, B., and Wolf, C. (2018). Impact of a housing tax credit on local housing markets: Evidence from france. *Annals of Economics and Statistics/Annales d'Économie et de Statistique*, (130):101–131.

Chapelle, G., Wasmer, E., Bono, P.-H., Oswald, F., Urvoy, C., et al. (2017). Impact du grand paris express sur le marché du travail et le marché du logement. Technical report, Sciences Po.

Duranton, G. and Turner, M. A. (2012). Urban growth and transportation. *Review of Economic Studies*, 79(4):1407–1440.

Einiö, E. and Overman, H. G. (2016). The (displacement) effects of spatially targeted enterprise initiatives: Evidence from uk legi.

Garcia-López, M.-À., Hémet, C., and Viladecans-Marsal, E. (2017). Next train to the polycentric city: The effect of railroads on subcenter formation. *Regional Science and Urban Economics*, 67:50–63.

Hilber, C. A., Carozzi, F., and Xiaolun, Y. (2019). The economic impacts of help to buy.

Kline, P. and Moretti, E. (2013). Local economic development, agglomeration economies, and the big push: 100 years of evidence from the tennessee valley authority. *The Quarterly Journal of Economics*, 129(1):275–331.

Mayer, T. and Trevien, C. (2017). The impact of urban public transportation evidence from the paris region. *Journal of Urban Economics*, 102:1–21.

Neumark, D. and Simpson, H. (2015). Place-based policies. In *Handbook of regional and urban economics*, volume 5, pages 1197–1287. Elsevier.

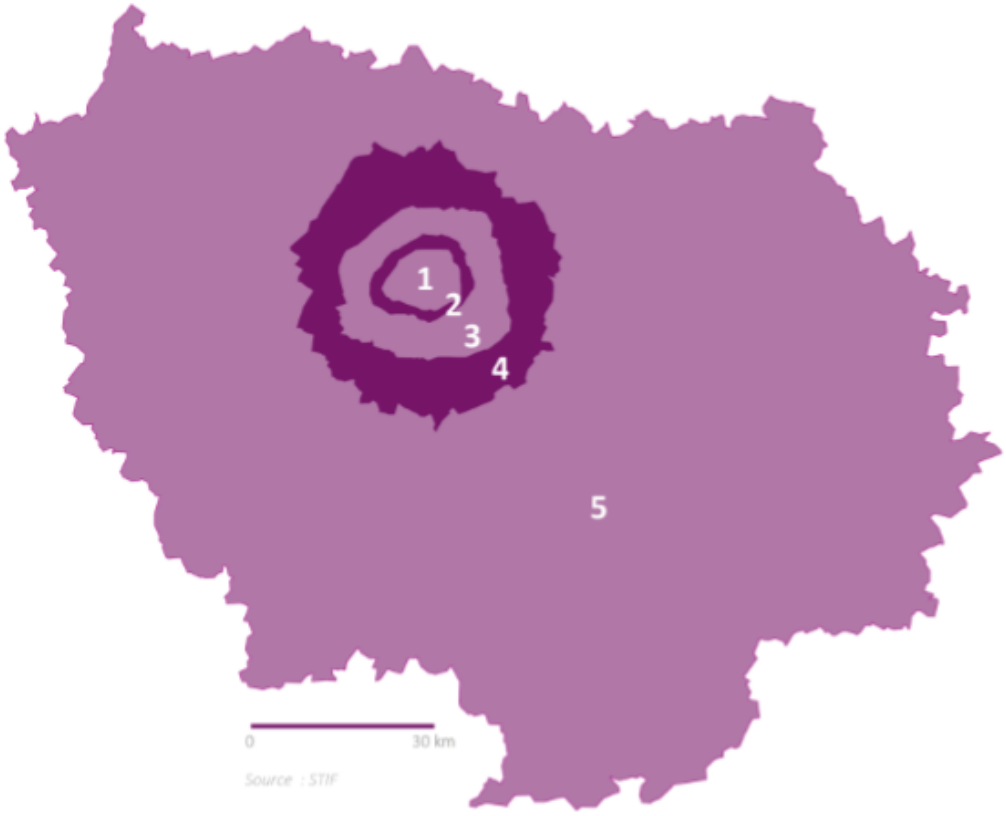
OMNIL (2018). Note repères 10 les forfaits toutes zones, quels impacts deux ans après leur mise en place ? *Online, posted 19 avril 2018.* `http://www.omnil.fr/IMG/pdf/note_reperes_n10_tarif_unique_vf.pdf`.

Postel-Vinay, F. and Robin, J.-M. (2002). Equilibrium wage dispersion with worker and employer heterogeneity. *Econometrica*, 70(6):2295–2350.

Wasmer, E. and Zenou, Y. (2002). Does city structure affect job search and welfare? *Journal of urban Economics*, 51(3):515–541.

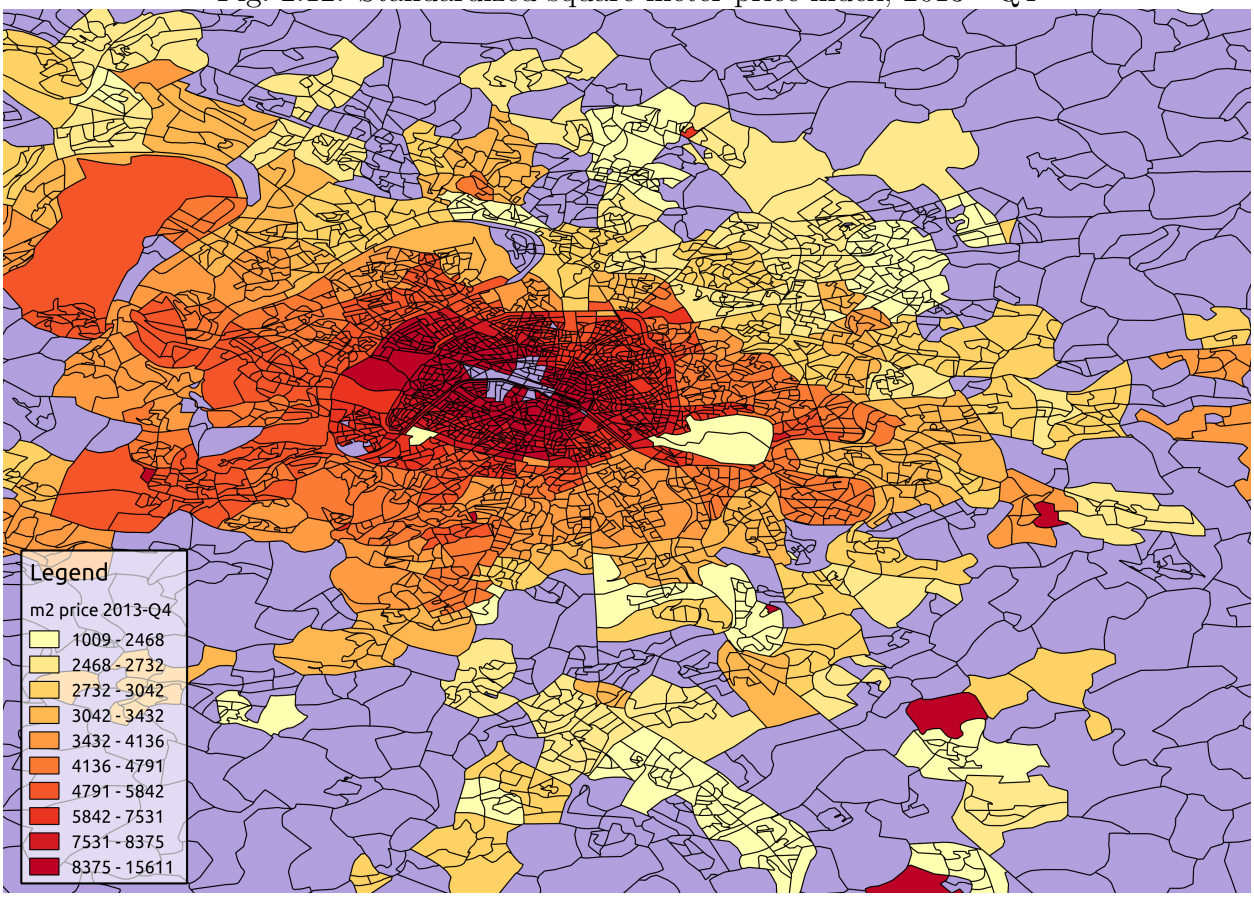
A. Data

Fig. 2.11. Public transport networks fare areas in the region Ile-de-France



Notes: This figure shows the fare areas for the public transport network of the region Ile-de-France.

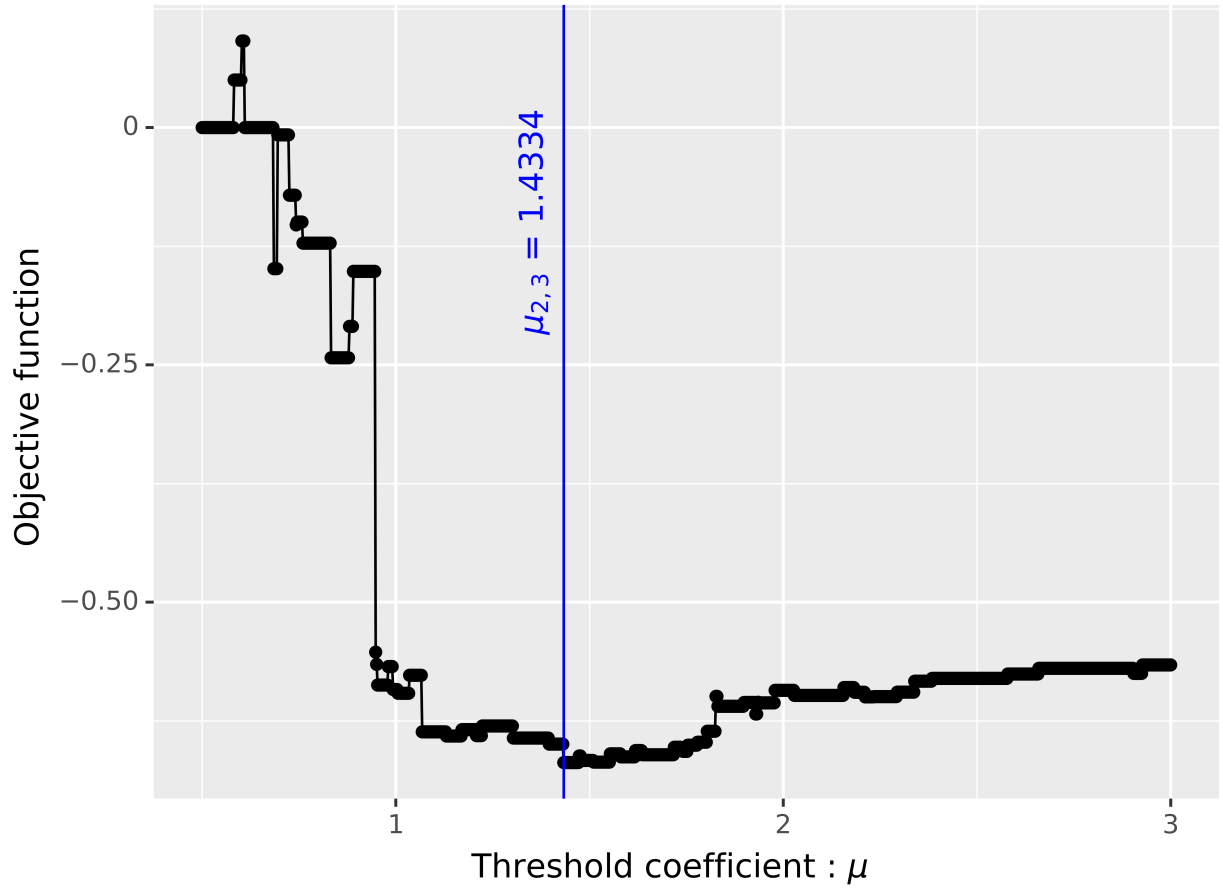
Fig. 2.12. Standardized square meter price index, 2013 - Q4



Notes: This figure shows the standardized square meter price index (at the municipality level) in the Parisian region. Missing values (not enough transactions were realized to generate a reliable measurement) are in purple.

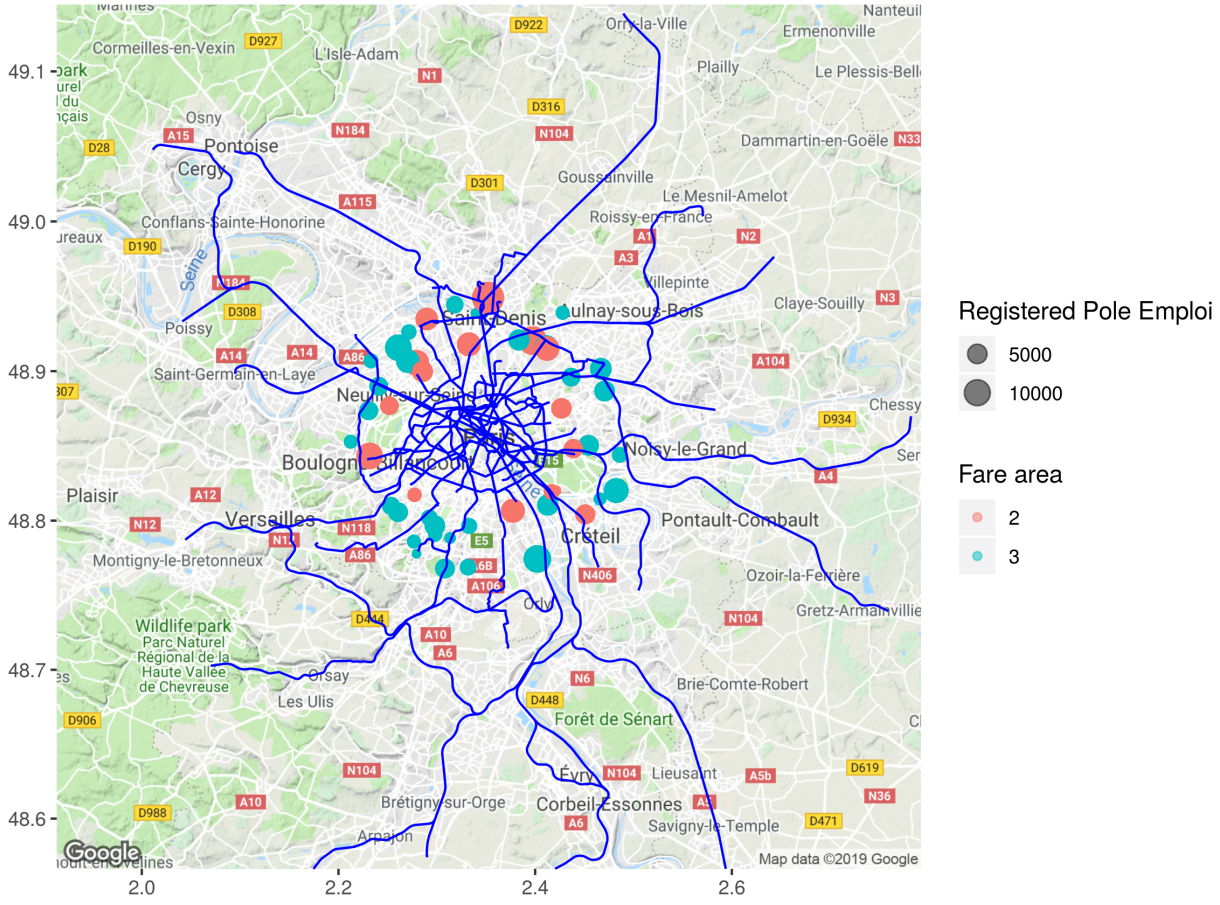
Source: data from the Grand Paris notary agency

Fig. 2.13. Minimization of the objective function in (2.2)



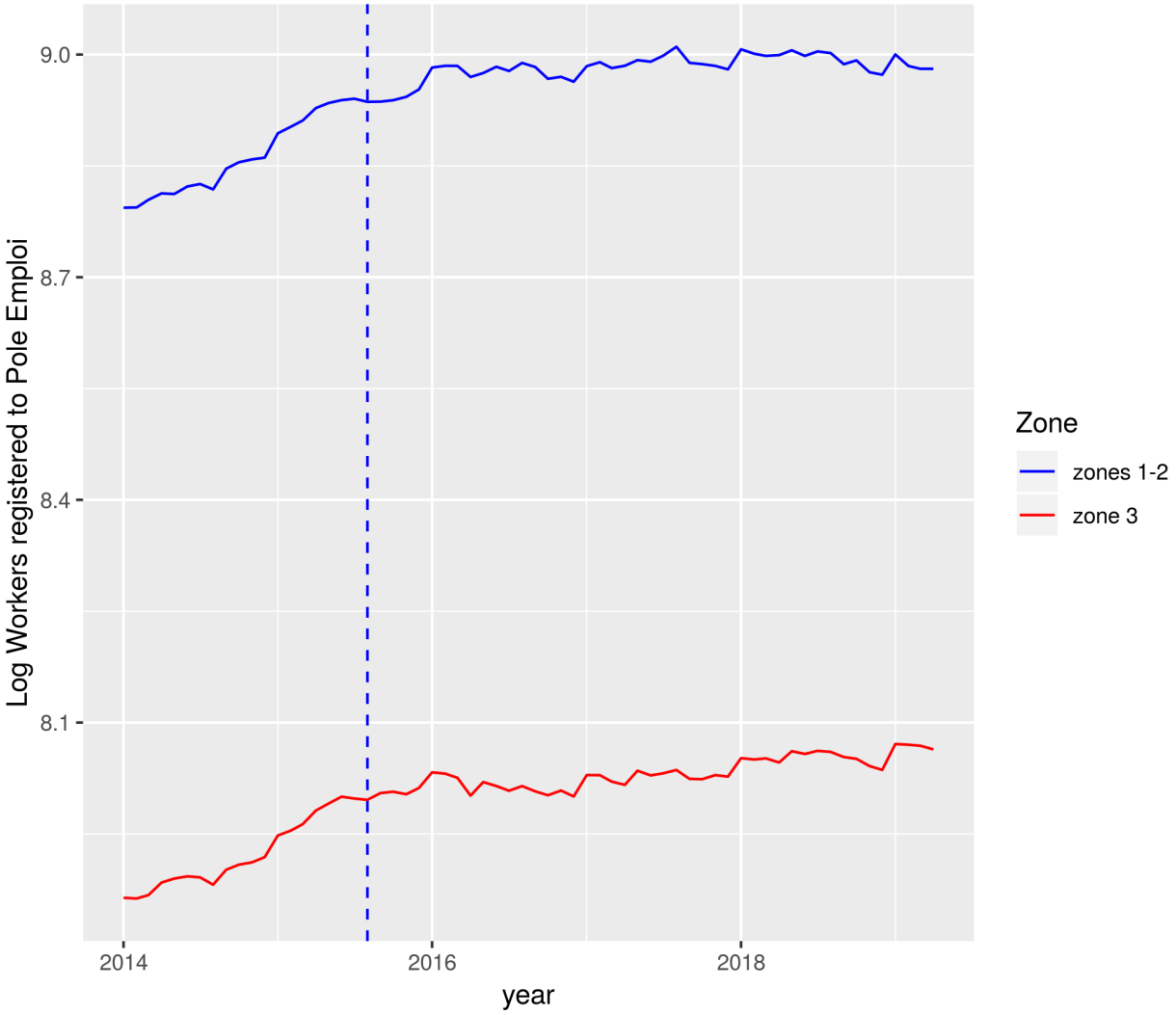
Notes: This figure shows the objective function being minimized in (2.2). The smaller the objective function, the more similar the treated and non-treated groups are in the pre-treatment period. The vertical blue line is the minimum of the objective function (found by a naive grid search with 1000 grid points evenly spaced on $[0.5; 3.0]$. The objective function was normalized such that its value is equal to 0 for $\mu = 0$.

Fig. 2.14. Municipalities included in the sample



Notes: This figure display the cities included in the sample. Each circle represents a city. The size of the circle is proportional to the number of workers registered to Pole Emploi (French unemployment center) as of April 2015. Cities in blue are in the fare area 3, while the ones in blue are in the fare area 2. The blue lines represent available rail public transport lines (metro and RER).

Fig. 2.15. Number of workers registered to Pole Emploi by fare area



Notes: This figure displays the average (log) number of workers enrolled to Pole Emploi for cities in the fare areas 2 (red) and 3 (in blue) for my selected sub-sample. The vertical blue line indicates the date at which the "dézonage" reform became effective (September 2015).

B. Pole Emploi Categories

Workers registered to Pole Emploi (the French unemployment agency) are assigned to one of the 5 existing categories (A, B, C, and E)[16]:

- A: unemployed worker registered to Pole Emploi and actively searching for a perma-

[16] source:https://www.insee.fr/en/metadonnees/definition/c2010

nent, temporary or seasonal job

- B: unemployed worker registered to Pole Emploi and actively searching for a permanent, temporary or seasonal job, having worked no more than 78 hours in the last month
- C: unemployed worker registered to Pole Emploi and actively searching for a permanent, temporary or seasonal job, having worked more than 78 hours in the last month
- D: unemployed worker registered to Pole Emploi, but not directly available for a vacancy because of an internship, a training program, a sick leave
- E: employed worker registered to Pole Emploi and searching for an alternative position

C. Robustness

C.1. Placebo tests

Table 2.7: Placebo test (all categories)

	Dependent variable: Log number of workers registered to Pole Emploi	
	(1)	(2)
β	0.002	0.007
	$p = 0.765$	$p = 0.255$
Note:		*p<0.1; **p<0.05; ***p<0.01

Notes: Each regression is based on 940 observations (balanced panel: n = 47, T = 20, N = 940). In the Placebo test, I pretend that the reform happened in September 2014. Columns (1) corresponds to regressions (2.3). Columns (2) corresponds to regressions (2.5). For column (1) the treatment (control) group includes cities in the fare area 3 (2), where the clear separation between the treated and non-treated is obtained by rounding the median fare area (measured at the IRIS level) to the nearest integer. Column (2) measures the treatment effect by allowing for differences in the intensity of treatment. Cities entirely located in the fare area 2 are assigned an intensity of 0; cities entirely in the fare area 3 are assigned an intensity of 1; cities overlapping the border are assigned a value in $(0, 1)$.

Table 2.8: Placebo test category A

	Dependent variable: Log number of workers registered to Pole Emploi in category A	
	(1)	(2)
β	0.004	0.009
	$p = 0.677$	$p = 0.310$
Note:		*p<0.1; **p<0.05; ***p<0.01

Notes: Each regression is based on 940 observations (balanced panel: n = 47, T = 20, N = 940). In the Placebo test, I pretend that the reform happened in September 2014. Columns (1) corresponds to regressions (2.3). Columns (2) corresponds to regressions (2.5). For column (1) the treatment (control) group includes cities in the fare area 3 (2), where the clear separation between the treated and non-treated is obtained by rounding the median fare area (measured at the IRIS level) to the nearest integer. Column (2) measures the treatment effect by allowing for differences in the intensity of treatment. Cities entirely located in the fare area 2 are assigned an intensity of 0; cities entirely in the fare area 3 are assigned an intensity of 1; cities overlapping the border are assigned a value in $(0, 1)$.

Table 2.9: Placebo test category B

	Dependent variable: Log number of workers registered to Pole Emploi in category B	
	(1)	(2)
β	0.009	0.012
	$p = 0.629$	$p = 0.529$

Note: *p<0.1; **p<0.05; ***p<0.01

Notes: Each regression is based on 940 observations (balanced panel: n = 47, T = 20, N = 940). In the Placebo test, I pretend that the reform happened in September 2014. Columns (1) corresponds to regressions (2.3). Columns (2) corresponds to regressions (2.5). For column (1) the treatment (control) group includes cities in the fare area 3 (2), where the clear separation between the treated and non-treated is obtained by rounding the median fare area (measured at the IRIS level) to the nearest integer. Column (2) measures the treatment effect by allowing for differences in the intensity of treatment. Cities entirely located in the fare area 2 are assigned an intensity of 0; cities entirely in the fare area 3 are assigned an intensity of 1; cities overlapping the border are assigned a value in $(0, 1)$.

C.2. Main regression and categories C - E

Table 2.10: Number of workers registered to Pole Emploi in category C

	Dependent variable: Log number of workers registered to Pole Emploi in category C			
	(1)	(2)	(3)	(4)
β	0.004		0.001	
	$p = 0.876$		$p = 0.977$	
β_{2015}		0.001		0.005
		$p = 0.963$		$p = 0.695$
β_{2016}		-0.002		-0.002
		$p = 0.912$		$p = 0.918$
β_{2017}		0.003		-0.004
		$p = 0.912$		$p = 0.871$
β_{2018}		0.014		0.012
		$p = 0.628$		$p = 0.674$
β_{2019}		0.010		0.003
		$p = 0.777$		$p = 0.940$
Note:				*p<0.1; **p<0.05; ***p<0.01

Notes: Each regression is based on 3008 observations. The panel contains 47 cities and 64 periods. Columns (1) and (2) correspond to regressions (2.3) and (2.4). Columns (3) and (4) correspond to regressions (2.5) and (2.6). For columns (1) and (2), the treatment (control) group includes cities in the fare area 3 (2), where the clear separation between the treated and non-treated is obtained by rounding the median fare area (measured at the IRIS level) to the nearest integer. Columns (3) and (4) measure the treatment effect by allowing for differences in the intensity of treatment. Cities entirely located in the fare area 2 are assigned an intensity of 0; cities entirely in the fare area 3 are assigned an intensity of 1; cities overlapping the border are assigned a value in $(0, 1)$.

Table 2.11: Number of workers registered to Pole Emploi in category D

	Dependent variable: Log number of workers registered to Pole Emploi in category D			
	(1)	(2)	(3)	(4)
β	0.037		0.008	
	$p = 0.440$		$p = 0.877$	
β_{2015}		0.021		0.012
		$p = 0.711$		$p = 0.804$
β_{2016}		0.087		0.059
		$p = 0.194$		$p = 0.373$
β_{2017}		-0.028		-0.065
		$p = 0.633$		$p = 0.271$
β_{2018}		0.033		-0.011
		$p = 0.559$		$p = 0.879$
β_{2019}		0.116		0.094
		$p = 0.116$		$p = 0.302$

Note: *p<0.1; **p<0.05; ***p<0.01

Notes: Each regression is based on 3008 observations. The panel contains 47 cities and 64 periods. Columns (1) and (2) correspond to regressions (2.3) and (2.4). Columns (3) and (4) correspond to regressions (2.5) and (2.6). For columns (1) and (2), the treatment (control) group includes cities in the fare area 3 (2), where the clear separation between the treated and non-treated is obtained by rounding the median fare area (measured at the IRIS level) to the nearest integer. Columns (3) and (4) measure the treatment effect by allowing for differences in the intensity of treatment. Cities entirely located in the fare area 2 are assigned an intensity of 0; cities entirely in the fare area 3 are assigned an intensity of 1; cities overlapping the border are assigned a value in $(0, 1)$.

Table 2.12: Number of workers registered to Pole Emploi in category E

	Dependent variable: Log number of workers registered to Pole Emploi in category E			
	(1)	(2)	(3)	(4)
β	0.064		0.087	
	$p = 0.253$		$p = 0.172$	
β_{2015}		0.046		0.050
		$p = 0.332$		$p = 0.195$
β_{2016}		0.120**		0.138**
		$p = 0.035$		$p = 0.021$
β_{2017}		0.068		0.074
		$p = 0.298$		$p = 0.339$
β_{2018}		−0.020		0.0002
		$p = 0.802$		$p = 0.998$
β_{2019}		0.055		0.037
		$p = 0.634$		$p = 0.743$

Note: *p<0.1; **p<0.05; ***p<0.01

Notes: Each regression is based on 3008 observations. The panel contains 47 cities and 64 periods. Columns (1) and (2) correspond to regressions (2.3) and (2.4). Columns (3) and (4) correspond to regressions (2.5) and (2.6). For columns (1) and (2), the treatment (control) group includes cities in the fare area 3 (2), where the clear separation between the treated and non-treated is obtained by rounding the median fare area (measured at the IRIS level) to the nearest integer. Columns (3) and (4) measure the treatment effect by allowing for differences in the intensity of treatment. Cities entirely located in the fare area 2 are assigned an intensity of 0; cities entirely in the fare area 3 are assigned an intensity of 1; cities overlapping the border are assigned a value in $(0, 1)$.

D. Existence and uniqueness of a solution to $U(l)$

This section shows that the operator T satisfies Blackwell's sufficient conditions for a contraction. I first show monotonicity (section D.1) before proving discounting (section D.2).

D.1. Monotonicity

Let us fix $x \in X$ and consider two function $U, W \in B(X)$ such that $\forall x \in X, U(x) \leq W(x)$. Let us also consider an interest rate r strictly positive and a job meeting rate $\lambda_l \geq 0$.

Monotonicity of T follows from the fact that the exponential and the logarithm functions are strictly increasing on their respective domains:

$$\begin{aligned}
[T(U)](x) - [T(W)](x) &= \frac{\lambda_l}{1+r}\big(U(x) - W(l)\big) + \frac{1-\lambda_l}{1+r}\big[log\big(\sum_{l'=1}^{L}\exp\{U(l') - m_{l,l'}\}\big) - \\
&\quad log\big(\sum_{l'=1}^{L}\exp\{W(l') - m_{l,l'}\}\big)\big] \\
&\leq \frac{1-\lambda_l}{1+r}\big[log\big(\frac{\sum_{l'=1}^{L}\exp\{U(l') - m_{l,l'}\}}{\sum_{l'=1}^{L}\exp\{W(l') - m_{l,l'}\}}\big)\big] \\
&\leq \frac{1-\lambda_l}{1+r}\big[log\big(\frac{\sum_{l'=1}^{L}\exp\{W(l') - m_{l,l'}\}}{\sum_{l'=1}^{L}\exp\{W(l') - m_{l,l'}\}}\big)\big] \\
&= 0
\end{aligned}$$

Where the first inequality comes from the assumption:

$$\forall x \in X, U(x) \leq W(x)$$

and the fact that $\log(a) - \log(b) = log(\frac{a}{b})$. The second inequality uses the fact that the logarithm and the exponential functions are strictly increasing on their domain.

Hence $\forall (U, W) \in B(X)^2$ such that $U \leq W$, $(TU)(x) \leq (TW)(x)$.

$$
\begin{aligned}
[T(U+a)](x) - [T(U)](x) &= \frac{\lambda_l}{1+r}(U(x)+a-U(x)) + \frac{1-\lambda_l}{1+r}\big[log\big(\sum_{l'=1}^{L}\exp\{a+U(l')-m_{l,l'}\}\big) \\
&\quad - log\big(\sum_{l'=1}^{L}\exp\{U(l')-m_{l,l'}\}\big)\big] \\
&= \frac{\lambda_l}{1+r}a + \frac{1-\lambda_l}{1+r}\big[log\big(\frac{\sum_{l'=1}^{L}\exp\{a+U(l')-m_{l,l'}\}}{\sum_{l'=1}^{L}\exp\{U(l')-m_{l,l'}\}}\big)\big] \\
&= \frac{\lambda_l}{1+r}a + \frac{1-\lambda_l}{1+r}\big[log\big(\frac{\sum_{l'=1}^{L}\exp\{a\}\exp\{U(l')-m_{l,l'}\}}{\sum_{l'=1}^{L}\exp\{U(l')-m_{l,l'}\}}\big)\big] \\
&= \frac{\lambda_l}{1+r}a + \frac{1-\lambda_l}{1+r}\big[log\big(\frac{\exp\{a\}\sum_{l'=1}^{L}\exp\{U(l')-m_{l,l'}\}}{\sum_{l'=1}^{L}\exp\{U(l')-m_{l,l'}\}}\big)\big] \\
&= \frac{\lambda_l}{1+r}a + \frac{1-\lambda_l}{1+r}\big[log(\exp\{a\})\big] \\
&= \frac{1}{1+r}a \\
&\leq \beta a
\end{aligned}
$$

where $\beta = \frac{1}{1+r} \in [0,1]$ for $r > 0$.

Hence $\forall U \in B(X)$ and $a > 0$, there exists $\beta \in [0,1]$ such that $[T(U+a)](x) - [T(U)](x) \leq \beta a$.

Chapter 3

Rental Housing Market and Directed Search

Abstract

This chapter presents new empirical facts on the rental housing market in the Paris metropolitan area. Using a novel dataset of online ads for the Parisian rental market and a hedonic model that includes apartment features and photos, two main stylized facts are established. Firstly, landlords who set a lower rent — at constant dwelling characteristics — attract more applicants, as predicted by a standard directed search model. Secondly, a new stylized fact regarding landlords' pricing strategy is established. A non-negligible proportion of landlords use a two-step pricing strategy: they set a high advertised rent, before lowering the advertised rent after a wait-and-see period.

Keywords: Rental housing market; Search; Hedonic pricing model
JEL Classification: D83, R21, R31, C81

Introduction

The main goal of this chapter is to present new empirical evidence on how functions the rental housing market. The Great Recession illustrated the importance of the housing market for the economy as a whole. Besides its importance for financial markets, the housing market is at the root of many social and economic problems. Ganong and Shoag (2017) show that the sharp increase in house prices in major US cities for the period 1980-2010 slowed down internal migration and reduced economic convergence among US States. That is, without the surge in house prices observed for the period 1980-2010, the difference in real income per capita across US States would have been smaller. Brown and Matsa (2019) document that job seekers in areas with depressed housing markets are less likely to apply for jobs requiring relocation, distorting labor market outcomes. Herkenhoff et al. (2018) present evidence that a lack of new houses created by land regulations substantially depressed US

output and productivity growth relative to historical trends. There is evidence that the housing market and the lack of affordable housing profoundly impacts the political process. Adler and Ansell (2020) argue that the sentiment of being excluded from a rising housing market were pivotal in shaping support for populism in the Britain's 2016 referendum on membership to the European Union and France's Presidential Election in 2017. The state of the housing market is also crucial for many other aspects in life. For instance, Wei et al. (2017) show that housing wealth is a key variable to determine one's attractiveness in the marriage market in China.

While there is a growing literature studying how houses are sold and bought, little is known about the rental market. There are at least two reasons why the economic literature has focused on the property selling market. Firstly, home ownership has been at the center of the political discourse for decades, with governments aiming at boosting home ownership rate in major economies. In the UK, Margaret Thatcher started her "right-to-buy" program in the 1980s, allowing Britons in social housing to buy their property. Fischer and Sard (2017) report that the US government spent 190 billion USD in 2015 (more than 1% of GDP) to help Americans buy homes. In France, governments have tried to promote home ownership with many programs starting in the 1980s (Laferrère et al. (2017)). Secondly, while data on house transactions are easily available in developed economies because of legislation encouraging the creation of such records — for instance, all transactions must be officially recorded by a notary in France — the rental housing market is characterized by a lack of data. In France, there is no comprehensive national dataset on tenancy agreements.[1]

This chapter helps to fill the knowledge gap on the rental housing market by constructing and utilizing a dataset of online ads for the Parisian market. Chapelle et al. (2020) perform a related web-scraping exercise for France as a whole. Comparing their observations with administrative data, they show that data collected from online ads offer a representative snapshot of the rental market. They also show that tenants in social housing receive an implicit subsidy approximately equal to 46% of their rent. Boeing and Waddell (2017) present an empirical analysis of the US rents collecting data from Craiglist, an advertisement website that dominates the US rental housing market. Hyland et al. (2013) also use data from online ads on prices and rents to measure the effect of energy efficiency ratings on sale and rental prices for properties in Ireland. The author find that energy efficiency has a positive effect on sales prices of properties.

The present chapter makes two contributions to a growing literature on the rental housing

[1]Tenants eligible to rent allowance must declare their rent to the 'Caisse des Allocations Familiales' (CAF), a French governmental agency. The dataset collected is not made public though. Another potential source is Clameur, which is a private agency collecting data on rent. While Clameur releases indexes on rents, the underlying dataset they have is kept private.

market. First, one key novelty of the present analysis is that information on both sides of the market was collected at a detailed level. The dataset includes detailed information on properties, as well as the number of contacts received by landlords for each ad. Generally, housing search behavior is unobserved. Two exceptions are Genesove and Han (2012) and Piazzesi et al. (2020). The former paper collects search activity on information from both buyers and sellers using survey data[2] to show that a standard random search model fits the data well. The latter paper is closer to the present study. The authors have access to email alert settings from the website trulia.com, a leading advertisement platform for home buyers in the San Francisco Bay Area. The authors use the data to build a random search model with segmented housing markets. The present paper differs from the above-mentioned studies because it focuses on the rental housing market and it uses a fine measurement of housing search intensity: the number of contacts received by the landlord for each online ad. As predicted by a standard directed search model, I show that houses that are cheaper than expected (controlling for observable characteristics) attract more searchers. Landlords discounting their apartment are rewarded by a higher arrival rate of tenants. By the same token, tenants searching for a discounted apartment have to "pay" a higher price by lowering their probability to find such a good deal next period.

A second contribution is to show that a non-negligible proportion of landlords use a two-step pricing strategy. Approximately 7% of landlords set a rent that is higher than what is predicted by a hedonic regression model, before lowering the advertised rent to a level in line with the predictions based on observable characteristics. To the best of my knowledge, this empirical fact is new in the literature. While the literature on the selling market has shown that "bidding wars" (when multiple buyers compete for a house and push sales price above list price) are widespread (Han and Strange (2014)), little is known about the pricing strategies used by landlords.

This paper is structured as follows. Section 1 presents key descriptive statistics for the housing market in France, with a special focus on the rental market in the Paris metropolitan area, and explains how the data was collected. Section 2 presents the main empirical facts of the chapter. Section 3 describes a directed search model that can explain some keys facts established in the previous section. The final section concludes.

[2]Questions asked were "How long did you actively search before you located the home you recently purchased?"; "How long was this home on the market?"; "Including the home you purchased, how many homes did you walk through and examine before choosing your home?"

1. Key facts on the French rental housing market and data collection

This section first presents key descriptive statistics for the rental housing market in France, before describing how the data was collected.

1.1. The rental housing market in France and in the Paris metropolitan area

This subsection presents important descriptive statistics for the rental housing market in France, with a special focus on the Paris metropolitan area.

In France, 60% of individuals are tenants. 58% of tenants live in the private sectors, while the rest have access to social housing (Laferrère et al. (2017)). In 2013, 93.5% of available dwellings in the private rental sector were owned by private landlords. Tenants are different from home owners in many ways. For instance, 80% of home owners live in houses while 75% of tenants live in apartments. Tenants in the private rental sector are more likely to be below 30 year old, single, divorced or single parents.

In 2013, 37% of tenants in France were searching for accommodation using ads on the Internet or in newspapers; 39% of them used real estate agencies; 19% relied on word of mouth (Chapelle et al. (2020)). Nowadays, it is highly likely that the proportion of tenants using online platforms to search for accommodation has increased compared to 2013. Internet usage in France increased from 68.9% in 2010 to 92.3% in December 2018[3], while the smartphone penetration rate went from 39% in 2013 to 75% in 2018.[4] In the US context, Piazzesi et al. (2020) document that over 90% of home buyers rely on the Internet in their search and 76% of them deemed real estate websites "very important" as a source of information.

Evidence indicates a shortage of accommodation in the private rental sector in the Paris metropolitan area. Since 2000, rents for the Parisian market have increased faster than the region Ile-de-France, which includes the city of Paris (Figure 3.1). Parisian rents have also increased more than the rest of France and more than household disposable income. These facts points towards a situation in which demand for renting has outpaced supply. Dietrich-Ragon (2013) discuss how it is common for one apartment advertised on the Internet to receive between 30 to 50 applicants the same day for a group visit. Applicants are usually expected to earn three times the monthly rent and must provide a thorough candidacy file. The fact that the Parisian rental market is tighter than the rest of France — from a tenant's

[3]Source: `https://www.internetworldstats.com/eu/fr.htm`

[4]Source: `https://www.statista.com/statistics/732147/smartphone-penetration-in-france/`

perspective — is also illustrated by the delay to obtain a social housing unit. While it takes on average 3 months to obtain a social housing property in a "loose" rental market, such as in the region Cantal or Creuse, the average waiting time skyrockets to 39 months for Paris. Hence, the Parisian housing rental market is to be viewed as sharing similarities with other large cities in which the supply of housing has not kept up with a surge in demand.

Fig. 3.1. Trends in rents and household disposable income in France: 2000 - 2017



Sources: Clameur and INSEE

Notes: Rent indexes calculated for privately rented dwellings by Clameur. Household disposable income is calculated by the INSEE.

1.2. Data collection

This subsection describes key features of the data collection process and important characteristics of the website from which data was collected.

The data on rents was collected in between May and July 2019 from the platform LouerAgile. LouerAgile is a website that aggregates online ads from the major actors of the rental housing market in the Paris metropolitan area. LouerAgile aggregates ads from real estate agencies (for instance, Orpi, Foncia or Logic-Immo) and ads managed and advertised directly by private landlords (for instance PAP or leboncoin). Each ad includes the rent and a detailed list of characteristics, including the number of rooms, bedrooms, the surface, the location.

Ads can be received by email or can be consulted directly on the website. Users of the

website can filter the characteristics of the ads received. I chose a very loose set of criteria to maximize the number of ads visible. I chose to receive ads with a maximum rent of 5000 euros per month, with a minimum surface of 8 square meters. I chose to include all ads from Paris. I also included a selected subset of municipalities in the close suburbs of Paris (Asnières-sur-Seine, Boulogne-Billancourt, Créteil, Ivry-sur-Seine, Montreuil, Nanterre and Saint-Denis).

In addition to the aforementioned elements, each ad also includes a textual description of the property written by landlords or by real estate agents. In the majority of instances, ads include a series of photos showing the accommodation. Potential tenants searching for a dwelling can contact landlords directly through the platform by clicking on a specific button (see Section A of the Appendix for screenshots of the website).

The website allows researchers to observe how both sides of the rental market interact. Indeed, LouerAgile has a feature that allows users the see how many individuals previously contacted the landlord/real estate agency. During the period of the web scraping, if the number of contacts received by the owner was strictly less than 10, it was not displayed.[5] When the threshold of 10 contacts was reached, the number of contacts was displayed publicly on the website.

There is an asymmetry regarding the information collected: while the information on properties is quite exhaustive, little is known regarding tenants. The only information that could be collected on them is the number of times they contacted landlords through the website.

How representative is the website? On its description page, the website claims that one third of potential tenants use its services to search for an apartment/house in Paris, which is a non-negligible fraction of potential tenants. However, I do not have access to data to verify this claim.

2. Stylized facts

This section first presents important descriptive statistics for the dataset obtained by web scraping, before analyzing the variables causing variations in rent using a linear regression model. The resulting hedonic regression model is then used to calculate the *deviation from the predicted rent*. Using the calculated deviation from the predicted rent, three empirical facts are then established: (i) dwellings that are cheaper than expected — compared to the value predicted by the estimated hedonic model — attract more applicants; (ii) approxi-

[5]Instead, the website would display a message inviting the user to contact the owner to be one of the first to do so.

mately 7% of the landlords use a two-step pricing strategy: they set a rent above what is predicted by the hedonic model, before decreasing the advertised rent; (iii) dwellings that end up being discounted are more expensive than predicted by the hedonic model; receive less applicants; are more likely to be administered by real estate agencies.

2.1. Descriptive statistics

This subsection offers key descriptive statistics for the Parisian rental housing market, using observations obtained by web scraping.

Table 3.1 displays important descriptive statistics for the variables that are used as explanatory variables in the subsequent hedonic model. In the sample, dwellings' surface range between 8 and 300 square meters[6], with a median surface of 34 square meters. The median number of rooms and bedrooms is equal to 2 and 1 respectively, while the median floor level is equal to 3. The median rent per square meter is equal to 34.5 euros.[7] Table 3.1 also indicates that 68% of ads are managed by a real estate agency; approximately 50% of dwellings are rented furnished; 41% of them have access to an elevator; less than 9% of them are endowed with a terrace or a balcony.

Table 3.2 offers a breakdown of these summary statistics by municipality. A first observation from Table 3.2 is the sample is predominantly a sample of accommodation in Paris, which represents 86% of all ads. Key facts from Table 3.2 are that, compared to the rest of the municipalities, dwellings in Paris are smaller, more expensive, more likely to be managed by real estate agencies and are more likely to be situated in a high floor.

Figure 3.2 displays some essential descriptive statistics for the Paris metropolitan area. Figure 3.2(a) indicates that the population is very dense is the two-digit arrondissements.[8] That is, the 10th to 20th arrondissements, which form the outer ring of Paris. The center of Paris, which are the one-digit arrrondissements, are much less densely populated. Figure 3.2(b) shows that the two-digit arrondissements also receive more ads on average. Yet, the two-digit arrondissements situated in the west of Paris receive more ads than what could be predicted by the distribution of population only. Figure 3.2(c) shows the well-established fact that in Paris the rent per square meter is a decreasing function of the distance to the city center (Brueckner et al. (1999)). Figure 3.2(d) indicates an opposite gradient for the median surface: as the distance to the city center increase, dwellings are larger, and

[6]Even though the French decree number 2002-120 (30th January 2002) prohibits renting a dwelling with a surface smaller than 9 square meters.

[7]As of March 2020, SeLoger — a website that collects and displays real estate ads — calculates a mean rent per square meter equal to 32 euros per square meter.

[8]An arrondissements is a subdivision of a large municipality (Paris, Lyon, and Marseille). It functions as a lower administrative division, with its own mayor.

the mean number of rooms and bedrooms increases (Figures 3.2(e) and 3.2(f)). There are weaker negative relationships between the distance to the city center and the percentage of dwellings that are furnished and managed by real estate agencies (Figures 3.2(g) and 3.2(h)).

Table 3.1: Dwelling characteristics

Characteristic	**Statistic**	**Value**	**Characteristic**	**Statistic**	**Value**
Surface (m^2)	Mean (SD)	40.9 (26.7)	**Elevator**	No	29906 (59.2%)
	Median [Min, Max]	34.0 [8.00, 300]		Yes	20622 (40.8%)
Rent/m^2	Mean (SD)	36.1 (10.1)	**Balcony/terrace**	No	46179 (91.4%)
	Median [Min, Max]	34.5 [5.62, 86.1]		Yes	4349 (8.6%)
Rental agency	No	16179 (32.0%)	**Floor**	Mean (SD)	3.14 (2.20)
	Yes	34349 (68.0%)		Median [Min, Max]	3.00 [0.00, 31.0]
# Rooms	Mean (SD)	1.93 (1.03)		Missing	11694 (23.1%)
	Median [Min, Max]	2.00 [1.00, 7.00]	**# Bedrooms**	Mean (SD)	1.51 (0.725)
Furnished	No	25446 (50.4%)		Median [Min, Max]	1.00 [1.00, 6.00]
	Yes	25082 (49.6%)		Missing	26295 (52.0%)

Source: Author's calculations based on data from LouerAgile

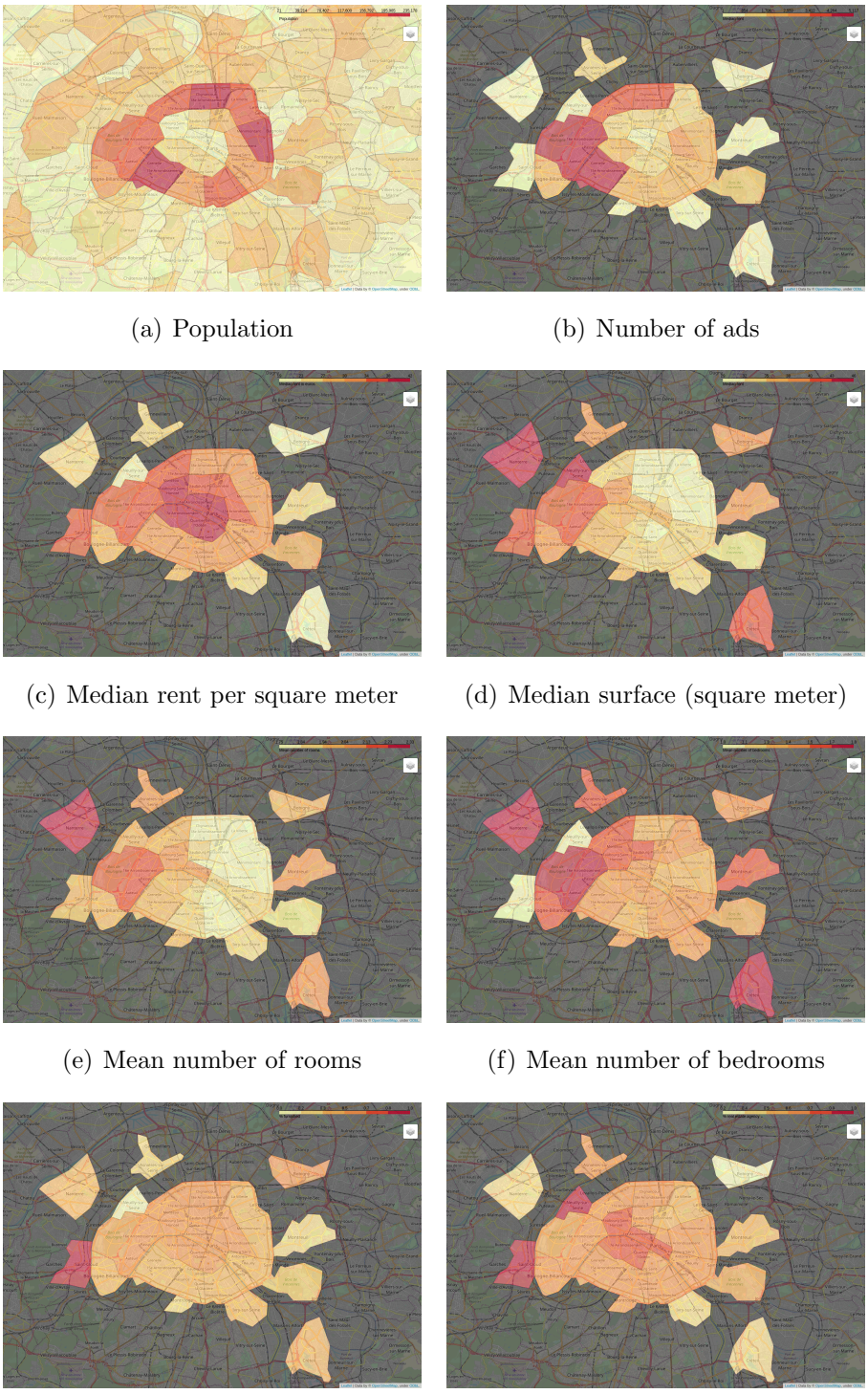
Notes: The sample includes 50528 observations. The variables "Rental agency", "Furnished", "Elevator", "Balcony/terrace" are categorical variables equal to 1 ("Yes") when the accommodation: was managed by a real estate agency; rented furnished; has access to an elevator (for an apartment); has a balcony or a terrace. Categorical variables are equal to 0 ("No") otherwise. Table 3.1 also reports the proportion of mission values for each variable (when applicable).

Table 3.2: Dwelling characteristics by municipality

		Asnieres sur Seine (n=1254)	Bobigny (n=172)	Boulogne-Billancourt (n=2094)	Creteil (n=515)	Ivry sur Seine (n=378)	Montreuil (n=831)	Montrouge (n=622)	Nanterre (n=583)	Paris (n=43435)	Saint-Denis (n=644)
Surface (m^2)	Mean (SD)	44.1 (22.9)	44.4 (26.9)	43.2 (24.3)	44.2 (23.9)	36.4 (18.6)	42.9 (23.7)	40.4 (20.0)	50.0 (35.0)	40.5 (27.0)	41.9 (20.1)
	Median [Min, Max]	40.0 [9.00, 160]	40.0 [10.0, 226]	38.0 [8.00, 209]	41.0 [8.00, 220]	33.0 [10.0, 130]	38.0 [9.00, 200]	37.0 [9.00, 145]	44.0 [9.00, 300]	33.0 [8.00, 215]	40.0 [9.00, 100]
Rent/m^2	Mean (SD)	26.1 (7.48)	19.3 (7.84)	31.6 (7.44)	22.8 (9.22)	27.6 (8.68)	25.3 (7.30)	28.1 (6.90)	23.9 (8.43)	37.6 (9.50)	23.3 (8.74)
	Median [Min, Max]	24.3 [7.81, 69.4]	18.9 [7.94, 49.4]	30.4 [8.50, 75.0]	20.9 [5.63, 55.0]	25.8 [9.62, 64.9]	23.8 [8.93, 64.7]	26.8 [10.0, 61.1]	23.0 [5.62, 65.0]	35.7 [18.8, 86.1]	21.4 [6.55, 68.9]
Rental agency	No	416 (33.2%)	130 (75.6%)	603 (28.8%)	287 (55.7%)	234 (61.9%)	329 (39.6%)	198 (31.8%)	346 (59.3%)	13258 (30.5%)	378 (58.7%)
	Yes	838 (66.8%)	42 (24.4%)	1491 (71.2%)	228 (44.3%)	144 (38.1%)	502 (60.4%)	424 (68.2%)	237 (40.7%)	30177 (69.5%)	266 (41.3%)
# Rooms	Mean (SD)	2.09 (1.04)	2.01 (0.991)	2.00 (1.02)	2.10 (1.12)	1.88 (0.809)	2.05 (1.04)	1.99 (0.926)	2.28 (1.27)	1.91 (1.03)	2.12 (0.945)
	Median [Min, Max]	2.00 [1.00, 6.00]	2.00 [1.00, 5.00]	2.00 [1.00, 7.00]	2.00 [1.00, 7.00]	2.00 [1.00, 4.00]	2.00 [1.00, 7.00]	2.00 [1.00, 5.00]	2.00 [1.00, 7.00]	2.00 [1.00, 7.00]	2.00 [1.00, 5.00]
# Bed rooms	Mean (SD)	1.60 (0.773)	1.50 (0.770)	1.54 (0.737)	1.70 (0.737)	1.42 (0.609)	1.59 (0.826)	1.50 (0.640)	1.79 (0.867)	1.49 (0.720)	1.59 (0.707)
	Median [Min, Max]	1.00 [1.00, 6.00]	1.00 [1.00, 4.00]	1.00 [1.00, 5.00]	2.00 [1.00, 4.00]	1.00 [1.00, 3.00]	1.00 [1.00, 5.00]	1.00 [1.00, 4.00]	2.00 [1.00, 5.00]	1.00 [1.00, 6.00]	1.00 [1.00, 4.00]
	Missing	594 (47.4%)	112 (65.1%)	981 (46.8%)	279 (54.2%)	233 (61.6%)	403 (48.5%)	282 (45.3%)	322 (55.2%)	22749 (52.4%)	340 (52.8%)
Furnished	No	881 (70.3%)	83 (48.3%)	1339 (63.9%)	326 (63.3%)	195 (51.6%)	601 (72.3%)	455 (73.2%)	346 (59.3%)	20827 (47.9%)	393 (61.0%)
	Yes	373 (29.7%)	89 (51.7%)	755 (36.1%)	189 (36.7%)	183 (48.4%)	230 (27.7%)	167 (26.8%)	237 (40.7%)	22608 (52.1%)	251 (39.0%)
Elevator	No	799 (63.7%)	142 (82.6%)	1089 (52.0%)	348 (67.6%)	323 (85.4%)	611 (73.5%)	362 (58.2%)	443 (76.0%)	25297 (58.2%)	492 (76.4%)
	Yes	455 (36.3%)	30 (17.4%)	1005 (48.0%)	167 (32.4%)	55 (14.6%)	220 (26.5%)	260 (41.8%)	140 (24.0%)	18138 (41.8%)	152 (23.6%)
Balcony/terrace	No	1073 (85.6%)	149 (86.6%)	1719 (82.1%)	470 (91.3%)	359 (95.0%)	684 (82.3%)	551 (88.6%)	467 (80.1%)	40127 (92.4%)	580 (90.1%)
	Yes	181 (14.4%)	23 (13.4%)	375 (17.9%)	45 (8.7%)	19 (5.0%)	147 (17.7%)	71 (11.4%)	116 (19.9%)	3308 (7.6%)	64 (9.9%)
Floor	Mean (SD)	2.68 (2.10)	1.60 (1.57)	2.91 (2.16)	2.42 (2.14)	2.24 (2.01)	1.98 (1.70)	2.77 (2.17)	1.95 (1.88)	3.21 (2.20)	2.42 (2.15)
	Median [Min, Max]	2.00 [0.00, 20.0]	2.00 [0.00, 6.00]	3.00 [0.00, 11.0]	2.00 [0.00, 11.0]	2.00 [0.00, 11.0]	2.00 [0.00, 10.0]	2.00 [0.00, 10.0]	1.00 [0.00, 10.0]	3.00 [0.00, 31.0]	2.00 [0.00, 11.0]
	Missing	380 (30.3%)	124 (72.1%)	474 (22.6%)	300 (58.3%)	200 (52.9%)	262 (31.5%)	188 (30.2%)	324 (55.6%)	9045 (20.8%)	397 (61.6%)

Source: Author's calculations based on data from LouerAgile

Notes: Table 3.2 displays a breakdown of the statistics available in Table 3.1 by municipality. The proportion of mission values for each variable by municipality is reported (when applicable).



(a) Population

(b) Number of ads

(c) Median rent per square meter

(d) Median surface (square meter)

(e) Mean number of rooms

(f) Mean number of bedrooms

(g) Percentage of dwellings furnished

(h) Percentage of ads managed by a real estate agency

Fig. 3.2. Ads characteristics at the municipality level

Notes: Panel 3.2(a) shows the population per geographical unit (municipality or arrondissement) as of January 2017. Panel 3.2(b) shows the number of ads in the sample per geographical unit. Panels 3.2(c), 3.2(d), 3.2(e) and 3.2(f) present the median rent per square meter, the median surface of the accommodation, the mean number of rooms and the mean number of bedrooms respectively. Panels 3.2(g) and 3.2(h) display the proportion of dwellings rented furnished and the proportion of apartments being managed by real estate agencies. Zoomed versions of each Panel are available in section B of the Appendix.

2.2. Hedonic pricing model

This subsection develops a hedonic pricing model that can predict rents based on observable features. In the next subsection, the variable of interest is the *deviation from the predicted rent*, not the rent itself. Because the subsequent analysis uses the residual rent, it is important to incorporate all the relevant information available to generate accurate predictions. One novelty of the present analysis compared to the existing literature is to incorporate the information contained in photos in the set of explanatory variables.

The basic idea of a hedonic pricing model, since the seminal contributions of Griliches (1961) and Rosen (1974), is that many goods differ in their attributes. Each attribute is priced by users/consumers. A typical hedonic regression model is the form:

$$\boldsymbol{y} = X\boldsymbol{\beta} + \boldsymbol{u} \tag{3.1}$$

where $\boldsymbol{y}$ is the a vector of observed prices (which may be logged), $\boldsymbol{\beta}$ a vector of coefficients associated to observable characteristics (to be estimated), and $\boldsymbol{u}$ a vector of error terms. I follow the literature and use a similar regression framework.

One innovation compared to the existing literature on hedonic models for the real estate market is that I include the information contained in photos to predict rents. I do this in a way that allows for the results to be easily interpretable. Previous studies have used outdoors image, especially images from Google Street View to predict house prices. For instance Zhang and Dong (2018) calculate a *street greenery index* in Beijing and show that the presence of visible street greenery can raise property prices by almost 10%. To the best of my knowledge, only two papers in the real estate literature have used indoor images in a hedonic regression setting. Ahmed and Moustafa (2016) use both indoor to outdoor images to estimate a hedonic model. Their results are difficult to interpret because they extract n important features using an unsupervised feature detector, and use the important features as explanatory variables.[9] In what follows, a two-step approach is used: photos are encoded to produce an interpretable aesthetic score; this score is used as an additional explanatory variable. A similar approach has been used by Poursaeed et al. (2018), who encode images into a "luxury" index. The authors find that more luxurious properties sell at a higher price, and that the predictive power of a hedonic regression model is improved by including the information contained in images.[10]

To transform photos into a numerical value that can be used as an input into a regression

[9] Another limitation is that the paper uses a limited sample of 535 house transactions from the state of California.

[10] For a review on how images can be used for research on the real estate market, see for instance Koch et al. (2019).

model, I use a convolutional neural network (CNN) based on the contribution of Talebi and Milanfar (2018). A CNN is a type of neural network especially tailored to process images (for a review of CNN, see for instance O'Shea and Nash (2015)). The CNN in Talebi and Milanfar (2018) was designed and trained to predict the distribution of human opinion scores when judging the aesthetic qualities of an image.[11] Scores are normalized to be between 1 and 10, with 10 being the highest aesthetic score associated to an image. Each ad is assigned an aesthetic score using a two-step approach, taking into consideration that one ad may have several photos. In a fist step, each photo was assigned an aesthetic score using a CNN based on the work of Talebi and Milanfar (2018).[12]. In a second step, each ad was assigned a global aesthetic score, calculated as the median score for its photos. Figure 3.3(a) indicates that the vast majority of ads featured at least one photo. Figure 3.3(b) shows the distribution of aesthetic score for each ad. Aesthetic score ranges between approximately 3 and 6, while the median score is close to 5. Figures 3.4 and 3.5 show a selected sample of photos from the top 1% and bottom 1% of the aesthetic score distribution respectively.

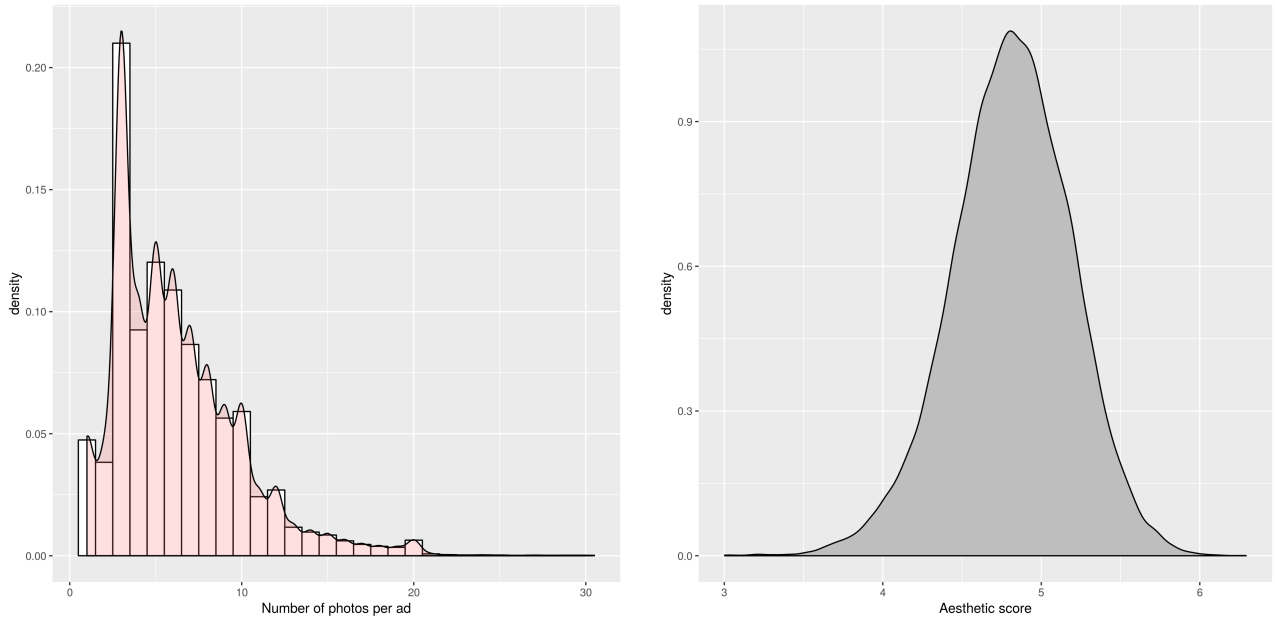
Table 3.3 presents the results of regressing advertised rents on a set of observable factors. The first column indicates that the monthly rent (in euros) is an increasing but concave function of surface (square meters). Dwellings having access to an elevator and/or rented furnished are also pricier. The second column adds the aesthetic score, as captured by the CNN using photos. Apartments/houses that score higher on the aesthetic scale are more expensive. According to the second column, a one standard deviation increase in the aesthetic score raises the monthly rent by approximately 47.8 euros. By comparison, the second column suggests that adding an elevator to the property raises the monthly rent by approximately 63 euros. The second column of Table 3.3 indicates that the CNN is successful in capturing aesthetic elements that are important for potential renters. For instance, it is likely that the CNN us successful in capturing the style and the quality of furniture, which matters for renters. The CNN may also capture other features, such as the luminosity of properties.

The third column adds postal code fixed effects, which increases the R-squared to 0.89. Once postal code fixed effects are included, an accommodation with balcony/terrace is identified as more expensive. The fourth column adds a real estate agency dummy variable. The associated coefficient is positive and statistically significant, indicating that real estate

[11]Their work is based on three datasets of photos that were rated by humans. For instance, they use the AVA dataset, which contains about 255,000 images obtained from the online community of amateur photographers: dpchallenge. Each photo in the sample is scored by an average of 200 people, in response to photography contests.

[12]The CNN in Talebi and Milanfar (2018) is designed to reproduce the *distribution* of aesthetic scores that a jury of human could produce when looking at a photo. To summarize the information contained in the distribution, I use the mean aesthetic score.

agency tend to set a higher rent compared to private landlords. The fifth column adds floor fixed effects, to control for the possibility that the floor level impacts prices. The last column adds a variable for the number of bedrooms. This column indicates that, keeping the surface constant, an apartment with more rooms and bedrooms will be rented at a higher price.



(a) Number of photos per ads

(b) Density of aesthetic scores

Fig. 3.3. Number of photos per ads and distribution of aesthetic scores

Sources: Author's calculations based on data from LouerAgile

Notes: Figure 3.3(b) shows the distribution of aesthetic score attributed to each ad, based on its photos. In a fist step, each photo was assigned an aesthetic score using a CNN based on the work of Talebi and Milanfar (2018). In a second step, each ad was assigned an aesthetic score equal to the median score for its photos. Ads with no photos are not included in Figure 3.3(b).

Fig. 3.4. Selected sample of photos from the top 1% in terms of aesthetic score



Source: Photos from LouerAgile.

Notes: Photos were assigned an aesthetic score using a CNN based on the work of Talebi and Milanfar (2018). Figure 3.4 displays a selected sample from the top 1% in terms of aesthetic score.

Fig. 3.5. Selected sample of photos from the bottom 1% in terms of aesthetic score



Source: Photos from LouerAgile.

Notes: Photos were assigned an aesthetic score using a CNN based on the work of Talebi and Milanfar (2018). Figure 3.4 displays a selected sample from the bottom 1% in terms of aesthetic score.

Table 3.3: Hedonic regression model

	Dependent variable:					
	Monthly rent (in euros)					
	(1)	(2)	(3)	(4)	(5)	(6)
Surface (m^2)	28.502***	28.061***	27.852***	27.783***	28.307***	28.385***
	(0.211)	(0.226)	(0.189)	(0.190)	(0.215)	(0.355)
Surface^2	−0.010***	−0.009***	−0.017***	−0.017***	−0.017***	−0.019***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)
Nb rooms	−3.767	2.803	26.291***	26.506***	24.966***	29.639***
	(3.144)	(3.343)	(2.791)	(2.792)	(3.133)	(6.749)
Elevator	68.365***	62.916***	32.140***	31.320***	34.020***	32.803***
	(3.110)	(3.315)	(2.798)	(2.808)	(3.243)	(5.163)
Furnished	210.312***	190.459***	147.181***	149.517***	161.718***	217.045***
	(2.961)	(3.207)	(2.692)	(2.779)	(3.080)	(4.880)
Balcony/terrace	−37.152***	−36.047***	25.838***	24.975***	15.967***	28.651***
	(5.275)	(5.512)	(4.617)	(4.624)	(5.053)	(7.206)
Aesthetic score		126.693***	105.884***	106.228***	110.645***	141.676***
		(4.163)	(3.466)	(3.467)	(3.920)	(6.364)
Rental agency				10.179***	5.021	16.532**
				(3.017)	(3.520)	(7.454)
Nb bedrooms						20.916***
						(7.175)
Constant	85.299***	−503.223***	−119.910***	−128.505***	−175.648***	−354.851***
	(4.845)	(19.952)	(20.104)	(20.263)	(23.024)	(38.700)
Postal code FE	No	No	Yes	Yes	Yes	Yes
Floor FE	No	No	No	No	Yes	Yes
Observations	50,528	42,307	42,307	42,307	33,180	17,439
R^2	0.831	0.840	0.890	0.890	0.899	0.882
Adjusted R^2	0.831	0.840	0.890	0.890	0.899	0.881
Residual Std. Error	323.499 (df = 50521)	315.749 (df = 42299)	262.149 (df = 42270)	262.116 (df = 42269)	258.405 (df = 33128)	299.566 (df = 17386)
F Statistic	41,393.470*** (df = 6; 50521)	31,775.790*** (df = 7; 42299)	9,493.990*** (df = 36; 42270)	9,239.973*** (df = 37; 42269)	5,792.983*** (df = 51; 33128)	2,494.237*** (df = 52; 17386)
Note:						*p<0.1; **p<0.05; ***p<0.01

Source: Author's calculations based on data from LouerAgile.
Notes: This table shows the point estimates for the regression model 3.1.

2.3. Residual rent and search behavior

This subsection shows that the *residual rent* — defined as the predicted rent minus the observed advertised rent — is correlated with tenants' search intensity. As predicted by a standard directed search model, apartments/houses that are cheaper than their predicted price attract more potential tenants.

Figure 3.6 plots the *residual rent* — as defined above — for each ad, against the number of contacts received by each ad through the platform. As previously explained, the number of contacts received through the platform was truncated during the data collection phase. This is why no ads with less than 10 contacts are displayed. As predicted by a standard directed search model, Figure 3.6 suggests a positive relationship between cheaper than expected accommodation (with a positive residual rent) and the number of contacts received.

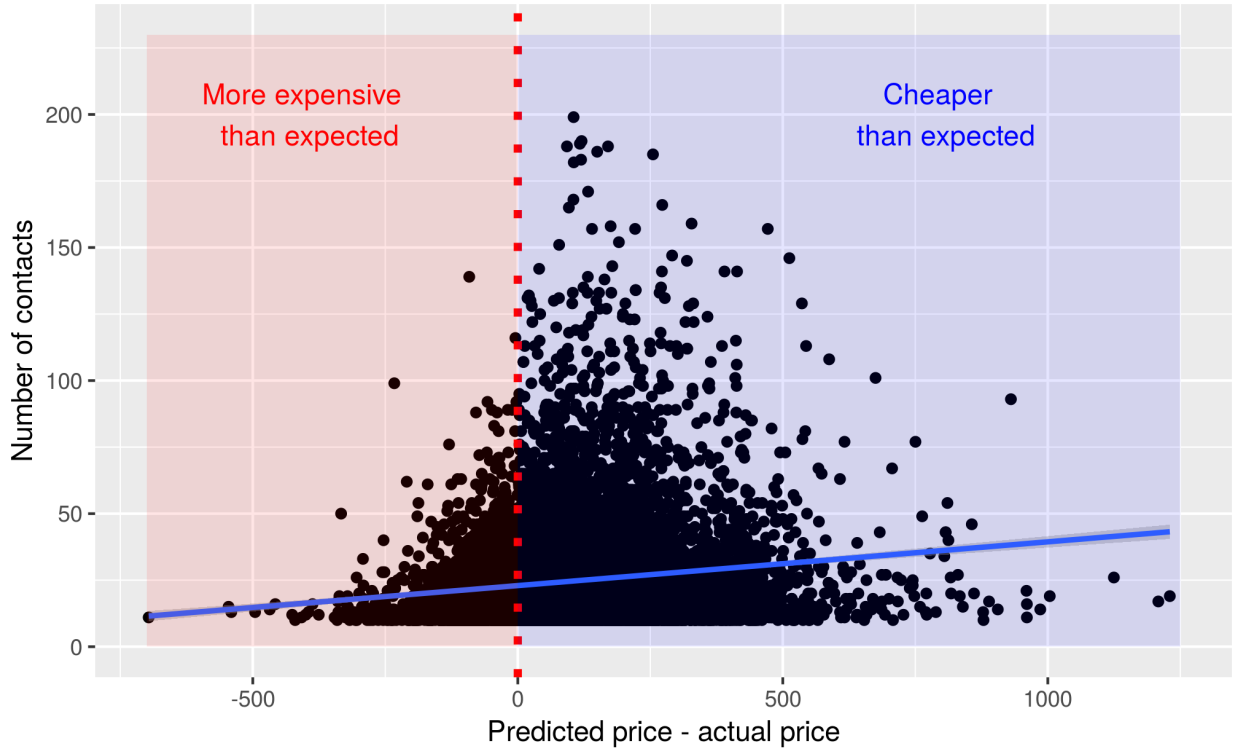
To precisely estimate the link between the residual rent and the number of contacts, taking into account the truncation threshold at 10 contacts, I use a truncated regression framework.[13] Formally, I estimate the following model:

$$\begin{aligned} Y_i &= X_i\boldsymbol{\beta} + \varepsilon_i < L, \quad \text{excluded} \\ Y_i &= X_i\boldsymbol{\beta} + \varepsilon_i \geq L, \quad \text{included} \end{aligned} \tag{3.2}$$

with Y_i the log of contacts received through the website, $L = 10$ and maintaining the assumption that, given X_i, Y_i is normally distributed distributed with mean $X_i\boldsymbol{\beta}$ and variance σ^2. Because the model is in log-linear form, Table 3.4 indicates that when a landlord reduces its advertised rent by 1 euro (compared to the predicted rent), she or he can expect the number contacts received through the website to increase by approximately 1.8%.

[13] It is well known that using OLS in that context results in a truncation bias (Hausman and Wise (1977)). In practice, I use R and the package *truncreg* (Croissant and Zeileis (2016)).

Fig. 3.6. Deviation from predicted rent and number of contacts



Source: Author's calculations based on data from LouerAgile

Notes: The x-axis is the *residual rent dispersion*. That is, the rent predicted by observable features according to the hedonic model estimated in Table 3.3, minus the actual observed rent. The y-axis is the total number of contacts received by the landlord for his/her ad through the website LouerAgile. The number of contacts is truncated. It is not observed for values strictly smaller than 10. The blue line is the naive OLS predicted line (not taking into account the truncation of the y variable).

Table 3.4: Truncated regression

	Dependent variable:
	log number of contacts
Constant	1.56601***
	(0.09016)
Predicted - actual rent	0.001770***
	(0.00014)
σ	1.095269***
	(0.027501)
Log-Likelihood	-6369.5 on 3 Df
Note:	*p<0.1; **p<0.05; ***p<0.01

Sources: Author's calculations based on data from LouerAgile
Notes: This table shows the outcome of estimating a truncated regression for the logarithm of contacts received by landlords through the online platform. A truncated regression is needed because the number of contacts was displayed only when the number was greater or equal to 10.

2.4. Pricing dynamics

This subsection offers new insights on landlords' pricing strategy. A non-negligible proportion of them first advertise a high rent before updating the rent to a value close to the one predicted by the previously estimated hedonic model.

Observing price changes is made possible by a feature of the website LouerAgile. The website sends emails to users to notify them when the advertised price for an accommodation experienced a decreased.[14] The email notifying a rent decrease contains the previous rent

[14]One may wonder why I rely on this functionality rather than building a panel by repeatedly observing the same units across time. While it is theoretically possible to scrape the website on a regular basis, screening the same ads several time to build a panel, in practice it is complicated. My initial strategy was to scrape all ads (already in the sample and newly posted) on a regular basis. This strategy defined de-facto an "extending window" strategy. However, during the collect the website changed layout, creating a discontinuity between recently posted and "old" ads. I updated the web scraping script to collect only the ads not already in the sample.

and the updated advertised rent, as well as a link to the full description of the property.

Several key facts on advertised rent discounting can be established. Firstly, approximately 7% of the ads in the sample experienced a decrease in the advertised rent. Secondly, the median percentage change in the advertised rent is approximately equal to 6.25% (Figure 3.7(a)). Thirdly, the median number of days before an advertised rent decrease is equal 27 days (Figure 3.7(b)).

I investigate how discounted ads differ from no-discounted ads using a three Welch's t-tests. Table 3.5 indicates that ads which were discounted were (reading the table from left to right): (i) more expensive than the rest of ads in the first place (the deviation from their predicted price was much more positive than for the control group); (ii) received less contacts than the rest of ads; (iii) were more likely to be managed by a real estate agency rather than being privately managed. Points (i) and (ii) make perfect sense in light of the results from the previous section. Ads that are more expensive than expected attract less potential tenants.

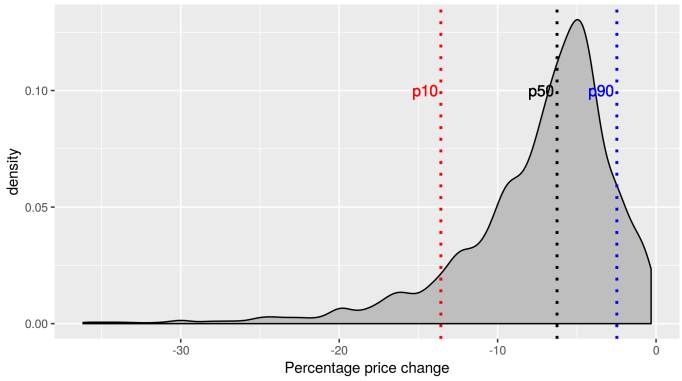
Point (iii) deserves a discussion. On the one hand, real estate agencies in France have an incentive to set a high rents for a least two reasons. First, they charge landlords a tenant-finding fee, which is usually equal to a fixed fraction of the monthly rent. Secondly, they charge landlords monthly management fees, which are also indexed on the rent [15]. Prior September 2014, real estate agencies also had an incentive to set a high rent because they charged tenants a fee for establishing a tenancy agreement, which usually was equal to multiple of the monthly rent.[16] On the other hand, real estate agencies have an incentive to set rents to a low level, because it increases the likelihood that they will find a tenant next period. Overall, the incentive to increase rents for real estate agencies (compared to accommodation with similar characteristics rented by private landlords) seems to dominate.

How can we interpret the fact that two-step pricing mechanism is used in priority by real estate agencies? At least two stories come to mind: one of learning, one of rent extraction. It may be the case that agents (private landlords or real estate agencies) do not know the valuation of their property. As a result, they use a descending auction-like strategy by first setting a high rent, and updating it when observing a low arrival rate of potential tenants. Because real estate agencies are professionals, they maybe more likely to react to market conditions and more willing to undergo a rent cut when few tenants are applying.

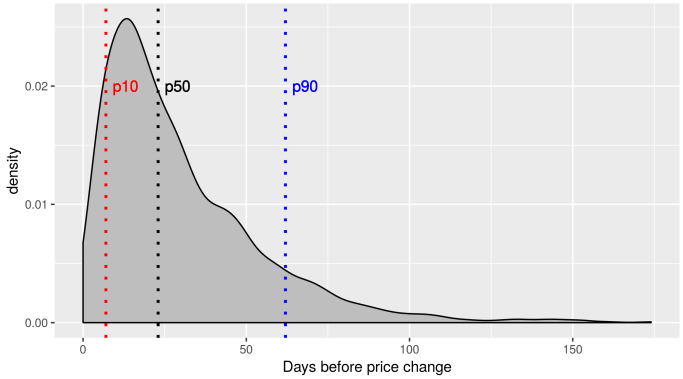
[15]Source: www.smartloc.fr, based on an analysis of the pricing practices of 4 major French real estate agencies as of January 2019.

[16]Since the law ALUR, entered into force on the 15th of September 2014, the costs of establishing a tenancy agreement are defined by law. In a "very tight" zone ("Zone très tendue"), which includes Paris, costs are equal to 12 € per square meter. In a "tight" zone (Zone tendu), which includes Lyon, Bordeaux and Toulouse for instance, costs are equal to 10 € per square meter. In a non tight zone, tenants pay 8 € per square meter.

However, it seems surprising that real estate agencies do not know what constitutes a "fair price" in the first place. Compared to private landlords, they certainly have an information edge because they manage several properties and have access to years of proprietary data. An alternative explanation is that because real estate agencies have an incentive to set a high rent, as discussed above, they are more likely to use the strategy that maximizes the seller's (landlord's) profit. It might be the case that the optimal seller's strategy is to use a descending auction-like strategy.



(a) Density of the percentage rent reduction advertised on the platform



(b) Density of days spent before rent reduction

Fig. 3.7. Advertised rent reductions and number of days before decline

Sources: INSEE and author's calculations based on data from LouerAgile
Notes: Panel 3.7(a) displays the density of advertised rent percentage change. Panel 3.7(b) displays the density of days before the rent change. The vertical red, black and blue line represent the 10th, 50th and 90th percentiles respectively.

Table 3.5: Welch t-tests

	predicted price - actual price	number of contacts per ad	real estate agency
t-statistic	27.136	3.255	−26.643
DF	3,975.673	4,109.786	4,592.625
p-value	$< 2.2e-16$	0.001	$< 2.2e-16$
mean in group of non-discounters	2.011	5.084	0.672
mean in group of discounters	−153.104	4.277	0.843

Source: Author's calculations based on data from LouerAgile

Notes: This table shows the results of three Welch's t-tests. A Welch's t-test assesses the hypothesis that two populations have equal means. It compares the mean for the population of ads which experienced a rent discount, with the mean for the rest of ads. The first column compares the deviation from predicted rents (using a hedonic regression model); the second column compares the number of contacts received by each ad; the last column compares the proportion of ads advertised through a real estate agency.

3. Model

The empirical analysis underlined two main facts for the rental housing market in the Paris metropolitan area. Firstly, apartments/houses more expensive than expected (controlling for observable features) attract less apartment searchers. Secondly, some landlords use a dynamic pricing strategy by first advertising a high rent, before decreasing the rent to a level in line with what is predicted by observable features. This section describes a directed search model that rationalizes the first finding. I also discuss the possible efficiency and welfare implications of the descending auction-like rent setting observed in the data.

Before laying out the model, two important questions must be answered: are the rents known by market participants and are they negotiated? The answer of the first question is obviously yes, as rents are publicly displayed on the online platform and easily accessible by any user. The answer to the second question is less straightforward, but evidence indicate that rents advertised on the online rental market are in most of the cases non-negotiated by either tenants or landlords before signing the tenancy agreement. Baietto-Beysson and Vorms (2012) report that rent negotiation during the signature of the first tenancy agreement is deemed "very unlikely" by French real estate agents.[17] On a more quantitative basis,

[17]"[Le] candidat à la location est en position de négocier avec le bailleur? Les professionnels estiment

Chapelle et al. (2020) show that observations on rents obtained by scraping French online ads nearly perfectly correlates with rent indices built using official rental lease contracts. This further indicates that rent negotiation during the signature of the first tenancy agreement is an unusual event, at least in the French context. Rent negotiations in ongoing contract are probably more common, especially if the rent is found to violate the Parisian rent control legislation, entered into force on July 2019 (see the section C of the Appendix). To make the model tractable, I ignore both the possibility of rent negotiations that may take place at the renewal of a tenancy agreement and the existence of rent control legislation. Ignoring the presence of the Parisian rent control legislation amounts to assuming that the upper bound of rents is not binding: landlords are not allowed to set a rent above 20 % of the median rent, calculated using a sample of similar properties in the same neighborhood.

To model the rental market, I use the competitive search equilibrium of Moen (1997). His theoretical framework is a model of the labor market, in which firms publicly announce wages and commit to them. Workers observe wages and direct their search efforts towards submarkets accordingly. The model rationalizes why job vacancies with higher wages (controlling for the "quality" of jobs) attract more applicants.

In what follows, I adapt his theoretical framework to match the stylized of the rental market previously established. One major limitation of the present model is that it does not take into consideration the "wait-and-see" pricing behavior used by some landlords. I subsequently discuss how the model could be extended to accommodate this dynamic pricing behavior.

3.1. Setting

Assume the existence of n submarkets ordered in increasing order according to their rent $\{r_1, r_2, ..., r_n\}$. Rents are public knowledge for both sides of the market and non-negotiable. The model is in continuous time and the discount factor is r. The rental market is populated by a continuum of infinitely-lived and risk-neutral tenants and landlords. The measure of tenants is normalized to one, while the measure of vacant apartments is pinned down by a free-entry condition. Landlords pay a sunk cost $k \geq 0$ when putting their vacant property on the rental market and they pay a flow cost c to maintain a vacancy. While properties are ex-ante identical, they differ ex-post. After putting their property on the rental market, landlords observe their value y_i, which is drawn from a probability distribution F with mass points $\{y_1, y_2, ..., y_m\}$. At any point in time, there is a measure of apartment searchers

que c'est très rare. Cela peut cependant se produire dans les marchés détendus, alors que dans les marchés tendus, où le bailleur est en mesure de choisir son locataire, le loyer pratiqué est presque toujours identique à celui de l'offre." (Baietto-Beysson and Vorms (2012), page 26)

denoted by h and a measure of vacant apartments denoted by v. Search frictions coming from the cost of filling and processing applications for both sides of the market, or information imperfections about properties, prevent the rental market from perfectly clearing. One important assumption is that there there is no information imperfection on rents. Rents are fully observed, and apartments searchers decide in which submarket to search. The flow of new tenancy agreements at any point in time is given by a matching function $m(h, v)$, assumed to be concave and homogeneous of degree one in (h, v), with continuous derivatives. The rate at which apartment searchers find an apartment and sign a tenancy agreement is given by $p\big(\theta\big) \equiv \frac{m(h,v)}{u} = m(1, \theta)$, where θ is equal to the ratio of the measure of vacant apartments v to the measure of apartment searchers h. The rate at which a landlords find a tenant for a vacant apartment is given $q\big(\theta\big) \equiv \frac{m(h,v)}{v} = m(\theta, 1)$. As it is standard in the search literature, let us assume that

$$\begin{aligned} \lim_{\theta \to 0} p\big(\theta\big) = \lim_{\theta \to +\infty} q\big(\theta\big) = 0 \\ \lim_{\theta \to +\infty} p\big(\theta\big) = \lim_{\theta \to 0} q\big(\theta\big) = +\infty \end{aligned} \tag{3.3}$$

The conditions listed in (3.3) indicate that when there is infinitely more vacant apartments than apartment searchers (when the market tightness goes to 0), the arrival rate of potential tenants to a vacant apartment goes to zero. Conversely, when there is infinitely more apartment searchers than vacant apartments, the apartment finding rate goes to 0.

A formal definition of a "submarket" is required. In what follows, a submarket i is defined as a measure of apartment seekers h_i and a measure of vacant apartments v_i searching for each others. Together, both sides of the submarket i define a market tightness $\theta_i \equiv \frac{v_i}{h_i}$. Apartment searchers and firms are free to move between submarkets, but it is assumed that tenants cannot search in two or more submarkets at the same time and that landlords cannot advertise the same dwelling across different submarkets.

3.2. The value of searching for tenants

The value of being an apartment searcher in the submarket i, denoted by H_i, writes:

$$rH_i = z + p\big(\theta_i\big)\big(T_i - H_i\big) \tag{3.4}$$

where z is the utility of being without an apartment while searching and θ_i is the market tightness in the submarket i. One could think of z as the value of living on your friend's couch or staying at the hotel, while searching for a tenancy agreement.

Let T_i denote the value of being a tenant in the submarket i. Let us assume that tenants have to move out from their current apartment at a rate s. For instance, the tenant-landlord relationship may break because their contract expires, or because the landlords may want want to sell or refurbish their apartment. The value of being a tenant in submarket i, denoted by T_i, writes:

$$rT_i = t_i - r_i - s\big(T_i - H_i\big) \tag{3.5}$$

where t_i denotes the flow utility of having a tenancy agreement in the submarket i, while paying a rent r_i. Combining equations (3.4) and (3.5) yields the following expression for the value of being an apartment seeker in the submarket i:

$$rH_i = \frac{z\big(r+s\big) + p\big(\theta_i\big)\big(t_i - r_i\big)}{r + s + p\big(\theta_i\big)} \tag{3.6}$$

for $t_i - r_i \geq z$. If $t_i - r_i < z$, agents do not accept the tenancy agreement and their expected value is $\frac{z}{r}$. Let $\tilde{t}_i = t_i - r_i$ denote the utility flow of being a tenant in submarket i, net of the rent. Let us assume that landlords set directly $\tilde{t}_i$. This assumption is without loss of generality if the value of t_i is common knowledge and landlords set rents taking t_i into consideration.

In which submarket agents should be searching for an apartment? Because agents are assumed to be identical, they should be indifferent between choosing any active submarket. That is, any submarket attracting potential tenants should yield exactly the same utility level, denoted by H. For any active submarket, equation (3.6) can be expressed as:

$$p\big(\theta_i\big) = \frac{rH - z}{\tilde{t}_i - rH}\big(r+s\big) \tag{3.7}$$

For a fixed value of H, Equation (3.7) uniquely defines a relationship between the tenant's net flow of utility $\tilde{t}_i$ and the labor market tightness. The market tightness in submarket i, denoted by $\theta(\tilde{t}_i; H)$, is continuous on $(rH, +\infty)$ and strictly decreasing in $\tilde{t}_i$. When the value of tenancy agreement is high from the tenant's perspective (measured by $\tilde{t}_i$), the tenant is willing to to wait longer before finding a proper apartment.

Limit cases are important to consider. If the net flow value of the tenancy agreement tends to rH, the apartment finding rate goes to $+\infty$. That is, searchers ending up in a submarket with high rents (low $\tilde{t}_i$) are compensated by a high apartment finding rate. A submarket with $\tilde{t}_i$ strictly less than rH does not attract any searcher, as the outside option of not renting is strictly preferred. When $\tilde{t}_i$ tends to $+\infty$, the labor market tightness θ goes to zero and the arrival rate of searchers to a vacant apartment goes to $+\infty$. That is,

landlords setting lower rents are compensated by a higher arrival rent of potential tenants, denoted by $q(\theta)$.

3.3. The value of renting for landlords

Let $V(y_i, r, \theta)$ denote the expected discounted value of a vacant apartment with value y_i, in a submarket with rent level r and market tightness θ. Let $J(y_i, \theta)$ denote the value of the same apartment when occupied by a tenant. The Bellman equation for $V(.)$ solves:

$$rV(y_i, r, \theta) = -c + q\big(\theta\big)\big(J\big(y_i, r\big)\big) - V(y_i, r, \theta)\big) \tag{3.8}$$

Equation (3.8) takes into consideration the flow cost of maintaining a vacancy c and the fact that the apartment may find a tenant with probability $q\big(\theta\big)$. The Bellman equation for the value of an occupied apartment solves:

$$rJ(y_i, r) = r_i - y_i - s\big(J\big(y_i, r\big)\big) \tag{3.9}$$

where equation (3.9) takes into consideration that the tenancy agreement is broken with probability s, in which case the vacancy is destroyed. An occupied apartment produces a flow value equal to the rent r_i minus the intrinsic value of the apartment (the value the landlord could get buy living in his or her apartment). Combining equations (3.9) and (3.8) yields a new expression for $V(.)$:

$$\big(r + q\big(\theta\big)\big)V(y_i, r, \theta) = q\big(\theta\big)\frac{r_i - y_i}{r + s} - c \tag{3.10}$$

Each landlord decides to enter the submarket i that maximizes the value of V, which can be expressed as choosing a rent r:

$$\max_r V\big(y_i, r, \theta\big(r; H\big)\big) \tag{3.11}$$

As it is common in the literature, I make the assumption that submarkets in which no trade occurs are shut down by a market maker. Taking into account the randomness on the intrinsic value of an apartment y_i, the expected value of vacant apartment writes:

$$\bar{V}\big(H\big) = \sum_{i=1}^{\tilde{i}} \Pr(y = y_i) \max_r V\big(y_i, r, \theta\big(r; H\big)\big) \tag{3.12}$$

Only vacancies that have positive asset values are maintained. Let $\tilde{i}$ denotes the highest accommodation type that can be maintained in equilibrium. Above this threshold, landlords

are not compensated enough to rent on the private sector and they withdraw their vacancy from the rental market.[18]

The expected value of opening a vacancy $\bar{V}\big(H\big)$ is decreasing in the outside value of tenants H. If tenants are better off not renting, the value of rents must fall to attract them. As a result, the expect value of opening a vacancy falls as well. Because there are no barriers to entry, landlords create vacancies until the expected value of doing so is equal to the creation cost k:

$$\bar{V}\big(H\big) = k \tag{3.13}$$

3.4. Competitive search equilibrium

As in Moen (1997), a competitive search equilibrium is a value H, a set of rent $\{r_i\}_{i=1}^{n}$, a set of market tightness $\{\theta_i\}_{i=1}^{n}$ jointly satisfying the following set of equations:

$$\bar{V}\big(H\big) = k \tag{3.14}$$

$$r_i = \arg\max_{r} V\big(y_i, r, \theta\big(r; H\big)\big), \quad i < \tilde{i} \tag{3.15}$$

$$rH = \frac{z\big(r+s\big) + p\big(\theta_i\big)\big(t_i - r_i\big)}{r + s + p\big(\theta_i\big)}, \quad i < \tilde{i} \tag{3.16}$$

$$h_i p\big(\theta_i\big) = \tilde{f}_i(1 - h_i)s, \quad i < \tilde{i} \tag{3.17}$$

$$h = \sum_{i=1}^{\tilde{i}} h_i \tag{3.18}$$

where $\tilde{f}_i \equiv \frac{\Pr(y=y_i)}{\Pr(y \geq \tilde{y}_i) - 1}$ is the truncated distribution (from above) of vacancy types.

As discussed in Moen (1997), at least one equilibrium exists. If there are multiple equilibria, they all yield the same value H to tenants. All equilibria satisfying conditions (3.14) to (3.18) are optimal, in the sense that they satisfy the Hosios (1990) condition. All search externalities are internalized by participants and their responses is optimal given the search frictions.

The value H is determined by the first equation. Given the value for H, the set of rents is determined by the second equation. Once the value for r_i are determined, a submarket tightness θ_i is pinned down by the third equation. The measure of apartment searchers within each submarket, denoted by h_i, follows from the fourth equation. The final equation

[18] In January 2015, 8.1% of dwellings were vacant in Paris (Boussad and Wolber (2018)). This proportion includes properties currently being sold or on the rental market, as well as properties on the short-term rental (for instance on Airbnb) and properties that are secondary residence.

sums up over the different submarkets to calculate the total measure of apartment searchers, denoted by h.

The key insight from the directed search model presented in this section is that, holding constant the quality/characteristics of accommodation, landlords face a trade-off between renting at a high price but waiting for a longer period of time before finding a tenant. Similarly, tenants face a trade-of between finding a cheaper dwelling, but having to search for a longer period of time before finding it. When the landlord's and tenant's indifference curves meet, an equilibrium is reached.

3.5. Advertised rent decreases: efficiency and welfare considerations

The model presented in the previous section made abstraction of the pricing strategy used by some of the landlords. This subsection considers what the rent-setting strategy uncovered in the empirical section could imply from a welfare and efficiency perspective.

A traditional result from the search literature is that in a market with search frictions, if (i) sellers post their prices in advance (with commitment) (ii) buyers can observe prices and search in the submarkets of their choice, the resulting decentralized search equilibrium is constrained efficient[19] (see Moen (1997) and Shimer (1996)). The fact that the combination of (i) and (ii) results in a constrained efficient outcome can be extended to many settings. Acemoglu and Shimer (1999) show that this efficiency result also holds if buyers have only observed a subset of posted prices. Menzio and Shi (2011) demonstrate that efficiency also holds in a model of directed search with aggregate productivity uncertainty. Schaal (2017) builds a directed search model with both idiosyncratic and aggregate uncertainty and demonstrates that the market outcome is also constrained efficiency. Hence, the search literature indicates that the combination of price posting on the seller's side (with commitment) and directed search on the buyer's side (the possibility to search in selected submarkets) usually leads to an efficient equilibrium when search frictions exist. It might be the case that the landlords' pricing strategy uncovered is the present paper is a sign of some sort of inefficiency on the housing rental market.

In a different strand of the literature, papers studying the *durapolist* problem provide us with important insights regarding the problem at hand. A durapolist is a monopolist selling a non-perishable good. The Coase's conjecture states that a monopolist selling a non-perishable good charges the competitive price, because the monopolist is de facto competing with its future self (Coase (1972)). Consumers internalize this information: if facing a price above the competitive price, they would simply wait for the next period. While the rental

[19] Constrained efficient in the sense that, taking market frictions as given, the decentralized market outcome and the central planner's solution are the same.

market is clearly not a durapoly (landlords do face competitors to attract tenants), some of Coasian forces may still apply. Fuchs and Skrzypacz (2010) show that when a seller faces with positive probability the arrival of a competitor (or of another buyer), some inefficient delays arises: the seller prefers to "wait-and-see", before decreasing her price. The delay we observe for some landlords may well be inefficient (they would prefer to trade without delay), but it is the best response to the uncertainty they face. While the analysis of Fuchs and Skrzypacz (2010) illustrates well the trade-offs that landlords face, it is a partial equilibrium analysis: the arrival rate of other sellers/buyers does not depend on the posted price. In a related contribution, Sandholm and Gilpin (2003) study an auction in which sellers make a series of take-it-or-leave-it offers to individual buyers in an pre-ordered fashion. They show that this type of auction is close to reaching the seller's maximum utility.

An interesting line of research would consist in combining the insights from the literature emphasizing the contractual environment, with the results from the search literature, which generally focuses on search frictions.

4. Conclusion

This paper offers new empirical evidence on how functions the rental housing market. Using a novel dataset of online ads from the Paris metropolitan area, which combines information on both property's characteristics and the tenants' search intensity, two new empirical facts are established. Firstly, as predicted by a standard directed search model, dwellings that are cheaper than predicted by their observable characteristics attract less searchers. Secondly, Approximately 7% of landlords use descending auction-like strategy when setting rents. I discuss the potential efficiency and welfare consequences of the landlord's strategy. This paper also contributes to the growing literature that uses photos to build more precise hedonic models in the real estate context.

The pricing mechanism used by a non-negligible fraction of landlords resembles some pricing mechanisms already studied in the theoretical literature. However, to the best of my knowledge, there is no theoretical exploration of a directed search model, with take-it-or-leave prices that may be updated as the seller learns more about the valuation of her property or about the market fundamentals. Given the importance of the housing market for the rest of the economy, such an investigation is essential. A natural follow-up of the current paper would consist in a theoretical exploration of a model with the aforementioned characteristics.

References

Acemoglu, D. and Shimer, R. (1999). Holdups and efficiency with search frictions. *International Economic Review*, 40(4):827–849.

Adler, D. and Ansell, B. (2020). Housing and populism. *West European Politics*, 43(2):344–365.

Ahmed, E. and Moustafa, M. (2016). House price estimation from visual and textual features. *arXiv preprint arXiv:1609.08399*.

Baietto-Beysson, S. and Vorms, B. (2012). Les observatoires des loyers. *Paris: La Documentation française, Agence nationale pour l'information sur le logement, Observatoire des loyers de l'agglomération parisienne, rapport pour le ministère de l'Égalité des territoires et du logement*.

Boeing, G. and Waddell, P. (2017). New insights into rental housing markets across the united states: Web scraping and analyzing craigslist rental listings. *Journal of Planning Education and Research*, 37(4):457–476.

Boussad, N. and Wolber, O. (2018). 2010-2015: en île-de-france, les tensions sur le logement freinent la décohabitation.

Brown, J. and Matsa, D. A. (2019). Locked in by leverage: Job search during the housing crisis. *Journal of Financial Economics*.

Brueckner, J. K., Thisse, J.-F., and Zenou, Y. (1999). Why is central paris rich and downtown detroit poor?: An amenity-based theory. *European economic review*, 43(1):91–107.

Chapelle, G., Wasmer, E., and Bono, P. H. (2020). An urban labor market with frictional housing markets: Theory and an application to the paris urban area. *Journal of Economic Geography*.

Coase, R. H. (1972). Durability and monopoly. *The Journal of Law and Economics*, 15(1):143–149.

Croissant, Y. and Zeileis, A. (2016). truncreg: Truncated gaussian regression models. *R package version 0.2-4, URL https://CRAN. R-project. org/package= truncreg*.

Dietrich-Ragon, P. (2013). Qui rêve du logement social? *Sociologie*, 4(1):19–42.

Fischer, W. and Sard, B. (2017). Chart book: Federal housing spending is poorly matched to need. *Washington, DC: Center on Budget and Policy Priorities. March*, 8.

Fuchs, W. and Skrzypacz, A. (2010). Bargaining with arrival of new traders. *American Economic Review*, 100(3):802–36.

Ganong, P. and Shoag, D. (2017). Why has regional income convergence in the us declined? *Journal of Urban Economics*, 102:76–90.

Genesove, D. and Han, L. (2012). Search and matching in the housing market. *Journal of Urban Economics*, 72(1):31–45.

Griliches, Z. (1961). Hedonic price indexes for automobiles: An econometric of quality change. In *The price statistics of the federal goverment*, pages 173–196. NBER.

Han, L. and Strange, W. C. (2014). Bidding wars for houses. *Real Estate Economics*, 42(1):1–32.

Hausman, J. A. and Wise, D. A. (1977). Social experimentation, truncated distributions, and efficient estimation. *Econometrica: Journal of the Econometric Society*, pages 919–938.

Herkenhoff, K. F., Ohanian, L. E., and Prescott, E. C. (2018). Tarnishing the golden and empire states: Land-use restrictions and the us economic slowdown. *Journal of Monetary Economics*, 93:89–109.

Hosios, A. J. (1990). On the efficiency of matching and related models of search and unemployment. *The Review of Economic Studies*, 57(2):279–298.

Hyland, M., Lyons, R. C., and Lyons, S. (2013). The value of domestic building energy efficiency—evidence from ireland. *Energy economics*, 40:943–952.

Koch, D., Despotovic, M., Leiber, S., Sakeena, M., Döller, M., and Zeppelzauer, M. (2019). Real estate image analysis: A literature review. *Journal of Real Estate Literature*, 27(2):269–300.

Laferrère, A., Pouliquen, E., and Rougerie, C. (2017). Les conditions de logement en france. edition 2017. vue d'ensemble. le logement en france depuis trente ans.

Menzio, G. and Shi, S. (2011). Efficient search on the job and the business cycle. *Journal of Political Economy*, 119(3):468–510.

Moen, E. R. (1997). Competitive search equilibrium. *Journal of political Economy*, 105(2):385–411.

O'Shea, K. and Nash, R. (2015). An introduction to convolutional neural networks. *arXiv preprint arXiv:1511.08458*.

Piazzesi, M., Schneider, M., and Stroebel, J. (2020). Segmented housing search. *American Economic Review*, 110(3):720–759.

Poursaeed, O., Matera, T., and Belongie, S. (2018). Vision-based real estate price estimation. *Machine Vision and Applications*, 29(4):667–676.

Rosen, S. (1974). Hedonic prices and implicit markets: product differentiation in pure competition. *Journal of political economy*, 82(1):34–55.

Sandholm, T. and Gilpin, A. (2003). Sequences of take-it-or-leave-it offers: Near-optimal auctions without full valuation revelation. In *International Workshop on Agent-Mediated Electronic Commerce*, pages 73–91. Springer.

Schaal, E. (2017). Uncertainty and unemployment. *Econometrica*, 85(6):1675–1721.

Shimer, R. (1996). Contracts in a frictional labor market. Technical report, mimeo.

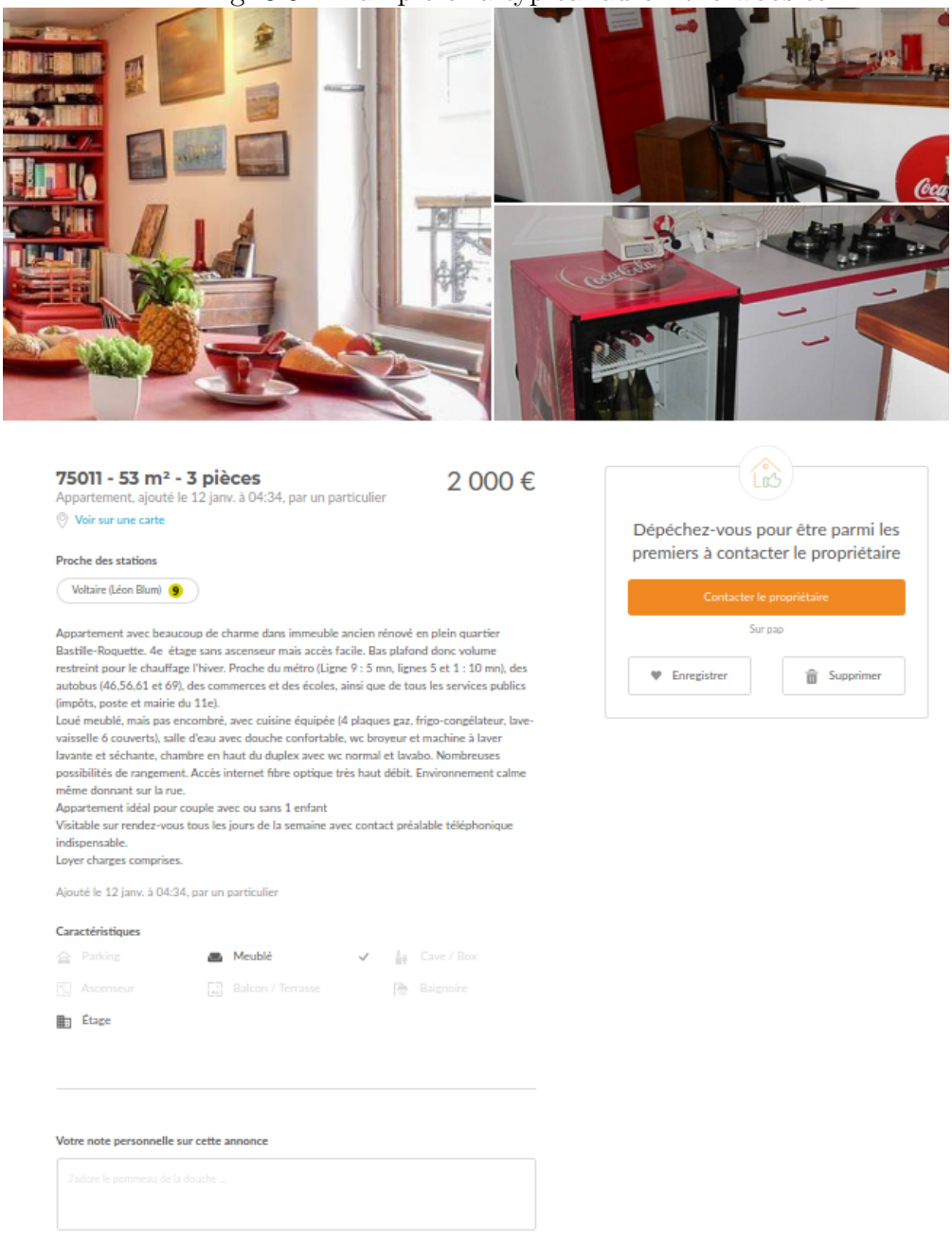
Talebi, H. and Milanfar, P. (2018). Nima: Neural image assessment. *IEEE Transactions on Image Processing*, 27(8):3998–4011.

Wei, S.-J., Zhang, X., and Liu, Y. (2017). Home ownership as status competition: Some theory and evidence. *Journal of Development Economics*, 127:169–186.

Zhang, Y. and Dong, R. (2018). Impacts of street-visible greenery on housing prices: Evidence from a hedonic price model and a massive street view image dataset in beijing. *ISPRS International Journal of Geo-Information*, 7(3):104.

A. Data collection

Fig. 3.8. Example of a typical ad on the website



75011 - 53 m² - 3 pièces 2 000 €

Appartement, ajouté le 12 janv. à 04:34, par un particulier

Voir sur une carte

Proche des stations

Voltaire (Léon Blum) 9

Appartement avec beaucoup de charme dans immeuble ancien rénové en plein quartier Bastille-Roquette. 4e étage sans ascenseur mais accès facile. Bas plafond donc volume restreint pour le chauffage l'hiver. Proche du métro (Ligne 9 : 5 mn, lignes 5 et 1 : 10 mn), des autobus (46,56,61 et 69), des commerces et des écoles, ainsi que de tous les services publics (impôts, poste et mairie du 11e).
Loué meublé, mais pas encombré, avec cuisine équipée (4 plaques gaz, frigo-congélateur, lave-vaisselle 6 couverts), salle d'eau avec douche confortable, wc broyeur et machine à laver lavante et séchante, chambre en haut du duplex avec wc normal et lavabo. Nombreuses possibilités de rangement. Accès internet fibre optique très haut débit. Environnement calme même donnant sur la rue.
Appartement idéal pour couple avec ou sans 1 enfant
Visitable sur rendez-vous tous les jours de la semaine avec contact préalable téléphonique indispensable.
Loyer charges comprises.

Ajouté le 12 janv. à 04:34, par un particulier

Caractéristiques

Parking · Meublé · Cave / Box
Ascenseur · Balcon / Terrasse · Baignoire
Étage

Votre note personnelle sur cette annonce

J'adore le pommeau de la douche ...

Dépêchez-vous pour être parmi les premiers à contacter le propriétaire

Contacter le propriétaire

Sur pap

Enregistrer · Supprimer

Source: Screenshot from the website LouerAgile.

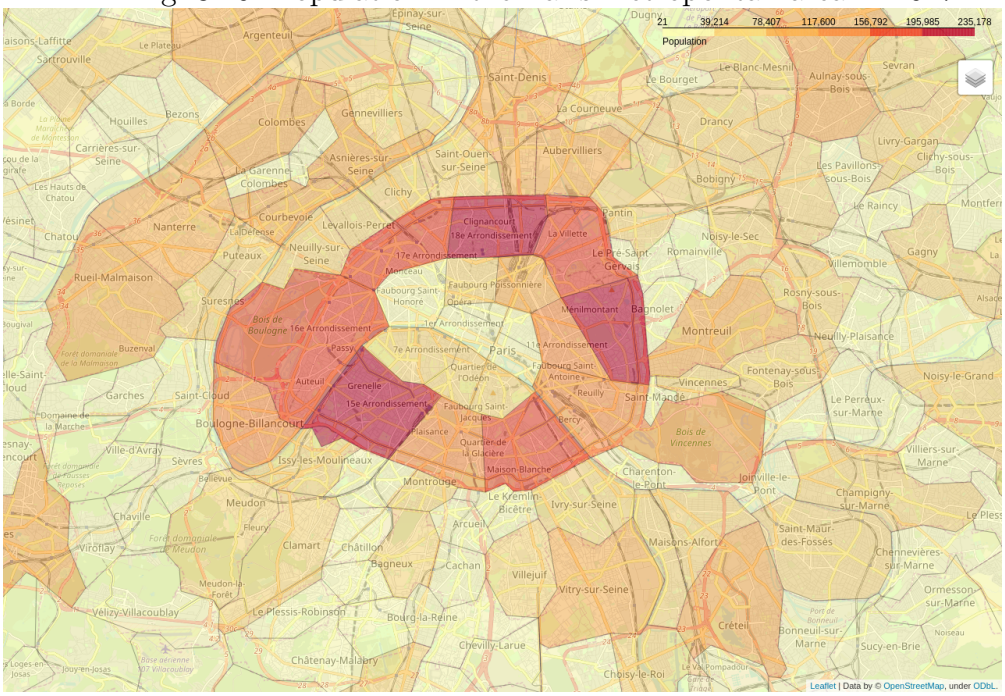
Fig. 3.9. Example of an ad in which the landlord was contacted 32 times



Source: Screenshot from the website LouerAgile.

B. Maps

Fig. 3.10. Population in the Paris metropolitan area in 2017



Source: INSEE.

Fig. 3.11. Number of ads



Source: Author's calculations based on data from LouerAgile.

Fig. 3.12. Median rent per m^2



Source: Author's calculations based on data from LouerAgile.

Fig. 3.13. Median surface of advertised accommodation (m^2)



Source: Author's calculations based on data from LouerAgile.

Fig. 3.14. Mean number of rooms



Source: Author's calculations based on data from LouerAgile.

Fig. 3.15. Mean number of bedrooms



Source: Author's calculations based on data from LouerAgile.

Fig. 3.16. Percentage of accommodation rented furnished



Source: Author's calculations based on data from LouerAgile.

Fig. 3.17. Percentage of ads managed by a real estate agency



Source: Author's calculations based on data from LouerAgile.

C. Rent control legislation in the Paris metropolitan area

The Paris rental market is governed by a rent control legislation. The "Arrêté fixant les loyers de référence, les loyers de référence majorés et les loyers de référence minorés pour la Ville de Paris" [20], entered into force on July 2019, establishes a set of rules regarding rent setting practices.

If it is the fist time the residence is on the rental market, landlords are free to fix the rent as long as it does not go above a certain "reference level", which depends on the property's characteristics (including its address, construction year, number of rooms, whether it is furnished or not).

When the tenancy agreement comes to an end, the landlord must notify the tenant of the new rent he or she wants to apply for the new tenancy agreement at least 6 months in advance. The new rent must no go above the maximum of the two following values:

- 50 % of the difference between the rent practiced for similar properties and the last rent paid before the renewal of the tenancy agreement
- 15 % of the renovation costs

If it is not the first time the property is on the rental market, the rent must not exceed the last rent, which can be increased by a factor equal to the increase of an index (the "Indice de référence des loyers"). Exceptions are allowed if renovation works were undertaken, or if the last rent was "undervalued".

[20] http://www.drihl.ile-de-france.developpement-durable.gouv.fr/IMG/pdf/arrete.pdf

Conclusion

The common denominator of this thesis is the exploration of markets with search frictions, with an emphasis on understanding how heterogeneity may impact economic outcomes.

In the first chapter, I revisit the question of the magnitude and persistence of costs of a job loss. While empirical evidence indicates that losing a job cost between 1 and 3 years of labor earnings in present value terms, most job search models have difficulties in replicating this magnitude. I explore how taking into account both the wage ladder and the "job ladder" can help in explaining why experimenting a job loss is so costly for workers. Because the speed at which workers climb the job ladder endogenously collapses during recessions, the model also helps in understanding parts of the *scarring effects of recessions*. A priority for future research is to see how the "job ladder" and the human capital channels interact along the business cycle. Another extension of the model would consist in adding precautionary saving motives and treating the interest rate as endogenously determined. While the former extension is easily feasible with the tools presented in Chapter 1, the latter would be more challenging as introducing a nonlinear utility function leads to a situation in which a simple sufficient statistics to determine job feasibility does not exist anymore: job feasibility now depends on wages.

In the second chapter, I also analyze the labor market. While in the first chapter I ignore the role of space, in the second chapter I take into account that labor markets are embedded within cities. Because of the spatial component of labor markets, the cost of commuting to the city center matters for local employment dynamics. In September 2015, a French reform changed the cost of public transport in the Paris metropolitan area. I analyze the reform using the spatial discontinuity created by the reform. I find that municipalities benefiting from a decrease in public transport costs experienced a positive employment effect. I use a frictional model of the labor market to explain some of mechanisms at play. Empirically, an investigation of the reform at a finer geographical level — at the IRIS level for instance — could produce new insights on the relationship between commuting costs and local employment dynamics. Theoretically, the model could be extended to include the reaction of the housing market to the reform.

In the third chapter, I focus on another key frictional market: the housing market. The motivation for this third chapter is that the rental housing market is understudied relative to the real estate market. I take advantage of the fact that most of the rental ads are now posted online to collect a novel dataset for the Parisian rental market. I establish that the rental housing market is well described by a directed search model, in which landlords advertise take-it-or-leave it offers and tenants search in submarkets. One empirical novelty is to shed light on a descending auction-like rent setting used by some landlords. Further research is needed to understand how these rent movements arise as an equilibrium outcome.

Institut d'études politiques de Paris
ÉCOLE DOCTORALE DE SCIENCES PO
Programme doctoral en économie
Département d'Économie
Doctorat en sciences économiques

Recherche, Appariement et Hétérogénéité

Julien Pascal

Thèse dirigée par

Jean-Marc Robin, Professeur d'économie, IEP de Paris et

University College London

Défendue le 30 Juin, 2020

Jury:

M. Fabrice COLLARD, Directeur de recherche, CNRS-TSE (*rapporteur*)

M. Olivier C. PIERRARD, Economiste senior, Banque centrale du Luxembourg

M. Xavier RAGOT, Directeur de recherche, CNRS-OFCE

M. Jean-Marc ROBIN, Professeur des universités, IEP de Paris

Mme Thepthida SOPRASEUTH, Professeur des universités, Université de Cergy-Pontoise

M. Etienne WASMER, Professor, New York University Abu Dhabi (*rapporteur*)

Résumé

Recherche, Appariement et Hétérogénéité

Cette thèse a pour dénominateur commun l'analyse des marchés frictionnels. Les modèles frictionnels s'éloignent des modèles d'équilibre général walrasiens, reposant sur l'idée d'un ajustement instantané entre l'offre et la demande. Alors que l'outil théorique du modèle walrasien est pertinent et utile dans de nombreux domaines, la compréhension de certains marchés est approfondie lorsque les frictions sont prises en considération. Un exemple canonique est celui du marché du travail. Un modèle d'équilibre walrasien ne peut rendre compte de l'existence *simultanée* d'un stock de chômeurs et d'un stock d'emploi vacants (Pissarides (2000)). De simples considérations permettent de se rendre compte de l'ampleur des frictions dans le marché du travail. Du côté des travailleurs, il est fréquent pour les nouveaux diplômés de devoir passer entre 4 et 8 entretiens pour obtenir un emploi dans une entreprise de consulting par exemple. Entre le moment de l'envoi du CV et l'offre final, il n'est pas rare de voir s'écouler plusieurs mois. Du côté des entreprises, recruter un travailleur coûte entre 10 et 17 semaines de salaire (Blatter et al. (2012)). Le processus de recrutement est long et incertain (voir par exemple Marchal (2019)).

Dans un autre domaine, celui du logement, quiconque a cherché à acheter une maison ou un appartement dans une zone tendue sait à quel point le processus peut être coûteux en temps et en émotion. Même le "marché du mariage" peut être pensé comme un problème de recherche et d'appariement (voir par exemple Chiappori et al. (2012)). L'idée centrale de cette thèse est que le processus d'appariement et les frictions qui en résultent ont des conséquences de premier ordre sur les phénomènes économiques et qu'elles méritent d'avoir une place centrale dans l'analyse.

Cette thèse peut aussi être vue comme une tentative pour améliorer nos connaissances à la fois théoriques et empiriques sur les marchés frictionnels. Le premier chapitre s'intéresse à une question qui peut paraître triviale à première vue : pourquoi perdre son emploi a-il autant de conséquences sur les revenus futurs des travailleurs ? Il y a ici un hiatus entre la plupart des modèles économiques qui prédisent que perdre un emploi ne devrait pas avoir de conséquences financières importantes et les observations empiriques basées sur des panels de travailleurs qui montrent que perdre un emploi génère des pertes de revenu importantes et persistantes. Comment expliquer aussi que le fait que perdre son emploi durant une période de récession génère une pénalité financière supplémentaire qui persiste plusieurs années, même dans la période de reprise économique? Cette question est importante du point de vue d'un planificateur qui cherche à réguler les fluctuations au cours du cycle économique, par exemple en utilisant des stabilisateurs automatiques. Les stabilisateurs automatiques sont des provisions légales qui réduisent la variance des variables agrégées au cours du cycle, par exemple une assurance chômage ou une taxe sur le revenu progressive (McKay and Reis (2016)). Dans le premier chapitre, je mets en exergue ce que j'appelle

le "risque d'appariement" (la qualité des complémentarités entre travailleurs et entreprises change au cours du cycle), qui vient s'ajouter aux risques déjà étudiés du "capital humain non transférables" (certaines compétences ne sont pas transférables entre deux emplois) et du "risque de signal" (un travailleur perdant sont emploi dévoilerait son "type" aux entreprises, qui par la suite refuseraient de l'embaucher). Une extension naturelle du premier chapitre viserait augmenter le modèle en permettant le choix de l'épargne. Il est facile d'imaginer que les travailleurs chercheraient à s'auto-assurer contre les chocs mentionnés ci-dessus, ce qui pourrait donner lieu à de riches mécanismes à l'échelle agrégée (voir Ravn and Sterk (2016) pour une exploration des liens entre risque d'emplois et modèles macroéconomiques).

Le deuxième chapitre est lui aussi une analyse du marché du travail. Alors que le premier chapitre met l'accent sur l'hétérogénéité au cours du cycle économique (les expansions et les récessions), le second chapitre souligne l'importance de "l'hétérogénéité spatiale", ou dit plus simplement le fait que le marché du travail soit inscrit dans des villes avec des contraintes de déplacements. La question principale dans ce second chapitre revêt un intérêt aussi bien théorique que pratique pour les planificateurs: est-ce qu'une baisse des coûts de transport peut modifier les dynamiques locales d'emploi? A travers le monde, des milliards d'euros ont été investis pour augmenter l'offre des transports en commun. Par exemple, le Grand Paris Express est un investissement de 40 milliards d'euros qui vise à étendre considérablement le réseau de transports en commun en région Ile-de-France. La société du Grand Paris estime que 115 000 emplois pourraient être créés dans les 25 ans après le début du chantier:

> "D'ici à 25 ans, le nombre d'emplois créés, directs et indirects, lié au Grand Paris Express est estimé à 115 000. De nouveaux emplois qui s'ajoutent aux quelque 685 000 résultant de la croissance naturelle de la région. Les bénéfices sociaux et économiques du projet sont estimés aujourd'hui à près de 80 milliards d'euros."
> Source: `https://www.societedugrandparis.fr/gpe/financement-192`

Est-il possible de valider ou de réfuter cette thèse ex ante ? Le second chapitre de cette thèse vise à éclairer les mécanismes qui lient dynamiques d'emploi et offres de transport, de manière à la fois de manière empiriques et théorique. D'un point de vue empirique, j'utilise une *discontinuité* créée par la création du "Forfait Toutes Zones" dans les transports en commun de la région parisienne en septembre 2015. Alors que le prix forfaitaire du transport a diminué en zone 3-5, il a augmenté marginalement en zones 1-2. D'un point de vue théorique, je développe un modèle du marché du travail d'appariement avec une dimension spatiale qui éclaire les mécanismes d'emploi après la réforme de cette tarification.

Le troisième chapitre change d'objet d'étude. Il ne s'agit plus d'analyser le marché du travail, mais un autre marché frictionnel d'importance capitale : le marché immobilier

locatif. La motivation principale pour la rédaction du troisième chapitre est que le marché immobilier locatif a été peu étudié, notamment à cause d'un manque de données, alors que les pratiques modernes sur internet offrent une quantité de données en continuelle expansion. Après une large collecte de données pour le marché locatif parisien, je mets en évidence que ce marché est bien décrit par un modèle d'appariement directionnel. Cependant, les stratégies d'établissement du loyer par certains propriétaires s'éloignent du modèle standard d'appariement directionnel. Est-ce que les mouvements de prix mis en évidence reflètent une inefficacité du marché locatif ? Ou bien, est-ce la réponse optimale des propriétaires face à l'incertitude à laquelle ils font face ? Est-ce que les changements de prix reflètent au contraire un mécanisme de recherche de rente par les propriétaires dans un marché ultra-tendu ? Bien que ces questions ne trouvent pas réponse définitive dans ce troisième chapitre, au vu des difficultés de modélisation et des difficultés économétriques qu'elles impliquent, ce chapitre pose des bases empiriques pour tenter d'y répondre dans le futur.

Chapitre 1: Chocs de revenu du travail et cycles économiques

Le premier chapitre offre un nouvel éclairage sur le rôle de la perte d'un emploi sur l'effet de scarification des récessions — l'observation empirique que les individus perdant leur emploi durant une période de récession perdent *plus* de revenu du travail que les travailleurs perdant leur emploi durant une période d'expansion, même après avoir pris en compte la durée du chômage et les caractéristiques des individus.

Une première observation est que les modèles du marché du travail à appariement ont de grandes difficultés à générer des coûts de destruction d'emploi raisonnables. Un simple calcul illustre cette idée: en prenant en compte le salaire moyen d'un travailleur américain et le nombre de jours moyen pour combler un emploi vacant, on peut montrer qu'une entreprise dépense approximativement 1000 USD pour embaucher un travailleur (Hall and Rogerson (2011)). Si on fait l'hypothèse que les salaires sont la solution d'une négociation de Nash symétrique (Nash (1951)), alors le coût de la perte de la destruction d'un emploi juste après l'embauche s'élève à 1000 USD également. Cependant, les analyses empiriques mettent en évidence que le coût de la destruction d'un emploi, après plusieurs années d'expérience, s'élève à des dizaines de milliers de dollars. Plus précisément, aux États-Unis, la valeur présente de la perte d'un emploi est comprise entre 1 et 3 *années* de revenu du travail. Comment concilier le fait que l'impact financier de la destruction d'un nouvel emploi est quasi négligeable, alors le coût de la destruction du même emploi après quelques années s'élève en milliers de dollars ? Ce fait met à rude épreuve les modèles du marché du travail actuels. L'économiste Robert Hall résume bien ce qui vient d'être énoncé ci-dessus:

"Le *modèle maître* fait face à un challenge important qui consiste à expliquer comment la participation d'un travailleur à un nouvel emploi passe d'un millier de dollars, à une centaine de milliers de dollars en seulement trois années d'expérience. Le gradient d'accumulation du capital humain spécifique à l'emploi est remarquablement pentu et constitue un réel challenge pour les analystes." Hall and Rogerson (2011), traduit en Français par mes soins.

Cette première observation est à mettre en relation avec un autre fait empirique qui représente un challenge pour les modèles du marché du travail à appariement : l'effet de scarification des récessions. Ce terme reflète l'observation empirique que la perte d'un emploi en période de récession est plus coûteux que la perte d'emploi en période d'expansion. L'effet de scarification des récessions a été mis en évidence aux Etats-Unis dans plusieurs travaux, notamment ceux de Jacobson et al. (1993), Stevens (1997) et Davis and Till (2011). Dans l'analyse de Davis and Till (2011), les auteurs montrent que les travailleurs perdant leur emploi en période de récession perdent entre 30 et 40% de revenus de *plus* que des travailleurs similaires perdant leur emploi durant une période d'expansion. Les travaux précédemment cités montrent aussi que la persistance des pertes de revenu est liée non pas une diminution des heures travaillées, mais à une perte de salaire horaire.

Pour expliquer l'amplitude et la persistance des chocs touchant le revenu du travail, deux voies principales ont été explorées. La première explication repose sur la théorie de l'accumulation de capital humain qui est spécifique à un emploi. Cette théorie offre une explication presque tautologique à l'amplitude et à la persistance des chocs touchant le revenu du travail: lorsqu'un travailleur perd son emploi, la capital humain spécifique qu'il a accumulé est détruit. Cela conduit à une perte de productivité, qui se traduit naturellement par une perte de salaire. Cette explication basée sur la notion de capital humain spécifique est par exemple présente dans les travaux de Ljungqvist and Sargent (1998).

La théorie du capital humain spécifique est attrayante par sa simplicité. Elle semble pouvoir expliquer le gradient mentionné par Robert Hall ci-dessus. Cependant, cette théorie a du mal à rendre compte de l'effet de scarification des récessions. En effet, pour que l'effet de scarification des récessions soit expliqué par la théorie du capital humain spécifique, il faudrait que celui-ci se *déprécie plus en période de récession qu'en période d'expansion.* A ma connaissance, nous ne disposons pas de preuves empiriques pour justifier cette hypothèse. D'autre part, des travaux empiriques ont mis en évidence le fait que les étudiants obtenant leurs diplômes pendant une récession subissent une perte de revenu du travail durant toute leur carrière professionnelle(par rapport à des étudiants similaires commençant leur carrière en période d'expansion). L'amplitude et la persistance des chocs de revenus pour les jeunes diplômés est similaire à ce que nous observons pour les travailleurs (Kahn (2010)). Cepen-

dant, les jeunes diplômés ne disposent pas de capital humain spécifique. Le cas des jeunes diplômés nous permet donc de nous demander si une part de l'effet de scarification ne serait pas liée au fonctionnement du marché du travail en récession. Nous reviendrons sur ce point plus bas.

La deuxième voie qui a été la plus explorée est celle de la théorie du signal. Cette théorie est basée sur l'idée que les travailleurs diffèrent selon de nombreuses caractéristiques et nombre d'entre elles ne sont pas observables par l'économètre. Certains caractéristiques ne sont même pas observables directement pas les employeurs. On pourrait par exemple penser à l'effort au travail, ou bien à la capacité à mener à terme un projet lorsque des obstacles de dernière minute se présentent. Les employeurs doivent deviner le "type" des travailleurs qu'ils embauchent, en étudiant leurs expériences professionnelles passées et en leur faisant passer plusieurs entretiens (Marchal and Rieucau (2010)). Lorsqu'un travailleur perd son emploi, cela laisse un "trou" dans son CV qui pourrait être par la suite interprété comme un signal négatif par les employeurs. Selon la théorie du signal, le coût de la perte d'un emploi n'est pas liée à une perte de productivité, mais à une productivité qui au contraire éclate au grand jour. La théorie du signal est par exemple étudiée dans Gibbons and Katz (1991).

La théorie du signal explique bien comment perdre son travail génère des effets larges et persistants. Cependant, elle peine à rendre compte de l'effet de scarification des récessions. En effet, la probabilité de perdre son emploi pendant une récession est plus élevée que durant une expansion. Durant une période de récession, beaucoup de "bons" travailleurs sont laissés sur le carreau (Nakamura (2008)). Perdre son emploi durant une période d'expansion pourrait donc être interprété comme un signal négatif par les employeurs. On peut donc se demander pourquoi les employeurs ne pénalisent pas plus les travailleurs perdant leur emploi pendant une expansion que ceux perdant leur emploi durant une période de récession.

Le premier chapitre vise donc à offrir une explication qui complète la théorie du capital humain spécifique et la théorie du signal. L'explication que je propose est basée sur la qualité de l'appariement au cours du cycle. Les travailleurs au cours de leur carrière non seulement progressent sur l'échelle des salaires en obtenant des promotions au sein de leur entreprise, mais ils progressent également "l'échelle des entreprises" en passant d'une entreprise à une autre. Lorsque les travailleurs perdent leur emploi, ils perdent la qualité de l'appariement qu'ils avaient accumulée jusque-là. Cela conduit à des pertes de revenu du travail substantielles et persistantes, en adéquation avec les observations empiriques. Cette explication permet aussi de rendre compte de l'effet de scarification des récessions, dans la mesure où la vitesse avec laquelle la qualité de l'appariement tend à augmenter avec l'expérience décroît en période de récession. En effet, durant une récession les entreprises sont moins nombreuses à ouvrir de nouveaux postes et les possibilités de mouvement vers

des emplois plus appropriés sont moindres. Certaines entreprises se mettent aussi à faire du "upskilling": elles rendent plus stricts les critères de sélection alors que l'emploi lui-même ne change pas (Modestino et al. (2019)). Perdre son emploi en période de récession revient donc en moyenne à rester plus longtemps mal apparié. Et un mauvais appariement entraîne une perte de salaire horaire, mais non une perte d'heures travaillées, comme dans les données empiriques.

Pour étayer cette explication basée sur la complémentarité entreprises-travailleurs, j'utilise un modèle du marché du travail avec des travailleurs et des entreprises hétérogènes, estimé avec des données sur les États-Unis. Le modèle est composé de deux blocs : une partie détermination de l'emploi et une partie détermination des salaires. La détermination de l'emploi est identique au modèle de Lise and Robin (2017). Ma contribution en termes de modélisation est de résoudre le problème de la détermination des salaires. Plus précisément, je montre comment les salaires peuvent être calculés dans un modèle d'appariement aléatoire avec des travailleurs et des entreprises hétérogènes et une incertitude agrégée sur la productivité des travailleurs (qui génère les cycles économiques). Mon approche algorithmique repose sur l'observation que (i) la détermination de la faisabilité de l'emploi le long du cycle économique peut-être résolue dans une première étape, indépendamment de la détermination des salaires; (ii) bien que les salaires dépendent de manière très générale des flux des stocks de travailleurs, la dimension du problème des salaires peut être réduite de manière drastique en utilisant l'information de la première étape et des hypothèses similaires à Krusell and Smith (1998). Plus généralement, ce chapitre montre que dans la résolution des modèles d'appariement aléatoire avec agents hétérogènes et incertitude agrégée, il est crucial de rendre indépendant le problème de la faisabilité de l'emploi de celui de la détermination des salaires. Le problème de la détermination des salaires peut de son côté rester extrêmement général (dépendre des flux et des stocks d'emplois) sans pour autant empêcher la résolution du modèle.

Une suite logique du premier chapitre viserait à caractériser précisément les contributions de la théorie du capital humain, de la théorie du signal et de la théorie de l'appariement au coût final de la perte d'un emploi. Je fais l'observation que le modèle présenté dans le premier chapitre permet d'intégrer aisément l'accumulation du capital humain spécifique des travailleurs. Une prochaine étape consisterait en l'estimation d'un modèle structurel incorporant des fluctuations dans la qualité de l'appariement et une accumulation de capital humain spécifique.

Chapitre 2: Dynamiques locales d'emploi et coûts des transports en commun

Le deuxième chapitre analyse comment la structure des coûts des transports en commun peut altérer les dynamiques locales d'emploi. Ce chapitre est aussi une analyse du marché du travail, mais l'approche se veut plus empirique que pour le premier chapitre. Ici, il n'est plus question d'hétérogénéité au cours du cycle économique, mais d'hétérogénéité spatiale: le marché du travail prend corps dans des villes, dans lesquelles les individus doivent se déplacer, ce qui implique des coûts aussi bien financiers que temporels et mentaux (Raymond and Gotman (1982)).

Ce chapitre est motivé par l'observation que plusieurs pays se sont lancés dans des projets ambitieux d'extension de leur réseau de transports en commun. Par exemple, le Grand Paris Express, d'un montant approximatif de 40 milliards d'euros, a pour objectif d'ajouter 200 km de voies ferrées dans la métropole parisienne. Au Royaume-Uni, près de 19 milliards de livres ont été dépensés pour la construction du London Crossrail, un réseau ferroviaire de type réseau express régional (RER) qui doit desservir le Grand Londres à partir de 2021. En Espagne, plus de 16 milliards d'euros ont été dépensés pour les 42,6 nouveaux kilomètres de la ligne 9 du métro de Barcelone. Alors que des gouvernements investissent dans l'extension du réseau, d'autres décident de changer la tarification des transports en commun. Par exemple, le réseau de transports en commun Luxembourgeois est devenu gratuit à partir de mars 2020.

Les projets et réformes mentionnés ci-dessus ont divers buts. D'une part, ils sont motivés par des objectifs sanitaires et environnementaux, en particulier réduire la pollution et les émissions de C02 en encourageant les usagers à laisser leur voiture au garage. D'autre part, la création d'emplois est aussi souvent mentionnée. Par exemple, la Société du Grand Paris estime que plus de 100 000 emplois seront créés par le Grand Paris Express.[1]. L'objet de ce chapitre est donc d'analyser l'impact sur le marché de l'emploi des réformes mentionnées ci-dessus. Mon approche empirique est basée sur l'utilisation d'une discontinuité géographique créée par une réforme tarifaire des transports en commun dans la métropole parisienne en septembre 2015.

En septembre 2015, un forfait à prix unique pour les transports en commun, le "Forfait Toutes Zones" (FTZ), a remplacé la plupart des forfaits précédents. Avant cette réforme, le prix d'un forfait hebdomadaire, mensuel ou annuel dépendait des zones traversées durant les trajets. Ces zones tarifaires forment des disques concentriques autour de Paris intra-muros. Avant la réforme, les forfaits pour les zones 1-2 (Paris intra-muros et certaines communes limitrophes de Paris) étaient moins chers que les forfaits 1-3 (pour les personnes faisant le

[1]Source: `https://www.societedugrandparis.fr/info/economie-et-emploi-201`

déplacement Paris - proche banlieue).

Le FTZ a créé une discontinuité dans les coûts de transports en commun. Si on considère par exemple les utilisateurs du forfait Navigo annuel pour les zone 1-3, la réforme s'est traduite par une baisse d'environ $14,3$ € par mois, alors que les utilisateurs du forfait Navigo annuel pour les zone 1-2 ont vu le coût mensuel moyen augmenter de $3,5$ €. La stratégie d'identification que j'utilise est basée sur l'hypothèse que deux municipalités proches de la frontière séparant les zones 2 et 3 sont soumises à des dynamiques d'emploi similaires (après avoir contrôlé les différences causées par des variables observables). Si cette hypothèse est valide, il est possible d'utiliser la méthode des doubles différences pour estimer l'effet causal d'une baisse du coût des transports en commun sur les dynamiques d'emploi.

Ce chapitre se rapproche de travaux empiriques précédents qui étudient les liens entre infrastructures de transport et emploi. Mayer and Trevien (2017) montrent que la création du réseau RER dans la métropole parisienne a augmenté l'emploi de 8.8% dans les villes bénéficiant d'un arrêt de RER. Garcia-López et al. (2017) trouvent que le réseau RER a façonné durablement la région parisienne, en conduisant à l'émergence de nouveaux pôles d'emplois suburbains dans les villes disposant d'arrêts de RER. Duranton and Turner (2012) montrent que lorsqu'une ville voit sont stock de bus augmenter de 10%, la population de la ville augmente en moyenne de 0.8%. Les trois travaux mentionnés ci-dessus reposent sur la méthode des variables instrumentales pour identifier l'effet causal d'un changement dans les politiques de transport sur l'emploi ou la population. Pour étudier l'impact de la réforme FTZ sur l'emploi, je n'utilise pas un "instrument", mais la discontinuité géographique créée par la réforme (voir Neumark and Simpson (2015)). La méthodologie que j'utilise est très proche de celle utilisée dans les travaux de Chapelle et al. (2018), Einiö and Overman (2016), Kline and Moretti (2013) ou Hilber et al. (2019).

Mes résultats empiriques sont cohérents avec les travaux mentionnés ci-dessus, dans la mesure où je trouve également un lien positif entre offre de transports en commun et dynamique locale d'emploi. Plus précisément, je constate que la baisse du coût des transports en commun dans la zone 3 a conduit à une baisse moyenne de 2% dans le nombre d'inscrits à Pôle Emploi pour les villes située dans cette zone géographique. L'effet positif sur l'emploi local est particulièrement fort pour les chômeurs de catégorie A, qui sont aussi ceux qui ont le plus de chance d'être des chômeurs de longue durée.

Dans ce chapitre, je mets en exergue que les projets de nouvelles infrastructures de transport (ou bien les politiques changeant la tarification) peuvent avoir des conséquences sur l'emploi. L'idée centrale est que le transport jusqu'au lieu de travail est coûteux. Si le lieu de travail est trop éloigné et donc perçu comme trop coûteux, le salaire ne compensera pas assez les frais de déplacement et certains travailleurs refuseront l'offre d'emploi (Zenou (2000)). La

création du FTZ peut alors être vue comme une prime à l'emploi géographiquement localisée, à destination des résidents des zones 3 à 5. Le salaire de réserve des chômeurs en zones 3 à 5 a donc diminué. Un autre mécanisme ayant pu conduire à une expansion de l'emploi dans ces zones est que les chômeurs ont peut-être étendu la zone géographique dans laquelle ils cherchent un emploi (les déplacements pour les entretiens sont maintenant moins chers). D'autre part, les décisions liées au lieu de résidence des travailleurs et nouveaux arrivants ont pu être altérées par la réforme. Ce point est important puisque j'observe une répétition de données transversales pour chaque municipalité, et non un panel d'individus. Je formalise ces trois canaux au sein d'un modèle du marché du travail qui repose notamment sur les travaux de Brueckner et al. (1999), Wasmer and Zenou (2002) et la littérature des choix discrets. Ce modèle aide à comprendre les liens entre coût des transports en commun et dynamiques locales d'emploi.

Chapitre 3: Marché immobilier locatif et recherche directionnelle

Le troisième chapitre change d'objet d'étude, tout en maintenant le focus sur les marchés frictionnels. Il s'agit ici d'étudier le marché immobilier. Ce chapitre est motivé par le fait que la marché immobilier est d'une importance capitale pour le reste de l'économie. La Grande Récession a été déclenchée entre autres par l'éclatement d'une bulle immobilière aux États-Unis (Farmer (2012)). Des études ont montré que l'offre insuffisante de logements aux Etats-Unis a créé des distorsions sur le marché du travail (Brown and Matsa (2019)) et a entraîné la baisse de la croissance de la productivité du travail (Herkenhoff et al. (2018)). Il a également été établi que le fonctionnement du marché de l'immobilier impacte les décisions de mariages (Wei et al. (2017)). Il influence aussi directement sur les processus politiques: les frustrations liées aux problèmes de logement ont favorisé le vote pour le Brexit en Grande-Bretagne et la montée du votre pour le Front National en France (Adler and Ansell (2020)).

Ce chapitre part de l'observation que de nombreux travaux traitent de la vente et de l'achat des bien immobiliers, mais que la littérature sur le marché immobilier locatif est moins abondante. Est-ce parce que le marché immobilier locatif est marginal ? Cela ne semble pas être le cas. En effet, en France 60% des individus sont locataires et 58% d'entre eux vivent dans des logements appartenant au marché locatif privé. Je fais donc l'hypothèse que le marché locatif est moins étudié car les données étaient jusqu'à présent manquantes. Par exemple, en France, alors que les transactions immobilières doivent être enregistrées par les agences notariales, ce n'est pas le cas pour les contrats de locations. Ce chapitre a

donc pour ambition d'établir de nouveaux faits empiriques sur un marché auparavant peu visible. Cela est rendu possible par l'émergence des sites internet d'annonces locatives, et par l'adoption massive de ces sites par les propriétaires et les locataires. Ce chapitre se place dans la lignée des travaux qui cherchent à utiliser des techniques de *web scraping* (collecte automatique des données du net) sur des données générées directement par les utilisateurs (Edelman (2012)).

Ce chapitre apporte deux contributions empiriques à la littérature étudiant le marché immobilier locatif. Premièrement, en utilisant une nouvelle base de données collectées via *web scraping*, je fais l'observation que le marché immobilier locatif est bien décrit par un modèle de recherche directionnelle, en adéquation avec ce qui est décrit par le modèle de Moen (1997). En effet, après avoir contrôlé la "qualité" des logements via un modèle hédonique, on observe que les logements ayant un loyer plus bas que prévu attirent plus de locataires potentiels. Au contraire, un loyer plus élevé que prévu implique un nombre de locataires intéressés moins important. Les propriétaires bougent sur une courbe d'indifférence illustrant le compromis loyer-probabilité de trouver un locataire. Une deuxième contribution empirique est de montrer qu'une partie non négligeable des propriétaires choisissent une stratégie de prix descendante pour les loyers publiés sur internet. Ce fait empirique est nouveau. Sans offrir une réponse définitive aux questions soulevées, j'offre une discussion sur les conséquences en termes d'efficacité et d'optimalité des mouvements de loyer mis en évidence.

References

Adler, D. and Ansell, B. (2020). Housing and populism. *West European Politics*, 43(2):344–365.

Blatter, M., Muehlemann, S., and Schenker, S. (2012). The costs of hiring skilled workers. *European Economic Review*, 56(1):20–35.

Brown, J. and Matsa, D. A. (2019). Locked in by leverage: Job search during the housing crisis. *Journal of Financial Economics*.

Brueckner, J. K., Thisse, J.-F., and Zenou, Y. (1999). Why is central paris rich and downtown detroit poor?: An amenity-based theory. *European economic review*, 43(1):91–107.

Chapelle, G., Vignolles, B., and Wolf, C. (2018). Impact of a housing tax credit on local housing markets: Evidence from france. *Annals of Economics and Statistics/Annales d'Économie et de Statistique*, (130):101–131.

Chiappori, P.-A., Oreffice, S., and Quintana-Domeque, C. (2012). Fatter attraction: anthropometric and socioeconomic matching on the marriage market. *Journal of Political Economy*, 120(4):659–695.

Davis, S. J. and Till, M. (2011). Recessions and the cost of job loss. Technical report, Working Paper. National Bureau of Economic Research.

Duranton, G. and Turner, M. A. (2012). Urban growth and transportation. *Review of Economic Studies*, 79(4):1407–1440.

Edelman, B. (2012). Using internet data for economic research. *Journal of Economic Perspectives*, 26(2):189–206.

Einiö, E. and Overman, H. G. (2016). The (displacement) effects of spatially targeted enterprise initiatives: Evidence from uk legi.

Farmer, R. E. (2012). The stock market crash of 2008 caused the great recession: Theory and evidence. *Journal of Economic Dynamics and Control*, 36(5):693–707.

Garcia-López, M.-À., Hémet, C., and Viladecans-Marsal, E. (2017). Next train to the polycentric city: The effect of railroads on subcenter formation. *Regional Science and Urban Economics*, 67:50–63.

Gibbons, R. and Katz, L. F. (1991). Layoffs and lemons. *Journal of labor Economics*, 9(4):351–380.

Hall, R. and Rogerson, R. (2011). Recessions and the cost of job loss: Comments and discussion. Technical report, Working Paper. National Bureau of Economic Research.

Herkenhoff, K. F., Ohanian, L. E., and Prescott, E. C. (2018). Tarnishing the golden and empire states: Land-use restrictions and the us economic slowdown. *Journal of Monetary Economics*, 93:89–109.

Hilber, C. A., Carozzi, F., and Xiaolun, Y. (2019). The economic impacts of help to buy.

Jacobson, L. S., LaLonde, R. J., and Sullivan, D. G. (1993). Earnings losses of displaced workers. *The American economic review*, pages 685–709.

Kahn, L. B. (2010). The long-term labor market consequences of graduating from college in a bad economy. *Labour economics*, 17(2):303–316.

Kline, P. and Moretti, E. (2013). Local economic development, agglomeration economies, and the big push: 100 years of evidence from the tennessee valley authority. *The Quarterly Journal of Economics*, 129(1):275–331.

Krusell, P. and Smith, Jr, A. A. (1998). Income and wealth heterogeneity in the macroeconomy. *Journal of political Economy*, 106(5):867–896.

Lise, J. and Robin, J.-M. (2017). The macrodynamics of sorting between workers and firms. *American Economic Review*, 107(4):1104–35.

Ljungqvist, L. and Sargent, T. J. (1998). The european unemployment dilemma. *Journal of political Economy*, 106(3):514–550.

Marchal, E. (2019). *Les embarras des recruteurs: enquête sur le marché du travail*. Éditions de l'École des hautes études en sciences sociales.

Marchal, E. and Rieucau, G. (2010). *Le recrutement*. La Découverte.

Mayer, T. and Trevien, C. (2017). The impact of urban public transportation evidence from the paris region. *Journal of Urban Economics*, 102:1–21.

McKay, A. and Reis, R. (2016). The role of automatic stabilizers in the us business cycle. *Econometrica*, 84(1):141–194.

Modestino, A. S., Shoag, D., and Ballance, J. (2019). Upskilling: do employers demand greater skill when workers are plentiful? *Review of Economics and Statistics*, pages 1–46.

Moen, E. R. (1997). Competitive search equilibrium. *Journal of political Economy*, 105(2):385–411.

Nakamura, E. (2008). Layoffs and lemons over the business cycle. *Economics Letters*, 99(1):55–58.

Nash, J. (1951). Non-cooperative games. *Annals of mathematics*, pages 286–295.

Neumark, D. and Simpson, H. (2015). Place-based policies. In *Handbook of regional and urban economics*, volume 5, pages 1197–1287. Elsevier.

Pissarides, C. A. (2000). *Equilibrium unemployment theory*. MIT press.

Ravn, M. O. and Sterk, V. (2016). Macroeconomic fluctuations with hank & sam: An analytical approach.

Raymond, H. and Gotman, A. (1982). Le transport en commun. une problématique de la vie quotidienne. In *Les Annales de la Recherche Urbaine*, volume 14, pages 3–27. Centre de Recherche d'Urbanisme.

Stevens, A. H. (1997). Persistent effects of job displacement: The importance of multiple job losses. *Journal of Labor Economics*, 15(1, Part 1):165–188.

Wasmer, E. and Zenou, Y. (2002). Does city structure affect job search and welfare? *Journal of urban Economics*, 51(3):515–541.

Wei, S.-J., Zhang, X., and Liu, Y. (2017). Home ownership as status competition: Some theory and evidence. *Journal of Development Economics*, 127:169–186.

Zenou, Y. (2000). Urban unemployment, agglomeration and transportation policies. *Journal of Public Economics*, 77(1):97–133.