COUPLED DYNAMICS OF LABOR AND FIRMS THROUGH COMPLEX NETWORKS

by

Omar A. Guerrero
A Dissertation
Submitted to the
Graduate Faculty
of
George Mason University
in Partial Fulfillment of
The Requirements for the Degree
of
Doctor of Philosophy
Computational Social Science

Committee:

_____ Chair of Committee

_____

_____

_____

_____

_____ Director of Graduate Studies

_____ Director, Krasnow Institute for Advanced Study

Date: _____ Summer Semester 2013
George Mason University
Fairfax, VA

Coupled Dynamics of Labor and Firms through Complex Networks

A dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy at George Mason University

by

Omar A. Guerrero
Master of Science
University of Essex, 2008
Bachelor of Arts
Tecnológico de Monterrey, 2007

Director: Robert Axtell, Chair
Department of Computational Social Science

Summer Semester 2013
George Mason University
Fairfax, VA

This work is licensed under a creative commons attribution-noderivs 3.0 unported license.

DEDICATION

This is dedicated to Sanna, for her support, comprehension, and love.

ACKNOWLEDGEMENTS

I would like to thank Robert Axtell for his guidance, teachings, involvement, and for believing in this project since the beginning. I would like to thank Daniel Houser for his support, valuable inputs, and rigorous teachings. I would like to thank Carlos Urzúa for his interest and support, which were instrumental in starting this work. I would also like to thank my parents for their unconditional support throughout these years. I would like to thank all the members of the Department of Computational Social Science, for providing the best academic and personal environment to develop my ideas and pursue all my interests. Finally, I would like to thank the Mexican National Council of Science and Technology (CONACYT), for its financial support during my graduate studies.

TABLE OF CONTENTS

Page

List of Tables viii
List of Figures ix
List of Equations x
List of Abbreviations xi
Abstract xii
1. Introduction 1
2. A Survey on Labor Dynamics and Networks 5
2.1. The Conventional View of Labor Dynamics 7
2.1.1. Search Models 7
2.1.2. Matching and DSGE Models 9
2.2. Labor Dynamics from Stochastic Processes 10
2.2.1. Gibrat's Law 11
2.2.2. Simon's Models 11
2.2.3. The Kesten Process 13
2.2.4. Econophysics and Growth Rates 13
2.2.5. Gabaix's Micro-Foundations 14
2.3. Labor Dynamics and Networks 16
2.3.1. Exogenous Networks 17
2.3.2. Endogenous Networks 18
2.4. Agent-Based Models of Labor Dynamics 20
2.4.1. Bergmann's Simple Microsimulation 20
2.4.2. Tesfatsion's Evolutionary Model 21
2.4.3. Tassier & Menczer Evolutionary Networks 23
2.4.4. Job Search Model 24
2.4.5. Fagiolo et al.'s Evolutionary Model 25
2.4.6. Neugart's Endogenous Matching Function 27

2.4.7. Richiardi's Equilibrium Model 28
2.4.8. Richiardi's Disequilibrium Model 29
2.4.9. Neugart's Sectorial Model 30
2.4.10. French Labor Market Model 31
2.4.11. Shocks and Endogenous Institutions 33
2.4.12. Jobs and Social Networks 34
2.4.13. EURACE 36
2.4.14. Endogenous Social Networks 37
2.5. Conclusions 39
3. Employment Growth through Labor Flow Networks 41
3.1. Introduction 41
3.2. Materials and Methods 43
3.2.1. Primary Dataset 43
3.2.2. Support Dataset 45
3.2.3. Census Dataset 46
3.2.4. Statistical Units 46
3.2.5. Labor Flow Networks 47
3.2.6. Null Models 48
3.2.7. Employment Growth and Firms 49
3.2.8. Distributions 51
3.3. Results 53
3.3.1. Brief Comparison 54
3.3.2. Firm Dynamics 56
3.3.3. Labor Flow Networks 59
3.3.4. Networks and the Economy 66
3.3.5. Sectors, Regions, and Communities 68
3.3.6. Employment Growth 72
3.3.7. A Taxonomy for Labor Flow Network Firms 77
3.3.8. Network Formation 82
3.4. Discussion 86
4. Three Models of Labor and Firm Dynamics 88
4.1. Introduction 88

4.2. Firm Dynamics 89
4.2.1. Model Description 90
4.2.2. Implementation and Execution 92
4.2.3. Calibration 94
4.2.4. Experimental Design and Output 97
4.2.5. Results 98
4.3. Labor Dynamics Through Endogenous Social Networks 103
4.3.1. Model Description 105
4.3.2. Implementation and Execution 113
4.3.3. Calibration 116
4.3.4. Experimental Design and Output 119
4.3.5. Results 120
4.4. Small-World Labor Flow Networks and Job Search 131
4.4.1. Model Description 133
4.4.2. Implementation and Execution 139
4.4.3. Calibration 140
4.4.4. Experimental Design and Output 144
4.4.5. Results 144
4.5. Conclusions 148
5. Conclusions 149

LIST OF TABLES

Table Page
Table 1: Employer-employee matched datasets 47
Table 2: Comparison between Finland and Mexico 55
Table 3: LFN properties..... 60
Table 4: LFN and employment growth..... 74
Table 5: Employment by types of firms 75
Table 6: Parameters of firm dynamics model 97
Table 7: Parameters of unemployment and endogenous social networks model 119
Table 8: Parameters of job search and LFNs model 143

LIST OF FIGURES

Figure Page
Figure 1: Stylized facts of firm dynamics. 58
Figure 2: Topology of the labor flow network from Finland 61
Figure 3: Topology of the labor flow network from Mexico 65
Figure 4: Correlations between network properties and economic variables 67
Figure 5: Communities of firms 70
Figure 6: Classification of firms and industrial participation 76
Figure 7: Growth, hires, and separations 78
Figure 8: Firm taxonomy 80
Figure 9: Firm classes and change in employment 81
Figure 10: Formation of the Mexican LFN on a daily basis 84
Figure 11: Training duration and search space 96
Figure 12: Output from firm dynamics 100
Figure 13: Technology and productivity 101
Figure 14: Income-related outputs 102
Figure 15: LFN outputs 103
Figure 16: Effects of socialization in agent utility 107
Figure 17: Social tie formation game 108
Figure 18: Frequencies of social ties by differences in wages 111
Figure 19: Arrival probability 112
Figure 20: Empirical distribution of skills 118
Figure 21: Outputs from firm dynamics 121
Figure 22: Technology and firms 122
Figure 23: Scaling with firm size 123
Figure 24: Income-related outputs 124
Figure 25: Macroeconomic outputs 126
Figure 26: Unemployment and duration 127
Figure 27: Assortativity through social ties 129
Figure 28: Social networks and labor outcomes 130
Figure 29: Social capital in firms 131
Figure 30: Initial topology of the LFN 137
Figure 31: Labor dynamics from job search 145
Figure 32: Small-world topology from job search 146
Figure 33: Network effect of microeconomic parameters 147

LIST OF EQUATIONS

Equation Page
Equation 1 49
Equation 2 51
Equation 3 52
Equation 4 53
Equation 5 90
Equation 6 106
Equation 7 135

LIST OF ABBREVIATIONS

Agent-Based Model ..... ABM
Dynamic Stochastic General Equilibrium ..... DSGE
Finland ..... FIN
Finnish Longitudinal Employer-Employee Dynamics ..... FLEED
Gazelle Firm..... GF
Gross Domestic Product ..... GDP
High Impact Firm..... HIF
Labor Flow Network..... LFN
Labor Flow Network Firm..... LFNF
Mexico ..... MEX
Mexican Institute for Social Security ..... IMSS
Positively Growing Firm ..... PGF

ABSTRACT

COUPLED DYNAMICS OF LABOR AND FIRMS THROUGH COMPLEX NETWORKS

Omar A. Guerrero, PhD

George Mason University, 2013

Dissertation Director: Robert Axtell

This dissertation bridges the gap between labor and firm dynamics through the study of complex networks in labor markets. With extensive use of large-scale employer-employee matched micro-data and agent-based modeling, we tap into the effects that networked structures (between individuals or between firms) exert in labor outcomes and employment dynamics. Some of the contributions of this work are: (*i*) the first characterization of a network of firms (connected through labor flows, i.e. *labor flow networks*) for an entire economy; (*ii*) the study of the relationship between labor flow networks and employment dynamics; (*iii*) agent-based models that generate rich stylized facts about labor, firm, and social dynamics from microeconomic behavior; (*iv*) providing the microeconomic foundations of the formation process of labor flow networks by coupling job search models with models about the formation of complex networks. We show that the study of labor dynamics can be enriched by coupling firm dynamics. Using

agent-based modeling is a natural way to deal with the heterogeneous experiences of workers and firms while maintaining a simple representation of the labor market. Despite their simplicity these models are grounded on empirical evidence obtained from large-scale micro-data and are capable of generating more stylized facts than equation-based models. This approach has great potential for the design and evaluation of labor policies. Therefore, countries and regulators would be greatly benefited from collecting large-scale labor micro-data, analyzing labor flow networks, and developing agent-based models of labor markets.

1. INTRODUCTION

For decades, social scientists and policy makers have struggled understanding the inner workings of labor markets. As one of the fundamental factors of production, labor plays a crucial role in the growth and performance of economies. Its dynamics are intrinsically connected to the growth of companies and other economic processes of countries. Despite the significant advancements in understanding labor dynamics, there are still many challenges to overcome. For example, most models designed for labor policy do not take into account the interplay between realistic firm dynamics and labor markets. Another example is the limited understanding of the role of networks (social or economic) in labor markets, which remains as an extremely stylized illustration about spillover effects and referral hiring.

Despite recent advancements in understanding real-world networks, and in the implementation of computational simulations, the study of labor markets is predominantly conducted through econometric and mathematical methods. In order to obtain clear results using these methods, the modeler has to oversimplify the decentralized and heterogeneous nature of the market, which often drives its dynamics. This dissertation contributes to advance the understanding of labor dynamics in four main ways: (*i*) by providing the first characterization of a network of firms (connected through labor flows, i.e. *labor flow networks*) for an entire economy; (*ii*) by finding evidence of

the relationship between labor flow networks and employment dynamics; (*iii*) by developing agent-based models that generate rich stylized facts about labor, firm, and social dynamics from microeconomic behavior; and (*iv*) by providing the microeconomic foundations of the formation process of labor flow networks through the coupling of job search models with models about the formation of complex networks.

Chapter one presents a literature survey of models about labor and firm dynamics. First, we overview the dominant approach to model labor dynamics: *search models* and *matching functions*. Then, we review *stochastic processes* that generate stylized facts of firm dynamics. Next, we survey *game-theoretic models* that bridge the gap between networks and labor markets. Finally, we review every *agent-based model* that has been developed so far to analyze labor markets from an economic perspective. The literature review in this chapter has an emphasis in agent-based models because that is the chosen methodological tool for this dissertation.

Chapter two introduces a new way to analyze labor and firm dynamics as networks of firms connected through flows of labor. The concept of a *labor flow network* (LFN) is developed and several stylized facts are documented for multiple large-scale employer-employee matched micro-datasets. We demonstrate that LFNs provide useful information to improve our understanding about employment growth. We argue that labor flow networks are neither the product of purely random processes, nor the intended consequence of strategic interactions between economic agents. Therefore, we advocate for the use to computational methods in order to understand how these robust structures

emerge as the side product of economic interactions in the labor market. This is an important aspect of labor dynamics because networks feed back into the market by reshaping the space through which agents find jobs. Hence, by understanding how the properties of labor flow networks affect labor and firm dynamics and vise versa, we can create models that provide us with a better understanding about labor markets and the impact of labor policies.

Chapter three explores different ways to employ agent-based models in order to study endogenous labor and firm dynamics. First, we propose an agentized version of a popular microeconomic model that generates a Zipf distribution of firm sizes. Our agent-based model recovers a heavy-tailed firm size distribution without the need of assumptions such as macroeconomic equilibrium and full information/rationality. We show that this model generates additional stylized facts of firm dynamics such as the growth rates distribution and the employer-size premium. Additionally, the model emerges some features of labor markets such as a skewed income distribution. However, the properties of the labor flow networks produced by this model are not as rich as the empirical ones, suggesting that their stylized facts are the product of more complex dynamics.

In the same chapter we present a model about social network formation in the workplace. In this model, agents form networks of personal contacts by following their preferences over effective work and socialization, and their endowment of skills. Information asymmetries generate diversity in the types of contacts that an agent has. At

the same time, it emerges empirical properties of social networks such as homophily and a skewed degree distribution. Agents make use of the their personal contacts in order to get information about open vacancies when they are unemployed. The model requires few parameters in order to generate realistic labor and firm dynamics such as unemployment volatility, a skewed distribution of unemployment durations, income distribution, and correlation between skills and earnings.

Finally, we develop a third model to study how *small-world* labor flow networks emerge from search behavior. This model bridges the gap between well-studied search models from economics and popular models from physics about the formation of complex networks. By conducting computational experiments, we demonstrate that unemployment insurance has significant effects on some topological properties of the labor flow network such as global connectivity and local structure.

2. A SURVEY ON LABOR DYNAMICS AND NETWORKS

From all the markets in any country, the one of labor is the largest in size and extension. It reaches other markets and sectors in complex ways that have intrigued social scientists for centuries. The dynamics of labor markets have been the object of study in different fields and, still today, a topic that is difficult to understand. Labor market dynamics are typically studied from two perspectives: employees and employers. The former is conventionally termed as labor dynamics, and the latter is better known firm dynamics. Ironically, labor and firm dynamics have been treated as separate fields in economics for decades. Gaining insight about their connections and workings can improve our understanding about economies as a whole. Traditionally, economists and sociologists have studied labor and firm dynamics, leading to a variety of methodological approaches from highly abstract mathematical formalisms to state of the art large-scale computational simulations. In this chapter we provide an overview of the most popular methods to study labor and firm dynamics, with an emphasis in agent-based models of labor markets.

One of the first formal approaches employed in the study of firm dynamics was the use of stochastic processes. They can reproduce important stylized facts about the grow process of companies (e.g. firm size and growth rates distributions). However, they provide little insight about how economic behavior generate them. During the last

decades, economists have tried to couple labor and firms under the framework of dynamic general equilibrium. Since the late 1970's, this endeavor has flourished in the family of *matching models*. Unfortunately, the dynamics generated by these models are generated by exogenous stochastic processes. Therefore, they are useful for analyzing how economic behavior leads the market towards equilibrium in the presence of exogenous shocks. However, they are not well equipped to analyze dynamics under which labor and firms interact in a continuous disequilibrium and find jobs through personal contacts. Game theorists developed alternative models of labor dynamics through the strategic formation of social networks. These models can generate stationary dynamics by assuming agent-level strategic equilibria. The network topologies generated by these models tend to be unrealistic and static, which makes them difficult to use in order to analyze real labor markets. During the last two decades, agent-based models (ABMs) have taken important steps towards the construction of dynamical models, where heterogeneous firms and workers engage in realistic interactions that generate realistic dynamics. However, most of these models are still far from being usable with a real labor market.

Here we provide an overview of the main types of models employed to analyze labor and firm dynamics. We focus more on ABMs of labor markets for two reasons: (*i*) agent-based modeling is the chosen modeling framework for this dissertation; and (*ii*) we believe that providing an extensive account of the existing ABMs of labor dynamics is an important task in order to advance the field. This chapter is structured in the following way: In Section 2.1 we describe the main idea behind the matching models. Section 2.2

presents popular stochastic processes that reproduce firm dynamics. Section 2.3 summarizes the developments in the field of labor dynamics and social networks, studied under the lens of game theory. In Section 2.4 we present a detailed account of the most visible agent-based economic models of labor markets. Section 2.5 provides the conclusions.

2.1. The Conventional View of Labor Dynamics

The study of labor dynamics in economics originated from the need to overcome the limitations of a static understanding of labor markets. Overcoming this problem required a major change the way that markets were conceptualized by economic theory: understanding markets as decentralized processes/institutions (without a Walrassian auctioneer) that would produce equilibrium outcomes from the interaction of rational economic agents.

2.1.1. Search Models

Search models were the first approach to understand markets as decentralized equilibrium entities. The first one (Stigler, 1962) originated from the study of information in markets. It established that people can only search in a limited number of companies for a job. Then, the optimal number of searches is determined by the expected marginal returns of one more search and its marginal cost. When those two elements equate, the agent has reached the optimal search level. Once the search has finished, the worker picks the best offer.

Although illustrative, multiple aspects of Stigler's model did not reflect central features of job search. For example, most workers stop searching once they find a good enough offer instead collecting a fixed number of offers. Additionally, Stigler's formulation is a one-shot search, lacking a truly dynamic nature. S*equential search models* were created to address the shortcomings of the early matching models. In this framework, a job seeker sequentially requests offers from different potential employers. Each offer arrives with a proposed wage. If the offered wage meets the agent's expectations, the agent takes the job and halts the search. In order to predict when the agent is going to take an offer, these models are centered around the formation of a *reservation wage*. The first models of this nature were developed by (Gronau, 1971; McCall, 1970; Mortensen, 1970; Phelps, 1970). In order to obtain analytically tractable predictions about the reservation wage[1], these models employ tools from operations research in order to solve a dynamical optimization problem. The conventional approach consists of assuming an infinitely lived representative agent in order to obtain a time-invariant reservation wage. Multiple extensions of this framework were develop to understand how economic factors affect search behavior and unemployment dynamics. For comprehensive surveys on these extensions see (Lippman & McCall, 1976; Mortensen & Pissarides, 1999a, 1999b; Mortensen, 1987; Rogerson, Shimer, & Wright, 2005).

[1] Section 4.4 provides a detailed specification of this type of models.

2.1.2. Matching and DSGE Models

Sequential search models argued that decentralized rational economic behavior could produce stable labor outcomes with clear implications about labor dynamics. The next challenge was to show that such outcomes could be coupled with macroeconomic equilibrium. The *matching models* (Diamond, 1982; Jovanovic, 1979; Mortensen, 1987; Pissarides, 1984) brought together both sides of the labor market in order to reconcile the ideas of decentralized matching and macroeconomic equilibria. Most of these models employ *matching functions*: aggregate functions that map unemployed agents and vacancies to a number of matches to take place between job seekers and available jobs (usually with a Cobb-Douglas form: $M=U^aV^b$, where M is the number of matches, U is number of unemployed, and V is number of vacancies). For a given number of matches, a matching function becomes the Beveridge curve (often assumed with $a+b\leq 1$). Therefore, using matching functions is a convenient way to incorporate employment dynamics in macroeconomic models.

Given that a representative agent can form a reservation wage and that a representative firm determines offers and vacancies, the matching function determines important aspects of employment dynamics (e.g. hiring probabilities, hazard rates, unemployment duration, etc.). Multiple empirical studies have estimated matching functions in order to validate them (see (Petrongolo & Pissarides, 2001) for a comprehensive survey). However, deriving a matching function from microeconomic

behavior remains a challenge that has been studied only recently (Burdett, Shi, & Wright, 2001; Lagos, 2000; Shimer, 2007; Smith & Zenou, 2003).

The state of the art models that incorporate labor dynamics are the so-called *dynamic stochastic general equilibrium* (DSGE) models. They are mainly used to understand the associations between labor dynamics and the real business cycle. Real business cycles are the product of exogenous "technological shocks". Therefore, dynamic fluctuations in DSGE models are the product of exogenous shocks. DSGE models that concentrate on labor markets (Costain & Reiter, 2008; R. Hall, 2005; Shimer, 2005, 2010) couple stochastic processes and dynamic equations (e.g. consumption, production, etc.). This adds a noise term to the "laws of motion" of the economy, which generate macroeconomic fluctuations.

2.2. Labor Dynamics from Stochastic Processes

Before the conventional view of labor dynamics was established, stochastic processes were used in economics to explain important features of labor and firm dynamics (e.g. firm size and growth rates distributions). These models originated from the study of skewed frequencies in biology (Kapteyn, 1903) and the distribution of plant species (Yule, 1925). Here we focus on models that have been influential to understand firm dynamics. More detailed surveys about this type of models can be found in (Coad, 2009; de Wit, 2005; Mitzenmacher, 2004; Sutton, 1997).

2.2.1. Gibrat's Law

The first stochastic process applied to firm dynamics (Gibrat, 1931) explains the emergence of the heavy-tailed firm size distribution. It describes the growth in size x of a firm as a multiplicative process. Let ε_t be the proportional growth in time t. Then, the change in x is described by $x_t=(1+\varepsilon_t)\, x_{t-1}$. Approximating $\varepsilon_t \approx \log(1+\varepsilon_t)$ and simplifying yields $\log x_t \approx \sum \varepsilon_i$. Then, for long enough t, the initial size becomes irrelevant. Therefore, if all firm sizes are *i.i.d.*, $\log x_t$ is normally distributed. This is a lognormal distribution (a heavy-tailed one since the decay is non-exponential).

Gibrat named this process the *law of proportional effects*, since it describes the proportional growth that companies experience. Using data from French manufacturing establishments, he demonstrated that their sizes follow a lognormal distribution. Therefore, Gibrat's theory about how the labor force is distributed across firms was a plausible explanation. This model assumes a fixed population of firms and generates an increasing variance of firm sizes as t becomes larger. This is a problem because the observed volatility of firm sizes from real data is stable. A modified version of Gibrat's model was proposed by (Kalecki, 1945) in order to address this issue.

2.2.2. Simon's Models

Although Gibrat's model was useful to generate a lognormal distribution, it lacked other aspects that are important about how companies grow. For example, Herbert Simon found that the largest firm sizes are Pareto distributed (H. Simon & Bonini, 1958). Additionally, Gibrat's model lacked of realistic dynamics where firms enter and exit the

market. Based on the work of (Champernowne, 1953), Simon developed multiple models to address these issues. Here we focus on three of them (H. Simon, 1955, 1960), product of a discussion with (Mandelbrot, 1959).

The first model assumes a continuously growing population of firms. Workers are initially assigned to different firms such that all firms have at least one worker. Each period, a new worker enters the market and joins a firm. Once employed, a worker remains in the same firm all the time. With a constant probability, the new worker creates a new firm instead of joining an existing one. The probability that the new worker joins firm i is proportional to the number of firms of size x_i. In the second model, firms can enter and exit the economy. The probability that a firm exits is proportional to the number of firms of the same size. When a firm closes, all of its workers exit the market. For every worker that leaves, a new worker enters and is hired by a surviving firm in the same way as in the first model. Therefore, the population size is constant. Finally, the third model describes how firms gradually decrease in size. The probability that a worker exits the market is proportional to the number of firms of the same size as her employer. When a firm loses all of its workers, it closes down. All other rules are the same as in the second model.

In the three models, initial size of firms becomes irrelevant for long enough t. Therefore, Gibrat's law holds. They also generate a stable Yule distribution, which is a discrete analogue to the Pareto distribution. Other variants of these models (Y. Ijiri & Simon, 1964, 1977; Yuji Ijiri & Simon, 1967) deal with entry of firms with

heterogeneous sizes and exits independent of size. They also generate heavy-tailed distributions. This kind of model differs from the purely multiplicative process in that there exists a minimum size for every firm. This acts as a bouncing threshold that allows the system to generate larger firms than what is possible through a purely multiplicative process. That is why Gibrat's model generates a lognormal distribution and Simon's models produce power laws. A critique to Simon's models is that it takes a really large t to achieve convergence to a stable distribution (Krugman, 1996). Gregory Kesten addressed this problem by combining a multiplicative process with an additive term.

2.2.3. The Kesten Process

A more general and simpler process was proposed by (Kesten, 1973; Levy & Solomon, 1996). It consists of a multiplicative process and an additive term x^*. The former is the proportional growth of firms (Gibrat's law) and the latter represents the minimum firm size. Therefore, the growth process takes the form $x_t = \max\{(1+\varepsilon_t)\, x_{t-1}, x^*\}$. Clearly, $x^*=0$ takes us back to Gibrat's model, and a large x^* is suitable for the upper tail of a skewed distribution. This process quickly converges to a Pareto distribution. There is a crucial assumption in this model: ε_t is normally distributed. In the case of labor dynamics this is a problematic assumption since firm growth rates have peaked distributions. This issue was later addressed by econophysicists.

2.2.4. Econophysics and Growth Rates

The *tent-shaped* distribution of growth rates (Amaral et al., 1997; Ashton, 1926; Little, 1962; Stanley et al., 1996) is a stylized fact that has been modeled only recently.

Likewise, the negative relationship between growth volatility and firm size (B. Hall, 1987; Hart & Prais, 1956; Hymer & Pashigian, 1962; Stanley et al., 1996) seems to be an important property that was not modeled before. The former has been formally described as a double exponential distribution and the latter as a power law relationship.

A modified version of Gibrat's model was proposed by (Buldyrev et al., 1997; Stanley et al., 1996). It consists of a modification to the term of proportional growth ε_t. If a firm size is bellow the minimum size x^* then $\varepsilon_t=k(1+z_t)$, where k is larger than one and z_t is an uncorrelated Gaussian random number, otherwise $\varepsilon_t=(1+z_t)/k$. This modification generates a double exponential distribution of growth rates and supports Gibrat's law.

Another contribution came from (Bottazzi & Secchi, 2006). Here, firms are conceptualized as tuples of workers that grow proportionally by one unit each period. The probability of adding a new worker to a specific firm is proportional to its size. This combinatorial approach yields a Bose-Einstein firm size distribution. The distribution of the growth rates is tent-shaped and is well fitted by a Laplace law. Although econophysics has contributed to the refinement of stochastic models, all the stochastic processes previously discusses lack of microeconomic foundations.

2.2.5. Gabaix's Micro-Foundations

A model proposed by (Gabaix, 1999) tried to bridge the gap between microeconomic behavior and the empirical regularity of the Pareto-distributed firm sizes with exponent equals 1: *Zipf law*. Although this model was originally developed to

understand the growth of cities, it has been used to study firms (Saichev, Malevergne, & Sornette, 2009).

The model consists of utility-maximizing representative workers and firms. Workers receive utility from a combination of a wage and a benefit plan from their employers. Additionally, they lose their job at any moment with a constant probability. Firms offer wages and benefit plans. The marginal product of a new hire determines the wage offered by firms. Offers of benefit plans are randomly determined every period. When workers lose their job, they search for the best combination of wages and benefits across the economy. If the economy is in equilibrium, there exists an equation that describes the growth process of every firm in terms of utilities and layoff probabilities. This equation is identical for every firm and is independent from their initial size, preserving Gibrat's law. Through exogenous shocks to the system, this process converges to a Zipf distribution.

This was the first model to make a direct connection between utility-maximizing behavior of economic agents and the aggregate stochastic process that was previously developed by in the literature of firm dynamics. However, the implications of the model are quite trivial because the dynamics of the model still depends on exogenous shocks. Labor economists and game theorists have dealt with this issue arguing that networked structures in labor markets might be responsible for a significant amount of the dynamics.

2.3. Labor Dynamics and Networks

The study of labor dynamics and networks dates back to the work of (Boorman, 1975; Corcoran, Datcher, & Duncan, 1980; Diamond, 1981; Granovetter, 1973; Montgomery, 1991; Myers & Shultz, 1976; Rees, 1966). These studies showed that labor dynamics are the product of information transmission and economic interactions through networks. For discussions about the importance of networks in labor markets see (Ioannides & Loury, 2004; Jackson, 2011).

Some economic models attempted to explain how networks contribute to labor dynamics. Most of them stemmed from the recent development of models about network formation through games and games on networks. This approach allows the study of employment rates and unemployment duration as the outcome of information transmission through job contacts. If the sources of information about open vacancies are interconnected in specific patterns, the study of such structure could yield useful insights about how, when, and where do workers find new jobs. Furthermore, these models overcome the problems of lack of explicit interaction and reliance on exogenous shocks to generate dynamics. In general, the study of labor dynamics and networks can be divided in two types: dynamics on fixed (or exogenous) networks and dynamics with endogenous network formation. Here we provide an overview of analytic models from both classes.

2.3.1. Exogenous Networks

Models with exogenous networks concentrate on the effects that a given network structure might have on labor outcomes. In general, these models study networks of contacts through which the information about vacancies flows. The contact network is assumed to be constant, which allows the modeler to compute the probability that a job seeker hears about a vacancy. Conventionally, a worker receives information about open vacancies with a fixed probability and decides to pass it on to her contacts in case she is currently employed.

The first of these model was proposed by (Montgomery, 1991). It studies how the propensity to be connected to people with similar skills –*inbreeding social bias*– affects labor outcomes. It shows that inbreeding bias plays a crucial role in wage dispersion because it affects how firms set wage offers through formal and informal recruiting methods. In a later model (Montgomery, 1994) the author studies the effect of weak and strong ties in employment transitions. He finds that, as long as the inbreeding bias is low, weak ties have a positive effect on employment rates and equality. A more recent study (Finneran & Kelly, 2003) explores how referral networks favor groups of workers. The authors find that when the structure of the network is hierarchical, there is a critical density level that makes it practically impossible to receive any referral offer to those at the lower levels of the network. This is a structural problem of the market since there is no intended discrimination from firms or inbreeding bias from workers. Similar results about critical aspects of contact networks in employment were found by (Krauth, 2003,

2004). The author focuses in the interplay between human capital and the richness of social networks within groups. He finds that structural attributes of agents in the network (e.g. the proportion of weak ties) have nonlinear effects in employment rates.

Although the previous studies enrich our understanding about the effect of social networks in labor outcomes, they lack in dynamics. The first model that provides a dynamical perspective of how network structures impact employment was developed by (Calvó-Armengol & Jackson, 2004, 2007). It shows that networks are responsible for a number of empirical regularities about labor markets: positive correlation between wages and employment status, higher unemployment probability associated to longer unemployment durations, higher drop-out rates among poorly connected individuals, and higher unemployment in groups with higher initial drop-out rates. All these results are the product of contagion effects through the network. Unemployed neighbors become competitors for vacancies when the information comes from a friend that they have in common. Therefore, the properties of the network have direct consequences in the dynamics of unemployment. Finally, a study by (Calvó-Armengol & Zenou, 2005) bridged the gap between the matching function and social networks. The authors found that, for dense networks, unemployment rates increase with the size of the network.

2.3.2. Endogenous Networks

It is often the case that social networks are endogenously generated by economic agents who then use them as means of interaction. To address this issue, there are models where the structure of the network is the product of strategic interactions. The first of this

type of models was proposed by (Boorman, 1975). In this model, agents have to allocate an effort level in order to maintain their social relations, while confronted with an environment where they can lose their jobs. Unemployment equilibrium levels are dependent on the turbulence of the labor market (the separation rate). A spin-off of Boorman's model was developed by (Calvó-Armengol, 2004). Here, workers decide to form social ties by computing the expected payoff. Expected payoffs take into account the current configuration of the network topology in order to compute the probability of getting job offers in the future. The analysis is focused on deriving the topologies that would result from a Nash equilibrium. The model shows the importance of direct and indirect edges. The former are beneficial because they increase the chances of getting a job. The latter are detrimental because second-degree neighbors are competitors for the jobs that are transmitted by a common acquaintance.

Despite the contribution of game-theoretic models, the connection between the micro behavior and the macro effects is still not well understood. This is partly due to the difficulty of modeling labor dynamics and networks from a rational expectations point of view. The networks that emerge from this process have trivial topologies (e.g. stars, circles, complete graphs) that do not resemble real world structures. By using computation, it is possible to study richer and more realistic settings. Agent-based models of labor markets have been developed with this purpose.

2.4. Agent-Based Models of Labor Dynamics

The agent-based literature about labor markets is broad. Many agent-based macroeconomic models have a labor-market module embedded. Neugart and Richiardi (2012) provide an overview of different types of ABMs of labor markets according to their modularity and the nature of the interaction. Here we present an extensive review of models that focus exclusively on labor markets. Furthermore, we discuss only those that aim to understand economic labor dynamics. We exclude sociological, organizational, and general macroeconomic models.

2.4.1. Bergmann's Simple Microsimulation

This microsimulation model (Bergmann, 1990) resembles in spirit to ABMs, and is one of the earliest computational models of a labor market. It consists of a simple simulation with individual records instead of autonomous agents. It is a precursor of the ZI-traders model (Gode & Sunder, 1993). The author describes a simple yet compelling model of unemployment dynamics providing pseudo-code and steps of verification. The objective is to show how simulation models can be used to study the macroeconomic effect of unemployment insurance.

The model consists of two types of agents: firms and workers. During each period, firms fire and hire workers. The amounts of hiring and firing are exogenous parameters provided as a turnover time series. Workers can be either employed or unemployed. Unemployed workers are randomly matched with firms that have vacancies. When a match takes place, the worker is hired. An additional component is the

unemployment insurance. When a worker loses her job, she receives income for a number of periods. This is coded in the following rule: as long as a worker is qualified for unemployment insurance, she does not take any job offer. It is evident that the model is extremely simple. However, the author shows that it can provide valuable insights about labor markets.

The author calibrated the model with a turnover time series where there was a recession. During the recession period, firms have larger turnover rates, generating higher unemployment in the population. She tested two scenarios: with and without unemployment insurance. The results show that, during recession times, unemployment insurance did not increase the unemployment rate (as it is commonly believed by labor economists). The reason is because the rejection of job offers by insured agents opens a chance to an uninsured agent to become employed. Therefore, the effects that can be deduced form standard microeconomic analysis do not have a one-to-one correspondence to the macroeconomic outcomes.

2.4.2. Tesfatsion's Evolutionary Model

This model has been re-used in several studies (Pingle & Tesfatsion, 2003; Tesfatsion, 1998a, 1998b, 2001). Originally, it is an early contribution from Leigh Tesfatsion, who designed one of the first ABMs of a labor market. Her objective is to endogenously generate some aspects of labor dynamics such as entry/exit rates. She provides an experimental design in order to test hypotheses about how market tightness affects entry/exit rates. It is also one of the first models to explicitly treat contractual

networks. It is evident from the language and the presentation that is was written in a context where agent-based models were still rare, and concepts from evolutionary computation were becoming popular in economics. Although it is not clear why the author uses an evolutionary approach, it provides a fresh perspective to a field that is dominated by the matching function paradigm, and a natural mechanism for contractual networks to emerge.

The model has three types of agents: workers, employers, and workers/employers. Workers have a quota of how many offers they can receive, and employers have a quota of how many vacancies they can open. These quotas are used as a control for the market tightness. Worker/employer agents are versatile actors who can take either role. Their population is used to control for three different market structures: a two-sided market (no versatile agents), a semi-fluid market (some population of versatile agents), and an endogenous-type market (all are versatile agents). Workers and employers are matched through the Gale-Shapley algorithm and they play a prisoner's dilemma game that provides the utility derived from their contract. Similarly, workers are randomly partnered with colleagues at the workplace to play a similar game. Agents have memory about their payoffs, which define their strategies. Tesfatsion uses genetic programming to evolve these strategies and generate dynamics.

Despite its simplicity, the implementation of the model requires competent programming skills. It is a relatively large model, consisting of 19 parameters. It confirms some results from the standard matching literature. The author suggests that measures

employed to evaluate efficiency of static labor market models perform poor when it comes to evolutionary markets. She argues that contractual networks account for the poor performance of static indicators since they reinforce the persistence of certain behaviors (such as hiring and firing) depending on the market structure. Unfortunately, the lack of empirical data about these networks limits the analysis. Finally, she finds that job capacity is the main source of market power in aggregate terms. That is, the higher the quote of job offers (openings) the more power a labor supplier (consumer) can exert.

2.4.3. Tassier & Menczer Evolutionary Networks

This model (Tassier & Menczer, 2001) investigates labor market dynamics as the product of an evolving job referral network. Similarly to Tesfatsione's, this model employs evolutionary algorithms to generate dynamics. It provides a first step to understanding how endogenous social networks affect job search.

The model has workers and jobs. Initially, agents are connected at random. Jobs are fixed objects that can be vacant or taken. Each job pays a constant wage that is randomly drawn from a given distribution. Agents that have a job are fired at an exogenous separation rate. Every period, agents have to decide how may friends to have, and how much to spend in searching for a job. Agents obtain information about vacant jobs through direct search (which involves effort) and direct neighbors. There are constant costs for the number of friends and for the search intensity. By the end of a period, if a worker's accumulated wealth is greater than the total costs, she survives. In case of survival, the agent duplicates herself as long as her wealth is above a given

threshold. Therefore, the population size is endogenous to the model. Other sources of variation are introduced by adding random noise to the marginal costs and to the reproduction threshold.

The main result is the emergence of a small-world social network. The model sheds some lights about the economic motivations that may cause these networks to emerge (physics models of network formations are poor in this respect). Additionally, the authors found that the social network allows agents to survive longer because some individuals have a larger margin to select job offers, and the rejection of an offer becomes an opportunity for an unemployed agent. Due to the high content of evolutionary computation, this is a relatively large model with 11 parameters to calibrate.

2.4.4. Job Search Model

The purpose of this model (Sapienza & Fontana, 2002) is to explore labor outcomes, under different theoretical frameworks of job search. It compares labor dynamics under limit-size search against optimal stopping rule frameworks. This is more of an exploratory work, showing how agent-based modeling can be used to incorporate heterogeneity and networked structures in models of labor dynamics.

The model has workers and firms. Workers have skills that can be improved by investing in education. If workers do not invest in education and remain unemployed, their skills decay. Workers that reach welfare levels below a given threshold have to leave the market. Firms offer wages equal to the marginal contribution of the worker to

the firm's production. The productivity of a firm is determined by the average skills of its workers. Unemployed workers are randomly matched with firms, and if the firms' offers meet the workers' reservation wage, they form a contract. Additionally, the authors track the underlying networks of firms and workers. The benchmark case is where workers can be matched with any firm and both types of agents engage in contracts without restrictions. Other specifications include bounded search and skills mismatch.

The results of the benchmark case show poor in dynamics. Unemployment rate, workers' wealth, profits, and other variables converge to single values. When selection is removed, labor outcomes are more dynamic and with regular patterns. When bounded search and skill mismatch are introduced, the model generates richer dynamics. In the long run, the model seems to converge to a steady state. All the results are described qualitatively. The model has five parameters and depends on multiple laws of motions.

2.4.5. Fagiolo et al.'s Evolutionary Model

This model (Fagiolo, Dosi, & Gabriele, 2004) has the purpose of emerging three stylized facts of labor markets from microeconomic principles: (*i*) the Beveridge curve, (*ii*) the wage curve, and (*iii*) Okun's law. It is a refined version of a model previously developed by (Gabriele, 2002). It uses an evolutionary framework in the sense that failing firms exit the market and are replaced by new entrants. This is one of the first ABMs that reproduced empirically valid regularities of the labor market.

The model consists of workers and firms. Although the authors specify multiple behavioral rules for firms and workers, here we will only review the general framework of the model. Firms have linear production functions where labor is the only input. Workers form reservation wages based on their experience, and firms propose wage offers based on profits and applications received. When workers and firms are matched job applications are queued. Once all contracts have been formed, production takes place and aggregate supply and demand are computed in order to provide firms with information for the next period. Firms with negative profits exit the market and are replaced by "average firms". A key assumption is that all workers lose their jobs in every period. Despite its simple description, this model relies heavily in stochastic variations to generate dynamics. For example, firms and workers have exogenous noise when they update their expectations. The authors model technological change as a law of motion for the evolution of the productivity parameters of firms. The selection of behavioral rules seems ad-hoc.

The results of the model are robust. Under different behavioral specifications and parameterizations it gives rise to the three stylized facts previously mentioned. The authors demonstrate that the aggregate output is sensible to behavioral parameters. Hence, it sheds some light on how this kind of models can be used to assess labor policies. The model uses seven parameters, but most of them are difficult to calibrate empirically since they describe the random noise. Albeit its ad-hoc specification at the micro level, this model represents an important contribution to understand how ABMs can be used to emerge empirical regularities form the bottom-up.

2.4.6. Neugart's Endogenous Matching Function

This study (Neugart, 2004) evaluates the accuracy of the concept of the matching function by modeling the matching process as a decentralized process where workers and firms engage in labor relationships. The objective is to emerge the Beveridge curve from bottom-up and assess if its returns to scale are close to the ones empirically estimated.

The model consists of workers and firms. Firms have linear production functions where labor is the only input. There is a given quota of how many vacancies a firm can open. Firms decide how many vacancies to open at a constant marginal cost. Workers can submit up to a given number of job applications at a constant marginal cost. In an evolutionary spirit, firms and workers with negative payoffs exit the market and are replaced by agents with randomly copied strategies from the surviving population.

Neugart's model has only five parameters. It is able to emerge a negative relationship between the number of available vacancies and the unemployment rate. Its Beveridge curve is L-shaped, meaning that the model can reach full employment on one hand, and generate excess labor supply, on the other. These results are primarily induced by the market tightness specified by the upper limit of how many jobs each firm can open. The main result is that the emerging matching function has decreasing returns to scale, while those of the conventional matching function are assumed to be constant. The author also tests the sensibility of the matching function's parameters to microeconomic parameters.

2.4.7. Richiardi's Equilibrium Model

This paper (Richiardi, 2004) presents an analytic model and an ABM of labor dynamics. Its purpose is to provide a framework under which labor dynamics are closely related to firm dynamics. This model generates empirical regularities from both labor and firm dynamics. The analytic model is formulated as a Markov chain and the long-run probabilities of transitioning between employment and unemployment are characterized by the model's parameters. Then, the author agentizes the analytic model in order to study disequilibrium dynamics and stylized facts of firm dynamics. An innovation of the model is that worker agents can open their own firms, generating birth and death processes in the population of firms.

The model consists of workers and firms. Workers can be employed by firms, unemployed, or be entrepreneurs who start a new firm. There is an exogenous cost of opening a new firm, which increases with the amount of labor required. Every period, firms open a random number of vacancies (determined by an exogenous growth process) and close those that remain unfilled at the end. Wages change every period through idiosyncratic shocks on firms. This incentivizes workers to quit. Workers form rational expectations about the wages in all possible employment states. Matching takes place when workers submit applications to different vacancies. In the ABM, worker's expectations are adaptive, and the idiosyncratic shocks are correlated across firms.

The results are very rich for a model with 10 parameters. The model emerges the Beveridge curve and Okun's law. It generates skewed firm size and age distributions.

Unemployment rates have realistic orders of magnitude. The dynamics of the artificial economy produce realistic fluctuations and are associated to the business cycle. The model relies on the exogenous shocks in order to generate its dynamical properties. Its implementation is clean and robust. Its characterization in the analytic version provides a good idea of the general workings of the system, while its agentized version provides realistic labor dynamics. This model is an important contribution since it provides a well-grounded theory about labor dynamics, and realistic results that can be used for policy purposes.

2.4.8. Richiardi's Disequilibrium Model

This paper (Richiardi, 2006) presents an ABM similar in spirit to (Richiardi, 2004), but more parsimonious and incorporating endogenous dynamics. The purpose of the paper is show that the matching function is a feature of systems in disequilibrium. Hence, it questions the assumption of aggregate matching functions in equilibrium models. The model innovates by incorporating of skills, risk, and on-the-job search. It also provides very intuitive and empirically grounded micro-foundations for each behavioral specification.

The model consists of workers and firms. Workers have different levels of productivity that are visible to firms. They can be active (searching for a new job) or inactive. Both employed and unemployed agents can be active or inactive. Hence, workers go through a two-step decision problem. Firms pay wages that are a constant share of the workers' productivity. Workers submit applications to vacancies. Therefore,

workers can receive multiple offers and firms can receive multiple applications for each vacancy. Each firm determines how many vacancies to open according to an "entrepreneurial" parameter (the author refers to it as the heterogeneity of business ideas). Firms may close if they reach negative profits or if all their employers quit.

The main result is that the desired properties of the matching function only exist when the system is out of equilibrium. The model generates the Beveridge, Okun, and wage curves. It also produces skewed firm size and income distributions. The author performs policy experiments where the system undergoes massive layoffs. He demonstrates that some of the statistical properties of stylized facts in labor markets are sensible to these shocks. This paper provides evidence of the potential problems of employment policies that assume market equilibria when they are aimed to remedy economic shocks.

2.4.9. Neugart's Sectorial Model

The purpose of this study (Neugart, 2006) is to evaluate the effect of training subsidy programs using an ABM. This model innovates by introducing multiple sectors and sector-specific skills. The model is simple and allows the author to design a clean way to identify and measure intended and unintended effects of labor policies.

The model has workers and firms. Firms belong to one of a finite number of sectors. Sectors are enumerated and sequentially arranged in a circle. The distance between sectors represents their similarity. Therefore skills between adjacent sectors are

more easily transferable. The aggregate labor supply and the demand are constant. There are no bargaining frictions, so workers take job offers and firms hire workers as long as their skills qualify. Unemployment is generated by exogenous sectorial shocks. When a shock takes place, all firms in the affected sector close. For each firm that closes, a new firm opens in a non-shocked sector, ready to hire job seekers. Workers can invest in their skills in order to make their skills compatible with another sector. The more they invest, the more sectors will be able to employ them. Agents decide their levels of investment through reinforcement learning. Experiments are performed by placing a subsidy to the cost of investing in human capital.

The main result of the model is that subsidies to investments in human capital reduce unemployment rates. However, there are crowding effects from these policies because those workers that were not treated by the subsidy are not able to get the jobs that the treated population takes. The model consists of 10 parameters. The implementation is straightforward and the experimental design allows the author to find clear evidence of the macroeconomic effects of training subsidy programs. This study is an important contribution for demonstrating how ABMs of labor markets can be employed in the design and evaluating of labor programs where micro-data studies cannot tell much about macroeconomic effects of labor policies.

2.4.10. French Labor Market Model

The purpose of this model (Lewkovicz, Domingue, & Kant, 2009; Lewkovicz & Kant, 2007, 2008) is to provide a computational tool that can be used to generate specific

features of the French labor market. More specifically, it is used to evaluate the age discrimination in hiring by French firms, and the effects of introducing of a new type of contract. The new contract intends to promote higher employability of the French youth by allowing a temporal hire of two years without accumulation of seniority. This model builds on the French tradition of microsimulation models of labor markets, such as the evolutionary models of (Ballot & Abraham-Frois, 1980; Ballot, 2002) and ones about the academic job market by (Caillou & Sebag, 2008). This paper contributes by providing an ABM that can be used for empirical purposes.

The model consists of workers, firms, and employers. Workers age with time, enter and exit the market, search for jobs, apply to vacancies, and work according to their productivity. Firms create vacancies according to their budget, hire through the employer agents (who rank applications), and fire underperforming employees. Employer agents are an important component of the centralized matching process that takes place in the French labor market. They play a crucial role in the publication of vacancies and the management of applications. Additionally, there is a government agent in charge of enforcing different types of contracts. When firms receive applications for a vacant position, they decide to provide an offer based on the ranking of the applications and the availability of types of contracts. There are two kinds of contracts: definite term contract and indefinite term contract. A third type of contract –*unified contract*– allows firms to hire employees for two years, without obligation to re-hire them, and without accumulation of seniority. When this contract ends, the firm has to re-hire the worker

with one of the other two types of contracts. Additionally, firms have a quota for the latter contract.

The results show that the model is capable of reproducing the empirical levels of unemployment for the French youth (which are twice as much as the rest of the working force). The authors introduce the third type of contract in the simulation in order to evaluate its effects on unemployment rates. Their main finding is that, in the short run, youth unemployment drops 6%. In the long run, total unemployment rises 0.3% more than what it was before the introduction of the unified contract. However, youth unemployment drops 5%, and 14% for middle age workers. An unexpected result is that unemployment rates for workers 50 years or older increases more than 3%. This raises questions about possible displacement effects due to the new contract.

2.4.11. Shocks and Endogenous Institutions

This work (Martin & Neugart, 2009) is based on Neugart's sectorial model (see Section 2.4.9), with the innovation of endogenous labor regulations. This study is a unique contribution to the field since it takes into account the regulations of the labor market as the product of a political process where the agents vote for the type of regulation they want. Like in Neugart's model, there are sectorial shocks. This generates interesting dynamics between the exogenous economic shocks and the transitions between two regulatory regimes. The paper also presents econometric analysis of the artificial data in order to identify different effects and draw inference about how different regulation schemes affect unemployment rates.

This model has workers, firms, and two political parties. The economic interactions follow the exact same protocol as in Neugart's sectorial model. In this model, workers are also voters. One political party is in favor of imposing severance package requirements to firms, and the other is against it. Workers cast their votes every number of periods, based on their average earnings during the legislatures of each party, and based on their ideological level. When a party is elected, it comes into power and implements its corresponding regulation. The severance package regulation is modeled as a quota to the maximum number of vacancies that firms can hold. This feature tries to capture the response of firms to higher costs of job destruction.

The main result of the model is that, on average, labor regulation has no statistically significant effect on unemployment rates. Perhaps the most interesting result is that the effect of economic shocks on the endogeneity of regulatory schemes depends in their duration. Longer shocks allow voters to relate their earnings to the party in power, producing votes against it. The model is significantly larger than Neugart's version (15 parameters). However, it is of a unique kind and questions the way in which labor policies are evaluated when they are created in response to economic shocks.

2.4.12. Jobs and Social Networks

This study (Tassier & Menczer, 2008) analyzes the effects of social network structures in group unemployment rates. Although it is close to the strand of ABMs that study segregation from a sociological perspective, it addresses the important and unexplored problem of the underlying structure of the matching process, which is critical

to understand labor dynamics. It innovates by bringing together the notion of referral hiring through social networks and information transmission about jobs through an underlying job network. It provides a framework to study how segregation and network randomness affect unemployment levels of different groups in a population.

The model consists of workers and jobs. Jobs are connected in a directed circular lattice that remains constant throughout the simulations. Workers are connected to each other and organized in two groups. Each group is connected in a circular lattice with variable randomness. When a worker is employed, she knows about vacant jobs through the job network. When a worker is unemployed, she searches for a job through her social network. There is a rewiring process of the social network. Each group rewires with different probabilities. Additionally, there is an exogenous probability that a new rewiring will end up connecting workers from different groups. Hence, the authors are able to control for the segregation levels in the social network and its randomness.

The results show that, in general, random networks induce higher employment levels than regular ones, except when there are high segregation levels. Another important finding is that employment levels of groups with non-random connections decrease when the population size decreases. The model has six parameters and sheds some light on how the underlying structure of the matching process affects group employment dynamics.

2.4.13. EURACE

The EURACE project (Deissenberg, van der Hoog, & Dawid, 2008) is a collaboration from multiple European research groups to conform a large-scale ABM of the European economy. As part of the project, the group from the University of Bielefeld developed a labor market model. This model has been used to study multiple features of labor markets (Dawid et al., 2008; Dawid, Gemkow, Harting, & Neugart, 2009), providing interesting policy insights. Since EURACE is a complete macroeconomic model, it is too large to be described in a few lines. Therefore, we provide a simple overview of the main features regarding the agents and the dynamics of the labor market model. The main contribution of these studies is to understand how the spatial distribution of regional labor markets affect economic growth and the effectiveness of policies that aim to upgrade workers' skills.

The model has workers and firms. Workers produce when they are employed and consume what firms produce through regional markets. Firms produce consumable goods using capital goods and labor. Firms and labor are spatially distributed in two regions. When regional markets run out of stocks, they place orders to the supplying firms. Firms adjust their production according to the inventory dynamics. Adjustments in production generate unemployment and vacancies. Firms advertise vacancies. Job seekers apply to advertised vacancies and form their reservation wages using their discounted last wage. Job seekers can be unemployed agents or on-the-job searchers. Workers have different types of skills and different levels for each type. Firms rank the received applications

according to types and levels of skills. The bargaining between firms and applicants go on for two rounds. Employed workers incur in commuting costs depending on how far they are from their employers.

The main results are that spatial distributions have a significant effect in economic growth and in the effectiveness of training policies. If the two regions are almost isolated from each other, a uniform distribution of the training policy yields better results (in terms of economic growth) than localized programs. When the interaction between regions is restricted by commuting costs, the results are the opposite, favoring a focalized targeting for an intensive skill upgrade in a single region rather than a medium-level general upgrade. The results also show significant effects between regions in terms of prices and income. The model consists of 16 parameters and is very complicated. It is a proof of concept of how ABMs can be used for policy testing.

2.4.14. Endogenous Social Networks

The purpose of this model (Gemkow & Neugart, 2011) is to study the effects of social networks in unemployment dynamics and inequality. The social network in this model is endogenous. Workers use their contacts as a way to insure themselves against unemployment. The authors employ behavioral models to motivate how agents decide the number of friends they will have. It is one of the first ABMs that connect the strategic decision of increasing social capital with labor outcomes.

The model consists of workers and firms. Firms demand labor stochastically, from a given distribution. The higher the variance of this distribution, the more turbulent the market becomes. Workers receive wages and maintain friendships at a constant marginal cost. Determining how many friends to have is achieved through the expected-weighted attraction algorithm. The matching between workers and firms follows the same protocol as EURACE, with the difference that firms prioritize workers that are referred by a current employee. There is a second matching process for those agents that have decided to make more social relations. Therefore, any agent can become acquainted with any other agent in the economy.

The results confirm what previous studies have showed about the importance of social relation: in more turbulent markets, social networks become less important because many acquaintances are also competing for jobs. The authors perform an experiment in which they exogenously raise the number of social contacts in order to study its effect on inequality. The treated groups experience higher employment levels, suggesting that social networks reinforce employment inequality. The model is large (15 parameters), mostly due to the learning algorithms. It provides a first guide on how to incorporate endogenous social networks into labor market models. This model was combined with the EURACE framework in (Dawid & Gemkow, 2013) in order to study how social networks affect wages. The authors found that richer social networks increase wage inequality inside groups but not between groups.

2.5. Conclusions

The study of labor dynamics dates back to the beginnings of the 20th century. Here, we reviewed what we consider to be the main families of theoretical models, according to their methodology. We have focused on economic models where the labor market is the object of study. We have emphasized in ABMs of labor markets since they have experienced more active development in the last years. All models have advantages and limitations; we believe that ABMs have an edge. Our main reason to believe this is twofold.

First, numerous studies have shown that networks have a significant effect in labor markets. Conventional models and stochastic processes that are not able to consider networks are ignoring a central component of labor markets. Game-theoretic models of networks are limited in the realism of the networks that they can analyze. For ABMs, networks and dynamics are natural components. This does not mean that they are always easy to model or that every ABM model is a good representation of real labor markets. Further developments of method and hardware are required to turn ABMs into useful scientific and policy tools. Second, computer technologies and large-scale micro-data are making it possible to develop more realistic and detailed models of the economy. The natural way to exploit them is through ABMs. In contrast with other types of models, agent-based modeling is evolving day by day, tagged to the latest development in hardware and programming paradigms. This opens a window of opportunities for future

ways of simulating and understanding real micro-empirically-grounded models of labor dynamics.

3. EMPLOYMENT GROWTH THROUGH LABOR FLOW NETWORKS

3.1. Introduction

Employment dynamics are the product of complex interactions taking place inside and between firms. In labor markets, human resources are continuously reallocated across firms, industries, and regions. In labor economics it is conventional to aggregate job hirings and job separations (both voluntary and involuntary) across companies to get pools of job changers and the unemployed (S. J. Davis, Haltiwanger, & Schuh, 1998). The sizes of these pools are then conceived of as being determined by rate processes over these pools (Mortensen & Pissarides, 1994). In reality, hiring and separation occur at individual companies and important information about the varieties of firm behavior is lost in the process of aggregating labor data into pools, with otherwise comparable firms experiencing quite different labor turnover. For instance, understanding how micro-dynamics affect aggregate variables (such as employment growth) from a disaggregate perspective is an ongoing challenge. We demonstrate that the science of complex networks can be helpful in tackling this problem.

Over the past fifteen years the important role of networks in human society has become readily apparent, from the topology of the internet (Albert, Jeong, & Barabási, 1999; Faloutsos, Faloutsos, & Faloutsos, 1999) to the rise of social media. In many areas of science the growing availability of micro-data has made possible the systematic study

of networks (e.g., citation networks (Redner, 1998)) while in other domains the growth of computing power has led naturally to network conceptions of social processes (e.g., epidemiology (Halloran, Longini, Nizam, & Yang, 2002; Longini et al., 2005)). In economics the study of networks has essentially revolved around strategic concerns and a game theoretic orientation has become the norm (Bala & Goyal, 2000; Galeotti & Goyal, 2010; Jackson & Wolinsky, 1996). Fewer studies analyze economic networks that are the side product of other kind of interactions, instead of being the intended consequence of strategic behavior. Networks of companies are an example of such structures. The first studies of large-scale complex networks of firms were made for the Japanese economy, including ownership networks (Souma, Fujiwara, & Aoyama, 2006) and costumer-supply networks (Konno, 2009; Saito, Berger, & Iwamura, 2007). More recent studies have done similar analyses for the US (Atalay, Hortaçsu, Roberts, & Syverson, 2011).

Here we blend these motivations for studying networks, using newly available micro-data and the ability to work with large-scale, complex networks computationally, to study labor dynamics. This study contributes to the field of labor economics by providing the first characterization of a labor flow network for an entire economy and by identifying the relationship between its topological properties and employment dynamics. It also contributes to the field network sciences by providing evidence of complex networks in labor markets and their relationship to their underlying social processes. Our findings are robust for different countries. With this study we introduce a new way to analyze labor markets by coupling data from labor and firm dynamics through the usage of complex networks analysis.

This chapter is structured in five main Sections. In Section 3.1 we introduce the reader to the current state of understanding labor dynamics and networks. Section 3.2 presents all the materials and methods that we used for our analysis. In Section 3.3 we show all the results of our study. Finally, Section 3.4 presents a discussion on how LFN are relevant to understand labor dynamics.

3.2. Materials and Methods

In this Section we describe the data sources and main methods employed for this study. Sections 3.2.1 and 3.2.2 describe employer-employee matched micro-datasets form Finland and Mexico respectively. In Section 3.2.3 we describe a complementary dataset from the Mexican economic census. Section 3.2.4 discusses the differences in statistical units between our various datasets. In Section 3.2.5 we define the concept of a *labor flow network* and describe how we construct them. We introduce the concept of null models in Section 3.2.6, commonly used in statistics for complex networks. Section 3.2.7 introduces a standard measure of employment growth for firms and presents conventional classifications of firms according to the intensity of their growth. Finally, Section 3.2.8 presents the statistical distributions used to characterize multiple of our empirical findings.

3.2.1. Primary Dataset

We used a comprehensive dataset about labor and companies in Finland. Most of our results derive from it. It contains the universe of employed individuals in Finland and their employers (both from the private and public sectors). FLEED's employer units are

enterprises, defined as *economic activity carried on by one or more persons for profit-making purposes*. Although this is a broader definition than the one conventionally used for firms, we will use it interchangeably since it does not change our ability to measure employment growth. Unless otherwise specified, all the analysis was conducted using this dataset.

This dataset is called Finnish Longitudinal Employer-Employee Dynamics (FLEED) and is provided by Statistics Finland. FLEED contains annual registries of every permanent resident in Finland that is employed. These records are constructed from administrative registries by extracting the social security number of the employed individuals and the identification number of their respective employers. Individuals and firms are anonymized through unique identifiers. FLEED consists of annual Panels with pairs of identifiers: employee and employer. Each Panel is constructed using the record available for each employed individual on the last day of the year. Therefore, if an individual is unemployed during the last day of the year, she will not appear in corresponding Panel. FLEED only captures annual movements of individuals and does not distinguish between workers who underwent unemployment spells and those who were job-changers. For most of the analysis, we used FLEED's Panels from 2005, 2006, 2007, and 2008. On average, each Panel contains 230,000 employer identifiers.

We merged FLEED with Statistics Finland's Business Registries in order to obtain accurate information about the size and age of each employer. These registries consist of annual Panels of the universe of firms in Finland. They provide information

about number of employees, year of birth, and year of death. In order to prevent identification of individual firms through their size, this variable was treated with a log-normally distributed random noise by Statistics Finland. We linked these records to FLEED's employers' identifiers.

3.2.2. Support Dataset

We used a sample dataset from Mexico that we obtained in order to evaluate the robustness of our results. Although it is similar in size to the primary set, it does not comprise the universe of Mexican firms and workers. Its nature and sampling method makes it prone to be biased in ways that the primary dataset is not. Nevertheless, it is a useful source of information to evaluate the robustness of our results. It was used exclusively to test the robustness of some of our results with the Finnish primary dataset.

We obtained this micro-dataset from the Mexican Institute for Social Security (IMSS after its acronym in Spanish). Like in the Finnish registries, the IMSS data contains records with anonymized pairs of individuals and employers. This set only contains records from the formal private sector. Approximately half of the employees in the private sector are not registered with the IMSS. Thus, they are considered informal workers. Additionally, nearly 16% of all Mexican workers are state-employed, so they are not in the IMSS records either.

The IMSS dataset has daily resolution. When an individual joins the formal private sector, a record is written in order to link her to the current employer and the

exact date she joined is recorded. If a worker joins a different employer, the new pair of identifiers is recorded with the date of the movement. The sample consists of 1% of all registered individuals in 2008. Once individuals have been sampled, their entire labor history was extracted from the database, i.e. for each individual, all the identifiers of her past employers are listed with the respective days in which she joined them. Therefore, an employer appears in the dataset as long it employs at least one individual from the sample. In total, our sample consists of 400,000 individuals, with an average of 10 records each. Roughly speaking, this dataset contains 270,000 employer identifiers.

3.2.3. Census Dataset

We obtained data about the firm size distribution of the universe of companies from Mexico. This dataset comes from the National Economic Census, and was generated by the Mexican National Institute of Statistics, Geography, and Informatics (INEGI for its acronym in Spanish). It consists of a geographical counting of the number of employees per enterprise (not establishment) in 2004. We obtained multiple binning configurations from this set, which allowed us to construct the empirical cumulative distribution of the sizes of all companies. INEGI data cannot be merged with the one coming from IMSS. However, it was useful to have an idea of the bias of the IMSS sample when employed to estimate firm sizes.

3.2.4. Statistical Units

Table 1 shows the main properties of FLEED and IMSS datasets. Each one uses different definitions of employers. In FLEED, employers are enterprises. Statistics

Finland defines enterprises as "any economic activity carried on by one or more persons for profit-making purposes". They can be either public or private. In the IMSS records, an employer is defined by the Federal Labor Law as "the person or entity that makes use of the services provided by at least one individual". For the purpose of this study, we will use different terminologies, such as firms, enterprises, establishments, and employers interchangeably. As we will show, the topological properties of the networks produced from both datasets are robust regardless of the nature of the statistical units.

Table 1: Employer-employee matched datasets

Property	Finland FLEED	Mexico IMSS
Employees from the universe	100%	1%
Total period	1988-2008	1980-2010
Number of firms for total period*	860,384	270,317
Study period	2005-2005	2005-2008
Number of firms during study period*	408,164	144,656
Frequency of records	Annual	Daily
Biases	None	Only private sector; informal workers not registered

* Firms that were registered at any moment during the corresponding period. This includes companies that disappeared before the end of the period. In the case of Finland, it also includes firms with zero employees.

3.2.5. Labor Flow Networks

Consider a network in which the firms are vertices and an edge is drawn between firms whenever a person has worked at one company and subsequently moved to another. For an economy as a whole we call this the *labor flow network* (LFN). We use FLEED data to construct the LFN of Finland. The motives and means of individuals to move

from one employer to another are diverse, ranging from economic incentives and unemployment spells to personal contacts and geographic relocation. This network implicitly captures most of these factors since it is constructed from the actual labor flows of the economy. We believe that studying its structure and its relation to other elements of firms' dynamics can improve our understanding of labor dynamics and the role of firms in employment growth.

The construction of a LFN is rather simple. For a selected period, we count the total flows of labor between every two firms in both directions. Although this is a directed network, we found that the most interesting insights come from studying its structural properties as an undirected graph. Therefore, our analysis uses algorithms for undirected networks (with exception of in-degree and out-degree centralities).

3.2.6. Null Models

When data is available for a particular network, it is the case that such a network is one realization of a social process. In our case, our main LFN is the reallocation of labor across Finnish firms between the last day of 2005 and 2008. If we were able to let the Finnish workers and companies to search again, assuming the same conditions of 2005-2008, it is possible that some properties of our LFN would not be found in the new LFN. In that case these properties are not robust and we cannot draw correct inferences from them. However, labor and firm dynamics at that scale are not easy to replicate under an experimental setting. For these situations, null models are useful to draw better inferences.

Null models were introduced by (Maslov & Sneppen, 2002) and, examples of how they are used with economic data can be found in (Fagiolo, Squartini, & Garlaschelli, 2013; Hidalgo & Hausmann, 2009). The main idea is to take the network provided by the data and randomize its structure while fixing some of the properties of the nodes (usually the degree). We need to create a sample of these randomized networks in order to estimate the parameters of interest. In our case, we generated 50 randomized LFNs. We indicate when an estimate was drawn from this procedure.

3.2.7. Employment Growth and Firms

In order to analyze employment growth we employ metrics from the small-business literature. Such measures typically extend for a defined period (3 or 4 years), and are based on changes in the sizes of companies. In this study, firms' sizes are measured as the number of employees. Since we are analyzing data from a European country, we use the metric developed in the OECD/Eurostat methodology (European Communities & OECD, 2007). Call S_t the size of a firm at time t, and S_0 its initial size. Then ΔS is the *average annual growth* of such a firm for the period between t=0 and t=T:

Equation 1

$$\Delta S = \sqrt[3]{\frac{S_T}{S_0}} - 1,$$

Growth is not defined for firms with zero employees initially, as is evident from Equation 1. We added one employee to all firms in order to compute Equation 1 for firms with zero employees. Another common problem is defining growth for companies that ceased to exist at $t<T$, and for those that were created at $t>0$. We resolved this issue by focusing our analysis on those firms that were born in 2005 or before and still existed in 2008. We call them *survivors*.

Following Equation 1, we measured growth of all *survivors* and classified them, according to three standard taxonomies employed in small-business research and added two more.

- **Positive growth firm (PGF):** *A survivor firm with positive growth.*

- **Labor flow network firm (LFNF):** *A PGF with at least one connection to another firm in the LFN.*

- **High-growth firm (HGF)** (European Communities & OECD, 2007)**:** *A PGF with average annualized growth greater than 20% per year.*

- **Gazelle firm (GF)** (European Communities & OECD, 2007)**:** *A HGF that is at most 5 years old.*

- **High-impact firm (HIF)** (Acs, 2011)**:** *A HGF with an employment growth quantifier (EGQ) of two or more.*

The EGQ is defined by (Acs, 2011) as: *the product of the absolute and percent change in employment, expressed as a decimal*. The definition of a HIF has been slightly modified (by adding the age requirement) to use the OECD definition of a GF.

3.2.8. Distributions

This study makes use of different statistical distributions in order to characterize some of the most relevant empirical regularities in the data. Here we introduce them and the methods employed to estimate their parameters.

3.2.8.1. Pareto Distribution and Power Law

The Pareto distribution is a well known, but not so commonly used distribution in economics. It heavy-tailed and its CDF takes the form:

Equation 2

$$\Pr[X \geq x_i] = \left(\frac{x_0}{x_i}\right)^{\alpha},$$

where α is known as the scaling parameter, and x_0 is the smallest unit in the data. We estimate the Pareto distribution using maximum likelihood.

Sometimes variable y decays with the growth of variable x at the same rate as in a Paretian function. In these cases, the relationship is commonly called a power law. We denote a power law with the following form:

Equation 3

$$y = \beta x^{-\alpha},$$

where β, is a normalizing parameter homologous to $(x_0)^{\alpha}$ in Equation 2. For this relationship, we estimate the parameters using OLS regressions on the logarithmic transformation of x and y.

3.2.8.2. Subbotin Distribution

This distribution is also known as power exponential or exponential power. More specifically, we use the skewed version of this distribution since some of the empirical regularities have an asymmetric nature. The PDF of the skewed Subbotin distribution takes the form (Ayebo & Kozubowski, 2003):

Equation 4

$$f(x)=\begin{cases} \dfrac{\alpha\kappa}{\sigma\Gamma(1/\alpha)(1+\kappa^2)}\exp\left\{-\dfrac{\kappa^\alpha}{\sigma^\alpha}(x-\theta)^\alpha\right\} & if \quad x\geq\theta, \\ \dfrac{\alpha\kappa}{\sigma\Gamma(1/\alpha)(1+\kappa^2)}\exp\left\{-\dfrac{1}{\sigma^\alpha\kappa^\alpha}|x-\theta|^\alpha\right\} & if \quad x<\theta, \end{cases}$$

where $\Gamma(\cdot)$ is the standard gamma function, θ is the location parameter, κ generates the skewness, σ is the scale parameter, and α is the shape parameter. We estimate the four parameters of the distribution using the grid method used by (Perline, Axtell, & Teitelbaum, 2006). The estimated parameters are the ones that minimize the Kolmogorov-Smirnov statistic.

3.3. Results

In this Section we present our main findings. In Section 3.3.1 we present a short comparison between Finland and Mexico in order to provide some background on the structural differences that might underlie our datasets. In Section 3.3.2, we confirm the heavy-tailed nature of firm dynamics in our datasets. We characterize the topology of the Finnish and the Mexican LFNs in Sections 3.3.3.1 and 3.3.3.2 respectively. Both networks have statistical properties of complex networks. Section 3.3.4 presents a correlation analysis of the relationship between economic properties of firms and their structural position in the LFN. We find that multiple of these patterns are well defined and highly non-linear, suggesting a multiplicity of regimes in different groups of firms. In Section 3.3.5 we employ community detection algorithms to let the Finnish LFN tell us

what are the most statistically robust communities of firms when it comes to labor reallocation. We find that the configuration of these communities does not match the one suggested by conventional ways of grouping firms (e.g. industries and municipalities). Section 3.3.6 presents a study of the firm growth and LFNs. We show that that the position of a firm in the LFN can provide strong indicators of future potential growth. In Section 3.3.7 we present a firm taxonomy using the network properties of companies. We are able to detect a few classes of firms with interesting features such as high shares in employment creation and destruction. Finally, in Section 3.3.8 we illustrate the dynamical process through which a LFN if formed. We show that it only takes a couple of months of labor flows to achieve a state in which the LFN shows the robust stylized facts documented in Sections 3.3.3.1 and 3.3.3.2.

3.3.1. Brief Comparison

As indicated in the Section 3.2, we used two datasets in order to test the robustness of our results. Lack of more comprehensive data for Mexico prevents us from doing any kind of comparative analysis. However, in order to provide some background regarding structural differences between Finland and Mexico, Table 2 presents macroeconomic information about Finland and Mexico during the period under study.

Although Mexico is considerably larger than Finland (both in population and GDP), the Nordic economy produces nearly twice as much as Mexico in per capita terms. The Finnish government provides unemployment benefits while in Mexico this is a private service that only a fraction of the formal sector acquires. Therefore, a Finnish

worker has incentives to remain unemployed longer. This is reflected in some of Finland's indicators, such as the persistent higher unemployment rates, lower self-employment rates, and considerably higher long-term unemployment rates. One of the most noticeable structural differences is the role of technology and R&D. The share of ICT to value added from the Finnish business sector is higher than that of Mexico. Additionally, Finland is known to be one of the countries with the highest investment in R&D as a percentage of its GDP (more than 3%). Finally, the structure of the manufacturing sector in terms of the firm size distribution seems to be the opposite between the countries. Given that both economies show remarkable differences, we expect that evidence of common features between both LFNs would be an indicator of robustness.

Table 2: Comparison between Finland and Mexico

	2005		2006		2007		2008	
	Finland	**Mexico**	**Finland**	**Mexico**	**Finland**	**Mexico**	**Finland**	**Mexico**
Population[1]	5,246.10	103,946.90	5,266.27	104,874.30	5,288.72	105,790.70	5,313.40	106,682.50
GDP[2]	161.10	1,293.79	174.53	1,439.30	191.28	1,530.84	202.34	1,627.07
GDP per capita[3]	30,707.92	12,460.54	33,140.17	13,740.55	36,167.38	14,485.97	38,080.46	15,267.18
Employment rate[4]	68.52	59.65	69.58	60.95	70.46	61.06	71.25	61.31
Self-employment rate[5]	12.67	35.54	12.90	34.46	12.65	34.34	12.85	33.94
Part-time employment[5]	11.20	16.82	11.41	16.96	11.71	17.57	11.50	17.58
Unemployment rate[6]	8.30	3.60	7.70	3.60	6.90	3.70	6.40	4.00
Long-term unemployment[7]	24.88	2.33	24.82	2.55	22.97	2.72	18.17	1.65
GDP per hour worked[8]	17.90	5.46	19.39	6.02	21.20	6.40	22.56	6.75
GDP on R&D[9]	3.48	0.41	3.48	0.39	3.47	0.37	3.72	n/a
Share of ICT in value added[10]	n/a	n/a	n/a	n/a	n/a	n/a	13.88	4.99
Firms in MFG, less than 10[11]	83.73	11.49	83.28	11.00	83.15	9.45	81.64	n/a
Firms in MFG, 50 to 249[11]	3.65	42.55	3.75	42.40	3.75	42.74	3.90	n/a
Firms MFG, more than 250[11]	1.01	27.67	1.03	28.83	0.99	30.01	1.06	n/a
Total number of firms	246,149	3,001,610	291,560	n/a	322,108	n/a	332,586	4,724,892

1 Thousands
2 Billion US dollars, current prices and PPPs
3 US dollars, current prices and PPPs
4 Share of persons of working age in employment
5 As a percentage of total employment
6 As a percentage of labor force
7 Persons unemployed for 12 months or more as a percentage of total unemployed

8 GDP per capita divided by the average of total hours worked annually by a person
9 Percentage of GDP invested in research and development
10 Percentage of the value added from the business sector that comes from the Information and Communication Technologies (ICT)
11 Number of firms in the manufacturing sector (MFG) with size in number of employees

3.3.2. Firm Dynamics

Some of the most robust evidence about the heavy-tailed character of firm dynamics is (*i*) the skewed firm size distribution, (*ii*) the tent-shaped growth rates distribution, and (*iii*) a power law relationship between the size of a firm and the volatility of its growth. These features have been documented for different countries and a wide variety of datasets. The oldest of these empirical regularities is the firm size distribution, first documented by French economist Robert Gibrat (1931). Using a sample of US firms, Ijiri & Simon (1977) showed that its tail followed a Pareto distribution. Using data from the universe of firms in the US, Axtell (2001) found that the entire distribution was a Zipf law . In fact, the last 20 years have seen nearly 50 different studies about the firm size distribution for more than 80 countries, using diverse data sources. In most cases they confirm the findings of the previously mentioned studies (see Appendix for a list of these studies).

Another empirical finding that has been thoroughly documented is the distribution of the growth rates of companies. First discovered by Stanley et al. (1996), this empirical regularity has been found in different countries. Its shape tends to be tent-shaped when plotted on a log-log scale. Amaral et al. (1997) found similar results growth in sales from

publicly traded firms in the US manufacturing sector. These results have been confirmed for several other countries (Bottazzi, Cefis, & Dosi, 2002; Bottazzi & Secchi, 2005; Fabritiis, Pammolli, & Riccaboni, 2003; Fagiolo, Napoletano, & Roventini, 2008). Most of these results argue that the distribution of the growth rates is a Laplace. Using data of the universe of firms from the US, Perline et al. (2006) showed that this distribution is asymmetric and follows the skewed Subbotin in Equation 4.

An additional result about firm dynamics is a power law relationship between the initial size of a company and the volatility of its growth rate. Measured as standard deviation, the volatility of growth of firms declines with their size. This was found by Stanley et al. (1996) for the US and has been documented for multiple datasets (Amaral et al., 1997; Matia et al., 2004). Other studies have showed similar results for firms and GDP in countries (Canning, Amaral, Lee, Meyer, & Stanley, 1998).

We analyzed our datasets to evaluate if the dynamics of firms from Finland and Mexico resemble those documented in the studies previously mentioned. The plots in Figure 1 show that our data sets capture those empirical regularities. Panel A shows the firm size distribution of Finland and Mexico. Both follow Pareto distributions (Equation 2) with scaling parameters α=2.698±0.005 and 4.429±0.014 respectively. Since the Mexican sample is biased, we estimated the firm size distribution for the universe of Mexican firms using the national economic census of 2004. The estimated scaling parameter is α =0.845±8.903e-05, quite close to a Zipf distribution.

Panel B in Figure 1 shows the distribution of the growth rates. Both countries present the tent-shaped pattern when the *y*-axis is plotted in logarithmic scale, being Mexico the one with a higher skewness for the Subbotin distribution in Equation 4 (κ=0.87). Finland growth rates are more symmetric with a κ=0.95 (considering that κ=1 yields the symmetric power exponential distribution). Both distributions are centered at zero, which conforms to earlier studies.

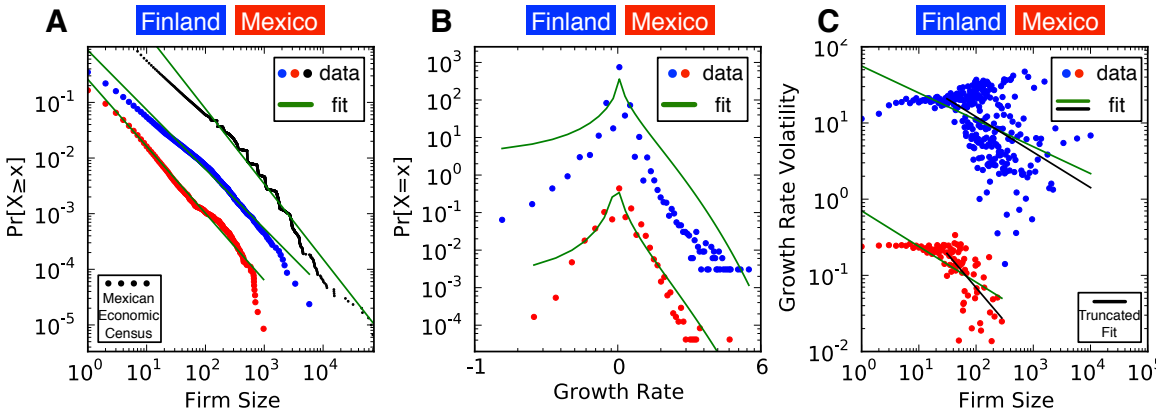

Figure 1: Stylized facts of firm dynamics.
The scattered plots in each Panel are stacked on top of the red (Mexican) data for clarity of the presentation. Volatility of growth rates in Panel C is measured as the standard deviation of the growth rates of firms of the corresponding size. The black solid fits in Panel C correspond to estimations using only data from firms with 50 or more employees.

We tested the power law relationship between the size of firms and the volatility of their growth rates. Earlier studies of this stylized fact construct logarithmic bins of the sizes of firms and compute the volatility for each bin. This produces clean data points with a high R^2 for the fitted model. We observed that for these studies, the slope between

the first and the second bins tend to be nearly horizontal. This implies that most of the decay in growth volatility takes place between medium and large firms. Additionally, most of the data used for earlier studies comes from publicly traded firms registered in COMPUSTAT, which biases the sample towards larger companies. We investigated the robustness of this pattern using our datasets and both cases confirmed our conjecture. As showed in Panel C of Figure 1, there is no decay of growth volatility for firms with approximately less than 50 employees. The power law relationship becomes clear for bigger companies. Fitting Equation 3 to the data we obtained a scaling parameter of α=0.353±0.044 (R^2=0.2) and α=0.468±0.061 (R^2=0.414) for Finland and Mexico respectively. These results are close to the ones found in earlier studies. However, they change considerably if we restrict the data to larger firms where the structure of the data shows a clear negative relationship. In this case, Finland has α=0.465±0.063 ($R^2=0.2$) and Mexico has α=0.884±0.157 (R^2=0.373). Our R^2 is considerably lower than those of earlier studies because we did not bin the data otherwise we obtain similar results.

3.3.3. Labor Flow Networks

Table 3 shows the first statistics from the LFNs of Finland and Mexico. Both of them are similar in size and density. The Finnish network is half as clustered as the Mexican one. The difference in the diameter is 12 firms, and the average shortest path length differs in almost one enterprise. This means that Finnish labor can be reallocated in far-away firms with fewer employee-employer matches than the Mexican one. The most central firm in Finland has a more preponderant role in labor flows and, on average;

each company sends and received workers to and from two other companies. We found that the most central company in Finland has a more preponderant role in labor flows than its Mexican homologous. This makes the process of labor flows more centralized. Nevertheless, most labor movements occur in a distributed way across the economy. This process is so extensive that in OECD countries, it accounts for more than 20% of total employment (Cahuc & Zylberberg, 2006). In relatively short periods, labor flows can connect most of the economy. Proof this is the existence of a gigantic component in both LFNs. Such a component connects about 80% of the employers that have at least one worker.

Table 3: LFN properties

	Number of firms (Number of edges)	Clustering coefficient (Network density)	Network diameter (Average path length)	Centralization index inflow (outflow)	Number of components (Firms in largest component)	Average in-degree (out-degree)	Average flow size (Standard deviation)
Finland	118,870 (250,222)	0.0340 (1.77e-5)	20.00 (5.00)	0.0130 (0.0138)	11,377 (78.80%)	3.08. (2.11)	1.40 (8.23)
Mexico	163,666 (298,238)	0.0658 (1.11e-5)	32.00 (5.84)	0.0076 (0.0082)	8,870 (87.79%)	2.86 (2.38)	1.26 (1.68)

These measurements consider only firms with at least one connection in the LFN.

3.3.3.1. Finnish Topology

Labor flows are heavy-tailed. This means that extreme events involving large flows occur more often than would be expected if labor reallocations were normally distributed. The degree distribution follows Equation 2 with α=3.19±0.003 (Panel A in Figure 2). The size of the labor flows between pairs of firms (the LFN links) can be fitted

to the same distribution with α=11.58±0.02 (Panel B in Figure 2). An average Finnish firm in the LFN receives 2.95±0.07 workers from 3.08±0.03 firms and sends 2.95±0.08 to 3.12±0.04 firms.

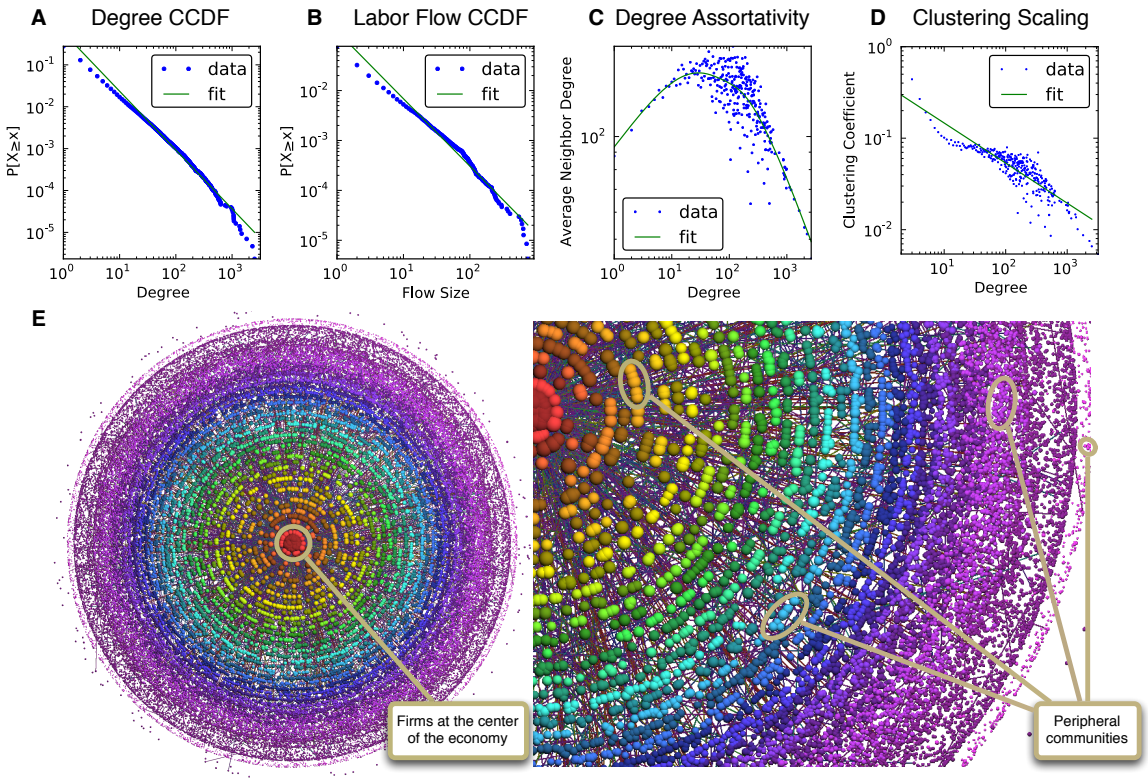

Figure 2: Topology of the labor flow network from Finland

Data from Panels A and B were fitted using maximum likelihood estimation. Due to the unusual magnitude of the scaling parameter estimated for Panel B, we do not think it is a power law. However, other skewed distributions do not produce better fits under the Kolmogorov-Smirnov criteria. We used kernel regression to identify critical regions in Panel C. Estimations in Panel D were made with OLS. Panel E shows the universe of firms in Finland. Only 1% of the edges are drawn. The size of the node represents the degree. The color identifies firms with the same *k*-core index. The image was produced with the visualization tool LaNet-vi and it shows the organization of the LFN into a core-periphery structure. Groups of firms are less tightly connected as we move from the center to the outside rings.

As in other complex networks, the topology of the LFN encapsulates structural information about the labor market. This relates to the connectedness that nodes have in terms of their neighborhood. In labor dynamics, we interpret connectedness as the accessibility that an individual has to other firms, given the position of her last employer in the LFN. Accessibly to firms can have different connotations, e.g. geographical, social, educational, industrial, etc. What is important is the close relationship between access to firms and access to vacancies or job opportunities. We believe that this is the essence of the underlying mechanism that drives labor reallocation dynamics and employment growth.

We analyzed the tendency of firms to be connected to firms with similar number of connections by looking at the average neighborhood degree. If the two variables are positively (negatively) correlated the network is said to be assortative (disassortative). Panel C in Figure 2 shows that when Finnish firms have 35 connections of more, the LFN becomes disassortative (Pearson r of -0.22 sampling null models). This is a peculiar property since evidence from several datasets shows that *social* networks tend to have an assortative character while *technological* networks are often disassortative (Newman, 2002).

A firm is structurally important in its neighborhood if it provides workers with means of mobility. Such means take the form of human capital, social capital, geographical proximity, or any other asset that is valued enough to reallocate the outgoing labor. A firm that facilitates such mobility is key in labor reallocation because it

becomes a middleman for its neighbors. Its absence restricts mobility to fewer firms, organized in smaller neighborhoods. We measured the clustering coefficients to estimate the structural importance of firms. Panel D in Figure 2 shows that clustering coefficients decrease with degree through a power law relationship. The estimated scaling parameter for the sampled null models is α=0.64±0.003 with an R^2=0.6; it is evidence of a *hierarchical structure* in a complex network (Ravasz, Somera, Mongru, Oltvai, & Barabási, 2002). Individuals in a hierarchical LFN have access to more firms, and more communities of firms, if their employers are in a higher level of the hierarchy. Here, clusters of nodes of a given level tend to be connected to a common cluster at a higher level. A cluster at an upper level becomes the broker of different communities from lower levels. This gives rise to a core-periphery structure of the economy.

We illustrate the core-periphery structure of the universe of Finnish firms in Panel E of Figure 2 through a *k*-core decomposition. The visual representation (Alvarez-Hamelin, Dall'Asta, Barrat, & Vespignani, 2005) shows the organization of the LFN into different communities. The firms in the center are the ones at the top of the hierarchy. The further a firm is from the center, the lower its degree, its hierarchical level, and the tightness of its community. Job-search-wise, individuals in peripheral firms would have a harder time finding vacancies due to poor connectivity. Therefore, we are interested in investigating if there is any relationship between LFNs and employment.

3.3.3.2. Mexican Topology

Our results from the Finnish LFN represent the universe of firms and employed individuals in this country. A natural question is how robust are our findings regarding other countries? The purpose of this exercise is to show evidence that the topological characteristics of LFNs are robust across economies. We constructed a LFN using the support dataset from Mexico. Similarly to the Finnish data, we counted annual flows of labor at the end of each year of the sample period. This gave us a network with more than 160,000 Mexican firms, with similar density to the Finnish one (see Table 3 for a comparison).

Despite the differences between Finland and Mexico, and between the two datasets, it is remarkable that most of the statistical patterns describing the topology of the LFNs are robust across countries. Panel A in Figure 3 shows that the degree distribution of the Mexican LFN is (2), with α=3.17±0.005. The scaling parameter for the fitted labor flow size distribution (Panel B of Figure 3) is quite high for a heavy tail: α=9.44±0.015. Panel C shows that this network is degree-disassortative, which is different from the Finnish LFN for firms with less than 35 connections. Finally, Panel D provides evidence of a hierarchical structure in the Mexican network.

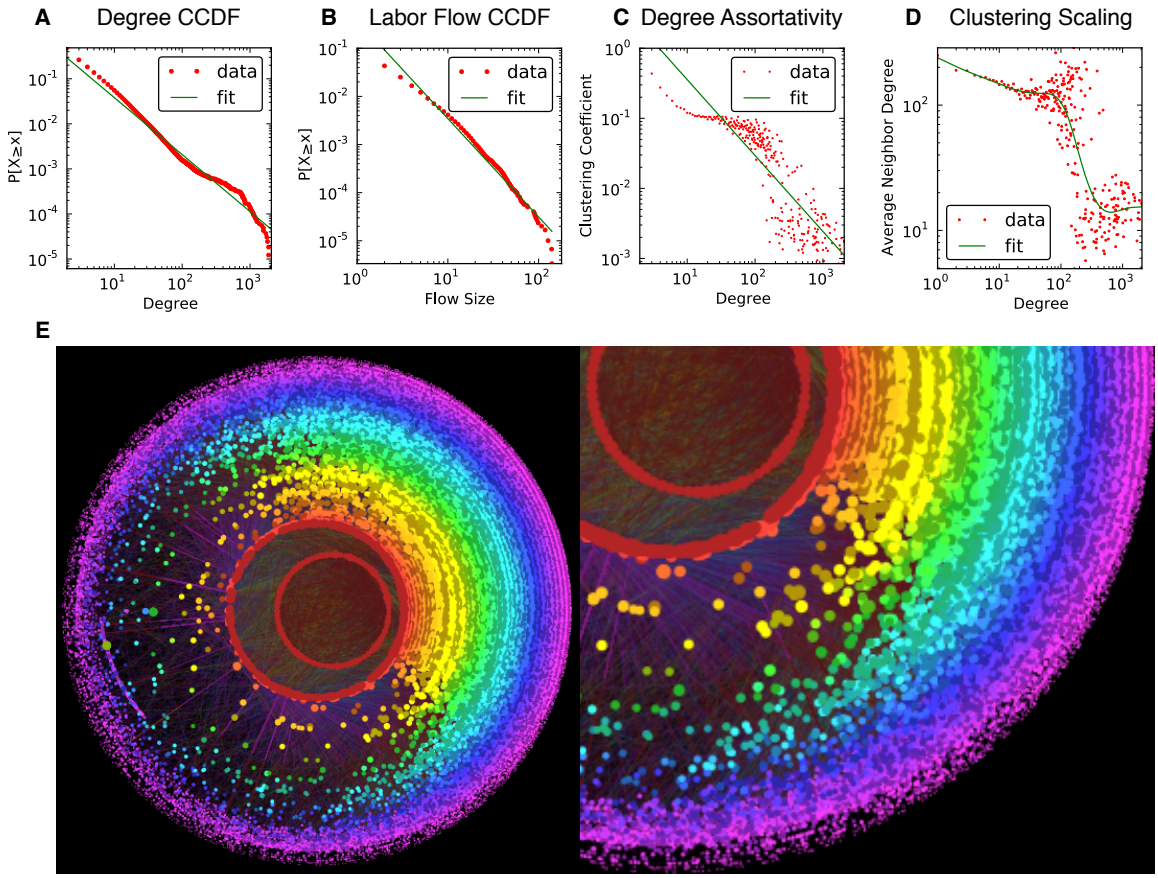

Figure 3: Topology of the labor flow network from Mexico

Data from Panels A and B were fitted using maximum likelihood estimation. Due to the unusual magnitude of the scaling parameter estimated for Panel B, we do not think it is a power law. However, other skewed distributions do not produce better fits under the Kolmogorov-Smirnov criteria. We used kernel regression to identify critical regions in Panel C. Estimations in Panel D were made with OLS. Panel E shows the sample of firms from Mexico. Only 1% of the edges are drawn. The size of the node represents the degree. The color identifies firms with the same *k*-core index. The image was produced with the visualization tool LaNet-vi and it shows the organization of the LFN into a core-periphery structure. Groups of firms are less tightly connected as we move from the center to the outside rings.

In a contemporary but independent study, Gianelle (2011) constructed a LFN for the industrial region of Veneto, Italy. It consisted of all the workers and employers from the private sector of Veneto in the decade of the 1990's. His network had approximately 380,000 vertices. His results confirm the robustness of our findings regarding the Pareto distribution of the degree distribution of LFNs. A different network notion —bi-partite

graphs of workers and firms— has been used for a similar analysis, finding Pareto degree distributions (Schmutte, 2010).

A different but related type of firms network is the supplier-customer ones. They have been studied comprehensively for Japan by (Konno, 2009; Saito et al., 2007), where it was found that a network of 800,000 Japanese firms connected through economic transactions has the scale-free and hierarchical properties that we have found in our LFNs. Although these networks are of a different nature, they share common features with LFNs, suggesting the important role of firms' dynamics in economic systems.

3.3.4. Networks and the Economy

An advantage of the FLEED dataset is that it can be merged with the Finnish business registries. This allowed us to go beyond simply characterizing the LFN topology in order to explore its relationship to the economic attributes of firms. We measured network properties for the 2005-2008 period in order to study their relationship to initial firm size and age. This helps to identify groups of firms that have particular structural roles in the LFN, firms that could be important for labor reallocation. All the results presented throughout the rest of our study are only for Finland.

We found that firm size and its degree are strongly correlated (Pearson r of 0.83), and that degree volatility increases as firms become larger (see Panel A in Figure 4). However, degree and firm age are more complicated. Panel B in Figure 4 shows that the correlation between age and the average degree of the group exists only for firms less

than 50 years of age (Pearson *r* of 0.79). This suggests that LFN formation is not determined by a pure Yule multiplicative process, a model commonly used to explain firms' growth (Gibrat, 1931; H. Simon, 1955; Yule, 1925) and more recently applied to scale-free networks (Barabási & Albert, 1999). It means that if the workers would tend to flow towards older firms, the latter would be hubs. This is clearly not the case for the Finnish LFN.

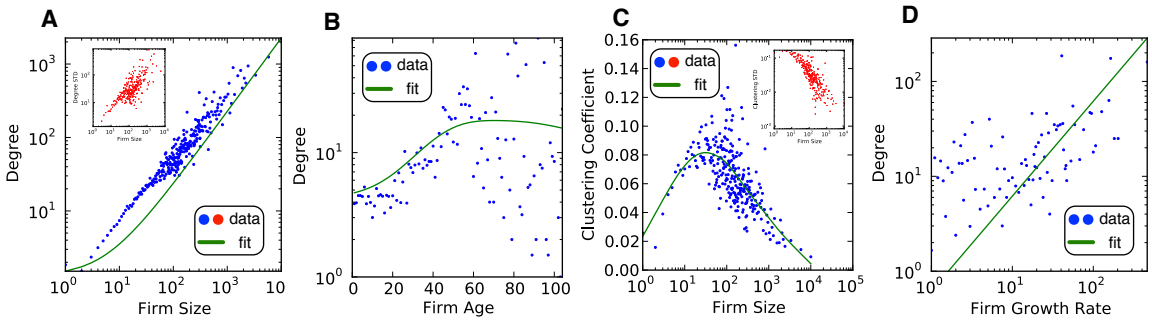

Figure 4: Correlations between network properties and economic variables
Critical regions in Panels B and C were identified using kernel regressions.

Firm size can tell us something about the structural importance of a firm in its neighborhood. Panel C in Figure 4 shows a positive correlation between size and average clustering coefficient for firms in groups with less than 28 employees (Pearson *r* of 0.67 after null model estimation). This relationship is negative for bigger firms (Pearson *r* of -0.72). Both small and large firms are middlemen in their communities. Large firms connect different groups of firms due to their higher position in the hierarchy. Small

firms connect other small and medium-sized individual firms that, otherwise, would not be part of the LFN. Finally, Panel D in Figure 4 suggests a positive relationship between degree and growth (Pearson *r* of 0.44).

3.3.5. Sectors, Regions, and Communities

Identifying the target population of an employment policy is crucial for its success. Conventionally, firms are classified into industries or geographical regions at some particular level of aggregation. It is natural to think that firms in the same class tend to interact more with each other. Labor-wise, we would expect that a worker employed by Chrysler in Detroit would more likely get a job in Michigan and/or in a firm that falls into a related industrial classification. This does not seem to be the general case in today's labor market. Skill-based technical change of industrialized economies has increased the value of transferable and non-cognitive skills (Heckman, Stixrud, & Urzua, 2006). Today, it is rare to make a career in a single firm. Therefore, job changes are becoming more common. In the U.S., job-to-job flows have increased nearly 60% during the last decades (Stewart, 2002). The magnitude of job-to-job flows is nearly twice the number of employment-unemployment transitions (Fallick & Fleischman, 2001; Nagypal, 2008) for this country. In the case of Finland, industries and regions are more connected than ever (Ilmakunnas & Maliranta, 2005; Maliranta & Nikulainen, 2008). Thus, defining communities of firms to be the target of employment policies becomes a major challenge.

Labor policies that use standard classifications can incur in two types of errors when defining their target population: (*i*) exclusion of important firms and (*ii*) inclusion of irrelevant firms. Imagine that the wood manufacturing industry requires specific skills to use a new technology, and that firms are not able to provide enough training to meet their demand for trained labor. *On-the-job training programs* are a common solution. Here, the government finances wood manufacturers to provide training that meets their needs. However, now imagine the LFN tells us that this sector receives a substantial amount of labor from the recycling industry. Government financing training programs for firms from the recycling sector could improve the program's impact. Similarly, such a program could be more efficient by discarding wood manufacturers that do not show evidence of facilitating labor mobility. Hiring employees from neighboring firms (or sectors) can meet their demand for skills.

Figure 5 presents the composition of the Finnish LFN by communities. Panel A and B illustrate how the conventional industrial classification do not match the arrangement generated by labor flows. The nodes in this chart represent entire industries at the three-digit classification level (larger nodes have more firms). The color gradient represents different industries at the two-digit classification level. If subsectors of a two-digit sector would exchange labor between them with higher propensity than with other sectors, then by grouping the nodes according to the "attraction" represented by their edges should produce a layout where sectors of the same color are visually clustered. This is clearly not the case since the color gradient seems well mixed throughout the entire graph. Similar results occur for municipal classifications. In Panel B we provide

two and three-digit municipal aggregations, with the gradient representing the geographical position form north to south. In both industrial and geographical cases, non-labor based classifications do not appear to be representative of the community structure that underlies labor dynamics. We provide more rigorous evidence of this claim by analyzing the structure of the LFN at the firm level employing community detection algorithms.

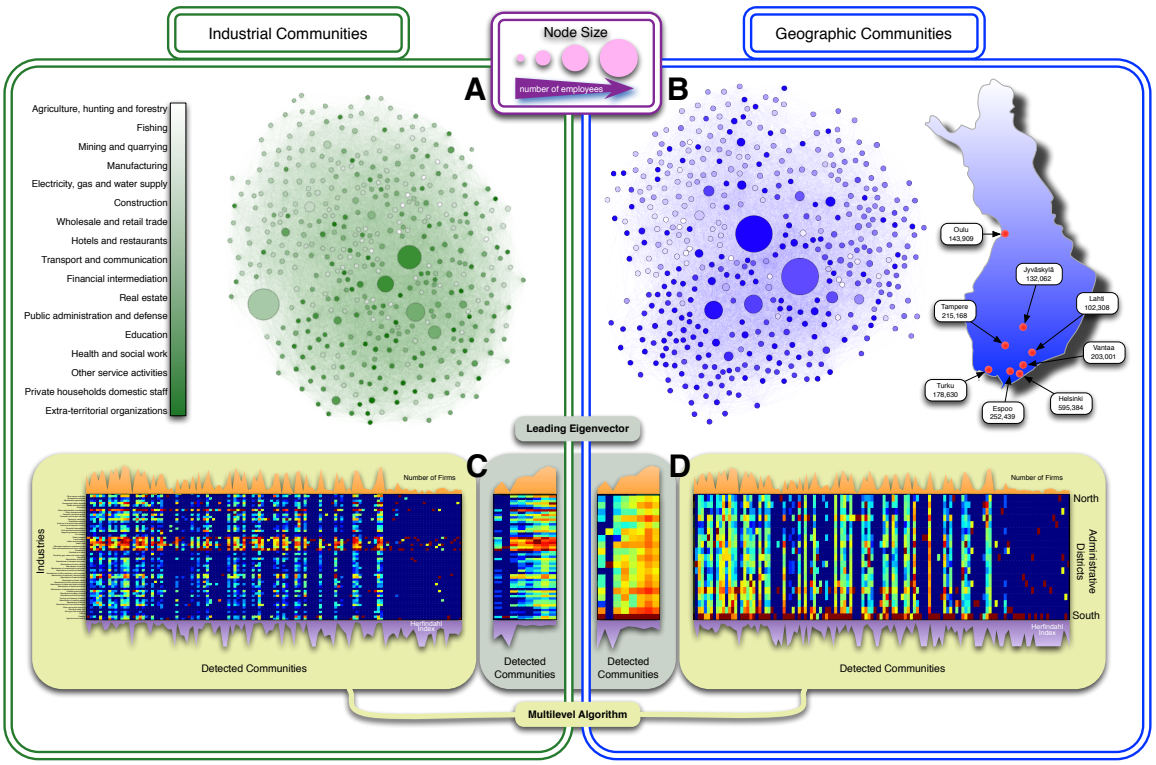

Figure 5: Communities of firms

Panels A and B provide a visual example of clusters in a reduced version of the LFN. The nodes represent industrial/geographical sectors as defined by the three-digit classifications from Statistics Finland. In Panel B we provide information about the population of the eight largest cities in the country in order to illustrate the high concentration in southern districts. For both Panel A and B, the color gradient corresponds to two-digit classifications. Their networks are laid out by the Force Atlas algorithm, which groups nodes according the strength of their ties. Panels C and D show the density matrices of the detected communities and the predefined industrial/geographical

sectors. Each column has been normalized to illustrate the diversity of sectors in each community as a heat map. Cells represent the share of firms that each industry/region has in its respective community. The normalized total number of firms in each community is plotted on top of the heat maps. An inverted series of the Herfindahl–Hirschman index is plotted in charts bellow the heat maps.

Network community detection methods can complement other ways of characterizing the target population. There are many methods to detect such communities in networks (Fortunato, 2010). A popular approach is the maximization of a modularity score, defined as *the difference between the number of edges inside a community and the expected quantity if such edges were placed at random* (Porter, Onnela, & Mucha, 2009). There are many algorithms that try to discover communities by partitioning the network and evaluating the modularity score. If the score is maximized and the partitions are not trivial, then we have evidence of communities.

In Panels C and D of Figure 5 we present the results of two community detection algorithms that perform especially well in large-scale complex networks (Blondel, Guillaume, Lambiotte, & Lefebvre, 2008; Newman, 2006). The approaches use different ways of maximizing modularity. Therefore, their results differ. Our purpose here is to show the disparity between the communities that are identified using the LFN and the ones that are defined by industrial and geographical classifications. The discussion of which algorithm is more suitable for labor policies is something that we will leave for future inquiries. Vertical patters indicate the presence of communities composed of firms from different sectors or regions (type *i* error). Horizontal patterns indicate that there are different communities inside the same sector or region (type *ii* error). If there were a

strong correspondence between detected communities and conventional categories, each community would show a high concentration index. We found that more than 70% of the communities detected via leading eigenvectors had a Herfindahl–Hirschman index (HHI) lower than 0.5, considered low. For communities detected using the multilevel algorithm, 79% had a HHI<0.6 for industries and 62% for geographical regions. We should point out that we used two-digit industrial and geographical classifications, which are the most aggregate. When using less aggregate classifications, the HHI drops because members of a community that used to belong to the same industry or region now belong to different ones.

3.3.6. Employment Growth

A common policy approach to promote employment growth is the creation of programs that are targeted to firms that show signs of potential growth. Identifying these firms is often a challenge since understanding the underlying causes of a company's growth remains difficult.

Conventionally, the firm-dynamics literature employs the size and age of firms to identify their potential growth. For example, small-business advocates tend to argue that the size of a firm is negatively correlated with its growth. Since the work of David Birch (1994; 1981, 1987), this has been the dominant paradigm in the small businesses literature. Birch argued that small rapidly growing businesses, which he called gazelle firms (GFs), were responsible for most employment creation. Acs (2011) found that when we account not only for proportional grow, but also for the effect on employment, there is

a subclass of gazelles that growth significantly more intensively; he calls them high-impact firms (HIFs). A posterior study (Haltiwanger, Jarmin, & Miranda, 2010) questioned the rationale behind such classifications by showing that the relationship between firm size and growth rate becomes irregular when controlling for age.

We propose the use of LFNs to identify potential growth of firms. Since LFNs capture structural information about the dynamics of labor, measures of the structural importance of a firm in the network may contain useful information. We ran a logistic regression of the probability that a firm experiences positive growth between 2005 and 2008, as a function of conventional predictors (initial size and age) and its LFN characteristics. The network covariates included measures of network centrality, clustering, and geographical assortativeness. In order to prevent spurious relationships, we computed the LFN metrics for the period 2002-2005.

Table 4 shows the network properties of firms can be used to identify the likelihood that a firm will experience positive growth. Once we introduce the network covariates, the marginal effect of the initial size becomes significant and consistent with the literature: larger firms are less likely to experience positive growth. Firms with higher in-degree and higher closeness are more likely to grow. Companies that are part of numerous shortest paths between two other firms are *less likely* to increase employment.

Table 4: LFN and employment growth

	(1) Pr[Δsize>0]	(2) Pr[Δsize>0]	(3) Pr[Δsize>0]	(4) Pr[Δsize>0]	(5) Pr[Δsize>0]	(6) Pr[Δsize>0]	(7) Pr[Δsize>0]
Size in 2005	1.92e-05	2.07e-06	-0.000413***	-0.000333***	-0.000392***	-0.000384***	-0.000374***
	(2.33e-05)	(2.41e-05)	(1.00e-04)	(9.25e-05)	(9.96e-05)	(9.87e-05)	(9.83e-05)
Age in 2005		0.00140***	0.00139***	0.00117***	0.00115***	0.00114***	0.00110***
		(0.000118)	(0.000118)	(0.000119)	(0.000119)	(0.000119)	(0.000119)
In-degree			0.00287***	0.00152***	0.0101***	0.00983***	0.00924***
			(0.000454)	(0.000387)	(0.000804)	(0.000799)	(0.000798)
Closeness				0.00788***	0.00698***	0.00666***	0.00591***
				(0.000359)	(0.000363)	(0.000366)	(0.000379)
Betweenness					-3.499***	-3.360***	-3.214***
					(0.335)	(0.331)	(0.331)
Clustering Coefficient						0.000816***	0.000846***
						(0.000139)	(0.000139)
Neighbors Same Municipality							-0.000297***
							(4.12e-05)
Observations	55,180	55,180	55,180	55,180	55,180	55,180	55,180

Logistic regressions with marginal effects of covariates. The model was performed for the Finnish dataset. Non-survivor firms were excluded. Closeness, betweenness, clustering coefficients, and neighbors in same municipality are in percentages. It must be noted that although closeness and betweenness are in percentage, their empirical range is quite restricted. For example, the firm with highest betweenness has a level of nearly 7%, thus the high marginal effect. Standard errors in parentheses. *** significant at 1%.

Neighborhood metrics of firms also yield useful information. When firms have higher clustering coefficients, it means that they live in better-connected communities. These firms are more likely to produce positive employment growth. Additionally, if a firm has a higher percentage of neighbors from the same municipality, it is *less likely* to experience positive growth.

Given the evidence that LFNs can be useful to identify employment growth, we proceed to study the contribution of firms to employment growth. We compare employment growth from different groups of firms according to Equation 1 and using standard classifications form the small-business literature and one that considers the LFN (see Section 3.2.7). Employment growth in a class of firms is measured as the sum of the net growth of all the companies of that group, in terms of the number of employees. This is a standard procedure in this type of exercise. Using net growth implies that the total

employment growth of an economy is measured exclusively through firms that experience positive growth. Therefore, the following results are focused on subgroups of such firms. We found that LFNFs contribute to 88.3% of employment growth in Finland while the second largest contributor –high-growth firms (HGFs)– produce 63.5%. LFNFs represent 7.1% of all firms in Finland. They are more common than HGFs by 0.9%. Employment-wise, LFNFs are more productive (see Table 5).

Table 5: Employment by types of firms

	Labor Flow Network Firms	High Growth Firms	Gazelle Firms	High Impact Firms
Average firm size	6.0	1.7	1.7	1.8
Average firm age	13.6	11.1	2.7	8.4
Employment growth share	88.3%	63.5%	21.1%	32.1%
Firms share	7.1%	6.2%	1.8%	3.1%
Employment growth per firm	3.0	2.5	2.9	2.5
Average incoming labor	4.3	2.3	3.0	2.4
Average outgoing labor	2.2	0.7	0.9	0.7
Average turnover	1.8	2.3	2.3	2.3

Firms were classified using the taxonomy presented in the Section 3.2.7. Employment growth was measured only for survivor firms. Shares are in terms of the total employment growth of the Finnish economy.

HGFs, GFs, and HIFs are predominantly composed of firms with initial size zero (about 80%), which explains part of their explosive growth. This proportion drops to 44% for LFNFs. GFs and HIFs are subsets of HGFs. Some LFNFs intersect each one of those groups and some do not belong to any of them (Panel A in Figure 6). This means that, on one hand, a subpopulation of HGFs, GFs, and HIFs is heavily composed of isolated companies that do not participate in any flow of workers with any other firm in the

economy. On the other, LFNFs might include firms that do not growth so intensively, but that take part in the labor reallocation process. As showed in Figure 6, an average LFNF has six employees and is 13.6 years old, which makes it larger and older than the average firm in one of the other classes. LFNFs receive more workers and also send more labor to other firms. The LFNF turnover rate is of 1.8 workers; 0.5 workers lower than other types of firms. A LFNF produces a net average of three jobs.

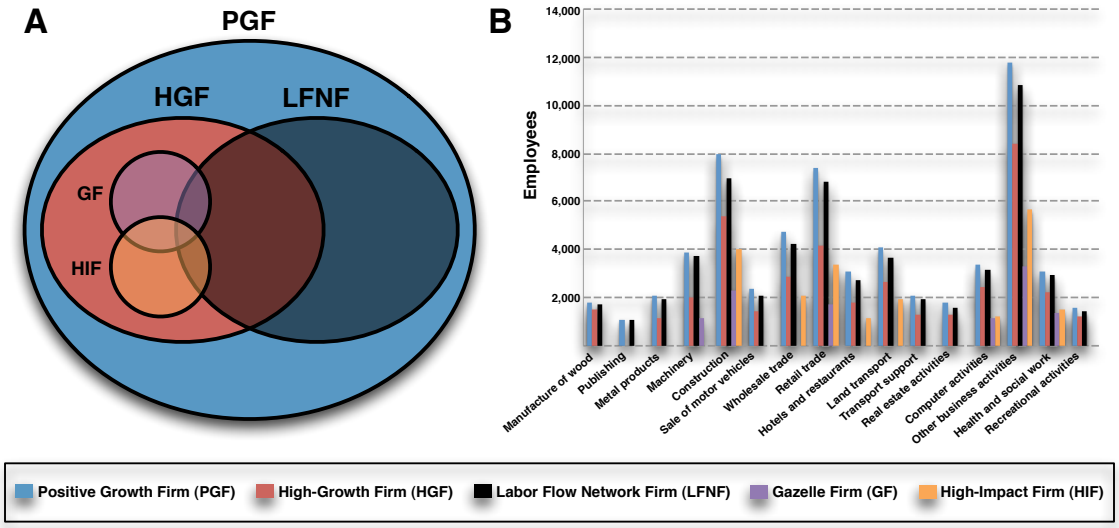

Figure 6: Classification of firms and industrial participation
The Venn diagram is approximately proportional to the number of firms in each category. The bar chart compares total employment of each group in each sector. Industries are classified using the two-digit European Union's classification of economic activities, NACE.

From all firms with positive growth (PGFs), only LFNFs are present in all industrial sectors (Panel B in Figure 6). They are, consistently, the subclass with most

firms and employment across sectors. Therefore, classifications that are based exclusively on growth intensity have the problem of excluding industries that might by important to employment growth through the LFN.

It is evident that firms embedded in the LFN have an important role in employment growth. Considering all the firms in a LFN (not only survivors), they represent 28% of all firms. They are responsible for 90% of employment growth and 91% of its destruction. This implies that the majority of the destroyed jobs are transformed into new ones, most of them filled with people who found their way through the LFN. We next demonstrate that agent-based models are able to generate these dynamics.

3.3.7. A Taxonomy for Labor Flow Network Firms

Davis et al. (2006) showed that US firms that experience shrinkage of their size have an important role in the hiring process. As showed in Figure 7, firms that experience positive growth have hiring shares (as a percentage of total employment) that are more than proportional to their growth rate. The same happens to separations for firms that experience negative growth. A surprising result is that companies with negative growth still have high hiring levels (about 2.5% of all employment), and positively growing firms have high separation levels. From a LFN perspective, this implies that shrinking firms are important because their participation in hiring and separation allows more mobility of labor, which facilitates a better allocation of human capital.

We have showed that LFNs provide information that facilitates filtering firms with potential growth. Our definition of LFNFs is one way to use such information. However, we can modify our classification to include negatively growing firms and study their role in labor dynamics.

Figure 7: Growth, hires, and separations

Source: (S. Davis et al., 2006)

The principle of this taxonomy is to employ some of the structural indicators that we have computed for each firm in the Finnish LFN. Specifically, we will focus in the direction of the flows with respect to the firm, diversity of links, amount of flows,

closeness to other firms, and clustering. We restrict our sample to *survivor firms* since we are interested in measuring employment growth (positive and negative). The principle for of this classification is to group firms into two categories for each network metric. One group includes all firms with a metric above a predefined threshold and the other has all the remaining firms. The criterion to define a threshold depends on the purpose of the study. In this case, we set a somehow arbitrary threshold as the average metric plus one standard deviation. The purpose here is to measure the contribution to employment creation and destruction of those firms with extreme network properties. In the case of number of connections, defining extreme events in terms of one standard deviation from the mean is not a very meaningful measure since the distribution of this metric is Pareto (see Panel A in Figure 2 and Figure 3). However, we will employ this threshold since we do not have a concrete question that would help to define a more relevant threshold at this point. Figure 8 shows the nomenclature system for the firm taxonomy. This classification is different from the one illustrated in Figure 6 since it includes companies with degree zero and firms with negative growth.

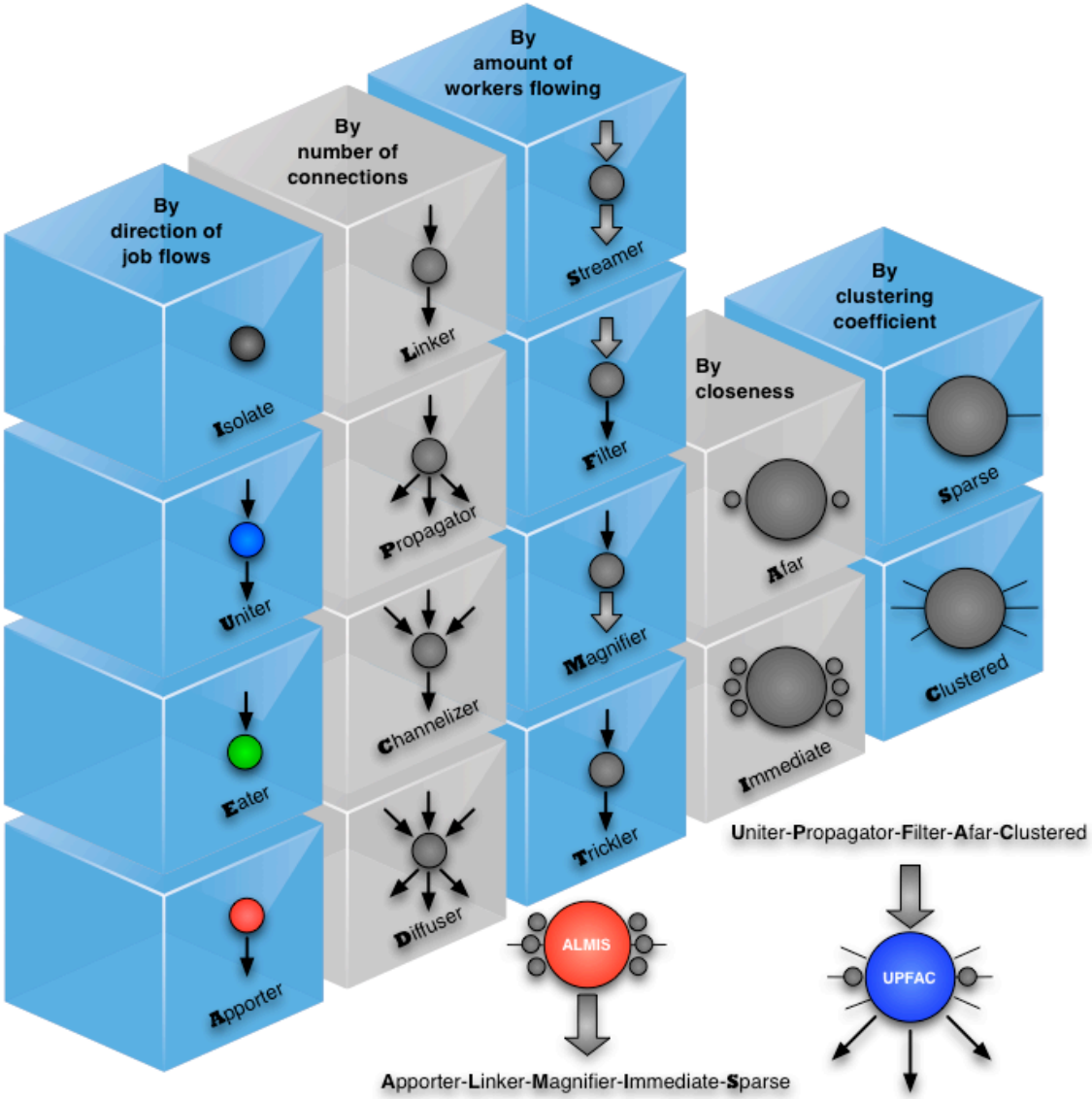

Figure 8: Firm taxonomy

We classify a firm by taking one cube from each stack, from left to right, and concatenating the names in that order. We join the first letters of each name in order to compose the name of the class. Similarly, we can construct an image representation of the class, as showed in the given examples. For each cube, thicker or more numerous elements adjacent to the nodes represent that the firms in that group have a metric above the given threshold.

From Figure 8 it is clear that not all possible nomenclatures are possible. For example, isolates have undefined levels of closeness or clustering. For simplicity, this

kind of firms will be grouped with firms that have metrics below the threshold, giving us a single class named ILTAS. The total number of classes that were present in the Finnish dataset was 48. We computed the absolute amount of positive and negative employment growth of each group. Figure 9 presents the total change in employment produced by each one of the classes obtained from this taxonomy.

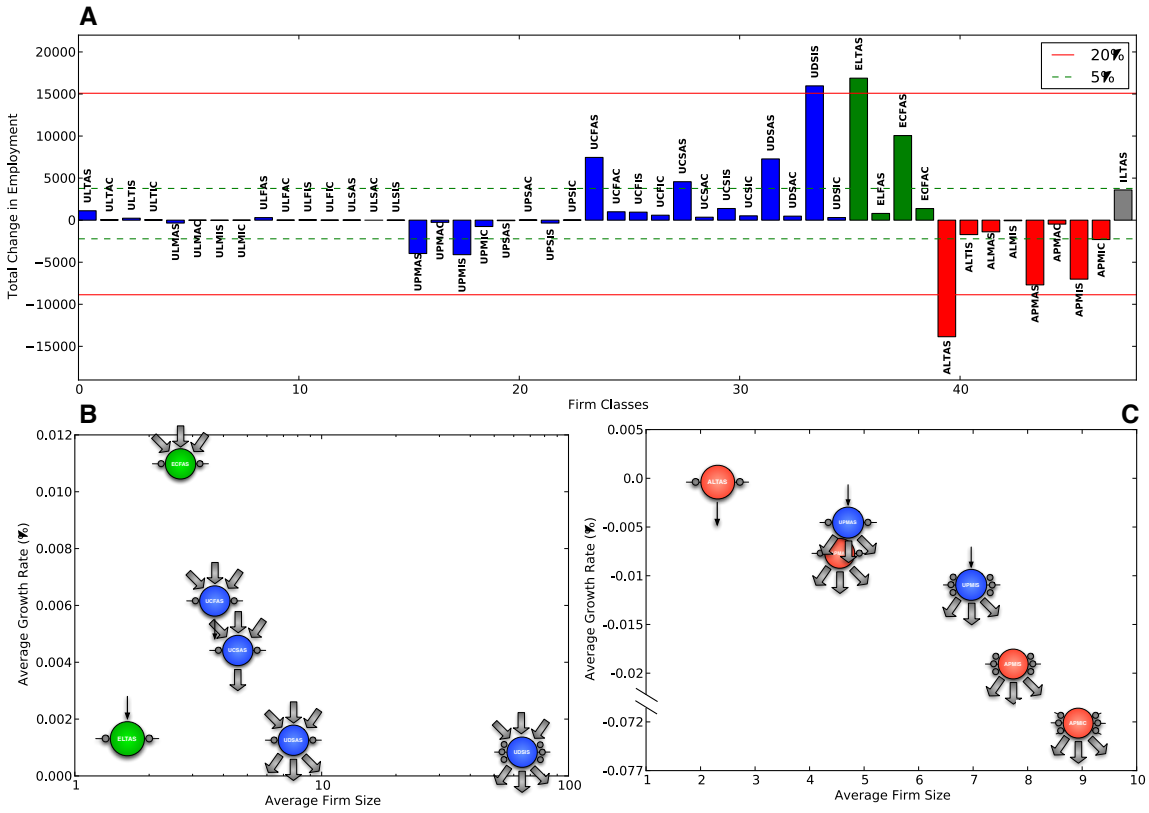

Figure 9: Firm classes and change in employment

Each bar in Panel A represents the net change in employment in a class of firms. The bars with names below zero represent negative change. The solid (dotted) lines indicate the 20% (5%) of the of the total positive and negative changes in employment among Finnish survivor firms. Blue bars are *uniters*, greens are *eaters*, reds are *apporters*, and grey are *isolates*. Panels B and C show average firm size and growth rate (in percentage) for classes of firms that contribute with more than 5% of employment creation and destruction respectively. Growth rates were computed using Equation 1 for each firm and then averaged for each class.

From Figure 9 it is clear that *uniters* are firms that have important roles both in employment destruction and creation. Panel A shows that the groups that contribute to the highest levels of employment creation are UDSIS and ETLAS. Although both create nearly the same amount of employment, their firms are quite different. In Panel B we can see that some of the largest firms are part of UDSIS while small firms dominate ETLAS. Firms from both classes experience a similar average growth rates. UDSIS is a particularly interesting class of firms because it receives and sends labor simultaneously, suggesting a relationship between the turnover process described in Figure 7, employment growth, and LFNs.

When it comes to employment destruction, ALTAS firms contribute with more than 20%. However, they release labor flows below the threshold for both diversity of destinations and size of flow. It is clear that this is driven by the distribution of outgoing edges, which is dominated by firms with lower degree. Panel C shows an interesting feature: the average growth rate of firms that are shrinking in size tends to increase when companies are closer to all other firms in the LFN. This suggests that workers that are separated from rapidly shrinking firms tend to find new jobs more frequently and in more firms.

3.3.8. Network Formation

Labor reallocation is dynamic in nature and seems to be an essential component of employment growth and its volatility. In OECD countries, reallocated labor accounts

for more than 20% of total employment (Cahuc & Zylberberg, 2006). Furthermore it reaches most of the economy in relatively short periods (evidence provided in Table 3).

We explore the dynamic character of the formation of the gigantic component in LFNs by analyzing the data through smaller time frames. Unfortunately, the primary dataset is composed of annual Panels and any LFN constructed for that window already shows the topological features described in this chapter. However, we can take advantage of the daily resolution of our support dataset from Mexico in order to characterize how periodic labor flows connect firms. We used daily timestamps from the IMSS records in order to construct LFNs for different periods. We started in January 1st 2005 and added one day to the time frame in order to construct a LFN. We repeated this process by adding two days, thee days, and so on. Figure 10 illustrates the result of this exercise.

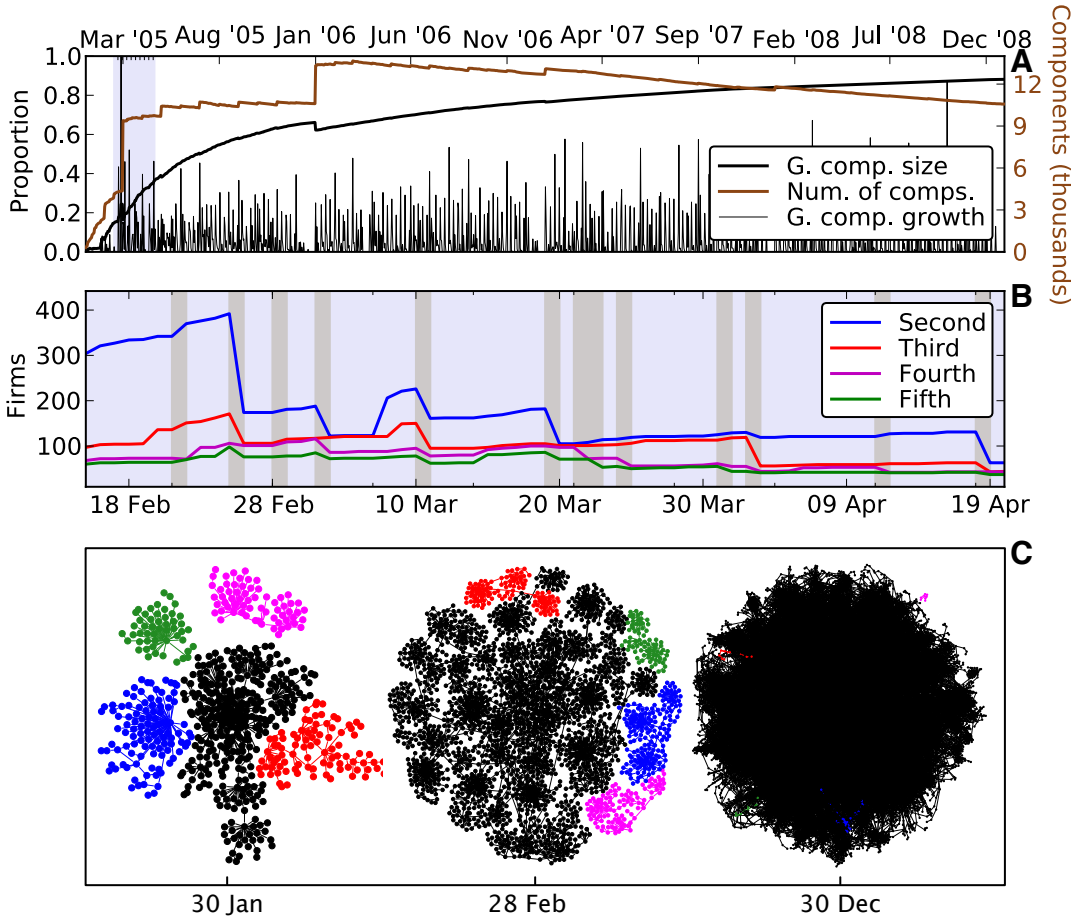

Figure 10: Formation of the Mexican LFN on a daily basis

Panel A shows the size dynamics of the largest component as a proportion of the total number of firms in the network, the normalized growth rates of the largest component that are above the average growth, and the total number of components (brown). The shaded region is zoomed in Panel B. It represents the evolution of the size of the second, third, fourth, and fifth largest components. Shaded regions correspond to growth periods of the largest component that are above the mean + standard deviation. Panel C shows snapshots of the five largest components in the LFN at different points in 2005.

Two major forces drive the growth process of the largest component: entry of new employers and unions of large components. Panel A in Figure 10 shows that the size of the giant component, as a proportion of the number of firms in the network, grows with decreasing marginal growth. Negative growth of the latter indicator is explained by the

entry of numerous employers as isolated nodes. The positive spikes are caused by the absorption of other large components. As we use longer periods for the network, the addition of new firms has a weaker effect, denoted by the concave shape of the series of total number of components in Panel A.

One year with two months is the time that it takes to the Mexican LFN to reach its maximum number of components. After that point, the major driving force is the labor flows that connect previously disconnected components. This is evidenced by a continuous decrease in number of components and persistent growth of the gigantic component. Between January and April, the largest component experiences extremely high growth. This is explained by the eventual flow of labor between the firms of the largest component and companies from any of the other largest components. We illustrate this process in Panel B of Figure 10 by plotting the growth process of the second, third, fourth, and fifth largest components during that period. By construction, the only way for a size-ranked component to decrease in size is by being replaced by a smaller one. This can only happen when this component becomes part of another one (in this case a larger one). Through this process, all the components that were initially the largest become a single one. In approximately one year, the gigantic component contains more than 50% of the companies in the network. This reveals an important feature of the LFN formation process: if firms live long enough, their communities or sectors will eventually become available to previously inaccessible labor. Panel C provides a visualization of the growth of the network through this process.

3.4. Discussion

The network character of labor flows between firms has been investigated and the usefulness of the labor flow network (LFN) concept for the study of firm and employment dynamics has been demonstrated. In many ways such networks have 'extreme' properties, in the sense that 'heavy tails' characterize many of their empirical features. The dynamics of labor flows in such networks are very far from the 'smooth' flows one might expect to occur if any worker could migrate to any company. Clearly, from the structure of empirical LFNs, migration is constrained and the extent to which LFNs deviate from complete graphs is indicative of the magnitude of the 'lumpy' and clustered labor flows that can occur in them. Macroeconomic dynamics that are qualitatively similar –also known as 'granular'– are produced by similar heavy-tailed distributions in firm size (Gabaix, 2011). Furthermore, many of the network properties of LFNs have explicit, underlying economic meaning. The regional structure of firm production, along with its demarcation into sectors, are both clearly embedded in LFNs. Perhaps most importantly, signs of employment growth are also present in LFNs and can be determined from network properties. Additionally, many of these properties are common to inter-sectorial networks, that lead to effects on aggregate production fluctuations of production (Acemoglu, Carvalho, Ozdaglar, & Tahbaz-Salehi, 2012). Finally, using the latest computational advances to model individual agents interactions and the emergence of LFNs, labor policies can be tested artificial laboratory. This new capability suggests that LFNs can be important analytical tools for an improved understanding of the performance and potential of modern economies. Countries would

be well-served to collect and make available the kinds of data that facilitate the construction of LFNs.

4. THREE MODELS OF LABOR AND FIRM DYNAMICS

4.1. Introduction

As discussed in Section 2.4, several ABMs of labor markets have been developed in the last decades. Here we present three models that contribute to this literature by bridging the gap between labor, firm, and network dynamics. The three models were constructed under the principle of making them simpler than the ones previously reviewed. Nevertheless, they provide emergent results that are richer than most of the existing models of labor markets. We contribute by developing a framework to couple realistic labor and firm dynamics in a parsimonious way. Additionally, our models incorporate in a natural way the interaction of agents through networks (economic and social), which is a standing challenge in the field of labor economics. We provide evidence of the validity of our models by reproducing multiple stylized facts of firms and labor. Additionally, by conducting controlled experiments we find evidence of the relationship between microeconomic parameters and the topology of labor flow networks.

This chapter is structured in the following way. Section 4.2 presents a model of firm dynamics. It is an agentized version of Gabaix's (1999) model, which is able to generate labor market stylized facts that are not obtained by Gabaix's formulation. In Section 4.3 we present a model of social network formation in the work place. This

model generates stylized facts of labor dynamics that are produced by the endogenous process of social network formation. Section 4.4 presents a model that provides microeconomic foundations to the formation of the small-world phenomenon in LFNs. It brings together well known models of job search behavior (Gronau, 1971; McCall, 1970; Mortensen, 1973) and small-world networks formation (Watts & Strogatz, 1998).

4.2. Firm Dynamics

It is difficult to deny the close relationship between the dynamic nature of firms and labor. As exposed in Section 1, firm dynamics is a widely studied topic, where several attempts to model some stylized facts have been made in the last century. However, little has been done to understand the link between the heavy-tailed nature of firm dynamics and the dynamics of labor.

We consider Gabaix's model (see Section 2.2.5) to be a good starting point because it provides the microeconomic building blocks to connect labor to firms. Although this model was originally developed to understand the growth of cities, it has been used to study firms (Saichev, Malevergne, & Sornette, 2009). Here we employ it to merge firm dynamics with labor. In an agentized version of this model, Guerrero and Axtell (2011) showed that Gabaix's microeconomic specification produces a degenerate firm size distribution when it is applied to realistic populations (discrete units and explicit interactions). Hence, the equilibrium mechanism through which agents maximize utility and firms maximize profits does not seem plausible. Here we present a simpler model

that recovers the heavy-tailed firm size distribution and allows us to explore the connections between firm dynamics and labor.

4.2.1. Model Description

In this model there is a fixed population of firms and workers. Workers receive utility from a combination of wage w and benefits plan b provided by their employers. The utility of worker i takes the form $u_i=b_i w_i$. Firms have production functions with constant returns. Their production is determined by the number of tenured workers T and the recently hired ones N. We assume a Cobb-Douglass technology $f(T,N)=(T)^{\alpha}(N)^{1-\alpha}$, where $0 \leq \alpha \leq 1$. Firm j determines the wage it will offer at period t by computing the marginal product of hiring a new employee:

Equation 5

$$w_{j,t} = \alpha\left(\frac{T_{j,t}+1}{N_{i,t}+1}\right)^{1-\alpha}$$

Equation 5 is a neat outcome because it captures a couple empirical aspects about firm dynamics. First, it says that between firms with equal N_t, the largest are the ones that offer higher wages, which relates to the employer-size premium (Brown & Medoff,

1989). Second, among firms with the same size, those who offer higher wages are the ones that have hired less people during period t (decreasing marginal returns of labor).

There are two sources of exogenous stochasticity in this model. The first is the benefits plan $b_{j,t}$, which is randomly determined for each firm j at every period t. The second is the separation rate λ. Every time that a worker loses her job, she has to find a new one (and she always gets hired in her best option).

In Gabaix's model it is assumed that, during a single period, job seekers and firms are repeatedly matched through a *tatonement process* until every firm reaches levels of N and Y where no more hires are made. Due to the impossibility of computing this equilibrium for discrete populations (Guerrero & Axtell, 2011), our implementation is a disequilibrium model where workers can only get hired once per period. With this adoption the model generates firm size distributions where one company tends to concentrate most workers. This is explained by the dynamics of Equation 5. When a request for a wage offer arrives to a firm in period t, this firm counts Y as the number of employees that it has hired so far during t. At the beginning of $t+1$, Y becomes zero. Then, the size difference between firms establish the initial conditions under which larger firms can offer higher wages and be more attractive to workers. The marginal decreasing returns of Y slow down this effect during each period. However, they do not prevent larger firms from being the first ones who hire each period. Therefore recovering a heavy-tailed distribution from this model is a challenge that we approach by studying two components that are mentioned but not treated in depth in Gabaix's study.

First, we consider the issue of new hires versus incumbent employees. More specifically, what does it take to become an incumbent worker? The dynamics of Y in Equation 5 imply that production of cities scales with their size. However, the scaling process suffers delays where Y needs to become part of N. This can be interpreted as the time needed for a city to update infrastructure to provide better conditions for productivity. In a labor context, it can be interpreted as the training of new workers. This is an essential part of our model. Here, Y is not set to zero every period. Instead, each firm keeps a record of the period when each worker was hired. Then, a worker is considered part of N if she has completed the minimum training duration τ required by the firm. This mechanism is enough to balance the dynamics of Equation 5 and facilitates the recovery of skewed distributions.

Second, we reconsider the ability of workers to request offers from all firms. This is a key assumption in Gabaix's model since it is necessary to achieve a global equilibrium. In our model, we do not have to maintain this assumption. In fact, limiting the search space of a worker can reduce the computing time of the model. Every time an agent is activated, she will randomly sample Θ firms to request offers.

4.2.2. Implementation and Execution

The model was implemented in the Python programming language version 2.7 (code provided in additional materials). It consists of three components:

- *Worker*: An object class of an agent that represents a worker. It contains the algorithms through which a worker searches for the best job offer.

- *Firm*: An object class of an agent that represent a firm. It contains the algorithms through which a firm computes offers and hires workers.

- *Model*: Contains the populations of workers and firms. It executes the model.

The model execution is serial, meaning that when an agent is active, all other agents have to wait for her to finish. Activation is random, defined by the layoff probability λ. In other words, a period is defined when a fraction λ of the population has been laid off. The group of agents to be activated is picked at random. The execution is summarized in Pseudocode 1.

Pseudocode 1: Execution of firm dynamics model

```
INITIALIZE workers and firms;
FOR each worker, get hired by random firm;
WHILE t<T DO:
    FOR each worker, separate with probability λ:
        Sample Θ firms;
        FOR each firm in sample:
            Count N and Y;
            Compute w;
            IF first time called durint t:
                Update b;
            Return offer (w, b);
        Get hired by firm with best offer;
    Increment t;
```

4.2.3. Calibration

We calibrated some parameters by combining empirical US data. Some parameters were calibrated with a single value and others as random variables coming from empirical distributions estimated from the data. Parameter heterogeneity is important for the system's stability because it prevents the degeneration of the emerging distributions.

4.2.3.1. Number of Workers

The number of workers is calibrated in relation to the number of firms. We use the average firm size of the US, estimated by Axtell (2001) as nearly 20 employees. Therefore, for experiments with different number of firms, the number of workers is always proportional.

4.2.3.2. Layoff Rate

Since the model assumes no unemployment, a layoff is equivalent to a job search. Hence, the labor dynamics in these models are produced by on-the-job search. This motivates an empirical calibration for λ. Fallick and Fleischman (2001) estimated that the percentage of employed individuals that search for a job every month is 4%. Assuming that this number is evenly distributed per week, we calibrate our model to have a weekly basis. Therefore, one period is completed every time that 1% of the employed population has been laid off.

4.2.3.1. Duration of Training

Using the NLSY79, Frazis and Spletzer (2005) estimated that the average duration of on-the-job training is 5.7 weeks. This number is not representative of the diversity of training requirements that firms have, especially since there is a strong relationship between employer size and decisions in training investment (Holtmann & Idson, 1991). We estimated the empirical distribution of training duration from the NLSY79 using information about the length in weeks of the on-the-job training for the cross sections between 1988 and 2010. Panel A in Figure 11 shows that the training durations are well approximated by a gamma distribution. We used this empirical model to provide firms with a way to establish the training requirement τ every time they hire a new employee.

4.2.3.2. Search Space

An important parameter is the number of firms to which an agent applies for a job. We used the NLSY79 dataset, which includes information about how many employers were contacted by the job seeker before getting her last job. On average, workers reach out to 12 employers. We estimated the empirical distribution of this quantity using the supplementary information section from the 1982 cross section. This quantity follows a gamma distribution (see Panel B in Figure 11). Therefore, every time that an agent is activated, she uses the empirical distribution to determine the number of firms to sample.

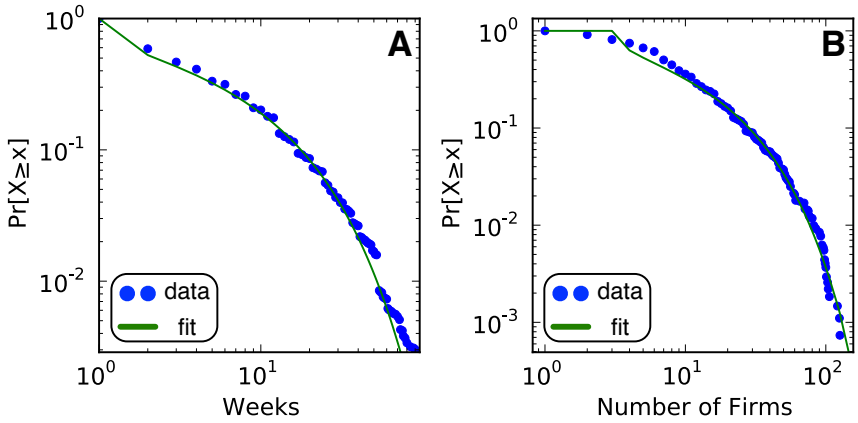

Figure 11: Training duration and search space

Both Panels show empirical and estimated cumulative density functions. The estimated distribution is a gamma for both cases. Panel A presents the distribution of training durations in weeks. Panel B illustrates the distribution of number of firms contacted by job seekers during their job search. Source: own calculations from the NLSY79.

Table 6 provides a description of the model's parameters. The values that were experimentally calibrated correspond to the baseline case. They were also tested for other values in order to check the robustness of the results.

Table 6: Parameters of firm dynamics model

Parameter	Value	Calibration	Change	Homogeneity
Number of firms	5,000	Experimental	Static	N/A
Average firm size	20	Empirical	Static	N/A
Number of periods, T	10,000	Experimental	Static	N/A
Layoff probability, λ	0.01	Empirical	Static	homogenous
Labor output elasticity, α	*Uniform* on (0,1)	Experimental	Static	heterogeneous
Firm benefits, b	*Uniform* on (0,1)	Experimental	Dynamic	heterogeneous
Training duration requirement, τ	*Gamma*	Empirical	Dynamic	heterogeneous
Number of firms to search, Θ	*Gamma*	Empirical	Dynamic	heterogeneous

This describes the nature of each parameter in terms of different characteristics expressed by the columns. The column *Value* denotes the values assigned to the parameters during the baseline simulations, or the probabilistic distributions from where the values are drawn. The *Calibration* column indicates if the source of the values is from empirical data or product of experimental procedures that are conventional in agent-based modeling. The *Change* column shows if the parameter is constant throughout a simulation or if it changes with each period. Finally, the *Homogeneity* column indicates if the value is the same for all agents.

4.2.4. Experimental Design and Output

We ran a sample of 36 simulations. The simulations were run independently from each other with initial seeds assigned by a unique time stamp. We employed the computational resources from the Department of Computational Social Science at George Mason University. The 36 simulations were distributed across four cores in a PC with 16 cores and 256 GB of memory.

The duration of each simulation (10,000 steps) is motivated by the well-known transient effects from ABMs, where the system shifts form its initial state to its steady state. We saved records for the last 500 steps. Each record was collected on a monthly basis (every four steps).

The records were stored as administrative micro-data. Each line of a record contains information about the state of an agent at a specific step of the simulation. Each row represents a variable of the agent. Records were organized in two tables: workers and firms. Both tables were stored in a hierarchical data file with format HDF5 and compressed with the library LZO. We employed the data management library PyTables to speed up the data recording process. It took approximately two weeks to perform all the simulations.

4.2.5. Results

The graphical results in this section are taken form the last four years of a single simulation. They have been replicated in each of the 36 simulations They are robust across the sample of simulations, so we will focus on the exposition of a representative case Although we were able to fit most of our outputs to statistical models, we will not elaborate on the details of any particular estimator since we are more interested on the general structure of the data that emerges from the model.

Our first results relate to firm dynamics. This model is able to recover the heavy-tailed firm size distribution from the equilibrium model of Gabaix. However, the

distribution here is a Weibull with shape parameter smaller than one (see Panel A in Figure 12), instead of a Pareto distribution with exponent one. It would appear that the assumptions of complete information and equilibrium are instrumental in Gabaix's model in order to arrive to a Zipf law, suggesting that the underlying microeconomic model by itself is not a plausible explanation about the emergence of this stylized fact. An alternative way to measure the size of a firm is by its output. In this case, firm sizes do follow a Pareto distribution with exponent between one and two (see Panel B in Figure 12).

This model is able to generate empirically valid growth rates of firms. Panel C in Figure 12 shows the tent-shaped distribution of the log-growth rates. We measured growth for one month, one year, and four years. From Panel C in Figure 12, it is clear that longer time frames yield wider distributions, which matches empirical studies. The distributions are symmetric, so they are well fitted by a Laplace law. Another stylized fact that we emerged is the scaling of growth volatility with company size. The standard deviation of the log-growth rates decays with firm size (in number of employees), following a power law (Panel D from Figure 12). The exponent of the relationship is around 0.35, which matches our empirical findings in Section 3. Notice that small firms experience weaker scaling effects just as shown in Section 3.3.2.

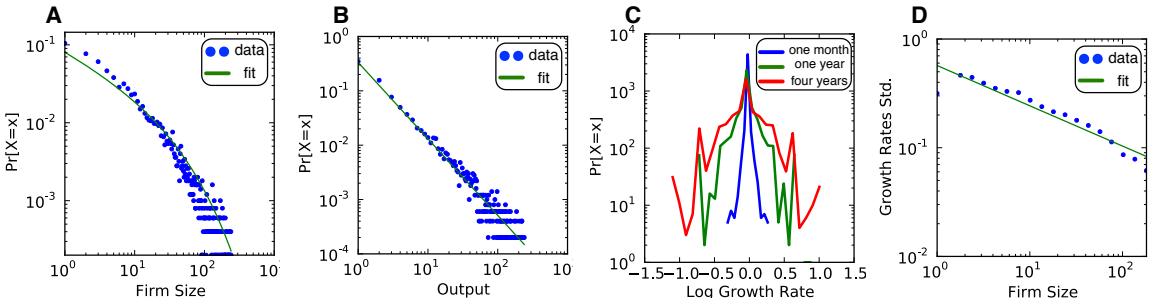

Figure 12: Output from firm dynamics

In Panel A we fitted a Weibull distribution. In Panel B we fitted a Pareto distribution. Panel C does not show the fits for the Laplace distributions due to oversaturation of lines. In Panel D we fitted a power law. The theoretical models in Panels A, B, and C were fit using MLE, while the one of Panel D was estimated via OLS.

This model allows us to investigate the role of technology in productivity. Here we refer to the production function of firms as technology. More specifically, we focus on the elasticity of tenured workers, which is captured by the parameter α from Equation 5. In Panel A from Figure 13 we can observe that firms that put more weight on tenured workers tend to be smaller. From a static perspective this seems counter intuitive since firms with higher α are able to offer higher wages than their competitors of the same size. Hence, they should attract more workers and become larger. The answer to this problem lies in the dynamics of the system. Since firms with higher α are more attractive at the beginning, they over saturate with new workers quickly. This allows firms with lower α to higher more people. Since the number of new hires is larger than α by orders of magnitude, this effect remains for most of the matching process. Therefore, larger firms are those that are able to produce more with new workers.

When we compute productivity of firms as output per capita, we find that technology has a nonlinear effect. Panel B in Figure 13 shows the U-shaped relationship

between α and output per capita. Companies with a well-balanced weight between tenured and new hires are the least productive. According to this model, small and large firms are the most productive.

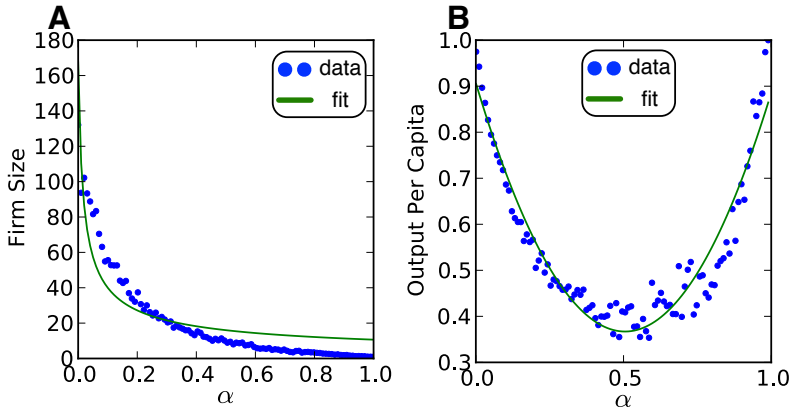

Figure 13: Technology and productivity

Panel A shows a fitted logistic function via the Levenberg-Marquardt algorithm. The data from Panel B was fitted with a polynomial of order two using OLS.

Due to the flexibility of agent-based modeling, we can turn and look at other micro-data that is generated by the model and that is relevant to labor markets. Panel A in Figure 14 shows the income distribution, which is highly skewe, but no heavy-tailed. It is best approximated by a beta distribution. In Panel B of Figure 14 we can observe the employer size premium. This stylized fact follows the super linear pattern documented by empirical studies. We approximate it with a concave function of the form x^a, where $0<a<1$.

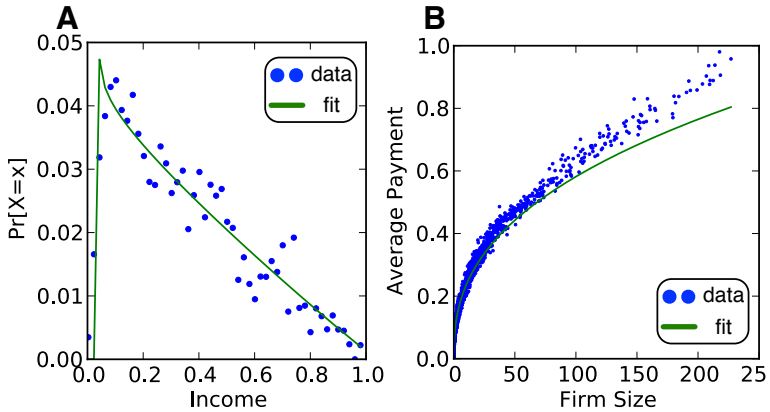

Figure 14: Income-related outputs

We fitted a beta distribution in Panel A using MLE. Panel B shows the fit of a concave function estimated with OLS.

Finally, we look at the patterns of labor flows generated by the model in the form of a LFN. Panel A in Figure 15 shows the heavy-tailed character of the LFN. Its degree distribution is a Weibull with shape parameter of less than one. The similarities between the firm size and the degree distributions suggest a connection between them. In Panel B from Figure 15 we show a high linear correlation (0.99 Pearson coefficient) between them, which matches our empirical findings from Section 3.

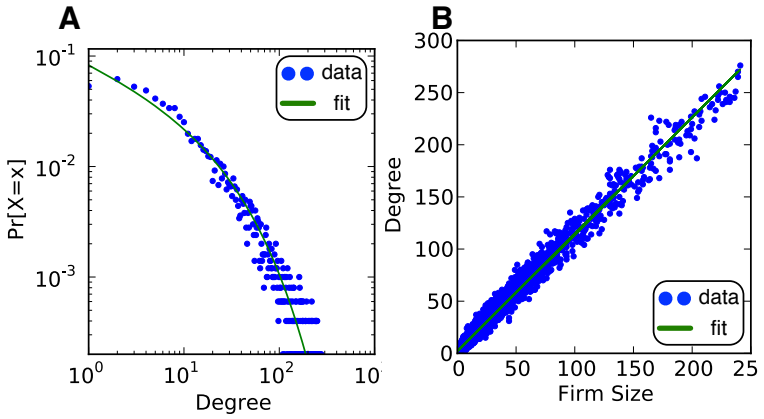

Figure 15: LFN outputs

In Panel A we fitted a Weibull distribution via MLE. In Panel B we fitted a linear function using OLS.

4.3. Labor Dynamics Through Endogenous Social Networks

Social capital is an important topic in the social sciences initially in sociology (Ben-Porath, 1980; Bourdieu, 1983; Burt, 1992; Coleman, 1988; Erickson, 1996; Flap & De Graaf, 1986; Lin, 1982; Loury, 1977; Portes, 1998; Putnam, 1993) and more recently in economics (Glaeser, Laibson, & Sacerdote, 2002; Guiso, Sapienza, & Zingales, 2004; Knack & Keefer, 1997; Narayan, 1999; Woolcock & Narayan, 2000). Although there is not a final consensus about its definition and measurement, it is generally agreed that the underlying principle is the idea that investment in social relations yield economic returns (Lin, 2001). While sociologists have focused on the structural aspects of social interactions that generate social capital (e.g. social networks), economists have paid closer attention to the social norms and institutions that promote social capital from a strategic perspective (e.g. trust games).

In the context of labor markets, social capital can be understood as the network of personal contacts that provides access to job opportunities. Pioneered by (Granovetter, 1973), the study of social networks in labor markets has focused in understanding the implications of personal contacts in labor outcomes such as unemployment (Montgomery, 1991), earnings (Mortensen & Vishwanath, 1994), inequality (Calvó-Armengol & Jackson, 2004, 2007; Ioannides & Loury, 2004), information transmission (Boorman, 1975; Calvó-Armengol & Zenou, 2005; Rees, 1966; C. Simon & Warner, 1992), and labor mobility (Munshi, 2003; Wahba & Zenou, 2005). These and other works have contributed to demonstrate that social networks matter significantly in labor outcomes. Unfortunately, the unavailability of large-scale data and the equilibrium-oriented approach to model network formation processes (Calvó-Armengol, 2004; Krauth, 2003, 2004) have been an obstacle to understand how the complex structure of social networks that arise form economic interaction affects macroeconomic variables.

This model contributes to understanding the effects of social networks in labor dynamics by endogenizing the formation process of personal contacts and generating labor dynamics through job offers that are made available through the social network. A distinctive feature of this model is that the topology of the network is not the intended consequence of strategic interactions but a side product of economic behavior and social interaction. Nevertheless, incentives and preferences play a crucial role in shaping the social network, which in turn affects labor outcomes.

4.3.1. Model Description

There is a fixed population of firms and workers. Firms hire unemployed workers and separate employees with a constant probability λ. Workers are endowed with k skills. Agents have two possible states: employed and unemployed. Their behavior depends on their active state.

4.3.1.1. Firms

We use the framework described in Section 4.2.1 to motivate how firms determine wage offers. We re-use Equation 5, with the adaptation that T and N are amounts of incumbent and newly hired *skills* respectively. Thus, any wage offer is per-skill unit. When a firm hires a worker it pays the wage offer multiplied by the amount of skills that the worker brings to the firm.

4.3.1.2. Effort in the Workplace

Workers that are employed have to face a two-step problem. The first step corresponds to the allocation of efforts. At the workplace, a worker has preferences over the amount of effective effort e that she puts into working and the excedent effort x that she assigns to other activities such as socializing or browsing the internet. This worker solves a standard utility maximization problem:

Equation 6

$$\max_{\{e,x\}} \quad U(e,x) = e^{\beta} x^{1-\beta}$$
$$s.t. \qquad e + x \leq k$$

Recall that k represents the endowment of skills of the agent. Then, it is clear that skillful agents are able to reach higher utility levels. The optimal effort $e^*=\beta k$ depends on the individual preference for working. Therefore, agents with higher β assign less excedent effort to socializing, relative to agents with the same skills.

4.3.1.3. Socializing in the Workplace

Once a worker has chosen e^* and x^* she has the possibility of increasing her utility beyond the limits that her skills allow her. This takes the form of an additive term γ that comes from socializing with other workers. If $\gamma>0$ it means that socialization has positive effects. Similarly, $\gamma<0$ means negative effects. These effects determine if a worker receives utility above or bellow the optimal.

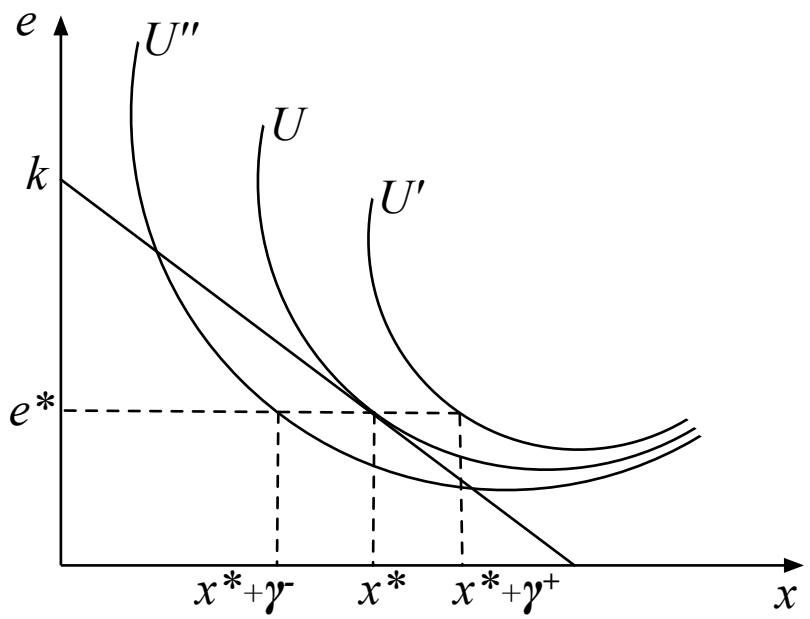

Figure 16: Effects of socialization in agent utility

The figure illustrates how the additive term γ affects the utility levels obtained by the worker. The superscript of the additive term indicates if it is positive or negative. The red line represents the skills constraint of the agent.

The outcome of socializing with another colleague determines if an agent receives utility beyond her possibilities. This brings us to the second step in the agent's behavior. Here, agent *i* has to decide if she will attempt to socialize with her colleague *j*. If so, *i* approaches *j*, then the latter decides to correspond or not. We can think of this process as a sequential game with agents: $A=\{1,2\}$; strategies $S=\{I, NI\}$ (where *I* stands for interact and *NI* for not interact). Payoffs are specified in the extensive form shown in Figure 17.

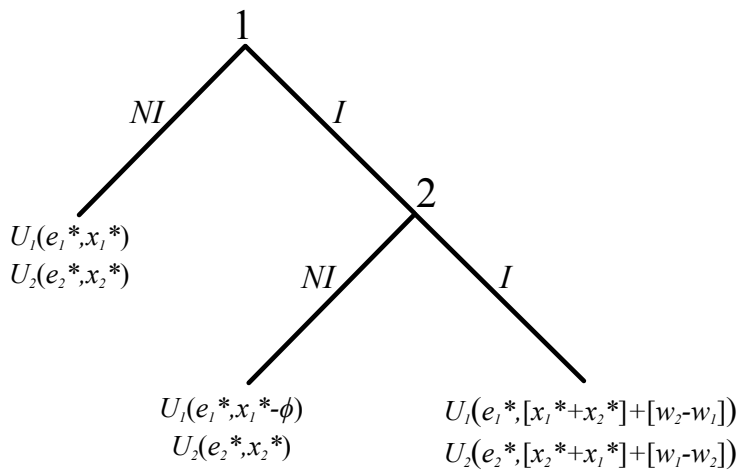

Figure 17: Social tie formation game

The diagram represents the sequential game that determines if a social tie is formed between two co-workers. Agent 1 is the first mover, followed by agent 2. Strategies *I* and *NI* stand for *interact* and *not interact* respectively. The utility functions have the functional formed presented in Equation 6, and are indexed for each agent because they can have different parameterizations. The terms e^* and x^* correspond to the optimal levels of effective and excedent efforts respectively. Parameter ϕ represents a fraction of excedent effort used by of agent 1 to unsuccessfully approach agent 2. It is equivalent to the negative additive term in Figure 16. Parameter w is the wage of each agent.

The payoff structure is rather intuitive. If agent i decides not to approach j, they will spend their corresponding excedent efforts in individual activities like reading their favorite blogs. Let $0<\phi$, then if the outcome is (*I*, *NI*) agent i receives less utility than in (*NI*, *NI*) because she spent ϕ of her excedent effort in reaching out to her colleague. If the outcome is (*I*, *I*), they receive the utility that corresponds to the sum of their interaction. The intuition for $[x_i^*+x_j^*]+[w_j-w_i]$ is simple. The first summand represents the combination of excedent efforts that both workers are dedicating to the interaction. The second summand corresponds to the expectations that agents have about the potential benefits that socializing with another agent might yield. Potential benefits are modeled as the similarity between i and j. In this model, similarity can be represented in terms of skills. However, workers cannot observe skills of other workers directly.

We use the agents' wages as a proxy of similarity since it is a natural way in which employees perceive the status of a colleague in the same firm. Furthermore, from Equation 5 we know that wages are correlated to skills and to the size of the firm *at the moment that the agent is hired*. This generates endogenous noise in the signaling of agents' skills because an agent with low skills that was hired when the firm was small has a chance for becoming acquainted with a more skilled worker who was hired when the firm became large. It is a natural way to endogenously correlate skills in the social network while allowing some degree of heterogeneity.

In the outcome (*I*, *I*), the interplay between $[x_i^*+x_j^*]$ and $[w_j-w_i]$ depends on the preferences of agents, the dynamics of their firm, and their endowment of skills. This provides the model with a variety of possible payoffs for outcome (*I*, *I*) that capture sociological aspects of human interaction such as homophily.

4.3.1.4. Micro-Equilibrium and Link Formation

Figure 17 shows a simple game that represents the interaction between co-workers. Finding the Nash equilibrium (NE) is straightforward via backwards induction. Since this mechanism leads to socialization, we would like to know under what circumstances such interaction would lead to the creation of social ties. If the outcome (*I*, *I*) is the only NE, then a link is formed between both workers. Hence, we only need to know the conditions under which none of them would have incentives to deviate.

Rearranging $[x_i^*+x_j^*]+[w_j-w_i]$ we obtain the additive term $\gamma_i = x_j^*+w_j-w_i$ from Figure 16. This determines if outcome (I, I) is a NE. Therefore, the condition for forming a social tie between two agents is that $\gamma>0$ for both i and j. In terms of the model's parameters this is: $x_1^*>w_2-w_1$ and $x_2^*>w_1-w_2$. The intuition is straightforward: an edge will be formed if the worse paid agent has strong enough preferences for excedent effort such that it compensates the dissimilarities between both workers. In other words, if the agent who would be most benefited from the social tie is socially skilled enough, she can convince her counterpart to engage in socialization. This is similarly to the way in which some employees try to build relations with their superiors in the organization.

Figure 18 shows that the implementation of the link formation process is robust for different distributions of skills and wages. We ran Monte Carlo simulations for 10,000 agents and 10,000 pairs of randomly matched agents. Then we to computed the frequency of edges that were formed for different levels of wage differences between the matched pairs. We tested different distributions of skills and wages: uniform (for integers between 1 and 30), binomial (with p=0.3), and geometric with (n=30 and p=0.5). Agents' preference parameter was β~U(0,1). It is clear that, regardless of the configuration of skills and wages in a firm, the endogenous social networks tend to be homophilic. However, there are multiple ties between agents with large differences of wages. The largest effect comes from the distribution of wages. Therefore, if the system endogenously generates a skewed income distribution, it induces a higher density of the social network.

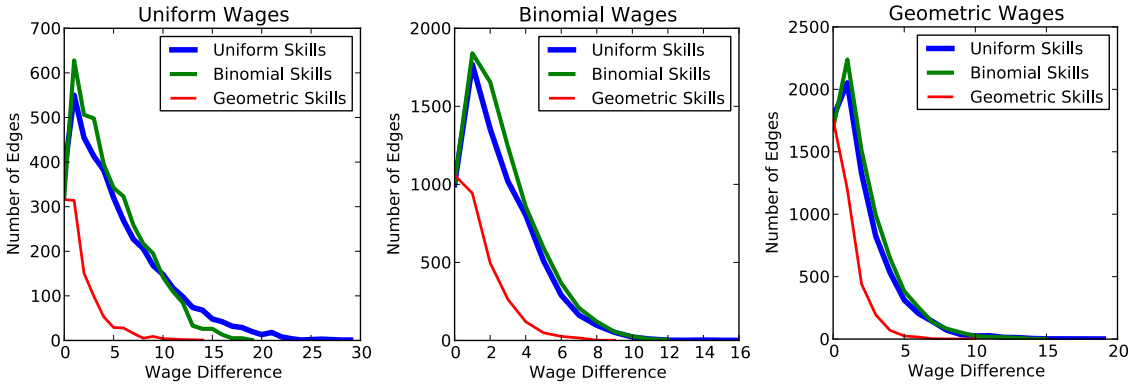

Figure 18: Frequencies of social ties by differences in wages

Each Panel presents Monte Carlo simulations of the link formation process. The *y*-axis shows the average number of social ties from 1,000 simulations. Each Panel has used a single distribution of wages. Inside each Panel, each line represents the results for a specific distribution of skills in combination with the wage distribution of the corresponding Panel.

4.3.1.5. Finding a Job

Every period, a worker with a state of unemployment requests job offers from each one of her personal contacts (excluding the ones working for her last employer), and submits one formal application to a random firm in the economy. With an arrival probability θ, the unemployed agent will receive a job offer from each of her requests. Conventional models endogenize this probability as a function of the intensity of the search (Jovanovic, 1979; McCall, 1970; Mortensen, 1973). For simplicity, we do not model the optimal search effort problem, as it involves incorporating costs and discounting parameters to the model. Instead, we construct the arrival probability as a function of the level of skills of the agent and her preferences for work effort β.

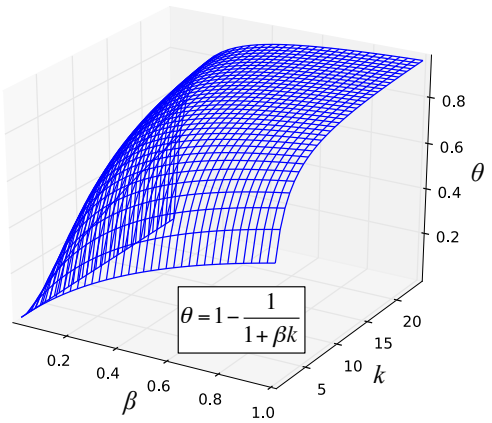

Figure 19: Arrival probability

Figure 19 shows a surface plot of the arrival probability. The construction of θ is rather intuitive. We use the preference parameter β as the effort level that an agent would exert in searching for a job. If an agent has stronger preferences for effective effort, she is more likely to search with higher intensity. However, the success with which an agent finds a job is bounded by her skills. Skillful agents can achieve higher probabilities. This specification provides heterogeneous arrival probabilities across the population of job seekers.

The social network plays a crucial role in successfully receiving job offers. The agent's degree has a positive relationship with the expected number of job offers. More importantly, less skilled agents rely more on the social network in order to become employed, which has been empirically documented (Holzer, 1988; Topa, 2001; Wahba &

Zenou, 2005). An agent who receives multiple offers will choose the best one. In this way, agents with rich social networks induce firm dynamics similar to the ones described in Section 4.2.

4.3.1.6. Network Density

A potential artifact that can be introduced by the network formation process from Section 4.3.1.3 is the eventual drop in unemployment rates due to a persistent increase of network density. A common practice is to introduce a decay rate for the edges and a threshold to remove them. This solution is computationally expensive, as it requires iterating through every edge periodically. Additionally, we would need to include parameters for the decay rate, the threshold, and the edge weights. We use a simpler approach by setting an upper bound to the density of the network. At the beginning of every period we check if the number of edges exceeds the upper bound. If this is positive, we remove the excedent edges by randomly sampling them. This provides a dynamically stable character to the social network.

4.3.2. Implementation and Execution

The model was implemented in the Python programming language version 2.7 (code provided in additional materials). It consists of three components:

- *Worker*: An object class of an agent that represents a worker. It contains the algorithms through which a worker searches for jobs in the social network.

- *Firm*: An object class of an agent that represents a firm. It contains the algorithms through which a firm computes offers and hires workers. It also contains the algorithm to match pairs of employees and determine if a social tie is formed.

- *Model*: Contains the populations of workers and firms. Determines the separation and job search rates of the agents. It contains the graph that keeps track of the social network. It executes the model.

The model execution is serial. Activation is uniform, meaning that all agents are active once per period. This is motivated by the calibration described in Section 4.3.3, where one period represents a month. During each activation, behaviors are executed depending on the state of the agent. Active employed agents always work and socialize. With probability λ each one loses her job. Unemployed agents always search for jobs.

A conventional norm in agent-based economic models is that agents in different states are activated in groups. For example, most agent-based models of labor markets group unemployed agents, employed workers, and firms into separate classes. Then all the agents are randomly activated *in each group*. However, agents with different states are not randomized.

In the context of parallel execution of the spatial prisoner's dilemma (Nowak & May, 1993), it has been shown that synchronized activations of agents can produce artifacts that generate misleading aggregate patterns (Huberman & Glance, 1993).

Although most agent-based models of labor markets are executed sequentially, in our experience, reserving separate activations for groups of agents also generates artifacts. To solve this problem, we implement *mixed activations*. In a mixed activation, all types of agents are pooled together and are called in random order. This guarantees a truly randomized experimental design. The model execution is summarized in Pseudocode 2.

Pseudocode 2: Execution of unemployment and endogenous social networks model

```
INITIALIZE workers and firms;
FOR each worker, get hired by random firm;
WHILE t<T DO:
    IF density of network > maximum density:
        Remove excess edges at trandom;
    FOR each worker:
        IF worker is employed:
            Compute e* and x*;
            Socialize with a random worker from the same firm;
            Separate with probability λ;
        ELSE:
            Compute r and θ;
            Get job offer from random firm with probability θ;
            FOR each firm of each employed neighbor:
                IF firm is different from last employer:
                    Get job offer with probability θ;
            IF number of offers > 0:
                Get hired by firm with best offer;
    Increment t;
```

4.3.3. Calibration

We calibrated the model using US empirical data. Some parameters are calibrated with a single value and others as random variables from the estimated empirical distributions.

4.3.3.1. Number of Workers

The number of workers is calibrated in relation to the number of firms. More specifically, we use the average firm size of the US, estimated by Axtell (2001) as nearly 20 employees. Therefore, for experiments with different number of firms, the number of workers is always proportional.

4.3.3.2. Layoff Rate

This is an unemployment model. Therefore, we calibrated the separation rate λ=1.71% according to the estimation for the US (Hobijn & Şahin, 2009). This estimation is the monthly average for 2006. Hence one period in our model represents one month.

4.3.3.3. Duration of Training

Training duration is a random variable drawn every time a firm hires an agent. This variable follows the gamma distribution estimated in Section 4.2.3.1 for the US. The variable was adjusted to monthly values.

4.3.3.4. Skills

The representation of skills has been an active topic in the human capital and labor economics literature. The standard approach was pioneered by Mincer's earnings equation (Mincer, 1974) where individual earnings are a function of the years of education and potential experience. Other studies have developed measures of skills using data on educational attainment from multiple countries (Barro & Lee, 1993, 1996). Here, we use some of these ideas to calibrate the skills of the workers as years of education attained. We used education data from the US population with 25 years or more of age in 2012. Figure 20 shows the bimodal character of different levels of education (Panel A). We estimated the distribution of years of education by assigning the approximate time that it takes for each level to be attained. We generated synthetic data using this information and estimated a kernel density with domain between zero and 22 years of education (Panel B). We use this distribution to assign the endowment of skills to the agents.

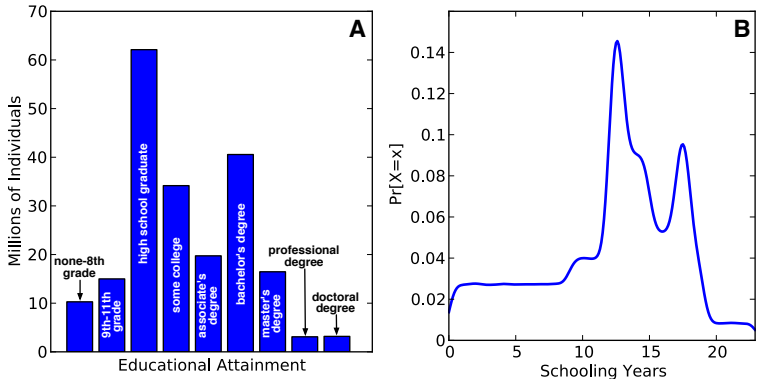

Figure 20: Empirical distribution of skills

Panel A shows the frequency of different educational attainment levels in the US population of 25 years old and over for 2012. Levels of educational attainment were translated into schooling years for Panel B. Panel B presents the Gaussian kernel density estimation for the schooling years of the US population. Source: own calculations using educational attainment data from the US Census.

4.3.3.5. Network Density

Some studies have attempted to estimate the structure of social networks under certain labor market environments (Munshi, 2003; Wahba & Zenou, 2005). However, they are far from being able to provide specific numbers about the density of social networks in entire economies. Despite recent efforts to quantify the job contact networks using online data (Gee & Jones, 2013), it is still not clear what is an empirically relevant value for this parameter. We have found that the network density has an effect in the magnitude of the unemployment rate in our model. Therefore, this parameter will be calibrated experimentally in order to target realistic unemployment rates.

Table 6 provides a description of the model's parameters. The values that were experimentally calibrated correspond to the baseline case. They were also tested for other values in order to check the robustness of the results.

Table 7: Parameters of unemployment and endogenous social networks model

Parameter	Value	Calibration	Change	Homogeneity
Number of workers	10,000	Experimental	Static	N/A
Average firm size	20	Empirical	Static	N/A
Number of periods, T	10,000	Experimental	Static	N/A
Separation rate, λ	0.0171	Empirical	Static	homogenous
Effort preference, β	*Uniform* on (0,1)	Experimental	Static	heterogeneous
Training duration requirement, τ	*Gamma*	Empirical	Dynamic	heterogeneous
Network density	0.001	Experimental	Static	N/A

This table describes the nature of each parameter in terms of different characteristics expressed by the columns. The column *Value* denotes the values assigned to the parameters during the baseline simulations, or the probabilistic distributions from where the values are drawn. The *Calibration* column indicates if the source of the values is from empirical data or product of experimental procedures that are conventional in agent-based modeling. The *Change* column shows if the parameter is constant throughout a simulation or if it changes with each period. Finally, the *Homogeneity* column indicates if the value is the same for all agents.

4.3.4. Experimental Design and Output

We ran a sample of 36 simulations. The simulations were run independently from each other, with initial seeds assigned by a unique time stamp. We employed the computational resources from the Department of Computational Social Science at George Mason University. More specifically, the 36 simulations were distributed across four cores in a PC with 16 cores and 256 GB of memory.

We saved records for the last 500 steps. Each record was collected on a monthly basis (every step). The records were stored as administrative micro-data. Each line of a record contains information about the state of an agent at a specific step of the simulation. Each row represents a variable of the agent. Records were organized into three tables: workers, firms, and social network. Each line in the social network table represents an

edge. They consist of the source node, target node, and period of collection. All tables were stored in a hierarchical data file, with format HDF5, and compressed with the library LZO. We used the data management library PyTables to speed up the data recording process. It took approximately two week to perform all the simulations.

4.3.5. Results

The graphical results in this Section are taken form the last four years of a single simulation. They have been replicated in each of the 36 simulations They are robust across the sample of simulations, so we will focus on the exposition of a representative case Although we were able to fit most of our outputs to statistical models, we will not elaborate on the details of any particular estimator since we are more interested on the general structure of the data that emerges from this model.

This model generates realistic firm dynamics. The firm size distribution is skewed but degenerated. The upper tail of the distribution does not decay smoothly. We suspect this is due to the smaller scale of the simulations (running significantly larger simulations becomes computationally unfishable in the current state of the model due to the network computations). However, the growth process of firms is replicated accurately. Panel A in Figure 21 shows the tent-shaped log-growth rates distributions for different time frames. In Panel B we confirm the power law relationship between firm size and the standard deviation of growth rates. Finally, Panel C shows the stylized fact of high correlation between firm size and firm degree in the LFN (0.97 Pearson coefficient).

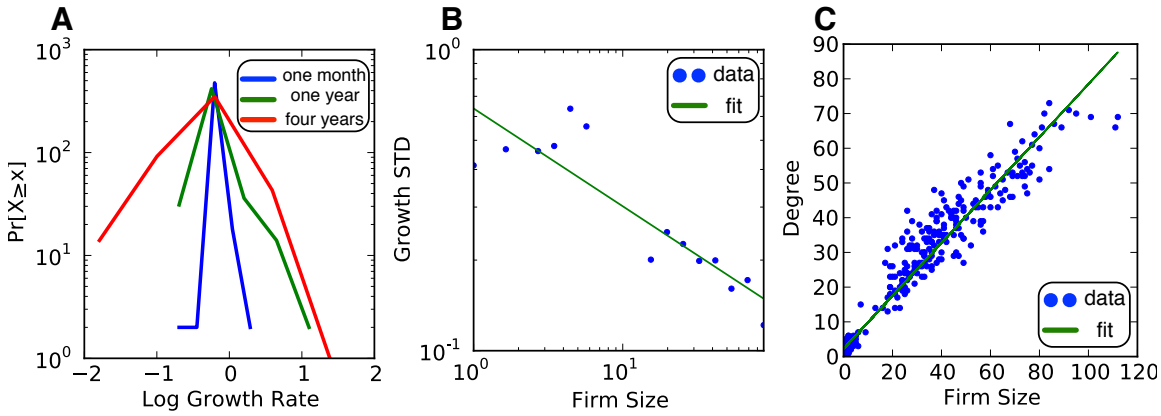

Figure 21: Outputs from firm dynamics

We fitted a power law to the data in Panel B using OLS. In Panel C we fitted a linear function using OLS.

Regarding the role of technology in productivity, the results of this model are opposite to the previous model. The reason is the specification of the production function. In the previous model, there is a direct relationship between firm size and output. In this model, this relationship is obscured by its dependence on skills, the referral hiring, and the preferences for effective work. In Panel A from Figure 22 we can observe that α has no effect on firm size when it is bellow 0.4. The largest companies are those that are able to produce more with already-trained skills. When we compute firm productivity, we find that there is a positive effect of α, with the exception of extreme values (see Panel B). This effect seems linear and divided in two regimes: below and above α=0.5. The latter is ten times stronger than the former (the slopes are 2 and 20 respectively).

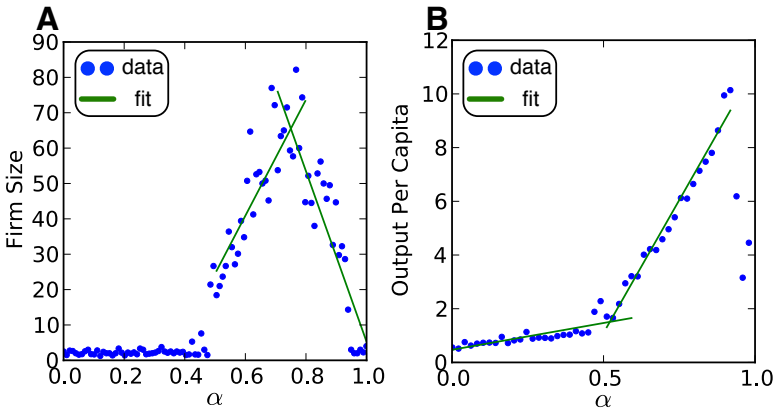

Figure 22: Technology and firms
We fitted Sections of the data in Panels A and B to separate linear models using OLS.

We investigate the scaling effects that firm size exerts on other characteristics. In Panel A from Figure 23 we can see that firm size has a positive effect in output per capita with diminishing marginal returns. As expected by construction, there is a high linear correlation between size and the amount of skills accumulated by a firm (Panel B). Panel C illustrates how the average skill level of an employee increases dramatically with firm size for employers that have between 0 and 20 workers. After that, the skill level per capita increases moderately with size. Finally, Panel D shows the employer size premium.

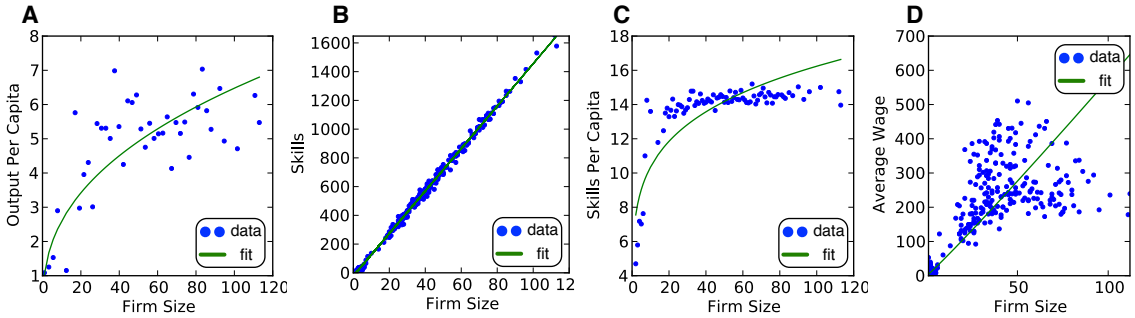

Figure 23: Scaling with firm size

Panel A shows the fit of a concave function. In Panel B we fitted a linear function. Panel C presents the fit of a concave function. Panel D shows the fit of a linear function. All the models were fitted via OLS.

The income distribution generated by the model is skewed. However, it does not fit conventional income distributions (lognormal, exponential, Pareto, etc.) (see Panel A in Figure 24). Panel B from Figure 24 shows the relationship between the endowment of skills and income. This is a realistic result for three reasons. The first is positive returns to education (which is coherent with the Mincerian earnings equation). Second, education has increasing marginal returns during the first years (consistent with evidence about the importance of early cognitive skills in labor prospects). Third, once a high educational level has been attained (e.g. a high school or college diploma), additional education has decreasing marginal returns (or even null).

An additional question that we explore is if preferences have some effect on income. Recall that the parameter β from Equation 6 denotes the preference of agents towards effective work. Panel C from Figure 24 shows an inverted-U relationship between β and income level. It means that individuals with well-balanced preferences for

working and socializing have higher earnings. It shows the importance of social skills in a labor market where social networks are key to obtain access to better job opportunities.

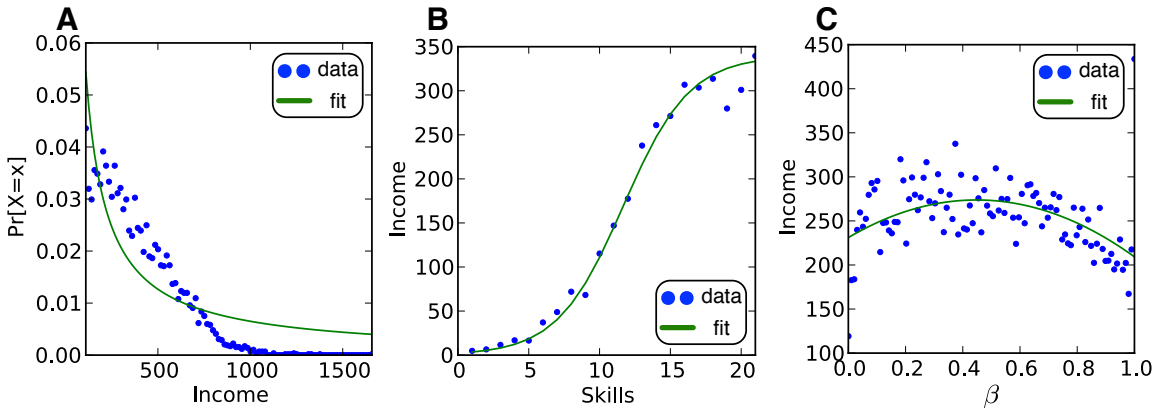

Figure 24: Income-related outputs

Panels A and B show the fit of logistic functions using the Levenberg-Marquardt algorithm. In Panel C we fitted a polynomial of order two via OLS.

Now we turn to labor outcomes that are important for macroeconomic policies. Our model is able to emerge two important stylized facts from the macroeconomic literature. The first if the Phillips curve (see Panel A from Figure 25), which exhibits a negative relationship between the change in real wages and the unemployment rate. The second is Okun's law (Panel B from Figure 25), which shows a negative relationship between growth rates of the economy and changes in unemployment rates. We should point out that both of these macroeconomic empirical regularities have been generated without the need of aggregate assumptions about the workings of the market, providing evidence of the empirical validity of the model.

Additionally, we are able to endogenously produce dynamic unemployment rates (Panel C from Figure 25). The average unemployment rate for a period of four years is approximately 1.8% with a standard deviation of 0.13%. We can control the magnitude of the unemployment rate by setting the density of the social network. Therefore, the model can be easily calibrated to produce desired levels of unemployment. Panel D in Figure 25 shows a time series of the average duration of unemployment spells. An unemployment spell lasts six months and varies by 12 weeks on average. We can induce higher duration levels by controlling the density of the social network. Once the transient has been passed, both time series are stable throughout the simulation.

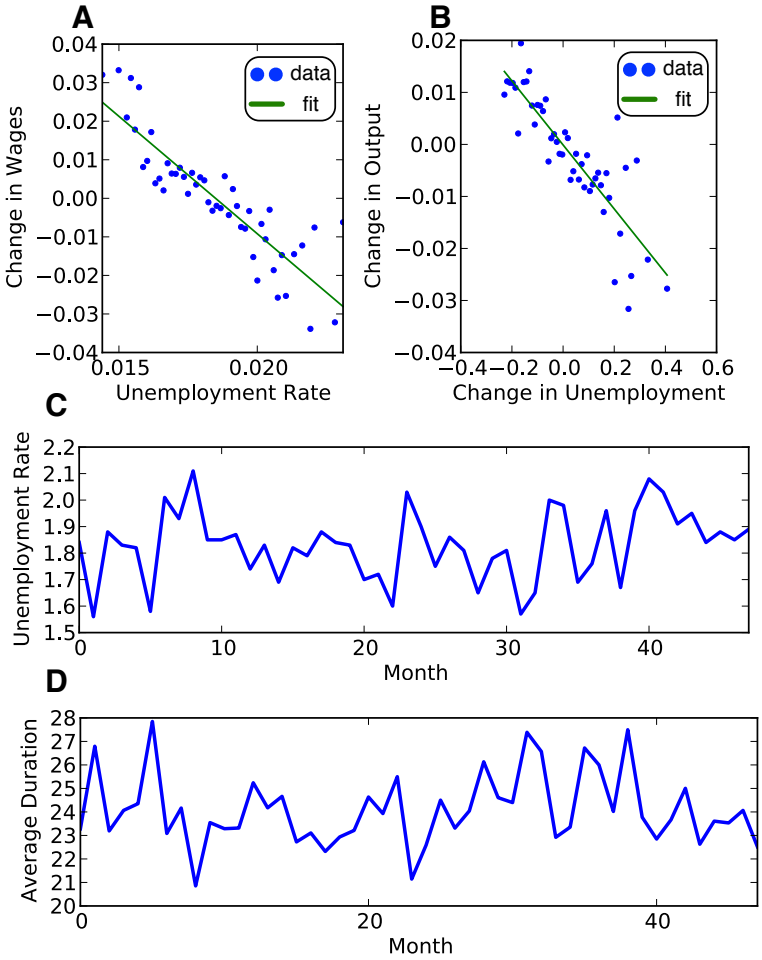

Figure 25: Macroeconomic outputs

Models in Panels A and B were fitted using linear functions via OLS. Panels C and D present time series for four years on a monthly basis.

We take a look at the employment status of agents and how it is correlated to their skills. Panel A in Figure 26 shows a positive relationship between the skills of agents and their employment status. Employment status is a binary variable where 1 means employed and 0 unemployed. The data points represent the average of the employment status for each skill cohort. Skills and average employment are linearly correlated with a Pearson coefficient of 0.82. The distribution of the length of unemployment spells is

skewed. Here we approximate it with a Pareto distribution. However, more data points would allow us to refine the model selection.

It is clear that the skill of an individual affects her chances to be employed. This is an emerging result since the model does not have explicit formulations where firms discriminate between workers. The explanation is in the access that agents have to job opportunities which is determined by the social network. Since skillful agents have more "resources" to socialize, they should be better equipped to confront situations of unemployment. Therefore, it is important to analyze the social network that emerged from the model and its implications in the labor market.

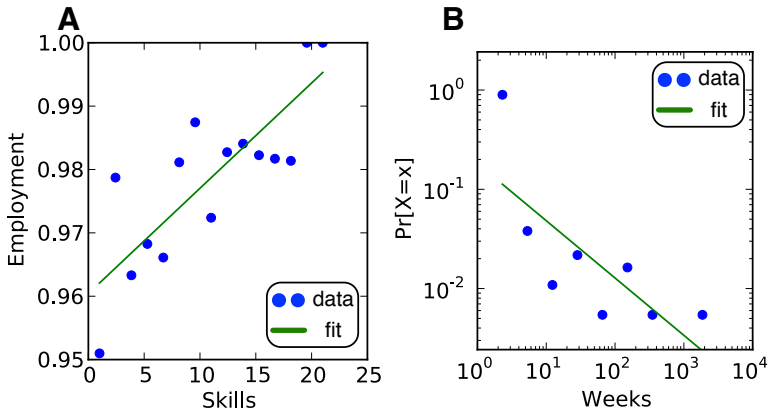

Figure 26: Unemployment and duration

The model in Panel A was fitted using a linear function using OLS. Panel B shows the fit of a Pareto distribution using MLE.

The model allows us to dissect the population and study how its social relations shape labor dynamics. Panel A from Figure 27 shows that the distribution of connections in the social network is skewed. Although it looks well approximated by a normal distribution, it does not pass a Chi-2 test, suggesting that the social network is not the result of a purely random process. The social network is assortative in degree, meaning that agents tend to be connected to other agents with similar number of connections. This conforms to empirical studies about social networks.

Panel C shows that agents tent to socialize with similarly skilled agents. This correlation is not perfect because of the mechanism described in Section 4.3.1.3. Therefore, despite of the noise introduced by the matching process and heterogeneous skill endowments, the system emerges a homophilic social network. Interestingly, the opposite result occurs for preferences. Panel D shows that agents with stronger preferences for working tend to acquaint with pro-social agents. This is explained from the uniform distribution of β: the chances of meeting an agent with a higher β than myself decrease as my own β increases. However, the plot in Panel D is not an exact map to the cumulative distribution of β due to the matching restrictions induced by the localized interactions at the workplace.

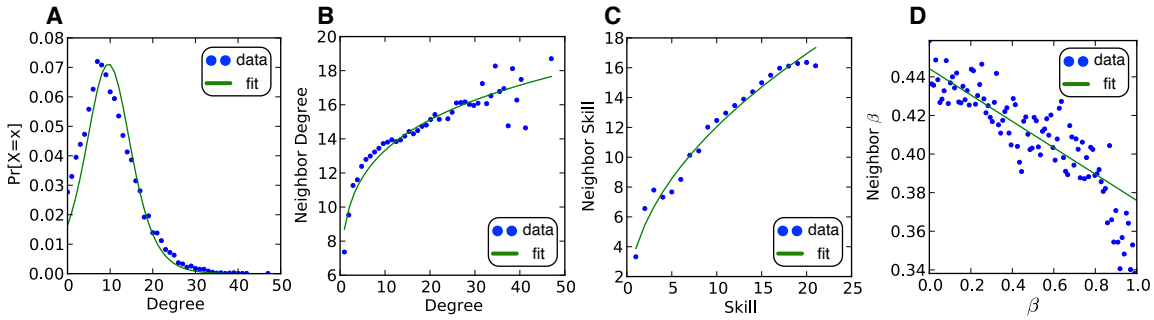

Figure 27: Assortativity through social ties

Panel A shows a fitted normal distribution via MLE. Panels B and C show fitted concave functions using OLS. Panel D presents a linear model fitted via OLS.

The properties of the social network are of particular interest to sociologists who study labor markets. Here we study how this network relates to labor outcomes, which is of interest to labor economists. In Panel A of Figure 28 we can see that agents with more social ties tend to earn more until a critical level of 10 connections is reached. After that, more links are negatively correlated to lower income. Beyond 30 connections, the relationship between degree and wages becomes ambiguous. This result is not very intuitive, since the conventional view of social capital is that more connections mean more access to better opportunities. This is clarified if we employ the clustering coefficient as our proxy of social capital. Remember that a lower clustering coefficient implies poorer connectivity between an agent's neighbors. Panel B shows that agents with higher clustering coefficients tend to earn lower wages. The reason is because second-degree neighbors can apply to the same companies. These firms are more likely to hire more of these individuals during the same period. Consequently, the offer received by an agent with high clustering coefficient will be lower when her second-degree

neighbors have been hired before her. Therefore, earnings depend strongly not on how many people an agent knows but on whom she knows.

The social network also has important consequences on employment status. Higher degree means lower chances of being unemployed (Panel C in Figure 28), while higher clustering coefficients generate higher unemployment rates. The causal mechanisms are similar as the ones described before. However, in the case of employment status both how many and whom you know are important.

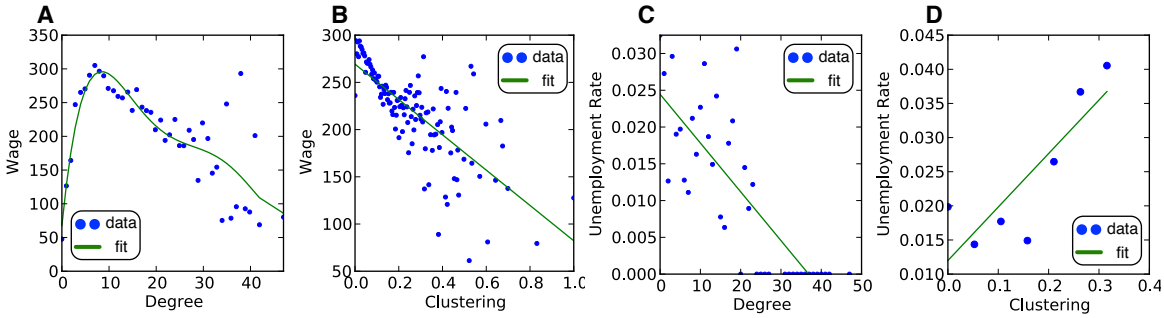

Figure 28: Social networks and labor outcomes
Panel A shows a polynomial of order five fitted by OLS. Panels B, C, and D present linear models fitted with OLS.

Finally, we explore another aspect about the social network and employment. We are interested in looking at the social capital that certain employers facilitate. Panel A in Figure 29 shows two regimes of firms that employ individuals with different degree levels. Firms with higher elasticity of tenured skills tend to employ individuals with more social ties on average. In Panel B we can observe that workers at larger firms have more

connections. This relationship is particularly steep for firms with less than 20 employees. After 20 employees, the marginal scaling effect of firm size in social capital weakens. This is caused by the socializing dynamics because a worker can only build social ties with people similar (in some degree) to them.

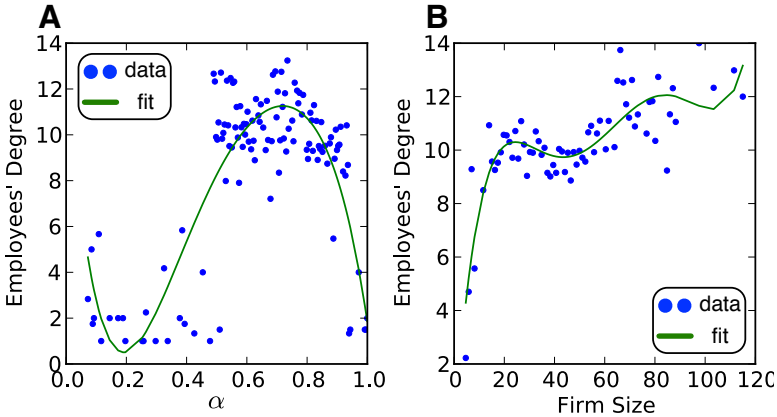

Figure 29: Social capital in firms
Panels A and B show polynomials of orders three and five respectively, fitted via OLS.

4.4. Small-World Labor Flow Networks and Job Search

Labor flows between companies generate endogenous changes to the way people find jobs. When individuals change jobs, they bring their personal contacts from previous jobs with them. This makes their new co-workers closer to those firms where their new colleague used to work. Structurally speaking, this is an endogenous process that continuously reshapes the search space of job opportunities. In some degree, the LFNs

captures this space. If workers search for jobs through the LFN, the firms that they can approach are determined by the structure of their connections.

Besides the empirical regularities documented in Section 3, we found two characteristics of LFNs that coincide with many other types of social networks: local structure and global connectivity. Local structure refers to the "tightness" of groups of firms. For a network, this is measured by the average clustering coefficient of its nodes. If an economy has high clustering coefficient, it has strong local structure. Intuitively, local structure means that labor tends to move within a well-defined group of firms. Global connectivity refers to how distant two firms are from each other in the LFN. Normally, this is measured as the average shortest path between all pairs of firms. Shorter paths mean that there is higher global connectivity. Intuitively, it means that a worker can get to any firm in the economy by being employed by a few other firms previously.

For our empirical LFNs, we found that they have strong local structure and high global connectivity. The average clustering coefficients for Finland and Mexico are 0.042 and 0.075 respectively. Their average path lengths are 4.78 and 5.06 respectively. If these LFNs were created by a purely random process (as in Érdös-Renyï networks), they would have an average clustering coefficient of 0.009 and 0.01 respectively. Their average path length would be of 4.48 and 4.6 for Finland and Mexico respectively (these quantities were estimated via null models). Strong local structure and high global connectivity has been known for decades by sociologists as the small-world phenomenon (Milgram, 1967; Travers & Milgram, 1969).

On one hand, economic theory has informed us about the incentive-driven mechanisms that produce labor mobility. However, these theories have ignored the underlying structure of economies through which people search for jobs. In fact the popular matching models of job search (Gronau, 1971; Lippman & McCall, 1976; McCall, 1970; Mortensen, 1970) assume that any worker can be randomly matched to any firm (an underlying complete graph). On the other, physics models about small-world networks (Kleinberg, 2000; Newman & Watts, 1999; Watts & Strogatz, 1998) have improved our understanding about their "mechanics", without deeper insight of the social principles that generate them.

In this Section, we present an agent-based model that connects physics models of small-world networks and economic job search models. It provides a framework to understand how the process of job search gives rise to the small-world phenomenon of LFNs, which in turn affects the search space of the agents. We provide evidence that behavioral aspects such as impatience and discounting behavior; policy dimensions such as unemployment benefits and search costs; and macroeconomic variables such as separation rates induce different outcomes of the topology of the LFN through its local structure and global connectivity.

4.4.1. Model Description

There is a fixed population of firms and workers. Workers can be either employed by one of the firms or unemployed and searching for a new job. There is a labor flow network of all firms, reflecting the most recent flows of workers between every two

firms. Job search takes place through this network. Therefore, the network is the product of the decisions of agents of searching for jobs, but it also affects the way in which they search.

4.4.1.1. Job Search

Job search behavior is modeled after the standard framework in the job search literature (Gronau, 1971; Lippman & McCall, 1976; McCall, 1970; Mortensen, 1970). More specifically, we use the framework developed by (DellaVigna & Paserman, 2005; Paserman, 2008), which considers exponential and hyperbolic discounting, search costs, layoffs, and unemployment benefits. In period t, unemployed agents have to face the decision of increasing the probability s_t of receiving a job offer in t+1 while incurring in a cost $C(s_t)$. At the end of t, they receive unemployment benefits b. On t+1, unemployed agents receive an offer with probability s_t, which they have to accept or decline. Offers are randomly drawn from a known distribution with cumulative form $F(\bullet)$. If an agent accepts an offer, she becomes employed with wage w and can become unemployed with probability q at any period after t+1. If the offer is declined, the agent remains unemployed for the next period, expecting to get a better offer. For infinitely long living agents, this problem is a life utility maximization one, with value function in the form:

$$\max_{s_t \in [0,1]} b_t - C(s_t) + \beta\delta\left\{ s_t \mathrm{E}\left[\max\left(V_{t+1}^{E}(w), V_{t+1}^{U}\right)\right] + \left(1 - s_t\right) V_{t+1}^{U} \right\},$$

where E is the expectations operator, V^U is the value function of being unemployed, and $V^E = w + \delta[qV^U + (1-q)V^E(w)]$ is the value function of being employed. V^U is assumed to be the same across time for analytical tractability[2]. Parameters β and δ are the subjective discount factors for short and long run respectively.

Under this specification, this problem has a solution in which agents set an optimal stopping rule by defining a reservation wage w^* and an optimal probability of receiving a job offer s^*. Both solutions are stationary and can be obtained from solving the following system of equations:

Equation 7

$$w^* = b - C(s^*) + \frac{\delta s^*}{1-\delta(1-q)}\left[\int_{w^*}^{\bar{w}} (u - w^*)\, dF(u)\right]$$

$$C'(s^*) = \frac{\delta\beta}{1-\delta(1-q)}\left[\int_{w^*}^{\bar{w}} (u - w^*)\, dF(u)\right]$$

Solving this system provides constant w^* and s^*, which are the connecting pieces between agent behavior and LFNs. Notice that the previous model is very comprehensive

[2] Paserman (2008) argues that this assumption only introduces a small bias and would be equivalent to say that agents are eligible to unemployment insurance for just their first unemployment spell. This assumption, together with strictly convex costs, allows the author to derive a stationary reservation wage.

since earlier models of job search can be derived with the right calibration. For example, if C=0 and β =1, we have the classic job search model (Lippman & McCall, 1976).

4.4.1.2. Network

The economy is structured as a LFN. We begin with a graph G where firms are nodes N={1,2,…,n} and a recent flow of labor between firms f and h constitutes an edge $e_{f,h}$=1. An edge does not exist if $e_{f,h}$=0, and it means that there have not been recent labor flows. G is a directed graph, although we will use it as an undirected graph to analyze some dynamics ahead.

Initially, all firms have 2k in-degree and 2k out-degree, and we assume $k\geq 2$ (a realistic degree obtained in our empirical findings in Table 3). In this setup, all the edges f are connected to the 2k neighbors, implying that $e_{h,f}$=1 and $e_{f,h}$=1. From its 2k neighbors, f is connected to two of them by *structural edges*. These edges constitute a circular lattice that connects all firms in the economy[3]. As in (Tassier & Menczer, 2001), we can justify this lattice as the geographical proximity of firms, which induces frequent labor flows. Hence, structural edges make every firm accessible from any other firm at any point in time.

The remaining 2k-2 neighbors are connected through *labor flow edges*. A labor flow edge can be rewired while a structural edge cannot. Firms can only rewire their

[3] This assures that the network remains connected at all times. Notice that the Watts & Strogatz model assumes $n>>k>>\ln(n)>>1$ in order to achieve permanent connectivity. Our empirical findings do not support $k>>\ln(n)$.

outgoing labor flow edges, since it indicates the most recent flows from them to the others. As in the construction of LFNs described in Section 3.2.5, we assume that a labor flow includes the possibility of unemployment spells between two jobs. Therefore, labor flow edges do not represent job-to-job changes exclusively. Figure 30 illustrates the network topology for k=2. For simplicity, we assume that edges are unweighted.

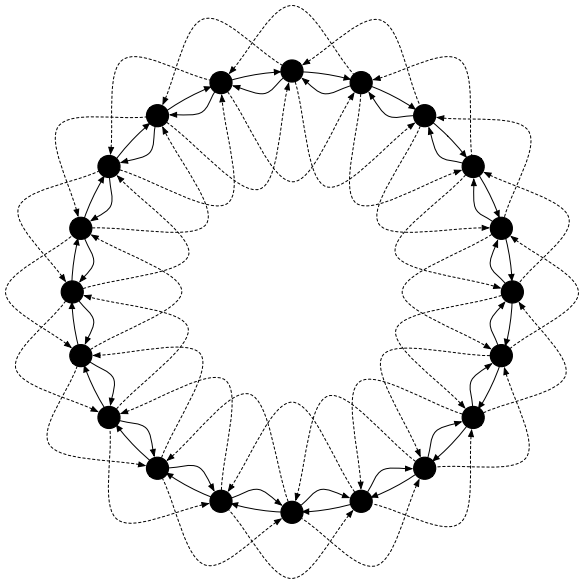

Figure 30: Initial topology of the LFN

The solid lines represent the *structural edges* that cannot be rewired. The dotted lines represent the *labor flow edges* that can be rewired to other firms. This figure illustrate the case of n=20 and k=2. However, it can be generalized for higher n and k.

4.4.1.1. Dynamics

Job search models in the economics literature assume that offers are *i.i.d.* draws from F, but does not assume any underlying structure of the origin of such offers. This is

where we incorporate the idea of the LFN as part of the job search process. As discussed in Section 3.2.5, LFNs implicitly capture the various motivations for agents to move between firms because they reflect the actual movements of workers. In a similar way, the LFN provides a reference of the space that restricts the job search of agents. Intuitively, an agent that left firm *f* starts her search in the firms that are closer to *f* in the LFN. Therefore, agents search for job opportunities in the order in which firms are accessible through the LFN. In other words, when an agent is laid off, she makes a list of all the firms, ordered by proximity in the LFN. Every time she receives an offer, this offer comes from the next firm in her list.

Let $d(f,h)$ be the shortest distance between firms *f* and *h* in number of edges. Agents search progressively through firms that are closer to their last employer. If there are multiple firms *h* with the same $d(f,h)$, they are reached in random order. When an agent that was employed at firm *f* accepts an offer from *h* such that $d(f,h)>1$, a random labor flow edge from *f* is rewired[4] to connect with *h*, indicating that the most recent labor flow from *f* was to *h*. This process generates network dynamics similar to those in (Kleinberg, 2000; Watts & Strogatz, 1998) but at different rewiring rates.

[4] Although in Section 3.2.5 we construct LFNs by adding edges to the network, here we prefer a rewiring approach because it allow us to have a better control over the network density in order to measure the effects job search behavior on the network topology.

4.4.2. Implementation and Execution

The model was implemented in the Python programming language version 2.7 (code provided in additional materials). It consists of two components:

- *Worker*: An object class of an agent that represents a worker. It contains the algorithms through which a worker searches for jobs in the social network.

- *Model*: Contains the populations of workers. Firms are simple nodes of the graph that represents the LFN. It executes the model.

The model execution is partially parallel. Groups of agents are run in separate threads[5]. Whenever an agent generates a labor flow, it locks the LFN object to produce the rewiring of her last employer. Agents' threads sleep at random intervals, providing the execution with asynchronous activations. Within a single thread, the order of execution is randomized. When awaken, an agent is activated by different rules, depending its state. If employed, an agent is laid off with probability q, which acts as the activation rate of this type of agents. Unemployed agents are activated by the endogenous probability s^*, which provides them with job offers. We summarize this process in Pseudocode 3.

[5] The model can be fully parallelized if each worker is assigned to a single thread. However, this imposes a restriction to the size of the system since Python only allows a to execute a couple hundred threads. Parallel processes are another option, however it involves complications when sharing the network across all processes.

Pseudocode 3: Job search and network rewiring process

```
INITIALIZE agents and network;
WHILE t<T DO:
    FOR each worker in parallel:
        IF worker is employed:
            Separate from firm f with probability q;
            IF separated:
                List all firms h in ascending order by d(f,h);
                Add b to wealth;
                Substract C(s*) from wealth;
            ELSE:
                Add w to wealth;
        ELSE:
            Receive offer with probability s*:
            IF offer is received:
                IF offer ≥ w*:
                    Accept offer;
                    IF e_{f,h}=0:
                        Pick a non-structural outgoing edge e_{f,g};
                        Set e_{f,g}=0;
                        Set e_{f,h}=1;
                ELSE:
                    Decline offer;
                    Remove h from list;
                    Increment unemployment spell;
            ELSE:
                Increment unemployment spell;
            Add b to wealth;
            Substract C(s*) from wealth;
    Increment t;
```

4.4.3. Calibration

We calibrated the model using US empirical data. Some parameters were calibrated with a single value and others as random variables from the estimated empirical distributions.

4.4.3.1. Number of Workers and Layoff Rate

Ideally, we would like to calibrate the number of agents in terms of the average firm size. However, the number of firms has an important topological effect in the LFN. For smaller networks, it is easier to achieve higher global connectivity since firms are not so far away from each other. For large enough networks, the number of agents that would need to be activated is computationally unfeasible due to the locking process when there is a rewiring. Therefore, we will assume that each firm has a small number of workers. This allows us to instantiate a large number of firms to test the topological effects with more precision.

Paserman (2008) calibrated his search model using NLSY data. He estimated the separation rate q for three different income groups. We calibrate q to be average across income groups: q=0.0101.

4.4.3.2. Network Density

The number of firms and the number of edges determine the density of the network. As explained in Section 4.4.1.2, the number of edges is fixed. Constant edges allow for a better control when measuring the effect of economic parameters on the LFN topology. The number of outgoing edges ($2k$) is not far from the empirical findings reported in Table 3 and is close to the benchmark case of well-known models of small-world networks dynamics.

4.4.3.3. Discount Factors

Discount factors of agents are subjective parameters that are normally estimated form data. Econometric and experimental studies of similar settings to a job search show similar results for the values of β and δ (Benhabib, Bisin, & Schotter, 2010; Benzion, Rapoport, & Yagil, 1989; DellaVigna & Paserman, 2005; Paserman, 2008). The values for δ range in [0.80, 0.9989], while the ones for β are in [0.4021, 0.8937]. Therefore, we calibrate these parameters to be uniformly distributed in their respective ranges.

4.4.3.4. Offers, Costs and Benefits

We assume that wage offers are uniformly distributed in [0,1]. This is an unconventional distribution for offers but it facilitates our calibration due to the need of normalizing w^* and s^*. Since we are interested in the effect of parameters on LFNs, the nature of the distribution does not alter the direction of our results.

Paserman (2008) specified the functional form $C(s)=cs^{1+\eta}$ for the cost of search. He estimated η=0.4 and c=82.43 USD using NLSY data. He reports an average wage previous to unemployment spells of $320.00, and average unemployment benefits of $134.00. We use these quantities to normalize the search cost and the unemployment benefits. Therefore, we obtain c=0.256 and b=0.419. Notice that we have calibrated each period to be representative of one week.

4.4.3.5. Endogenous Parameters

Finally, the endogenous parameters w^* and s^* are obtained for each agent by solving the non-linear system in Equation 7. Due to the convexity assumption of the costs function, we cannot obtain an explicit solution for w^* and s^*. Therefore, they are approximated numerically. Table 6 provides a description of the model's parameters. The values that were experimentally calibrated correspond to the baseline case. They were also tested for other values in order to check the robustness of the results.

Table 8: Parameters of job search and LFNs model

Parameter	Value	Calibration	Change	Homogeneity
Number of workers	3,000	Experimental	Static	N/A
Average firm size	3	Experimental	Static	N/A
Number of groups	50	Experimental	Static	N/A
Number of periods, T	10,000	Experimental	Static	N/A
Out-degree, $2k$	4	Experimental	Static	homogenous
Wage offers, F	*Uniform* on [0, 1]	Experimental	Static	homogenous
Separation rate, q	0.0171	Empirical	Static	homogenous
Marginal search cost, c	0.256	Empirical	Static	homogenous
Search cost convexity, η	0.4	Empirical	Static	homogenous
Unemployment benefits, b	0.419	Empirical	Static	homogenous
Short run impatience, β	*Uniform* on [0.40, 0.8937]	Empirical	Static	heterogeneous
Long run impatience, δ	*Uniform* on [0.80, .09989]	Empirical	Static	heterogeneous

This table describes the nature of each parameter in terms of different characteristics expressed by the columns. The column *Value* denotes the values assigned to the parameters during the baseline simulations, or the probabilistic distributions from where the values are drawn. The *Calibration* column indicates if the source of the values is from empirical data or product of experimental procedures that are conventional in agent-based modeling. The *Change* column shows if the parameter is constant throughout a simulation or if it changes with each period. Finally, the *Homogeneity* column indicates if the value is the same for all agents or not.

4.4.4. Experimental Design and Output

We ran a sample of 36 simulations. The simulations were run independently from each other, with initial seeds assigned by a unique time stamp. We employed the computational resources from the Department of Computational Social Science at George Mason University. More specifically, the 36 simulations were distributed across four cores in a PC with 16 cores and 256 GB of memory.

We saved records for the last 500 steps. Each record was collected on a monthly basis (every four steps). The records were stored as administrative micro-data. Each line of a record contains information about the state of an agent at a specific step of the simulation. Each row represents a variable of the agent. Records were organized into two tables: workers and LFN. Each line in the LFN table represents an edge. They consist of the source node, target node, and period of collection. All tables were stored in a hierarchical data file, with format HDF5, and compressed with the library LZO. More specifically, we used the data management library PyTables to speed up the data recording process. It took approximately two week to perform all the simulations.

4.4.5. Results

First we look into some of the labor dynamics generated by the model. Panel A from Figure 31 shows that the degree of the LFN follows a skewed normal distribution. Panel B presents the distribution of the duration of unemployment spells, which is lognormal. The average unemployment spell lasts for 21 weeks, which is close to the previous model. However, unemployment rates in this model are significantly higher.

Panel C shows a typical time series of the unemployment rates. On average, the unemployment rate is 18%, with a standard deviation of nearly half a percent point. Panel D shows a typical time series of the average duration of unemployment spells.

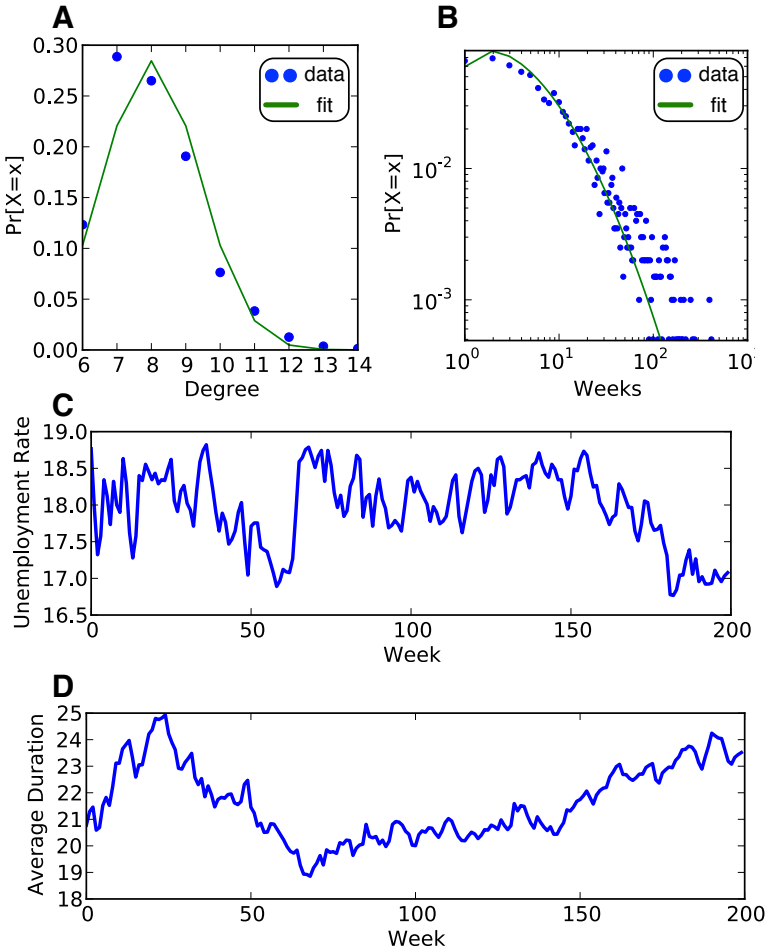

Figure 31: Labor dynamics from job search

Panels A and B show normal and lognormal distributions respectively, fitted via OLS. Panels C and D present time series for four years on a monthly basis.

The most important results connect the economic behavior of job search with the topology of the LFN. For a period of four years, we took measures of the LFN that provide evidence about its small-world feature: average shortest path length and clustering coefficient. Panel A in Figure 32 shows that the average path length is considerably shorter for the job search model than the initial lattice. The former is nearly 500 edges, while the latter is around 5.5. We generated a null model for the initial lattice, where we randomized all the labor flow edges. In this case, the average path length was approximately 4.7 edges. In Panel B we show that the clustering coefficients of the initial lattice and the model are of the same order of magnitude (0.5 and 0.13 respectively). In contrast, the randomized model has an average clustering coefficient two orders of magnitude lower: 0.0015.

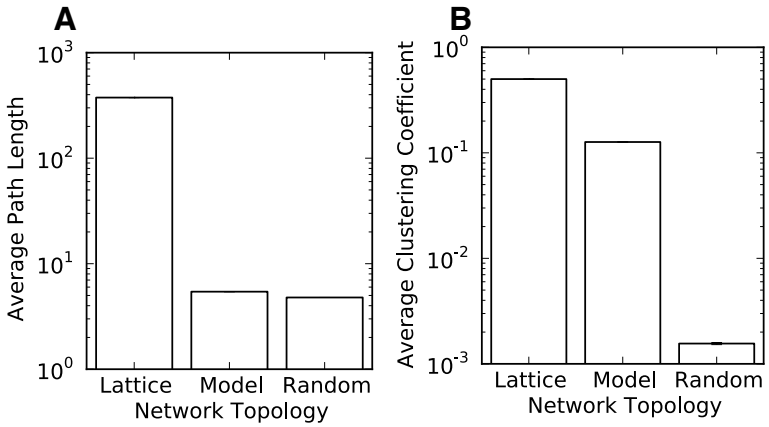

Figure 32: Small-world topology from job search

All Panels show values of the corresponding network metrics, averaged from the last four years of a simulation. All the time series are very stable, as shown by the insignificant standard error term, only visible in some bars. Both panels correspond to the comparison between the initial lattice, the baseline model, and the null model.

We explore the connection between the model parameters and the topology of the network. We ran treatment simulations in which we removed the unemployment benefits. We compare the baseline and treatment groups in Panels A and B from Figure 33 for average path length and average clustering coefficient respectively. Although both cases seem similar in magnitude, the one-sided t-tests show that they are statistically different. When unemployment benefits are zero, agents are willing to accept lower offers. Then, firms that are closer to their last employer hire them, inducing a shorter rewiring of the LFM. This generates longer paths and higher clustering. It is clear from the same figure that the effect goes in the same direction when the separation rete and the marginal search cost are higher.

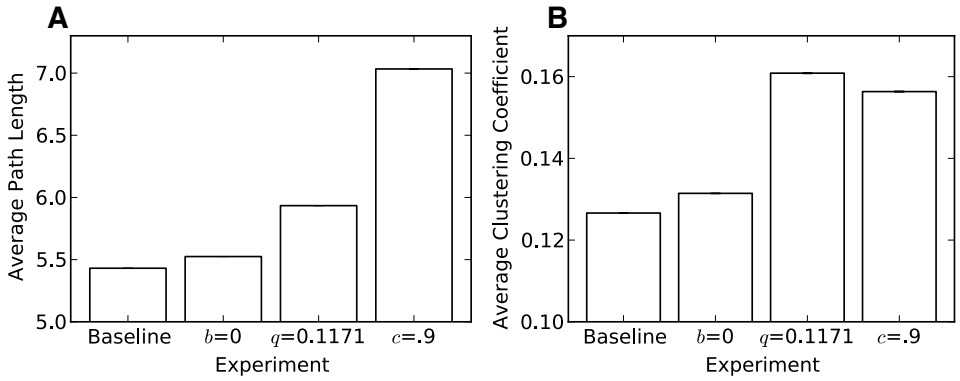

Figure 33: Network effect of microeconomic parameters

All Panels show values of the corresponding network metrics, averaged from the last four years of a simulation. All the time series are very stable, as shown by the insignificant standard error term, only visible in some bars. Both panels correspond to the comparison between the baseline model and the three treatments: null unemployment benefits, higher separation rates, and higher marginal search cost. Single-tailed *t*-tests reject the hypothesis that any of the treatments is equal to the baseline.

4.5. Conclusions

In this chapter we showed how ABMs are useful to connect labor and firm dynamics in order to deepen our understanding of labor markets. We presented three different models that are simpler than the average labor market ABM, and generate richer results than most models about labor markets. ABMs are a natural tool to incorporate networks and economic interactions. Therefore, we believe that it will be the way to go when it comes to construct comprehensive models for designing and evaluating labor policies. Further developments in this kind of models will increase our knowledge and expertise about modeling complex dynamics. Researchers and policymakers would be greatly benefited from engaging in an agenda that pursues this goal.

5. CONCLUSIONS

This dissertation studied the role of networks as mediators between labor and firm dynamics through an extensive empirical analysis of large-scale employer-employee matched micro-datasets and state of the art agent-based models. We provided an overview of the literature of models of labor and firm dynamics, and a detailed account of agent-based models that have been developed to study labor markets. We introduced the concept of labor flow networks as a new way to understand labor and firm dynamics from a decentralized and distributed perspective. We showed that conceptualizing economies as networked structures provide important information about employment growth and the way in which labor is continuously reallocated throughout the economy. We employed agent-based models in order to study the endogenous formation of different types of networks and their effect on labor and firm dynamics. We found that this type of models is naturally suited to couple complex dynamics, decentralized processes, networks, and rich micro-data.

This dissertation contributes to advance the understanding of labor dynamics in four main ways: (*i*) by providing the first characterization of a network of firms (connected through labor flows, i.e. *labor flow networks*) for an entire economy; (*ii*) by finding evidence of the relationship between labor flow networks and employment dynamics; (*iii*) by developing agent-based models that generate rich stylized facts about

labor, firm, and social dynamics from microeconomic behavior; and (*iv*) by providing the microeconomic foundations of the formation process of labor flow networks through the coupling of job search models with models about the formation of complex networks. Although they are quite general, these contributions embed other contributions in terms of methods and results that are important for the understanding of labor market dynamics. We argue that studying labor markets as networks through the lens of computational models offers an unparalleled opportunity to understand the complex ways in which labor dynamics affect multiple aspects of the economy. By employing state of the art computational tools and the everyday growing large-scale micro-data, social scientists and policy makers can view labor markets from a fundamentally different perspective: one where heterogeneous agents interact explicitly in decentralized forms, endogenously generating structures that limit their mobility and affect policy-relevant indicators such as unemployment rates, duration, and wages. We believe that the scientific and pragmatic sides of studying labor market can be significantly enriched with this approach. Governments, scientists, and regulators would be greatly benefited from performing this type of analysis and developing these kind models.

APPENDIX

Authors	Title	Published	Countries	Datasets
Almus and Nerlinger	Testing "Gibrat's Law" for Young Firms - Empirical Results for West Germany	2000	West Germany	ZEW-foundation Panel
Amaral et al.	Scaling behavior in economics I. Empirical results for company growth	1997	USA	Compustat
Angeline and Generale	On the Evolution of Firm Size Distributions	2008	Italy	Mediocredito
Ibid.	Ibid.	Ibid.	80 countries	WBES
Angeline and Generale	Firm size distribution: do financial constraints explain it all? Evidence from survey data	2005	Italy	Mediocredito
Ibid.	Ibid.	Ibid.	80 countries	WBES
Axtell	Zipf Distribution of U.S. Firm Sizes	2001	USA	Census, Compustat
Bottazzi	Firm Diversification and the Law of Proportionate Effect	2001	Italy	MICRO.1 databank developed by the Italian Statistical Office (ISTAT)
Bottazzi and Secchi	Growth and Diversification Patterns of the Worldwide Pharmaceutical Industry	2005	USA, UK, France, Germany, Spain, Italy, Canada	Pharmaceutical Industry Database (PHID)
Bottazzi et al.	Corporate growth and industrial structures: some evidence from the Italian manufacturing industry	2002	Italy	MICRO.1 databank developed by the Italian Statistical Office (ISTAT)
Bottazzi et al.	Invariances and Diversities in the Evolution of Manufacturing Industries	2003	Italy	MICRO.1 databank developed by the Italian Statisitical Office (ISTAT)
Cabral and Mata	On the Evolution of the Firm Size Distribution: Facts and Theory	2003	Portugal	IF4
Ibid.	Ibid.	Ibid.	Portugal	Survey conducted by the Portuguese Ministry of Employment– Quadros do Pessoal (QP)
Chen et al.	Firm Size Distribution in Fortune Global 500	2008	33 countries	Fortune Global 500
Chen et al.	The Zipf's Law in the Revenue of Top 500 Chinese Companies	2009	China	China Enterprise Confederation
Cirillo	An analysis of the size distribution of Italian firms by age	2009	Italy	CEBI, first developed by the Bank of Italy and now maintained by Centrale dei Bilanci Srl
Cirillo and Husler	On the upper tail of Italian firms' size distribution	2008	Italy	CEBI, first developed by the Bank of Italy and now maintained by Centrale dei Bilanci Srl
Corbellini et al.	Fitting Pareto II Distributions on Firm Size: Statistical Methodology and Economic Puzzles	2010	Italy	AIDA dataset, processed and managed by Bureau van Dijck Electronic Publishing
Corbellini et al.	Statistical regularity of firm size distribution: the Pareto IV and truncated Yule for Italian SCI manufacturing	2007	Italy	Micro1 data set selected by ISTAT (Italian National Institute of Statistics)
Crosato et al.	Technology and Firm Size Distribution: Evidence from Italian Manufacturing	2007	Italy	Micro1 survey, assembled by ISTAT
De Fabritiis et al.	On size and growth of business firms	2003	EU, North America	Pharmaceutical Industry Database (PHID)
Fujiwara et al.	Do Pareto-Zipf and Gibrat laws hold true? An analysis with European firms	2004	45 European countries	Bureau van Dijk's AMADEUS
Gaffeo et al.	On the size distribution of firms: additional evidence from the G7 countries	2003	Canada, France, Germany, Italy, Japan, UK, USA	Datastream International (DI) data-set
Gaffeo et al.	On the mean/variance relationship of the firm size distribution: evidence and some theory	2008	38 European countries	Bureau Van Dick's AMADEUS
Ganugi et al.	Firm size distributions and stochastic growth models: a comparison between ICT and Mechanical Italian Companies	2003	Italy	SCI, the Official Survey of ISTAT
Gupta et al.	Gradually truncated log-normal in USA publicly traded firm size distribution	2006	USA	Compustat

Authors	Title	Published	Countries	Datasets
Ibid.	Ibid.	Ibid.	USA	Compustat
Ibid.	Ibid.	Ibid.	Brazil	Fundacao Instituto de Pesquisas Conta beis da Universidade de Sao Paulo (USP)
Ibid.	Ibid.	Ibid.	Brazil	Fundacao Instituto de Pesquisas Conta beis da Universidade de Sao Paulo (USP)
Hart and Oulton	Growth and size of firms	1996	UK	OneSource Information Services Ltd.
Hart and Oulton	Zipf and the size distribution of firms	1997	UK	OneSource Information Services Ltd.
Hernandez	An analogy of the size distribution of business firms with Bose–Einstein statistics	2010	Argentina, Bolivia, Brazil, Chile, Colombia, Ecuador, Mexico, Mozambique, Peru, Venezuela, South Africa, USA, World (Fortune)	Magazines and fortune
Hernandez et al.	Company size distribution for developing countries	2006	Argentina, Bolivia, Brazil, Chile, Colombia, Ecuador, Mexico, Mozambique, Peru, Venezuela	Magazines
Laitinen	Stochastic Growth Processes in Large Finnish Companies: Test of Gibrat's Law of Proportionate Effect	1999	Finland	ETLA data base
Marsili	Technology and the Size Distribution of Firms: Evidence from Dutch Manufacturing	2005	Netherlands	Central Bureau of Statistics Netherlands (CBS)
Ibid.	Ibid.	Ibid.	Netherlands	Community Innovation Survey (CIS-2)
Marsili	Stability and Turbulence in the Size Distribution of Firms: Evidence from Dutch Manufacturing	2006	Netherlands	Production Statistics survey, carried out annually by the Central Bureau of Statistics Netherlands
Okuyama et al.	Zipf's law in income distribution of companies	1999	Japan	Diamond Inc.
Ibid.	Ibid.	Ibid.	Japan, USA, Italy	Moody's Investors Service Inc.
Ramsden and Kiss-Haypal	Company size distribution in di erent countries	2000	Austria, Belgium, Brazil, Canada, China, Denmark, France, Germany, GB, Hungary, Israel, Italy, Japan, Netherlands, Norway, South Africa, Switzerland, Spain, Sweden, USA	Multiple
Rossi-Hansberg and Wright	Establishment Size Dynamics in the Aggregate Economy	2007	USA	Census from Statistics of US Businesses (SUSB)
Ibid.	Ibid.	Ibid.	USA	US Census Bureau's Business Information Tracking Series (BITS)
Saito et al.	Do Larger Firms Have More Interfirm Relationships?	2007	Japan	Tokio Shoko Research Ltd.
Stanley et al.	Zipf plots and the size distribution of firms	1995	USA	Compustat
Zhang et al.	Zipf distribution in top Chinese firms and an economic explanation	2009	China	A Report on the Development of China's Enterprises provided by China Enterprise Confederation
Kaizoji et al.	Re-examination of the size distribution of firms	2005	USA	Bloomberg Ltd.
Ibid.	Ibid.	Ibid.	Japan	Bloomberg Ltd.

REFERENCES

Acemoglu, D., Carvalho, V., Ozdaglar, A., & Tahbaz-Salehi, A. (2012). The Network Origins of Aggregate Fluctuations. *Econometrica*, *80*(5), 1977–2016. doi:10.3982/ECTA9623

Acs, Z. (2011). High-Impact Firms: Gazelles Revisited. In M. Fritsch (Ed.), *Handbook of Research on Entrepreneurship and Regional Development: National and Regional Perspectives* (pp. 133–174). Edward Elgar Publishing.

Albert, R., Jeong, H., & Barabási, A. (1999). Internet: Diameter of the World-Wide Web. *Nature*, *401*(6749), 130–131.

Alvarez-Hamelin, J., Dall'Asta, L., Barrat, A., & Vespignani, A. (2005). K-Core Decomposition : A Tool for the Visualization of Large Scale Networks. In *Proceedings* (Vol. abs/cs/050). Presented at the World Wide Web Internet And Web Information Systems.

Amaral, L., Buldyrev, S., Havlin, S., Leschhorn, H., Maass, P., Salinger, M., … Stanley, M. (1997). Scaling Behavior in Economics: I. Empirical Results for Company Growth. *Journal de Physique I*, *7*, 621–633. doi:10.1051/jp1:1997180

Ashton, T. (1926). The Growth of Textile Businesses in the Oldham District, 1884-1924. *Journal of the Royal Statistical Society*, *89*(3), 567–583. doi:10.2307/2341720

Atalay, E., Hortaçsu, A., Roberts, J., & Syverson, C. (2011). Network Structure of Production. *Proceedings of the National Academy of Sciences*, *108*(13), 5199–5202. doi:10.1073/pnas.1015564108

Axtell, R. (2001). Zipf Distribution of U.S. Firm Sizes. *Science*, *293*(5536), 1818–1820. doi:10.1126/science.1062081

Ayebo, A., & Kozubowski, T. (2003). An Asymmetric Generalization of Gaussian and Laplace Laws. *Journal of Probability and Statistical Science*, *1*(2), 187–210.

Bala, V., & Goyal, S. (2000). A Noncooperative Model of Network Formation. *Econometrica*, *68*(5), 1181–1229. doi:10.1111/1468-0262.00155

Ballot, G. (2002). Modeling the Labor Market as an Evolving Institution: Model ARTEMIS. *Journal of Economic Behavior & Organization*, *49*(1), 51–77. doi:10.1016/S0167-2681(02)00058-6

Ballot, G., & Abraham-Frois, G. (1980). *Marché du Travail et Dynamique de la Répartition des Revenus Salariaux*. Université de Paris X-Nanterre. U.F.R. de Sciences Economiques, Paris.

Barabási, A., & Albert, R. (1999). Emergence of Scaling in Random Networks. *Science*, *286*(5439), 509 –512. doi:10.1126/science.286.5439.509

Barro, R., & Lee, J. (1993). International Comparisons of Aducational Attainment. *Journal of Monetary Economics*, *32*(3), 363–394. doi:10.1016/0304-3932(93)90023-9

Barro, R., & Lee, J. (1996). International Measures of Schooling Years and Schooling Quality. *American Economic Review*, *86*(2), 218–223. doi:10.2307/2118126

Ben-Porath, Y. (1980). The F-Connection: Families, Friends, and Firms and the Organization of Exchange. *Population and Development Review*, *6*(1), 1–30. doi:10.2307/1972655

Benhabib, J., Bisin, A., & Schotter, A. (2010). Present-Bias, Quasi-Hyperbolic Discounting, and Fixed Costs. *Games and Economic Behavior*, *69*(2), 205–223. doi:10.1016/j.geb.2009.11.003

Benzion, U., Rapoport, A., & Yagil, J. (1989). Discount Rates Inferred from Decisions: An Experimental Study. *Management Science*, *35*(3), 270–284. doi:10.1287/mnsc.35.3.270

Bergmann, B. (1990). Micro-to-Macro Simulation: A Primer With a Labor Market Example. *The Journal of Economic Perspectives*, *4*(1), 99–116. doi:10.2307/1942834

Birch, D. (1981). Who Creates Jobs? *The Public Interest*, *65*(8), 23.

Birch, D. (1987). *Job Creation in America: How Our Smallest Companies Put the Most People to Work*. Free Pr.

Birch, D., & Medoff, J. (1994). Gazelles. In L. Solomon & A. Levenson (Eds.), *Labor Markets, Employment Policy and Job Creation* (pp. 159–167).

Blondel, V., Guillaume, J., Lambiotte, R., & Lefebvre, E. (2008). Fast Unfolding of Communities in Large Networks. *Journal of Statistical Mechanics: Theory and Experiment*, *2008*(10), P10008. doi:10.1088/1742-5468/2008/10/P10008

Boorman, S. (1975). A Combinatiorial Optimization Model for Transmission of Job Information Through Contact Networks. *The Bell Journal of Economics*, *6*(1), 216–249. doi:10.2307/3003223

Bottazzi, G., Cefis, E., & Dosi, G. (2002). Corporate Growth and Industrial Structures: Some Evidence from the Italian Manufacturing Industry. *Industrial and Corporate Change*, *11*(4), 705–723. doi:10.1093/icc/11.4.705

Bottazzi, G., & Secchi, A. (2005). Growth and Diversification Patterns of the Worldwide Pharmaceutical Industry. *Review of Industrial Organization*, *26*(2), 195–216. doi:10.1007/s11151-004-7296-5

Bottazzi, G., & Secchi, A. (2006). Explaining the Distribution of Firm Growth Rates. *The RAND Journal of Economics*, *37*(2), 235–256. doi:10.2307/25046240

Bourdieu, P. (1983). The Forms of Capital. In J. Richardson (Ed.), *Handbook of theory and Research for the Sociology of Education* (pp. 241–258). Westport, CT: Greenwood Press.

Brown, C., & Medoff, J. (1989). The Employer Size-Wage Effect. *Journal of Political Economy*, *97*(5), 1027–1059. doi:10.2307/1831885

Buldyrev, S., Amaral, L., Havlin, S., Leschhorn, H., Maass, P., Salinger, M., … Stanley, M. (1997). Scaling Behavior in Economics: II. Modeling of Company Growth. *Journal de Physique I*, *7*(4), 635–650.

Burdett, K., Shi, S., & Wright, R. (2001). Pricing and Matching with Frictions. *Journal of Political Economy*, *109*(5), 1060–1085. doi:10.1086/322835

Burt, R. (1992). *Structural Holes: The Social Structure of Competition*. Cambridge, MA: Harvard University Press.

Cahuc, P., & Zylberberg, A. (2006). *The Natural Survival of Work: Job Creation and Job Destruction in a Growing Economy*. The MIT Press.

Caillou, P., & Sebag, M. (2008). Modelling a Centralized Academic Labour Market: Efficiency and Fairness. In *Proceedings*. Presented at the ECCS 2008.

Calvó-Armengol, A. (2004). Job Contact Networks. *Journal of Economic Theory*, *115*(1), 191–206. doi:10.1016/S0022-0531(03)00250-3

Calvó-Armengol, A., & Jackson, M. (2004). The Effects of Social Networks on Employment and Inequality. *American Economic Review*, *94*(3), 426–454. doi:10.1257/0002828041464542

Calvó-Armengol, A., & Jackson, M. (2007). Networks in Labor Markets: Wage and Employment Dynamics and Inequality. *Journal of Economic Theory*, *132*(1), 27–46. doi:10.1016/j.jet.2005.07.007

Calvó-Armengol, A., & Zenou, Y. (2005). Job Matching, Social Network and Word-of-Mouth Communication. *Journal of Urban Economics*, *57*(3), 500–522. doi:10.1016/j.jue.2005.01.001

Canning, D., Amaral, L., Lee, Y., Meyer, M., & Stanley, H. (1998). Scaling the Volatility of GDP Growth Rates. *Economics Letters*, *60*(3), 335–341. doi:10.1016/S0165-1765(98)00121-9

Champernowne, D. (1953). A Model of Income Distribution. *The Economic Journal*, *63*(250), 318–351. doi:10.2307/2227127

Coad, A. (2009). *The Growth of Firms: A Survey of Theories and Empirical Evidence*. Edward Elgar Publishing.

Coleman, J. (1988). Social Capital in the Creation of Human Capital. *American Journal of Sociology*, *94*, S95–S120. doi:10.2307/2780243

Corcoran, M., Datcher, L., & Duncan, G. (1980). Information and Influence Networks in Labor Markets. In G. Duncan & J. Morgan (Eds.), *Five Thousand American Families: Patterns of Economic Progress* (Vol. 8, pp. 1–38). Survey Research Center. The University of Michigan.

Costain, J., & Reiter, M. (2008). Business Cycles, Unemployment Insurance, and the Calibration of Matching Models. *Journal of Economic Dynamics and Control*, *32*(4), 1120–1155. doi:10.1016/j.jedc.2007.04.008

Davis, S., Faberman, J., & Haltiwanger, J. (2006). The Flow Approach to Labor Markets: New Data Sources and Micro-Macro Links. *Journal of Economic Perspectives*, *20*(3), 3–26. doi:10.1257/jep.20.3.3

Davis, S. J., Haltiwanger, J., & Schuh, S. (1998). *Job Creation and Destruction*. MIT Press.

Dawid, H., & Gemkow, S. (2013). How Do Social Networks Contribute to Wage Inequality? Insights from an Agent-Based Analysis. *Working Paper*. Retrieved from http://www.wiwi.uni-bielefeld.de/fileadmin/vpl1/LaborNetworks_jan2013.pdf

Dawid, H., Gemkow, S., Harting, P., Kabus, K., Neugart, M., & Wersching, K. (2008). Skills, Innovation, and Growth: An Agent-based Policy Analysis. *Journal of Economics and Statistics*, *228*(2+3), 251–275.

Dawid, H., Gemkow, S., Harting, P., & Neugart, M. (2009). On the Effects of Skill Upgrading in the Presence of Spatial Labor Market Frictions: An Agent-Based Analysis of Spatial Policy Design. *Journal of Artificial Societies and Social Simulation*, *12*(4), 5.

De Wit, G. (2005). Firm Size Distributions: An Overview of Steady-State Distributions Resulting from Firm Dynamics Models. *International Journal of Industrial Organization*, *23*(5–6), 423–450. doi:10.1016/j.ijindorg.2005.01.012

Deissenberg, C., van der Hoog, S., & Dawid, H. (2008). EURACE: A Massively Parallel Agent-Based Model of the European Economy. *Applied Mathematics and Computation*, *204*(2), 541–552. doi:10.1016/j.amc.2008.05.116

DellaVigna, S., & Paserman, M. (2005). Job Search and Impatience. *Journal of Labor Economics*, *23*(3), 527–588. doi:10.1086/430286

Diamond, P. (1981). Mobility Costs, Frictional Unemployment, and Efficiency. *Journal of Political Economy*, *89*(4), 798–812. doi:10.2307/1833035

Diamond, P. (1982). Aggregate Demand Management in Search Equilibrium. *Journal of Political Economy*, *90*(5), 881–894. doi:10.2307/1837124

Erickson, B. (1996). Culture, Class, and Connections. *American Journal of Sociology*, *102*(1), 217–251. doi:10.2307/2782191

European Communities, & OECD. (2007). EUROSTAT - OECD Manual on Business Demography Statistics.

Fabritiis, G., Pammolli, F., & Riccaboni, M. (2003). On Size and Growth of Business Firms. *Physica A: Statistical Mechanics and its Applications*, *324*(1-2), 38–44. doi:10.1016/S0378-4371(03)00043-8

Fagiolo, G., Dosi, G., & Gabriele, R. (2004). Matching, Bargaining, And Wage Setting In An Evolutionary Model Of Labor Market And Output Dynamics. *Advances in Complex Systems*, *07*(02), 157–186. doi:10.1142/S0219525904000135

Fagiolo, G., Napoletano, M., & Roventini, A. (2008). Are Output Growth-rate Distributions Fat-tailed? Some Evidence from OECD Countries. *Journal of Applied Econometrics*, *23*(5), 639–669. doi:10.1002/jae.1003

Fagiolo, G., Squartini, T., & Garlaschelli, D. (2013). Null Models of Economic Networks: the Case of the World Trade Web. *Journal of Economic Interaction and Coordination*, *8*(1), 75–107. doi:10.1007/s11403-012-0104-7

Fallick, B., & Fleischman, C. (2001). The Importance of Employer-to-Employer Flows in the U.S. Labor Market. *Board of Governors of the Federal Reserve System*, (18).

Faloutsos, M., Faloutsos, P., & Faloutsos, C. (1999). On Power-Law Relationships of the Internet Topology. In *Proceedings* (pp. 251–262). Presented at the Conference on Applications, Technologies, Architectures, and Protocols for Computer Communication, ACM. doi:10.1145/316188.316229

Finneran, L., & Kelly, M. (2003). Social Networks and Inequality. *Journal of Urban Economics*, *53*(2), 282–299. doi:10.1016/S0094-1190(02)00513-2

Flap, H., & De Graaf, N. (1986). Social Capital and Attained Occupational Status. *Netherlands Journal of Sociology*, *22*(2), 145–161.

Fortunato, S. (2010). Community Detection in Graphs. *Physics Reports*, *486*(3–5), 75–174. doi:10.1016/j.physrep.2009.11.002

Frazis, H., & Spletzer, J. (2005). Worker Training: What We've Learned from the NLS79. *Monthly Labor Review*.

Gabaix, X. (1999). Zipf's Law For Cities: An Explanation. *Quarterly Journal of Economics*, *114*(3), 739–767. doi:10.1162/003355399556133

Gabaix, X. (2011). The Granular Origins of Aggregate Fluctuations. *Econometrica*, *79*(3), 733–772. doi:10.3982/ECTA8769

Gabriele, R. (2002). Labor Market Dynamics and Institutions: an Evolutionary Approach. *Working Paper*.

Galeotti, A., & Goyal, S. (2010). The Law of the Few. *American Economic Review*, *100*(4), 1468–1492. doi:10.1257/aer.100.4.1468

Gee, L., & Jones, J. (2013). Social Networks and Labor Markets: How Strong Ties Relate to Job Transmission on Facebook's Social Network. *Working Paper (submitted)*.

Gemkow, S., & Neugart, M. (2011). Referral Hiring, Endogenous Social Networks, and Inequality: An Agent-based Analysis. *Journal of Evolutionary Economics*, *21*(4), 703–719. doi:10.1007/s00191-011-0219-3

Gianelle, C. (2011). Exploring the Complex Structure of Labour Mobility Networks. Evidence from Veneto Microdata. *Working Paper*.

Gibrat, R. (1931). *Les Inégalités Economiques*. Paris: Sirey.

Glaeser, E., Laibson, D., & Sacerdote, B. (2002). An Economic Approach to Social Capital. *The Economic Journal*, *112*(483), F437–F458. doi:10.1111/1468-0297.00078

Gode, D., & Sunder, S. (1993). Allocative Efficiency of Markets with Zero-Intelligence Traders: Market as a Partial Substitute for Individual Rationality. *Journal of Political Economy*, *101*(1), 119–137. doi:10.2307/2138676

Granovetter, M. (1973). The Strength of Weak Ties. *American Journal of Sociology*, *78*(6), 1360–1380. doi:10.2307/2776392

Gronau, R. (1971). Information and Frictional Unemployment. *American Economic Review*, *61*(3), 290–301. doi:10.2307/1813427

Guerrero, O., & Axtell, R. (2011). Using Agentization for Exploring Firm and Labor Dynamics. In S. Osinga, G. Hofstede, T. Verwaart, G. Fandel, & W. Trockel (Eds.), *Emergent Results of Artificial Economics* (Vol. 652, pp. 139–150). Springer.

Guiso, L., Sapienza, P., & Zingales, L. (2004). The Role of Social Capital in Financial Development. *American Economic Review*, *94*(3), 526–556. doi:10.1257/0002828041464498

Hall, B. (1987). The Relationship Between Firm Size and Firm Growth in the US Manufacturing Sector. *The Journal of Industrial Economics*, *35*(4), 583–606. doi:10.2307/2098589

Hall, R. (2005). Employment Fluctuations with Equilibrium Wage Stickiness. *The American Economic Review*, *95*(1), 50–65. doi:10.2307/4132670

Halloran, M., Longini, I., Nizam, A., & Yang, Y. (2002). Containing Bioterrorist Smallpox. *Science*, *298*(5597), 1428–1432.

Haltiwanger, J., Jarmin, R., & Miranda, J. (2010). Who Creates Jobs? Small vs. Large vs. Young. *NBER Working Paper*, (16300).

Hart, P., & Prais, S. (1956). The Analysis of Business Concentration: A Statistical Approach. *Journal of the Royal Statistical Society. Series A (General)*, *119*(2), 150–191. doi:10.2307/2342882

Heckman, J., Stixrud, J., & Urzua, S. (2006). The Effects of Cognitive and Noncognitive Abilities on Labor Market Outcomes and Social Behavior. *Journal of Labor Economics*, *24*(3), 411–482. doi:10.1086/504455

Hidalgo, C., & Hausmann, R. (2009). The Building Blocks of Economic Complexity. *Proceedings of the National Academy of Sciences*, *106*(26), 10570–10575. doi:10.1073/pnas.0900943106

Hobijn, B., & Şahin, A. (2009). Job-Finding and Separation Rates in the OECD. *Economics Letters*, *104*(3), 107–111. doi:10.1016/j.econlet.2009.04.013

Holtmann, A., & Idson, T. (1991). Employer Size and On-the-Job Training Decisions. *Southern Economic Journal*, *58*(2), 339–355. doi:10.2307/1060178

Holzer, H. (1988). Search Method Use by Unemployed Youth. *Journal of Labor Economics*, *6*(1), 1–20. doi:10.2307/2534865

Huberman, A., & Glance, N. (1993). Evolutionary Games and Computer Simulations. *Proceedings of the National Academy of Sciences*, *90*(16), 7716–7718. doi:10.1073/pnas.90.16.7716

Hymer, S., & Pashigian, P. (1962). Firm Size and Rate of Growth. *Journal of Political Economy*, *70*(6), 556–569. doi:10.2307/1828779

Ijiri, Y., & Simon, H. (1964). Business Firm Growth and Size. *The American Economic Review*, *54*(2), 77–89. doi:10.2307/1810899

Ijiri, Y., & Simon, H. (1977). *Skew Distributions and the Sizes of Business Firms* (Vol. 24). North Holland.

Ijiri, Yuji, & Simon, H. (1967). A Model of Business Firm Growth. *Econometrica*, *35*(2), 348–355. doi:10.2307/1909116

Ilmakunnas, P., & Maliranta, M. (2005). Worker Inflow, Outflow, and Churning. *Applied Economics*, *37*(10), 1115–1133. doi:10.1080/00036840500109621

Ioannides, Y., & Loury, L. (2004). Job Information Networks, Neighborhood Effects, and Inequality. *Journal of Economic Literature*, *42*(4), 1056–1093. doi:10.2307/3594916

Jackson, M. (2011). An Overview of Social Networks and Economic Applications. In J. Benhabib, A. Bisin, & J. Jackson (Eds.), *The Handbook of Social Economics* (Vol. 1A, pp. 511–586). North Holland Press.

Jackson, M., & Wolinsky, A. (1996). A Strategic Model of Social and Economic Networks. *Journal of Economic Theory*, *71*(1), 44–74. doi:10.1006/jeth.1996.0108

Jovanovic, B. (1979). Job Matching and the Theory of Turnover. *The Journal of Political Economy*, *87*(5), 972–990. doi:doi:10.2307/1833078

Kalecki, M. (1945). On the Gibrat Distribution. *Econometrica*, *13*(2), 161–170. doi:10.2307/1907013

Kapteyn, J. (1903). *Skew Frequency Curves in Biology and Statistics*. P. Noordhoff.

Kesten, H. (1973). Random Difference Equations and Renewal Theory for Products of Random Matrices. *Acta Mathematica*, *131*(1), 207–248. doi:10.1007/BF02392040

Kleinberg, J. (2000). The Small-World Phenomenon: An Algorithm Perspective. In *Proceedings* (pp. 163–170). Presented at the 32nd Annual ACM Symposium on Theory of Computing, New York, NY: ACM. doi:10.1145/335305.335325

Knack, S., & Keefer, P. (1997). Does Social Capital Have an Economic Payoff? A Cross-Country Investigation. *The Quarterly Journal of Economics*, *112*(4), 1251–1288. doi:10.1162/003355300555475

Konno, T. (2009). Network Structure of Japanese Firms. Scale-Free, Hierarchy, and Degree Correlation: Analysis from 800,000 Firms. *The Open-Access, Open-Assessment E-Journal*, *3*. doi:10.5018/economics-ejournal.ja.2009-31

Krauth, B. (2003). A Dynamic Model of Job Networking and Persistent Inequality. In P. R. Cowan & D. N. Jonard (Eds.), *Heterogenous Agents, Interactions and Economic Performance* (pp. 193–204). Springer Berlin Heidelberg. Retrieved from http://link.springer.com/chapter/10.1007/978-3-642-55651-7_12

Krauth, B. (2004). A Dynamic Model of Job Networking and Social Influences on Employment. *Journal of Economic Dynamics and Control*, *28*(6), 1185–1204. doi:10.1016/S0165-1889(03)00079-4

Krugman, P. (1996). *The Self Organizing Economy* (1st ed.). Wiley-Blackwell.

Lagos, R. (2000). An Alternative Approach to Search Frictions. *Journal of Political Economy*, *108*(5), 851–873. doi:10.1086/317674

Levy, M., & Solomon, S. (1996). Power Laws are Logarithmic Boltzmann Laws. doi:doi:10.1142/S0129183196000491

Lewkovicz, Z., Domingue, D., & Kant, J. (2009). An agent-based simulation of the French labour market: studying age discrimination. In B. Edmonds & N. Gilbert (Eds.), *Proceedings*. Presented at the 6th Conference of the European Social Simulation Association, Surrey.

Lewkovicz, Z., & Kant, J. (2007). A Multi-Agent System to Model the Labor Market: Simulating a New Job Contract Introduction. In *Proceedings* (pp. 151–162). Presented at the 4th Conference of the European Social Simulation Association, Toulouse.

Lewkovicz, Z., & Kant, J. (2008). A Multiagent Simulation of a Stylized French Labor Market: Emergences at the Micro Level. *Advances in Complex Systems*, *11*(02), 217–230. doi:10.1142/S0219525908001581

Lin, N. (1982). Social Resources and Instrumental Action. In P. Marsden & N. Lin (Eds.), *Social Structure and Network Analysis* (pp. 131–145). Beverly Hills, CA: Sage.

Lin, N. (2001). *Social Capital : A Theory of Social Structure & Action*. Port Chester, NY: Cambridge University Press. Retrieved from http://site.ebrary.com/lib/georgemason/docDetail.action?docID=5001727

Lippman, S., & McCall, J. (1976). The Economics of Job Search: A Survey. *Economic Inquiry*, *14*(2), 155–189. doi:10.1111/j.1465-7295.1976.tb00386.x

Little, I. (1962). Higgledy Piggledy Growth. *Bulletin of the Oxford University Institute of Economics & Statistics*, *24*(4), 387–412. doi:10.1111/j.1468-0084.1962.mp24004001.x

Longini, I., Nizam, A., Xu, S., Ungchusak, K., Hanshaoworakul, W., Cummings, D., & Halloran, M. (2005). Containing Pandemic Influenza at the Source. *Science*, *309*(5737), 1083–1087.

Loury, G. (1977). A Dynamic Theory of Racial Income Differences. In P. A. Wallace & A. Le Mund (Eds.), *Women, Minorities, and Employment Discrimination*. Lexington, MA: Lexington Books.

Maliranta, M., & Nikulainen, T. (2008). Labour Flow Paths as Industry Linkages: A Perspective on Clusters and Industry Life Cycles. *The Research Institute of the Finnish Economy*, (1168).

Mandelbrot, B. (1959). A Note on a Class of Skew Distribution Functions: Analysis and Critique of a Paper by H. A. Simon. *Information and Control*, *2*(1), 90–99. doi:10.1016/S0019-9958(59)90098-1

Martin, C., & Neugart, M. (2009). Shocks and Endogenous Institutions: An Agent-based Model of Labor Market Performance in Turbulent Times. *Computational Economics*, *33*(1), 31–46. doi:10.1007/s10614-008-9149-z

Maslov, S., & Sneppen, K. (2002). Specificity and Stability in Topology of Protein Networks. *Science*, *296*(5569), 910–913. doi:10.1126/science.1065103

Matia, K., Fu, D., Buldyrev, S., Pammolli, F., Riccaboni, M., & Stanley, H. (2004). Statistical Properties of Business Firms Structure and Growth. *Europhysics Letters*, *67*(3), 498–503. doi:10.1209/epl/i2004-10070-4

McCall, J. (1970). Economics of Information and Job Search. *The Quarterly Journal of Economics*, *84*(1), 113–126. doi:10.2307/1879403

Milgram, S. (1967). The Small World Problem. *Psychology today*, *2*(1), 60–67.

Mincer, J. (1974). *Schooling, Experience, and Earnings* (Vol. 2). NBER.

Mitzenmacher, M. (2004). A Brief History of Generative Models for Power Law and Lognormal Distributions. *Internet Mathematics*, *1*(2), 226–251. doi:10.1080/15427951.2004.10129088

Montgomery, J. (1991). Social Networks and Labor-Market Outcomes: Toward an Economic Analysis. *American Economic Review*, *81*(5), 1408–1418. doi:10.2307/2006929

Montgomery, J. (1994). Weak Ties, Employment, and Inequality: An Equilibrium Analysis. *American Journal of Sociology*, *99*(5), 1212–1236. doi:10.2307/2781148

Mortensen, D. (1970). A Theory of Wage and Employment Dynamics. In E. Phelps, A. Alchian, & G. Archibald (Eds.), *Microeconomic Foundations of Employment and Inflation Theory* (pp. 124–66). W. W. Norton & Company.

Mortensen, D. (1973). A Theory of Wage and Employment Dynamics. In E. Phelps (Ed.), *Microeconomic Foundations of Employment and Inflation Theory*. W. W. Norton & Company.

Mortensen, D. (1987). Job Search and Labor Market Analysis. In O. Ashenfelter & R. Layard (Eds.), *Handbook of Labor Economics* (Vol. 2, pp. 849–919). Amsterdam: Elsevier.

Mortensen, D., & Pissarides, C. (1994). Job Creation and Job Destruction in the Theory of Unemployment. *The Review of Economic Studies*, *61*(3), 397–415. doi:10.2307/2297896

Mortensen, D., & Pissarides, C. (1999a). New Developments in Models of Search in the Labor Market. In O. Ashenfelter & D. Card (Eds.), *Handbook of Labor Economics* (Vol. 3, pp. 2567–2627). Amsterdam: Elsevier.

Mortensen, D., & Pissarides, C. (1999b). Job Reallocation, Employment Fluctuations and Unemployment. In J. Taylor & M. Woodford (Eds.), *Handbook of Macroeconomics* (pp. 1171–1228). Amsterdam: Elsevier.

Mortensen, D., & Vishwanath, T. (1994). Personal Contacts and Earnings: It is Who You Know! *Labour Economics*, *1*(2), 187–201. doi:10.1016/0927-5371(94)90004-3

Munshi, K. (2003). Networks in the Modern Economy: Mexican Migrants in the U. S. Labor Market. *The Quarterly Journal of Economics*, *118*(2), 549–599. doi:10.1162/003355303321675455

Myers, C., & Shultz, G. (1976). *The Dynamics of a Labor Market: A Study of the Impact of Employment Changes on Labor Mobility, Job Satisfactions, and Company and Union Policies*. Greenwood Press.

Nagypal, E. (2008). Worker Reallocation over the Business Cycle: The Importance of Employer-to-Employer Transitions. *Working Paper (submitted)*.

Narayan, D. (1999). *Bonds and Bridges: Social Capital and Poverty*. World Bank Publications.

Neugart, M. (2004). Endogenous Matching Functions: An Agent-Based Computational Approach. *Advances in Complex Systems*, *07*(02), 187–201. doi:10.1142/S0219525904000147

Neugart, M. (2006). Labor Market Policy Evaluation with an Agent-based Model. *Discussion Paper*, (SP I 2006-113).

Neugart, M., & Richiardi, M. (2012). Agent-Based Models of the Labor Market. *Working Paper*, *125*.

Newman, M. (2002). Assortative Mixing in Networks. *Physical Review Letters*, *89*(20), 208701. doi:10.1103/PhysRevLett.89.208701

Newman, M. (2006). Finding Community Structure in Networks Using the Eigenvectors of Matrices. *Physical Review E*, *74*(3), 036104. doi:10.1103/PhysRevE.74.036104

Newman, M., & Watts, D. (1999). Renormalization Group Analysis of the Small-World Network Model. *Physics Letters A*, *263*(4–6), 341–346. doi:10.1016/S0375-9601(99)00757-4

Nowak, M., & May, R. (1993). The Spatial Dilemas of Evolution. *International Journal of Bifurcation and Chaos*, *03*(01), 35–78. doi:10.1142/S0218127493000040

Paserman, M. (2008). Job Search and Hyperbolic Discounting: Structural Estimation and Policy Evaluation. *The Economic Journal*, *118*(531), 1418–1452. doi:10.1111/j.1468-0297.2008.02175.x

Perline, R., Axtell, R., & Teitelbaum, D. (2006). *Volatility and Asymmetry of Small Firm Growth Rates over Increasing Time Frames* (No. 285). Washington, D.C.: SBA Office of Advocacy.

Petrongolo, B., & Pissarides, C. (2001). Looking into the Black Box: A Survey of the Matching Function. *Journal of Economic Literature*, *39*(2), 390–431. doi:10.1257/jel.39.2.390

Phelps, E. (Ed.). (1970). *Microeconomic Foundations of Employment and Inflation Theory*. New York: Norton.

Pingle, M., & Tesfatsion, L. (2003). Evolution of Worker-Employer Networks and Behaviors Under Alternative Non-Employment Benefits: An Agent-Based Computational Study. In A. Nagurney (Ed.), *Innovations in Financial and Economic Networks* (pp. 256–285). Edward Elgar Publishers.

Pissarides, C. (1984). Search Intensity, Job Advertising, and Efficiency. *Journal of Labor Economics*, *2*(1), 128–143. doi:10.2307/2535020

Porter, M., Onnela, J., & Mucha, P. (2009). Communities in Networks. *Notices of the American Mathematical Society*, *56*(9), 1082–1097.

Portes, A. (1998). Social Capital: Its Origins and Applications in Modern Sociology. *Annual Review of Sociology*, *24*(1), 1–24.

Putnam, R. (1993). The Prosperous Community: Social Capital and Public Life. *The American Prospect*, *4*(13), 35–42.

Ravasz, E., Somera, A., Mongru, D., Oltvai, Z., & Barabási, A. (2002). Hierarchical Organization of Modularity in Metabolic Networks. *Science*, *297*(5586), 1551–1555. doi:10.1126/science.1073374

Redner, S. (1998). How Popular is Your Paper? An Empirical Study of the Citation Distribution. *The European Physical Journal B - Condensed Matter and Complex Systems*, *4*(2), 131–134. doi:10.1007/s100510050359

Rees, A. (1966). Information Networks in Labor Markets. *American Economic Review*, *56*(1/2), 559–566. doi:10.2307/1821319

Richiardi, M. (2004). A Search Model of Unemployment and Firm Dynamics. *Advances in Complex Systems*, *07*(02), 203–221. doi:10.1142/S0219525904000159

Richiardi, M. (2006). Toward a Non-Equilibrium Unemployment Theory. *Computational Economics*, *28*(4), 421–446. doi:10.1007/s10614-006-9050-6

Rogerson, R., Shimer, R., & Wright, R. (2005). Search-Theoretic Models of the Labor Market: A Survey. *Journal of Economic Literature*, *43*(4), 959–988. doi:10.1257/002205105775362014

Saichev, A., Malevergne, Y., & Sornette, D. (2009). *Theory of Zipf's Law and Beyond* (1st ed.). Springer.

Saito, Y., Berger, T., & Iwamura, M. (2007). Do Larger Firms Have More Interfirm Relationships? *Physica A: Statistical Mechanics and its Applications*, *383*(1), 158–163. doi:10.1016/j.physa.2007.04.097

Sapienza, M., & Fontana, M. (2002). Search in Artificial Labour Makets: A Simulation Study. In F. Luna & A. Perrone (Eds.), *Agent-Based Methods in Economics and Finance* (pp. 283–304). Springer. Retrieved from 10.1007/978-1-4615-0785-7_11

Schmutte, I. (2010). Free to Move? A Network Analytic Approach for Learning the Limits to Job Mobility. *Working Paper*.

Shimer, R. (2005). The Cyclical Behavior of Equilibrium Unemployment and Vacancies. *The American Economic Review*, *95*(1), 25–49. doi:10.2307/4132669

Shimer, R. (2007). Mismatch. *American Economic Review*, *97*(4), 1074–1101. doi:10.1257/aer.97.4.1074

Shimer, R. (2010). *Labor Markets and Business Cycles*. Princeton University Press.

Simon, C., & Warner, J. (1992). Matchmaker, Matchmaker: The Effect of Old Boy Networks on Job Match Quality, Earnings, and Tenure. *Journal of Labor Economics*, *10*(3), 306–330. doi:10.2307/2535091

Simon, H. (1955). On a Class of Skew Distribution Functions. *Biometrika*, *42*(3-4), 425. doi:0.1093/biomet/42.3-4.425

Simon, H. (1960). Some Further Notes on a Class of Skew Distribution Functions. *Information and Control*, *3*(1), 80–88. doi:10.1016/S0019-9958(60)90302-8

Simon, H., & Bonini, C. (1958). The Size Distribution of Business Firms. *The American Economic Review*, *48*(4), 607–617. doi:10.2307/1808270

Smith, T., & Zenou, Y. (2003). Spatial Mismatch, Search Effort, and Urban Spatial Structure. *Journal of Urban Economics*, *54*(1), 129–156. doi:10.1016/S0094-1190(03)00040-8

Souma, W., Fujiwara, Y., & Aoyama, H. (2006). Change of Ownership Networks in Japan. In *Practical Fruits of Econophysics* (pp. 307–311). Presented at the Third Nikkei Econophysics Symposium, Springer. doi:10.1007/4-431-28915-1_56

Stanley, M., Amaral, L., Buldyrev, S., Havlin, S., Leschhorn, H., Maass, P., … Stanley, H. (1996). Scaling Behaviour in the Growth of Companies. *Nature*, *379*(6568), 804–806. doi:10.1038/379804a0

Stewart, J. (2002). Recent Trends in Job Stability and Job Security: Evidence from the March CPS. *BLS Working Paper*, (356).

Stigler, G. (1962). Information in the Labor Market. *Journal of Political Economy*, *70*(5), 94–105. doi:10.2307/1829106

Sutton, J. (1997). Gibrat's Legacy. *Journal of Economic Literature*, *35*(1), 40–59. doi:10.2307/2729692

Tassier, T., & Menczer, F. (2001). Emerging Small-World Referral Networks in Evolutionary Labor Markets. *IEEE Transactions on Evolutionary Computation*, *5*, 482–492.

Tassier, T., & Menczer, F. (2008). Social Network Structure, Segregation, and Equality in a Labor Market with Referral Hiring. *Journal of Economic Behavior & Organization*, *66*(3–4), 514–528. doi:10.1016/j.jebo.2006.07.003

Tesfatsion, L. (1998a). Preferential Partner Selection in Evolutionary Labor Markets: A Study in Agent-Based Computational Economics. In V. Porto, N. Saravanan, D. Waagen, & A. Eiben (Eds.), *Evolutionary Programming VII* (pp. 15–24). Berlin: Springer-Verlag.

Tesfatsion, L. (1998b). Ex Ante Capacity Effects in Evolutionary Labor Markets with Adaptive Search. *Working Paper*, *48*.

Tesfatsion, L. (2001). Structure, Behavior, and Market Power in an Evolutionary Labor Market with Adaptive Search. *Journal of Economic Dynamics and Control*, *25*(3–4), 419–457. doi:10.1016/S0165-1889(00)00032-4

Topa, G. (2001). Social Interactions, Local Spillovers and Unemployment. *The Review of Economic Studies*, *68*(2), 261–295. doi:10.2307/2695929

Travers, J., & Milgram, S. (1969). An Experimental Study of the Small World Problem. *Sociometry*, 425–443.

Wahba, J., & Zenou, Y. (2005). Density, Social Networks and Job Search Methods: Theory and Application to Egypt. *Journal of Development Economics*, *78*(2), 443–473. doi:10.1016/j.jdeveco.2004.11.006

Watts, D., & Strogatz, S. (1998). Collective Dynamics of "Small-World" Networks. *Nature*, *393*(6684), 440–442. doi:10.1038/30918

Woolcock, M., & Narayan, D. (2000). Social Capital: Implications for Development Theory, Research, and Policy. *The World Bank Research Observer*, *15*(2), 225–249. doi:10.1093/wbro/15.2.225

Yule, G. (1925). A Mathematical Theory of Evolution, Based on the Conclusions of Dr. J. C. Willis, F. R. S. *Philosophical Transactions of the Royal Society of London*, *213*, 21–87. doi:10.1098/rstb.1925.0002

CURRICULUM VITAE

Omar A. Guerrero graduated from the interdisciplinary high school program at Tecnológico de Monterrey, Mexico City, Mexico, in 2002. He received his Bachelor of Arts in Economics from the same institution in 2007. He received his Master of Science in Agent-Based Computational Economics and Finance from the University of Essex, Colchester, UK, in 2008.