ABSTRACT

Title of Dissertation: ESSAYS ON ENTREPRENEURSHIP, LABOR MARKETS, AND BUSINESS CYCLES

Gonzalo García-Trujillo
Doctor of Philosophy, 2021

Dissertation Directed by: Professor S. Borağan Aruoba
Professor John Haltiwanger
Department of Economics

This dissertation studies the relationship between entrepreneurship, labor markets, and business cycles. Chapter 1 studies how the labor market conditions affect the formation and growth potential of new businesses over the business cycle. I develop a dynamic occupational choice model with labor market frictions and joint firm and worker dynamics, in which heterogeneous individuals choose between being employed/unemployed workers, subsistence self-employed, or entrepreneurs with potential to grow. Using U.S data, I provide support for the following predictions. First, unemployment makes more people willing to start a business because of the lower outside option. Second, a lower job finding rate reduces the value of the fallback option when the business fails (harder to find a job), deterring entry from employment, especially for high-skill workers for whom the wage loss is greater. Third, high-growth startups are mostly started by high-skill workers. Then, with the calibrated model, I study the dynamic response of the economy to a negative productivity shock. I find that (i) entry due to stopgap motives increases while en-

try into entrepreneurship declines, and (ii) the composition of founders shifts toward fewer employed high-skill workers, making new cohorts to have fewer high-growth startups. Both features hinder job creation recovery, keeping the labor market depressed longer and the entrepreneurship entry persistently low.

Chapter 2 studies the short- and medium-term effects of employment spells in startups on future earnings of employees. Using Chilean Unemployment Insurance data, we find that those who moved toward startups earned over the next five years 21% less than those who transitioned to established firms. However, when taking selection into account, the difference is reduced to -14%, which implies that an important part of the observed effect comes from sorting. We further decompose this effect and find that 11 percentage points of the overall 5-year effect come from lower average wages. The remaining 3% arises from more unemployment periods. We also find that the contemporary differential, measured as the effect over the first year after the transition, is smaller than the medium-term effect. This suggests the existence of a scarring effect on earnings from working for a startup. Finally, we also find heterogeneous effects over the business cycle.

Chapter 3 provides a characterization of the business formation process in the U.S. from an individual-level perspective using the Survey of Participation Program (SIPP). The goal of this paper is twofold. First, it performs a detailed description of businesses and owners in the U.S. in terms of their characteristics. It also performs a detailed analysis of the survival probabilities and labor income differences between before and after self-employment spells. Second, it illustrates the potential of the SIPP as a source of information to study the early stage of business formation.

ESSAYS ON ENTREPRENEURSHIP, LABOR MARKETS, AND BUSINESS CYCLES

by

Gonzalo García-Trujillo

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2021

Advisory Committee:
Professor S. Borağan Aruoba, Co-Chair
Professor John Haltiwanger, Co-Chair
Professor Felipe Saffie
Professor Sergio Urzúa
Professor Liu Yang

Dedication

A Maria Paz, mi madre Ilse, y mi familia.

Acknowledgments

I am extremely grateful to all the people who were continuously supporting me throughout this adventure. This thesis and the completion of my doctoral process would not have been possible without the guidance of my dissertation committee and the encouragement and love of family and friends.

I am deeply indebted to my advisors, Borağan Aruoba and John Haltiwanger. Borağan Aruoba gave me the opportunity to start working with him at an early stage in the program. I owe much of my doctoral formation to him, both as a researcher and as an instructor. I'm truly thankful for his continuous guidance, support, trust, and patience over all these years. John Haltiwanger introduced me to the fascinating and complex world of entrepreneurship and labor markets. I will be eternally grateful to him for opening my eyes to this world and for the guidance, support, motivation, and encouragement that I received from him over all these years. I thank him for his invaluable advice in each step of this dissertation and job market, teaching me with patience, and supporting all my endeavors. What I learned from him has made a great impact on my professional formation, and I will always be thankful to him for that.

I especially thank Felipe Saffie. His support and advice during the last six years have been a fundamental piece of this journey. His advice in terms of research and in other dimensions of the PhD life has been a guide from beginning to end. He always encouraged me to keep pushing and help me to find ways to improve my work. Walking this path would not have been possible without his support. I´m also very grateful to Sergio Urzúa, for many discussions and continuous advice and support over the years. His advice was unique and insightful, and helped me to think in my research beyond the macroeconomic perspective. I also want to thank Liu Yang for serving as the Dean's representative for my dissertation defense.

I want specially thank John Shea for his exceptional advice and support. My work benefited enormously from his suggestions and revisions. I benefited even more from his support as Graduate Director on the many occasions that I needed his help. Also, I want to thank Guido Kuersteiner for his invaluable guidance as placement director. His advice and support were fundamental during this especially challenging and pandemic job market.

At different stages of my doctoral process, my work also benefited a great deal from discussions with Luminita Stevens, Sebnem Kalemli-Ozcan, Thomas Dreschel, and Katherine Abraham, for which I'm very grateful. I also thank Cindy Clement for her invaluable guidance during my time as an instructor.

I am eternally grateful to my mentor and friend Rodrigo Fuentes. He supported and guided me towards getting a PhD, and continuously encouraged me during the process with endless conversations. I also especially thank Rodrigo Heresi for helping me navigate the graduate program over the years. Their advice and friendship were invaluable.

My friends and colleagues at the University of Maryland were also key pieces in this adventure. I am grateful to Camila Galindo, Donggyu Lee, Seth Murray, Macarena Kutscher, Nathalie Gonzalez, Alvaro Silva, Yi Liu, Mateo Uribe-Castro, Fernando Saltiel, Matt Staiger, and Alejandro Graziano for numerous conversations, hours of rehearsals, and friendship. I also specially thank Joonkiu Choi, Veronika Penciakova, and Karam Jo for many discussions and support at different stages of this dissertation. I also thank to the economics department staff members, Emily Molleur, Amanda Statland, Jessica Gray, Terry Davis, and Mark Wilkerson. Their help, support, and words of encouragement helped me continuously over these years. I especially thank to Vickie Fletcher, for her continuous support and encouragement from the first until the very last day.

Lastly, but most importantly, thanks to my family and friends. Thanks to my wife, Maria Paz, for her love, support, and always being by my side. This dissertation is both hers and mine. Thanks to my mother, Ilse, for her unconditional love. I learned from her to persevere and never give up in life. Thanks to my uncle Alberto for his fatherly love and his example. Thanks to my aunts Angelica and Magally for their profound love and support. Thanks to my cousins for sharing every joy in my life. Thanks to my beloved grandfather, Alberto, for sowing in me the importance of studying. Thanks to my lifelong friends, the "Pelopinchos", for always being there in every step of my life, including this dissertation. Lastly, thanks to my neighbors and friends, the "Shannons", their friendship was an immense support over these last six years.

Preface

Startups and young businesses are the drivers of the aggregate employment creation dynamics. However, most startups exit and, among those who survive, just a small fraction grow large and end up boosting job creation. These up-or-out dynamics of startups imply a high risk of failure, which can result in large cumulative earnings losses both for business founders and startup employees. The cost of a business failure can vary across the type of individuals and over the business cycle, influencing the composition of founders, workers, and businesses in terms of their characteristics. This dissertation presents three works that study business formation from an individual-level perspective, considering the decisions of starting businesses and working for startups as occupational choice problems that individuals solve in the labor market.

Chapter 1 studies how the labor market dynamics over the business cycle affect the decisions to start businesses at the individual level and, by shifting the entry composition of business founders, the entry and growth potential of startups. In the first part of the quantitative analysis, I show theoretically and in a reduced-form analysis, using individual- and firm-level data for the U.S., that the labor market dynamics during downturns affect the entry and growth potential of startups through two channels. The first channel is related to the dynamics of the unemployment rate. In downturns, a higher unemployment rate means that more people solve the occupational problem with a lower outside option, increasing the business entry, but mostly as a subsistence self-employed. The second channel is related to the dynamics of the job finding probability. A lower job finding probability reinforces the entry from unemployment because unemployed individuals know that it is going to be harder to find a job. However, the opposite happens with employed individuals, who are discouraged from quitting their jobs to start a risky business because if the business fails, they will have to search for a job longer, increasing the cost of a business failure. Then, the lower job finding probability in downturns discourages employed individuals with good entrepreneurial ideas from starting businesses. This channel especially discourages the entry of highly educated employed individuals because their wage loss is larger if they have to search for a job longer. In the model, consistently with the data, the highly educated individuals are precisely those most likely to start high-growth startups because of their higher outside option and a positive correlation between educational attainment and business growth potential. So, the lack of them makes the cohorts of startups have disproportionately fewer high-growth startups.

In the second part of the quantitative analysis, I use a calibrated model to quantify the macroeconomic implications of these two channels during and after a period like the Great Recession. There are three main results from this analysis. First, the increase in the entry from unemployment makes subsistence self-employment countercyclical, while the decrease in the entry from employment makes entrepreneurship countercyclical. These dynamics are consistent with the flows between labor market states from the individual-level data. Also, subsistence self-employment plays an important role as a shock absorber mechanism by smoothing the jump in the unemployment rate. Second, the entrepreneurial sector in the model corresponds to employer businesses and subsistence self-employment to non-employer businesses. Therefore, the lower entry from employment, especially by highly educated individuals, leads to a decline in the entry of employer firms and a shift in their composition toward fewer high-growth businesses, consistently with the firm-level data. The third key result is that the missing generation and lower growth potential of entrepreneurial startups slows down the aggregate job creation recovery in the aftermath of the economic downturns. In the model, this declines further the job finding probability and keeps it low for more time, deterring, even more, the business creation by employed and highly educated individuals. Therefore, firm and worker dynamics interact in equilibrium to amplify and increase the persistence of the effects arising from an aggregate productivity/demand shock.

Chapter 2 switches the focus of the analysis towards the individuals deciding to become employees at startups. In this paper, which is joint work with Nathalie Gonzalez and Álvaro Silva, we study and quantify the short- and medium-term effects of employment spells in startups on future earnings of employees. To perform the empirical analysis, we use the Chilean Unemployment Insurance data. We start our empirical work by estimating the differential in earnings for the next 5 years after a job transition between workers who transition to a startup and workers who transition to an established firm. The results show that those who moved toward startups earned over the next 5 years on average 20.7% less than those who transitioned to existing firms. However, this result includes a sorting component because workers who transition to a startup are potentially different from workers transitioning to an established firm. To isolate the effects, we use a non-parametric approach to match each treated individual with two controls using age, gender, country of birth, and year-month of transition, and previous earnings, as the relevant individuals' observable variables. By doing this, we construct a ''triplet" formed by a treated individual (a worker who transitioned to a startup) and two controls (workers who transitioned to an existing firm). This is our baseline specification. The results show that for our matched sample, the 5-year earning effect of working at a startup is -13.8%, which means that 6.9 percentage points of the unmatched difference correspond to sorting. We also find that the differential effect attributed to only a decline in average wages is -10.5%, leaving 3.3 percentage points attributable to more unemployment.

Next, we explore whether the contemporaneous effects are different than the 5-year effects. To do so, we estimate the differences in earnings during the first 12 months after a transition. We find that the negative effect goes from -13.8% in the original matched sample to -11.6%, which suggests that the negative differential increases as the individuals' careers progress. Finally, we study potential cyclical differences in the estimated effects. We re-estimate our baseline specification for transitions that occurred in 2009-10, the period in which the Great Recession hit Chile, and 2012-13, which can be thought of as a normal time. The results for the matched sample show an earning differential of -13.9% for the period 2012-13, larger in magnitude than the -10.0% estimated for the Great Recession.

In Chapter 3, considering the limitations to identify business owners in the

U.S. administrative data, I use the Survey of Participation Program (SIPP) to provide a rich characterization of the business formation process in the U.S. from an individual-level perspective. Although the SIPP is much smaller than typical administrative data sets, its advantage is that it is easily accessible, follows individuals for up to four years, records their employer and occupational changes and spell durations, and provides detailed information on their demographic and job characteristics, and earnings. It also provides information that allows us to characterize the early stage of a new business, like start and end dates, legal form of organization, profits, size (categories), and hours spent in the business.

A descriptive analysis shows that 80% of all new ventures correspond to unincorporated businesses. Founders of incorporated businesses are relatively more educated and less likely to come from unemployment than those starting unincorporated businesses. Also, unincorporated business owners have approximately 40% less previous earnings than incorporated business owners. Next, I compute the probability of survival for different time windows, educational attainment of business owners, and previous labor market status. The analysis shows that incorporated businesses are more likely to survive than unincorporated businesses, and the survival probability increases when the owner is highly educated. The results also suggest that the previous labor market status of the business founder does not affect the probability of survival. Finally, I study the earnings differential between before and after the self-employment spell. The results show an important source of heterogeneity related to the business cycle, with the difference in earnings being strongly procyclical.

Table of Contents

List of Tables

List of Figures

Chapter 1: Startups, Labor Market Frictions, and Business Cycles

1.1 Introduction

It is a well known fact that the entry rate of employer businesses in the U.S. falls during economic downturns. Moreover, an increasing number of recent studies have documented that employer businesses born during recessions start on average smaller and grow less over their entire life-cycle.[1] Since the entry and growth of young businesses drive aggregate employment creation in the U.S., these two facts shape the recovery of the labor market in the aftermath of recessions.[2] Despite this key role, relatively little is known regarding the forces driving the entry and growth potential of startups over the cycle. From a firm-level perspective, recent studies have proposed the fall in the aggregate demand and tighter financial constraints as mechanisms to explain the decline in the entry and the shift of startups toward businesses with lower growth potential that we observed in recessions.[3] However,

[1] For the United States, this fact has been documented by Sedláček & Sterk (2017), Moreira (2016), and Smirnyagin (2020) using firm-level data from the Longitudinal Business Database (LBD).

[2] Haltiwanger, Jarmin & Miranda (2013), and Decker, Haltiwanger, Jarmin & Miranda (2014) provide empirical evidence on the importance of young firms in aggregate job creation.

[3] Sedláček & Sterk (2017) argue that a fall in the aggregate demand reduces the return to the expenditure in advertisement needed to accumulate customer base, making entrants choose to start businesses with lower growth potential. Smirnyagin (2020) argues that financial frictions slow the rate at which firms reach their target size, making larger projects less profitable for entrepreneurs when financial conditions deteriorate.

from an individual-level perspective, whether this kind of startup dynamics can also emerge from changes in the characteristics of the business founders over the business cycle still remains an open question.

In this paper, I study how the labor market dynamics over the business cycle affect the decisions to start businesses at the individual-level and, by shifting the entry composition of business founders, the growth potential of startups. I develop a framework in which unemployment rate and job finding probability drive the decision of heterogeneous individuals to start businesses, giving rise to a selection mechanism with respect to their previous labor force status and educational attainment. Then, by shifting the composition of business founders toward more previously unemployed and fewer highly educated individuals in downturns, labor market dynamics induce startups to have a lower potential to grow. First, I provide empirical support for three key mechanisms in the model: (i) unemployed workers are more likely to start a business due to their lower outside option; (ii) a lower job finding probability discourages (encourage) employed (unemployed) workers from starting businesses, especially highly educated workers; (iii) employed and highly educated workers are more likely to start high-growth startups. Then, I use a calibrated model to study how the labor market dynamics shape the entry and composition of startups during and after a period like the Great Recession, and quantify the effects of these dynamics on the recovery of the aggregate job creation.

The contribution of this paper to the literature is twofold. First, this paper is the first to relate the decline in the growth potential of startups during downturns to the characteristics of founders. Empirically, I do this by using both individual level-data (CPS, SIPP) and firm-level data (SBO) to document the change in the composition of business founders over the cycle and link this with the ex-post performance of startups. Theoretically, I develop a tractable model that features a selection at entry mechanism with respect to the previous labor force status and educational attainment of business founders, giving rise to an endogenous composition of founders. Second, I study labor market dynamics as a driver for the entry and growth potential of startups over the business cycle, adding to previous literature that has proposed aggregate demand and financial constraints channels as possible explanations.

To formalize the analysis, I build a dynamic occupational choice model with labor market frictions and joint firm and worker dynamics, in which heterogeneous individuals choose between being employed/unemployed workers or risky businesses owners, either as subsistence self-employed or as entrepreneurs with potential to grow. This occupational structure implies that (i) individuals must sacrifice valuable jobs to start a business when employed, and (ii) business owners can find another job if the business fails.[4] The existence of this fallback option reduces the cost of a business failure, but how fast individuals can find another job depends on the job finding probability. Therefore, the unemployment rate and job finding probability become drivers of the entry and composition of business founders.

In this environment, labor market downturns affect the entry and composition of business founders through two channels. First, more people solve the occupa-

[4] Similar to Hombert, Schoar, Sraer & Thesmar (2020) and Choi (2017), the uncertainty whether the business will survive makes not only the current but also the expected future value of the outside option a determinant of the occupational problem.

tional problem from unemployment with a lower opportunity cost, increasing the business entry, but mostly as a stopgap activity while they keep searching for a job. Second, a lower job finding probability reinforces the entry from unemployment, but it discourages employed workers from quitting their jobs to start businesses because they know that if the business fails, it will be harder to find another wage job. In other words, it reduces the value of the fallback option. I will refer to this mechanism as the "fear to fail" effect. This effect is stronger for high-skill workers because they face a larger wage loss if they have to search for a job longer. Because these are individuals with a high outside option, who decide to start businesses only with good entrepreneurial ideas, the cohorts of businesses born in downturns will contain disproportionately fewer potentially high-growth entrepreneurs. In addition to these two channels operating through the labor market, lower aggregate demand during recessions directly discourages business entry for unemployed and employed workers, so only individuals with the most promising ideas start businesses. I refer to these channels as the "labor force composition" channel, the "labor market tightness" channel, and the "profitability" channel, respectively.

I start the analysis by presenting some key features of the aggregate firm and labor market dynamics. First, using U.S. firm-level data from the Business Dynamics Statistics (BDS), I document the procyclical entry of employer businesses, the procyclical and persistent initial average size of startups, and the lower growth potential of cohorts of businesses started in downturns. Then, using U.S. individual-level data from the Current Population Survey (CPS), I show how the entry and composition of business founders changes over the cycle. I show that the total entry into self-employment increases in recessions, while the share of businesses started by individuals previously in employment (unemployment) decreases (increases) during recessions. I also show that entry rates by previously employed individuals declines across all educational levels, but it is more pronounced for highly educated workers.

I then present the model, which has five key ingredients. First, the occupational structure allows the decisions of entry, exit, and quits of individuals to be endogenous outcomes. Second, heterogeneity in labor skills gives rise to a distribution of outside options for business founders, generating a selection at entry mechanism with respect to the previous labor force status and educational attainment of the founders. Third, labor market frictions are the key ingredient that generates variations in the entry and composition of business founders through two equilibrium objects, the unemployment rate and the job finding probability. Fourth, the availability of two technologies to start a business allows to capture the different motivations behind this decision. Entrepreneurship offers a better technology while the subsistence alternative allows to search for a job with better efficiency and without paying the fixed operational cost. Fifth, convex hiring costs help to discipline the firm dynamics over age. In this environment, individuals choose their occupations conditional on their current labor force status (matched or unmatched), labor skills, entrepreneurial ability, and aggregate productivity. The model generates endogenous entry/exit of entrepreneurs, endogenous quits to start businesses, endogenous job finding probability, and exogenous separations.

The quantitative analysis is divided into two parts. In the first part, I give empirical support to the three mechanisms that drive the channels in the model.

To test the predictions related to the "labor force composition" and "labor market tightness" channels, I use data from the Survey of Income and Participation Program (SIPP). To take the model to the data, similar to Levine and Rubinstein (2018), I proxy the subsistence technology with unincorporated self-employment and entrepreneurship with incorporated businesses with owners spending more than 35 hours per week in this activity. First, I estimate the transition probabilities from employment and unemployment into both types of self-employment. The results support *prediction 1*: unemployed workers are more prone to start businesses because of their lower outside option. Then, I estimate the effect of the job finding rate on the previous four transition probabilities. The results support *prediction 2*: a fall in the job finding probability discourages employed workers from starting businesses, with a stronger decline for high-skilled workers. Finally, using data from the Survey of Business Owners (SBO), I provide empirical support for *prediction 3*: high-skill employed workers are more likely to start high-growth businesses.

In the second part of the quantitative analysis, I calibrate the model to reproduce a set of selected labor market and firm dynamics features of the U.S. pre-Great Recession period. To discipline the labor market dynamics, the model is calibrated to match the masses and flows between labor market states. The model also matches the educational distribution of business owners and relative wages from the data, which pin down the entry rate by educational level. To discipline the firm dynamics, the model matches a set of moments capturing the relative size of startups, relative size of businesses by the education level of the owners, survival probabilities, growth, and employment shares. While not directly targeted, the model captures well the entry by type of businesses, with most of the business being started in the form of subsistence self-employment. The model also captures well the entry composition by previous labor force status, with subsistence self-employment having a larger entry from unemployment than the entrepreneurial alternative. It also fits well the entry composition by the founders' education, with entrepreneurs being relatively more educated than subsistence self-employed. These are important features in the model since they shape the entry and growth potential of startups.

Then, with the calibrated model, I solve two perfect foresight transition dynamics exercises to better understand the mechanisms driving the entry and compositional dynamics during downturns. First, I feed the model with an exogenous aggregate productivity path that triggers an unemployment rate dynamic that mimics the one exhibited by the U.S. during and after the Great Recession. The model reproduces well the labor market dynamics and the persistent decline in the entry of employer businesses, mostly driven by the decline in transitions from employment into entrepreneurship. The entry composition of employer businesses shifts toward more previously unemployed and fewer high-skill business founders. The disproportionately decline in the entry of highly educated workers makes startups start at a smaller average size and contain fewer potentially high-growth businesses. Both features hinder job creation recovery, keeping the labor market depressed longer, and the entry into entrepreneurship persistently low. I also perform a decomposition analysis to quantify the relative importance of the entry and composition margins in the persistent decline of aggregate job creation. At the time of the recession, most of the fall in job creation is accounted for by the decreasing labor demand of existing firms and the missing generation of new employer businesses, with a minimal effect from the compositional change. However, as the economy starts recovering, the compositional shift becomes more important, preventing job creation from a faster recovery. In particular, the educational composition change seems to account for most of the slower job creation while the "labor force composition" channel just explains a small fraction of the difference. Still, the "labor force composition channel plays a key role in driving the entry from unemployment, which is mostly done as a stopgap activity in the form of subsistence self-employment.

Second, to isolate the effects associated with the "labor market tightness" channel from the "profitability" channel, I compute the impulse response functions for a one-time unexpected negative shock to aggregate productivity, and I perform a counterfactual analysis making the individuals believe that the job finding probability holds constant throughout the entire transition. The results show an amplification effect arising from the former channel. When I mute the decline of the job finding probability in the occupational problem, entry into entrepreneurship falls by less, and it is not disproportionately larger for highly educated individuals anymore. This means that the growth potential of startups also falls less, and aggregate job creation recovers faster. The results show that the "labor market tightness" channel accounts for a 33% of the decline in the entrepreneurship rate, and a 30% of the increase in the unemployment rate at the peak of the recession period. Therefore, firm and worker dynamics interact in equilibrium to amplify the effects and persistence of an aggregate productivity/demand shock: a lower job creation of startups declines further the job finding probability, deterring, even more, and more persistently, the entry of startups, especially high-growth businesses. This mechanism generates a slower recovery in the entry of employer businesses and aggregate job creation, consistent with the labor market dynamics in the aftermath of the Great Recession in the U.S.[5]

One possible concern in the analysis is that the decline in entry and change in the composition of business founders during the Great Recession could have accelerated existing trends. The literature studying the long-run trend of entrepreneurship has documented a secular decline in business dynamism for the United States post-2000.[6] In this paper, as other papers studying the cyclical dynamics of startups, I work with the firm-level data linearly detrended to deal with the existence of a downward trend in the business entry. However, on the labor market side, I assume that the unemployment rate and job finding probability are stationary variables. This is the usual assumption in the literature. Nonetheless, in the data, we see that after the recession in 2001, the job finding probability never recovered the previous level, and after the Great Recession, it just recovered the level from the mid-2000 in 2017. So, this might suggest the presence of some downward trend for the job finding probability too. In terms of the micro facts documented in the reduced-form analysis of the quantitative section of this paper, this might lead to overestimating the role of the cyclical dynamics of the unemployment rate and job finding in driving the entry and composition of startups. Regarding the macroeconomic implications presented in the second part of the quantitative analysis, the results from the first

[5]The labor market tightness in the U.S. returned to its pre-recession level just in October 2014, much slower than the recovery of aggregate demand.

[6]See Decker et al. (2014) and Decker, Haltiwanger, Jarmin & Miranda (2020).

transition exercise might also be affected. If during and after the Great Recession, the unemployment rate and the job finding probability experienced variations not only in their cyclical but also in their trend components, then the sequence of aggregate productivity shocks used to reproduce the labor market dynamics would also include a trend component. If so, the resulting changes in the entry of startups, the composition of business founders, and aggregate job creation would reflect that. Again, this would generate an overestimation of the role of the business cycle in shaping the entry and growth potential of startups. As part of the next steps in the agenda for this paper, I will investigate further this concern by disentangling the cyclical and trend components in the analysis.

The remaining structure of this document is as follows. Section 1.2 summarizes the related literature and contributions. Section 1.3 presents evidence about the aggregate dynamics of business formation, business founders' composition, and the relation between ex-post firm performance and business owners' characteristics. Section 1.4 develops the model and provides intuition for the model mechanisms in a partial equilibrium analysis. Section 1.5 presents the empirical analysis and the results from the calibrated structural model. Section 1.6 concludes.

1.2 Related Literature and Contributions

My work contributes to four main strands in the literature. First, my work relates to the literature on firm dynamics studying business cycles as a source of variation in the entry and growth potential of startups. Recently, Sedláček & Sterk

(2017) and Moreira (2016) argue that the worsening demand conditions at the time of birth lead to a selection at entry mechanisms that makes firms to start smaller in recessions and grow less over their entire life-cycle. Similarly, Smirnyagin (2020) and Vardishvili (2020) argue that financial frictions and the potential entrants' ability to delay entry prevent large projects from being started in recessions. In related work, Ates & Saffie (2020) argue that the credit shortages associated with recessions lead to "Fewer but Better" firms. In their framework, financial intermediaries assign scarce funds to the most promising ideas, decreasing the number of entrants but increasing their average productivity. These works are consistent with the results from Pugsley, Sedláček & Sterk (2020) who, using Census microdata, find that most of the differences in growth speed among startups are determined by ex-ante heterogeneity rather than persistent ex-post shocks. However, these works assume that potential entrants are ex-ante identical, with businesses heterogeneity only arising from the different types of businesses that entrants decide to start. I contribute to these works in two dimensions. First, I argue two business founders' characteristics as a new source of ex-ante business heterogeneity: the previous labor force status and educational attainment. Second, I study labor market dynamics rather than aggregate demand or financial constraints mechanisms as the driver for the decision to start businesses and their future performance.

In a similar line, Sedláček (2020) and Siemer (2014) study the effects of the deficit of startups during the Great Recession on the U.S. aggregate employment dynamics in the medium- to long-run. They find that even when the immediate impact of a drop in the firm entry on aggregate employment is small, in the later years, the negative effect of the missing generation of firms strengthens because of the lack of older firms growing large in the future. My model also generates this pattern, but the shift of composition toward fewer high-skill founders leads to a deficit of high-growth startups, which slows down, even more, the labor market recovery. This feature in the model increases the persistence of the decline in the entry of employer businesses after the recessions.

Second, my work also contributes to the literature studying entrepreneurship as an occupational choice in the labor market. The studies addressing the decline in U.S. business dynamism have proposed the increasing value of the outside option of entrepreneurs to explain the downward trend in the U.S. entrepreneurship rate in the last three decades. Engbom (2019) argues that the aging of the workforce has increased the opportunity cost of potential entrants because older people have usually found better jobs. Salgado (2020), and Kozeniauskas (2018) find that the decline in the entrepreneurship rates has been relatively larger for highly educated individuals, a fact they explain by the increasing returns to high skill labor due mostly to skill-biased technological change. I also argue the outside option as a driver for entry decisions, but I focus the analysis on the business cycle, and I propose the dynamics of the job finding probability as a determinant for the entry of highly educated individuals in downturns.

Regarding the role of labor market frictions on the decision to become an entrepreneur. Galindo Da Fonseca (2019) argues the difference in the outside option to self-employment between employed and unemployed workers as a key driver for differences in the entry decisions and firm size. Using Canadian administrative tax data, he finds that differences in outside options cause unemployed workers to be more likely to become self-employed than wage workers, but to create smaller firms that are more likely to exit. Similarly, Poschke (2019) argues that labor market frictions generate a positive relationship between unemployment and self-employment rates. He develops a cross-country analysis for a stationary equilibrium using a Diamond-Mortensen-Pissarides model extended with an occupational choice structure and firm heterogeneity.

Related to the "fear to fail" mechanism, Hombert et al. (2020) empirically study how the extension of the unemployment insurance to self-employment implemented in France in 2002 affected the decision to start businesses. They find a sharp increase in the entry rate, which they attribute to the smaller "fear to fail" effect. Gaillard & Kankanamge (2019) address this fact using a structural model with risky entrepreneurship and search frictions. They show that by allowing entrepreneurs, upon a business failure, to go to unemployment and claim UI benefits, the entry rate increases because the cost of a business failure decreases. Choi (2017) proposes outside options of business founders as a key source of heterogeneity in the early growth trajectory of young firms. He shows that entrepreneurs with higher outside options as paid workers tend to take larger business risks, and thus exhibit a more up-or-out type of firm dynamics. My work contributes to this line of research in three dimensions. First, I perform a business cycle analysis rather than a steady state comparison, which allows me to study the dynamic effects of a countercyclical "fear to fail" on entry decisions. Second, my framework explicitly captures the dissimilar characteristics and motivations between subsistence self-employed and entrepreneurs, and their opposite entry dynamics. Third, the firm dynamics feature in the model allows the study of the persistent effects from the change in the entry and composition of business founders on the aggregate job creation.

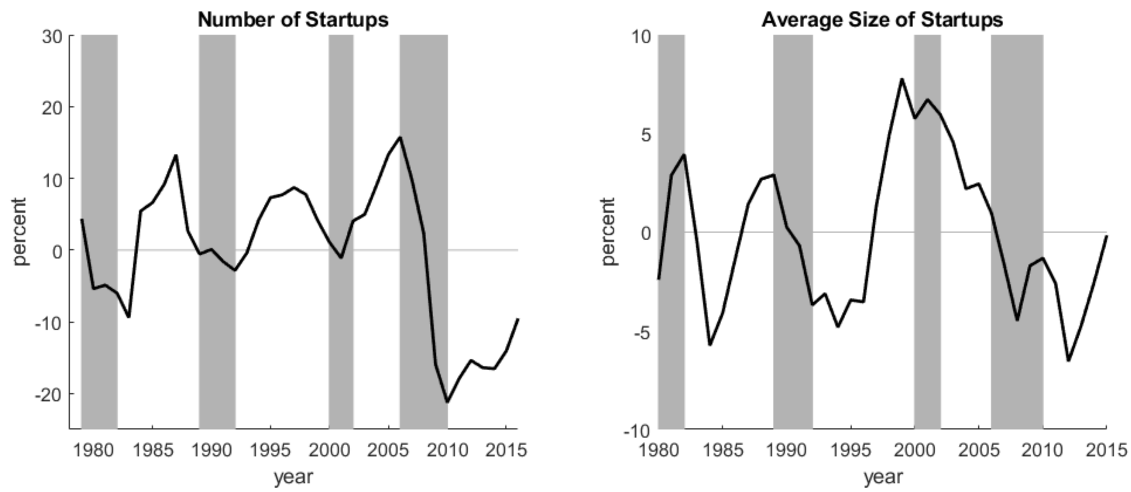
Few empirical works have studied how the characteristics of entrepreneurs vary over the business cycle. Levine & Rubinstein (2018) distinguish between entrepreneurs and other self-employment and, using data from NLSY79, show that entrepreneurship is procyclical, while self-employment is countercyclical. Fairlie & Fossen (2019) and Fossen (2020), using CPS data, show that entry into self-employment increases during recessions in the U.S., mostly due to the larger inflows from unemployment. The results in my paper are consistent with their findings. Also, to the best of my knowledge, my paper is the first to develop a structural model to study the drivers of the changes in entry and composition over the cycle in terms of previous labor force status and educational attainment of business founders.

Fourth, my work also contributes to the growing literature introducing a frictional labor market into models with firm dynamics and business cycles. Elsby & Michaels (2013) introduces a notion of firm size using decreasing returns in a random search and matching model with endogenous job destruction and aggregate uncertainty. To overcome the challenge of setting wages by Nash bargaining in a multi-worker firm framework, they use the marginal worker in an environment without the entry of firms. Schaal (2017) builds on the block recursivity approach from Menzio & Shi (2010) to extend a model of directed search on the job to a multi-worker firm environment that allows for endogenous entry and exit. In his model, constant hiring costs are needed to obtain the property of block recursivity. Audoly (2020) extends the Rank-Preserving Equilibrium approach from Moscarini & Postel-Vinay (2013) to build an on-the-job search model with convex hiring costs that features endogenous entry and exit, but with a constant returns to scale technology.[7] In my framework, endogenous entry and convex hiring costs are needed ingredients. The endogenous entry makes the selection of business evolve over the business cycle, and convex hiring costs help to discipline young firms' dynamics.[8] My model abstracts from on-the-job search, which simplifies the wage setting procedure, but still, the decreasing returns, convex hiring cost, and labor skills heterogeneity make the wage setting procedure challenging.[9] To keep the model tractable, I assume that entrepreneurs do not directly hire workers but instead buy labor units from labor agencies that match with workers under searching frictions in a one-worker-one-firm fashion. This assumption allows me to use a standard search and matching mechanism, avoiding the complexities of a framework where multi-worker firms face search frictions in a labor market with heterogeneous labor skills. Then, I set wages following a Nash bargaining solution computed for the representative worker of each level of labor skills as in Nakajima (2012).

[7] Engbom (2019) and Audoly (2020) use a constant return to scale technology and convex hiring costs. Then, the firm size is pinned down by the convex hiring cost and the entry-exit continuous process. Every firm will eventually exit, making firms not grow infinitely.

[8] In the literature, the parameters related to the stochastic technological process are also used to discipline the firm size distribution. In my framework, they are primarily used to match the high exit rates from the individual-level data.

[9] Shimer (2006) shows that in models with on-the-job search, the requirements of a convex set of possible payoffs for a unique Nash equilibrium is not satisfied. If wages are set before quit decisions, and the contract also lasts for at least some periods in the future, then the turnover that a firm will face is affected by the agreed wages. This generates multiple Nash equilibrium, violating the uniqueness of the standard models.



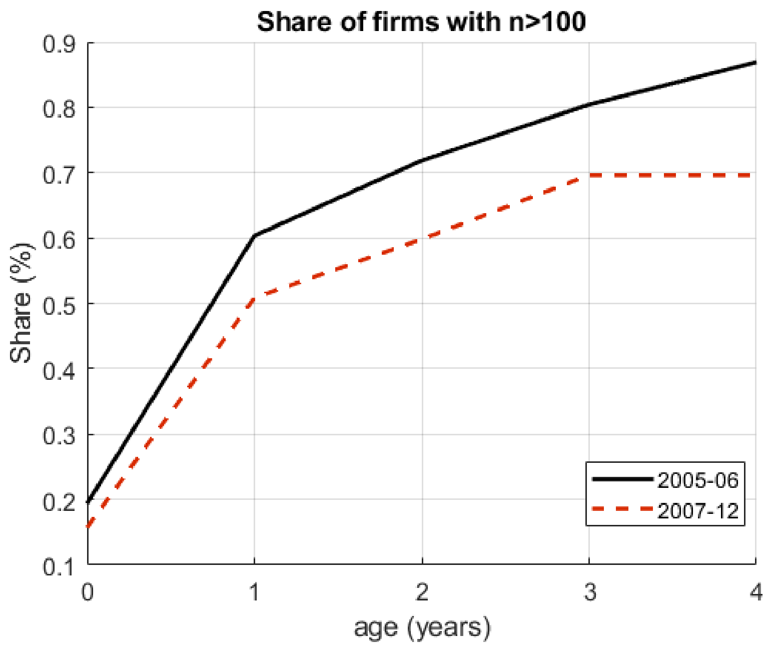
Source: author's calculations with BDS data. Figures plot percentage deviations with respect to the mean over 1979 - 2016.

Figure 1.1: Entry and Average initial Size

1.3 Firm and Labor Market Aggregate Dynamics

1.3.1 Firm Dynamics

First, in the spirit of Sedláček & Sterk (2017), I present evidence for procyclical entry of employer businesses and persistent average size of startups using firm-level data for the U.S. I use Business Dynamics Statistics (BDS) data between 1979 and 2016. BDS is the publicly available version of the confidential micro-level data from the Longitudinal Business Survey (LBD), which covers 98 percent of private employment. The BDS is an annual database, which allows us to follow cohorts of new firms for up to five years after they enter the economy. Thereafter, the BDS groups firms into age categories spanning five years, i.e., by windows of firms aged 6-20, 11-15, and 16-20 years.



Source: author's calculations with BDS data.

Figure 1.2: Share of employer firms with more than 100 employees over age

Figure 1.1 presents the number of startups and their average size between 1979 and 2016, with the latter smoothed using a 3-year moving average. Both variables are covariance stationary for this period, so their cyclical components are presented as percentage deviations with respect to the mean over 1979-2016. Both variables are procyclical, with slow recoveries after recession periods. This suggests that the negative effect of recessions on business formation go beyond the period of the recession itself. This decline in the entry during recessions reduces the contribution of startups to aggregate job creation, which is reinforced by the decline in their initial size.

Next, to explore the growth performance of businesses born at different stages of the business cycle, Figure 1.2 presents the evolution of the share of businesses with more than 100 employees over firm age for two periods, 2005-2006 and 2007-2012. I choose these two periods to compare the performance of the cohorts born just before the Great Recession with those born during and in the aftermath of it. The share of startups (0-year-old businesses) with more than 100 employees is smaller in the later period due to the smaller average initial size of startups during downturns. This difference persists and even increases as businesses age, suggesting a lack of high-growth entrepreneurs in cohorts born in downturns. A similar analysis can be done by analyzing the autocorrelations between the initial average size of startups and their future size. Appendix A.1 presents this analysis for horizons up to 5 years. Consistently, the autocorrelations for startups show that the future size of businesses is heavily correlated with their initial size. Firms that are born small during downturns are likely to remain smaller over their lifecycle. Then, I perform the same autocorrelation analysis but for a longer time horizon, using the 5-year windows cohort data from the BDS. We can see that the strong dependence of future size on initial conditions is present even 20 years later.

The findings from this section can be summarized as: (i) business creation is procyclical and highly persistent; (ii) the average size of startups is procyclical and persistent; (iii) the persistence from (i) and (ii) implies that the recovery of business entry and the average size of startups is slower than the recovery of aggregate economic conditions; (iv) the cohorts of businesses born in downturns seem to grow less, possibly due to the lack of high-growth entrepreneurs.

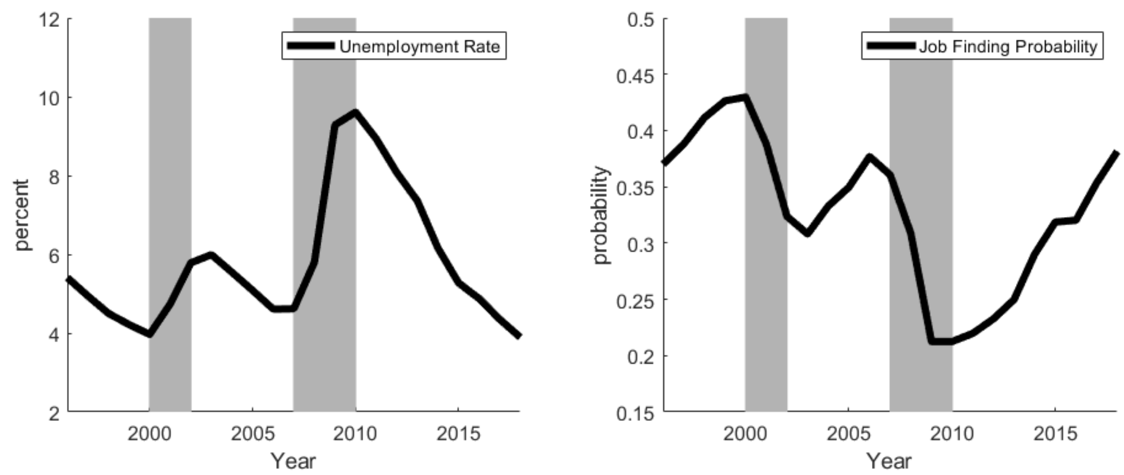
1.3.2 Labor Market Dynamics: Entry and Composition

Research studying business formation and growth of young businesses uses primarily firm-level data, which makes no possible to add information about the owners of those businesses into the analysis. Because the goal of this paper is precisely to study how the entry and composition of startups in terms of business founders characteristics change over the cycle, I need to relate the decision of entry to the business founder characteristics. To overcome this problem, I complement the firm-level analysis with individual-level data for the U.S. from the Current Population Survey (CPS) for the period 1996-2018. The CPS is the primary source of monthly labor force statistics in the U.S., and its sample size is about 60,000 households. First, I document the dynamics of the unemployment rate and job finding probability, and then I turn to analyze the entry and compositional dynamics in terms of previous labor force status and educational attainment of business founders.

Figure 1.3 shows that the unemployment rate increases heavily and persistently in recessions. The job finding probability mirrors this path but in the opposite direction.[10] This implies that more people solve the occupational problem from unemployment, and individuals make their decisions facing a persistently lower job finding rate during economic downturns.

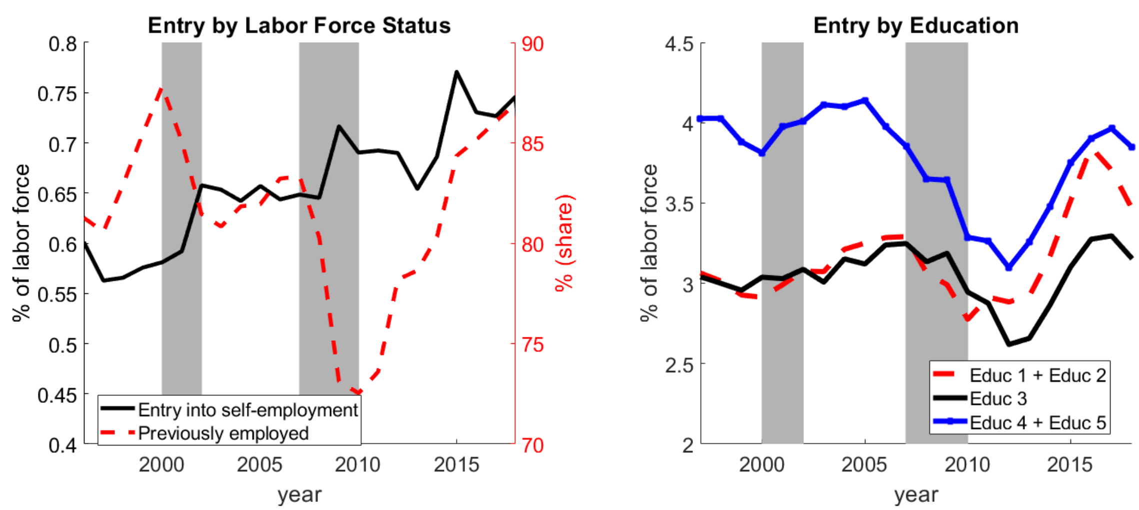
In Figure 1.4, the left panel presents the total entry into self-employment and the composition of entrants in terms of their previous labor force status for the period 1996-2019. The full set of nine transition rates between employment, unemployment

[10] Gray-shaded regions indicate NBER recession periods.



Source: author's calculations with CPS data.

Figure 1.3: Unemployment Rate and Job Finding Probability



Notes: Left panel: author's calculations with CPS. Entry into self-employment is calculated as the total number of monthly transitions from employment and unemployment into self-employment as share of the labor force. Entry by previously employed individuals corresponds to the share of entrants coming from employment only. Right panel: Author's calculations with CPS ASEC. It includes only individuals who worked during the previous year and corresponds to transitions rates from year t to year t+1. "Educ 1 + 2", "Educ 3", and "Educ 4 + 5" represent individuals with high-school or less, incomplete college, and college or graduate studies, respectively.

Figure 1.4: Entry and Composition of Business Founders

and self-employment is included in the Appendix A.2. Entry into self-employment exhibits an upward trend with a suggestive countercyclical behavior.[11] This kind of countercyclical behavior of entry in the CPS contrasts with the procyclical entry rate from the BDS. Self-employment in the CPS accounts for both employer and nonemployer businesses, while the universe of firms in the BDS only corresponds to employer firms. The dissimilar patterns in the entry of these two types of businesses suggest heterogeneous motivations to start businesses over the cycle at the individual level. The left panel also shows the share of new businesses started by previously employed individuals with respect to the total entry from employment and unemployment. The composition shifts sharply toward more people starting businesses from unemployment. This change is driven by the increase in the number of people solving the occupational problem from unemployment ("labor force composition" channel), and by the decrease in transitions from employment due to the decline in both the job finding probability ("labor market tightness" channel) and aggregate demand ("profitability" channel).

The right panel, using annual data from the CPS ASEC, presents the time path of the transition rates into self-employment by educational attainment, considering only individuals who reported to work during the previous calendar year.[12] Here, a transition into self-employment is identified as a person that reported waged work as the main activity in year "t" and then reported self-employment in the year "t+1". I present the entry rates by education at an annual frequency to focus the analy-

[11] In the data, this upward trend is matched by an upward trend in the exit rate, which makes the self-employment rate roughly constant over time.

[12] Current Population Survey Annual Social and Economic Supplement

sis on transitions cleaned from most of the stopgap activity. This makes business entry more comparable to the firm-level data.[13] The dashed red line corresponds to individuals with complete or incomplete high-school, the continuous black line to individuals with incomplete college, and the blue line with dots to individuals with complete college or graduate studies. We observe a decline in entry across all educational levels during and after the Great Recession, with a relatively larger fall for highly educated individuals, which almost double the fall of the other education levels. This pattern is consistent with the idea that the decrease in the job finding probability discourages employed workers from quitting their jobs to start a business during downturns, especially for highly educated individuals. Thus, this can be seen as suggestive evidence to support the existence of the labor market tightness channel. Regarding the composition, the larger decline in the entry for higher education levels suggests the composition shifts toward fewer high skill individuals during and after recessions.

These results are consistent with previous findings in the literature. Fairlie & Fossen (2019) document that transitions into self-employment increase during recessions in the U.S., mostly driven by the rise in transitions from unemployment to self-employment. Galindo Da Fonseca (2019), using Canadian tax data, shows that differences in outside options imply that unemployed workers are more likely to become self-employed than wage workers, but they create smaller firms and are

[13]Here, the annual transitions are constructed considering the occupations that are hold for the longest time during each year. This kind of analysis cannot be done with the monthly basic CPS data because individuals only report their occupation for a period of four months in the year. In Section 1.5, the formal empirical analysis is performed using monthly frequency data from the Survey of Income and Participation Program (SIPP).

more likely to exit. The findings from these papers are consistent with both the "labor force composition" and the "labor market tightness" channel proposed in my framework. Regarding the entry rate by education levels, to the best of my knowledge, the only work with a similar analysis is Kozeniauskas (2018). He documents a decline in the entrepreneurship rate trend across all education levels, but more pronounced for higher education levels. My analysis can be seen as an extension, which is only focused on the inflows into entrepreneurship. There is no evident decreasing trend for the entry side at any education level, which suggests that the trend of entrepreneurship might be being driven by the exit rates.

Overall, these empirical patterns suggest a role for the "labor force composition" and the "labor market tightness" channels as forces shaping the entry and composition of new businesses through the unemployment rate and the job finding probability dynamics, respectively. Entry and composition dynamics also exhibit a high persistence after the Great Recession, which is consistent with the firm-level analysis. This plays in favor of the idea that the declining employer business entry and the lower growth potential of startups might be driven by a selection mechanism arising from the labor market dynamics rather than by demand or financial conditions, which recover faster in the aftermath of recessions.

1.4 The Model

This section outlines a dynamic occupational choice model with labor market frictions and firm dynamics, in which heterogeneous individuals choose between being employed/unemployed workers, subsistence self-employed, or entrepreneurs with potential to grow.

1.4.1 Environment

Time is discrete, and the horizon is infinite. The economy is populated by a unitary mass of risk-neutral individuals, who are heterogeneous in labor skills and entrepreneurial ability.

Occupational choice structure Individuals start every period either matched or unmatched to an employer business, and the occupational decisions they can make depend on this condition. Unmatched workers can choose to be unemployed workers to search for a job with the best available search efficiency, or to start (or continue running) a business either as subsistence self-employed getting access to an inferior technology but searching for a job with an intermediate search efficiency or as an entrepreneur with access to the best possible technology but with a relatively low search efficiency. Those who choose to own a business must also decide how much labor units to hire. Matched workers can choose to stay on the job or quit to start a business at the cost of permanently losing the match. Workers becoming unmatched can start a business immediately in the current period. If they decide to keep the job, they can also become unmatched workers at the end of the period through exogenous separations. In this environment, individuals choose their occupations conditional on their current labor force status (matched or unmatched), labor skills, entrepreneurial ability, and aggregate productivity. This occupational choice structure allows the model to account for endogenous entry/exit in both technologies, and endogenous quits of workers to start businesses. The distinction between the two possible technologies is the key ingredient that captures the different motivations to start a business. Individuals starting businesses because of good entrepreneurial ideas are more likely to choose the entrepreneurial technology, while those starting business because of a lower outside option (e.g., unemployed workers) are more likely to start the business as subsistence self-employed.

Labor Skills Heterogeneity in labor skills is also key ingredient in the model. Labor skills are indexed by $h \in H$, are time-invariant for a given individual, and are distributed among the population according to a logarithmic distribution. This produces a wage distribution for workers, which generates a heterogeneous pool of potential entrants. In other words, the outside options of business owners are heterogeneous, which leads to a selection at entry mechanism: highly skilled individuals only decide to start a business when the entrepreneurial idea is good enough to compensate for the forsaken wage. At the same time, the labor skills determine the value of the fallback option in the event of a business failure. This is the potential wage that the individual would receive in the job found in a business closure event.

Entrepreneurial ability of incumbents (current business owners) For current subsistence self-employed and entrepreneurs, the idiosyncratic entrepreneurial ability of an individual i, z_i, evolves stochastically over time according to the following AR(1) process:

$$\begin{aligned} \ln z_i' &= \mu_{zh} + \rho_z \ln z_i + \sigma_z \epsilon_z' \\ \epsilon_z &\sim N(0,1) \end{aligned}$$

The stochastic realizations of the entrepreneurial ability will generate endogenous exit. In the presence of a fixed operational cost and a valuable outside option, individuals will decide to close the business if the draw of z is below some certain threshold $z^\star$. The constant μ_{zh} is used to introduce permanent heterogeneity in the model, by allowing the mean of the entrepreneurial ability to correlate with the labor skills. In particular, this parameter allows the model to capture the larger size of businesses owned by high-skill individuals that we see in the data.

Entrepreneurial ability of potential entrants Potential entrants are uncertain about their initial productivity if they decide to start a business either as a subsistence self-employed or as an entrepreneur. At the beginning of every period, before the occupational choice is made, employed and unemployed individuals receive a signal ξ about the post-entry initial productivity that they would have if they decide to start a business. The signal is drawn from a Pareto distribution $q \sim Q(q) = (\underline{q}/q)^{\xi}$. Conditional on entry, the distribution of the idiosyncratic productivity in the first period of operation is given by

$$\begin{aligned} \ln z_0 &= \rho_z \ln q + \sigma_q \epsilon_q \\ \epsilon_q &\sim N(0,1) \end{aligned}$$

Given this process, the value of starting a business is increasing in the signal q, which means that there will be a threshold $q\star$ above which the prospective entrants will decide to enter. Differently from Hopenhayn (1992), where potential entrants are identical, and there is a unique cutoff $z\star$ above which individuals decide to start businesses, the heterogeneity in the outside option in my model generates dispersion in the initial productivity of the entrants. This is a novel feature of my framework. On top of that, as in Clementi & Palazzo (2016), the uncertainty regarding the initial productivity generates heterogeneity even within individuals with the same outside option. Because of the labor market frictions, when an individual decides to quit a job to start a business, the match with the employer firm is severed. This makes that even if the entrepreneurial ability draw makes the value of being a business owner lower than the value of being a waged worker, the individual might get stuck as a business owner until get a job offer because now the relevant outside option is unemployment. This feature triggers the "fear to fail" effect in the model.

If the job finding probability is lower, then the "fear to fail" effect becomes bigger discouraging the entry from employment. This effect is stronger for high-skilled workers for whom the wage loss will be larger. This is the mechanism by which the "labor market tightness" channel works in the model.

Aggregate Productivity There is no aggregate uncertainty. Aggregate productivity A is assumed to take value 1 in the stationary equilibrium. Then, to analyze cyclical fluctuations in the entry and composition of business founders, two perfect foresight transitional dynamics exercises are performed. First, to validate the model dynamics, the model is fed with an exogenous aggregate productivity path. Then, a counterfactual exercise is performed to quantify the effects of the labor market dynamics on the entry and composition of startups. Here, I apply a one-time unanticipated shock to the path of the aggregate state, which is thereafter deterministic and perfectly known by everyone (MIT shock). In both cases, to run these exercises, I first compute the stationary equilibrium and then the perfect foresight transition dynamics consistent with the exogenous path of A using a shooting algorithm.

Labor market frictions To introduce searching and matching frictions in a tractable way, I use the assumption that firms managed by subsistence self-employed or entrepreneurs do not directly hire workers but instead buy labor efficiency units n from labor agencies. These labor agencies are subject to search frictions to hire workers.[14] Labor agencies produce labor efficiency units after a worker-firm match is realized. A labor agency transforms the indivisible h units of a worker's skilled labor into n labor efficiency units, with a one-to-one technology ($n = h$), and sells these units in a competitive market to the entrepreneurs. This implies that the size of businesses can effectively be measured in units of n. I also assume that worker ability is observable to the labor agencies, and thus these firms can direct their search to a particular worker type. These two assumptions allow for two desirable simplifications. First, because the labor agencies facing the search frictions are one-firm-one-worker matches, a standard search and matching mechanism can be used to avoid the complexities from a framework where a multi-worker firm faces search

[14] In the same fashion as Galindo Da Fonseca (2019).

frictions.[15] Second, the challenge of dealing with heterogeneous worker skills is simplified by the directed search assumption. Each type of worker is seen as a different segment of the labor market to which labor agencies direct their search, and where different wages are set according to Nash Bargaining, generating a distribution of wages $w_h(A)$ (as in Mueller (2017) and Hagedorn, Manovskii & Stetsenko (2016)).

1.4.2 Timing and model overview

Figure 1.5 summarizes the structure of the model. The timing is as follows.

- The economy starts each period with the occupational distribution from the end of the previous period.

- Signals for initial productivity of potential entrants and idiosyncratic shocks to entrepreneurial ability for incumbents are realized.

- Individuals choose occupations. Businesses also decide labor hiring. The occupational distribution for the current period, after applying the decision rules, is given by $\Psi(h, z_t, n_t, o_t)$.

- Wages are set following a Nash bargaining rule and also the market of labor efficiency units clears.

- Production and payments are carried out.

- Job findings and separations are realized according to $f(\theta_{h,t})$ and s. The occupational distribution for the beginning of the next period is determined.

[15] Elsby & Michaels (2013) introduce a notion of firm size directly into a search and matching model with endogenous job destruction.

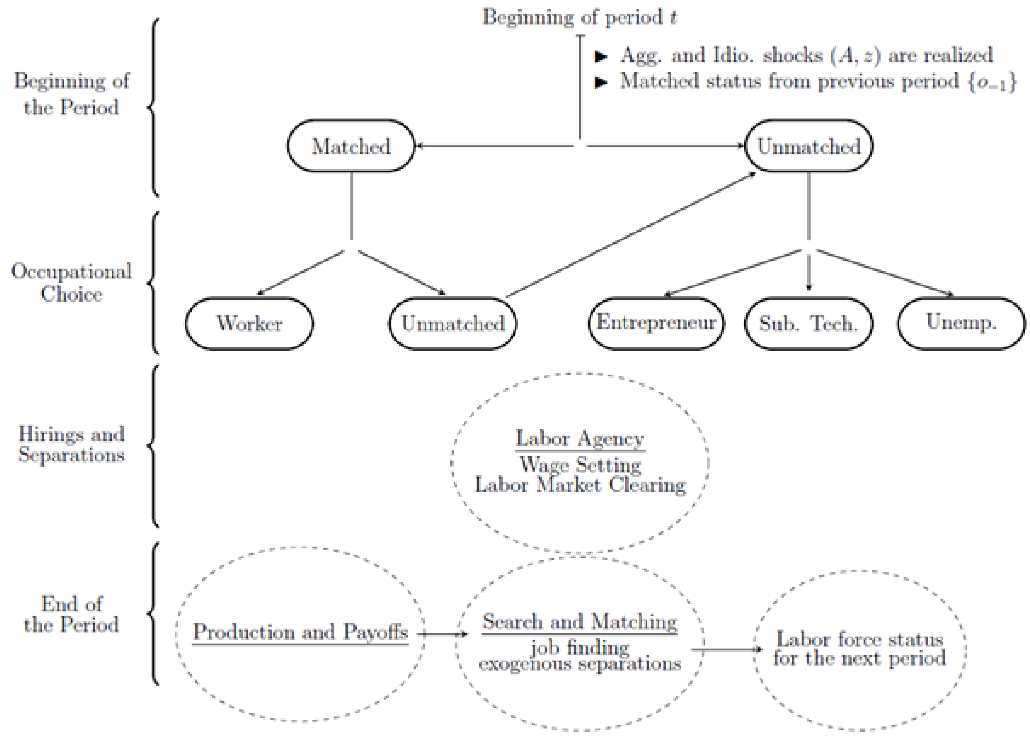


Figure 1.5: Model overview

Next, I turn to describe the individuals' decision problems, how the labor market works, and the equilibrium definition.

1.4.3 Individual's Decision Problem

First, I describe the occupational choice problem faced by matched and unmatched workers. Then, I develop expressions for the value functions associated with each possible occupation: wage worker, unemployed worker, subsistence self-employed, and entrepreneur.

1.4.3.1 Occupational Choice Problem

Individuals start each period being matched or unmatched to an employer business. The occupational choice problem for an individual with labor skills (h), signal of initial entrepreneurial ability (q), business size in $t-1$ (n_{-1}), conditional on aggregate productivity (A), is given by:[16]

Matched worker

$$W(h,q,n_{-1};A) = \max \mathbf{E}_{z/q}\Big[V^E(h,z,0;A), U(h,z,n_{-1};A)\Big] \tag{1.1}$$

where $V^E(h,z;A)$ and $U(h,z;A)$ correspond to the value functions of being a wage worker and an unmatched worker, respectively.

A matched worker can decide to stay as a waged worker or to quit to solve the problem of an unmatched worker. The only way in the model to start a business, either as subsistence self-employed or as an entrepreneur, is by quitting the job first.

Unmatched Worker

$$U(h,q,n_{-1};A) = \max \mathbf{E}_{z/q}\Big[V^U(h,z,0;A), V^S(h,z,n_{-1};A), V^F(h,z,n_{-1};A)\Big] \tag{1.2}$$

where V^U is the value of being an unemployed worker, V^S is the value of being subsistence self-employed, and V^F is the value of being an entrepreneur.

An unmatched worker can choose to be an unemployed worker or start a

[16]If the individual was a business owner in $t-1$, then the signal q is just the productivity realization z. If the individual was not a business owner in $t-1$, then $n_{-1}=0$.

business either as a subsistence self-employed or an entrepreneur.

Value of being a Wage Worker

$$V^E(h,q,0;A) = w_h(A) + \beta\Big[\Big(1-s\Big)\mathbf{E}_{q'}W(h,z',0;A') + s\mathbf{E}_{q'}U(h,z',0;A')\Big] \tag{1.3}$$

where $w_h(A)$ is the wage of a worker type h and s is an exogenous separation probability.[17] A wage worker receives a wage according to his type h, which is determined by Nash Bargaining with labor agencies, as will be explained below.

Value of being an Unemployed Worker

$$V^U(h,q,0;A) = Y^{ss} * w_h^{ss} + \beta\Big[f(\theta_h)\mathbf{E}_{q'}W(h,z',0;A') + \Big(1-f(\theta_h)\Big)\mathbf{E}_{q'}U(h,z',0;A')\Big] \tag{1.4}$$

where $Y^{ss} * w_h^{ss}$ is the unemployment benefit that a worker type h receives, which is proportional to the equilibrium wage, and $f(\theta_h)$ is the job finding probability for a type h worker.[18] θ_h corresponds to the labor market tightness in segment h.

[17] Same separation probability s across all labor skills is a conservative assumption considering that high-skill workers have a lower separation probability in the data.

[18] The model allows for heterogeneity in θ_h, however for simplicity, it is calibrated to the same value in the stationary equilibrium. In the data, $f(\theta_h)$ for high-skill workers is hire, but the magnitude of the decline in recessions is proportionally equivalent across all education levels.

Value of being an Entrepreneur

$$
\begin{aligned}
V^F(h,z,n_{-1};A) &= \pi(z,n_{-1};A) \qquad (1.5)\\
&+ \beta\Big[\varphi^F f(\theta_h)\mathbf{E}_{z'/z}W(h,z',n;A')\\
&+ \Big(1-\varphi^F f(\theta_h)\Big)\mathbf{E}_{z'/z}U(h,z',n;A')\Big]\\
s.t.&\\
\pi(h,z,n_{-1};A) &= \max_n\Big\{zAn^\alpha-\rho(A)n-\phi-g(\chi)\Big\}\\
g(\chi) &= \frac{\kappa}{\gamma}\chi^\gamma n;\quad \chi=\max\Big\{0,\frac{n-n_{-1}}{n_{-1}}\Big\}
\end{aligned}
$$

where ϕ is a fixed operational cost as in Hopenhayn (1992), φ^F is the search efficiency of entrepreneurs, n is the number of labor efficient units hired to produce the final good and ρ is its price, and (κ,γ) are the parameters in the hiring costs function. Relative to a model with linear recruitment costs, convex costs $\gamma>1$ generate a pronounced labor market propagation, featuring sluggish adjustments of the job-finding rate and of the vacancy-unemployment ratio.

Entrepreneurs produce by using labor efficiency units n as input according to a technology that depends on their entrepreneurial ability z and aggregate productivity A. I assume decreasing returns, so that $0<\alpha<1$. Because starting and closing a business correspond to occupational decisions of unmatched individuals, the entry and exit are endogenous outcomes in the model.

Value of being a Subsistence Self-Employed

$$
\begin{aligned}
V^S(h,z,n_{sub_{-1}};A) &= \pi_{sub}(z,n_{sub_{-1}};A) \\
&+ \beta\Big[\varphi^S f(\theta_h)\mathbf{E}_{z'/z}W(h,z',n;A') \\
&+ \Big(1-\varphi^S f(\theta_h)\Big)\mathbf{E}_{z'/z}U(h,z',n;A')\Big]
\end{aligned} \tag{1.6}
$$

$$
\begin{aligned}
s.t. & \\
\pi_{sub}(z;A) &= \max_{n_{sub}}\Big\{A_{sub}A^{\nu}zn^{\alpha_{sub}} - \rho(A)n_{sub} - g(\chi_{sub})\Big\} \\
g(\chi_{sub}) &= \frac{\kappa}{\gamma}\chi_{sub}^{\gamma}n_{sub}; \quad \chi_{sub} = \max\Big\{0, \frac{n_{sub}-n_{sub_{-1}}}{n_{sub_{-1}}}\Big\}
\end{aligned}
$$

where $A_{sub} < 1$ is a scale parameter that allows the model to discipline the inferior technology and A^{ν}, with $\nu < 1$, is used to allow for a different sensitivity for the subsistence activity with respect to the aggregate shock. The parameter α is the span of control measure. φ^S is the search efficiency of subsistence self-employed, with $1 > \varphi^S > \varphi^F$.

Given the trade off in terms of technology and search efficiency between the subsistence self-employment and entrepreneurship alternatives, individuals with different motivations will choose different options. In the model, individuals that fall into unemployment and are relatively low-skilled workers are more likely to start businesses as a subsistence self-employed, while employed individuals that with high labor skills are more likely to start businesses as an entrepreneur.

1.4.4 Labor Market Frictions

1.4.4.1 Matching function, Job Finding rate, and Job Filling rate

Similar to Bils, Chang & Kim (2009) and Mueller (2017), I assume worker ability is observable to the labor agencies and thus labor agencies can direct their search to a particular worker type. I assume a finite number of types $h \in H$, and for each there are unemployed workers searching for a job and a continuum of labor agencies searching for workers of type h. Workers and labor agencies are matched in each segment $h \in H$ according to the following matching function:

$$M_h(v_h, S_h) = m u_h^{\psi} v_h^{1-\psi} \tag{1.7}$$

where m is the matching efficiency parameter common across all segments, u_h is the number of unemployed workers in segment h and v_h is the number of vacancies posted by the intermediate firms in segment h.

The labor market tightness in each segment h is given by $\theta_h = \frac{v_h}{S_h}$. Then, the job finding probability for an unemployed worker of type h is given by $f(\theta_h) = \frac{M_h}{S_h} = m\theta_h^{1-\psi}$. The job filling rate is given by $q^h(\theta_h) = \frac{M_h}{v_h} = m\theta_h^{-\psi}$.

1.4.4.2 Labor Agencies

Labor Agencies direct their search by posting vacancies in each segment $h \in H$. The value of posting a vacancy in segment h is:

$$V_h(A) = -c + \beta[q^h(\theta_h)J_h(A') + (1 - q^h(\theta_h))V_h(A')] \tag{1.8}$$

The value of a filled vacancy in segment h is:

$$J_h(A) = \rho(A)h - w_h(A) + \beta[(1 - s)J_h(A') + sV_h(A')] \tag{1.9}$$

The zero profit condition of posting vacancies in segment h is $V_h = 0, \forall h \in H$.

1.4.4.3 Market Clearing Conditions / Wage Determination

The selling price $\rho(A)$ is determined by the market clearing condition of the labor good:

$$\int h\, d\Psi^E(h, z_t; A) = \int n(h, z_t, n_{t-1}; A)\, d\Psi^F(z_t, n_{t-1}; A) \quad \forall t \tag{1.10}$$

The total production of labor efficiency units across all segments $h \in H$, which is equivalent to the integral of h over the type distribution of wage workers, must be equal to the total demand for labor efficiency units which corresponds to the integral of the demand of subsistence self-employed and entrepreneurs $n(z_t, n_{t-1}; A)$ over their type distribution over (z, n_{-1}).

Wages are determined by centralized Nash Bargaining separately in each segment h. In each segment h, individuals differ by entrepreneurial productivity. Thus, following Nakajima (2012), I use a representative agent of each segment for the wage bargaining. Then, wages are determined by splitting the joint surplus from an employment relationship in the segment h according to a Nash Bargaining process:

$$[\tilde{W}(h,A) - \tilde{U}(h,A)] = \frac{\eta}{1-\eta}[J_h(A) - V_h(A)] \tag{1.11}$$

where $\tilde{W}(h,A)$ and $\tilde{U}(h,A)$ correspond to the value functions of the representative agent of type h, which are determined as the probability-weighted averages of the individual value functions over the distributions of z and o_{-1}.

1.4.5 Equilibrium

Define $i = \{E, U, S, F\}$ and o_t as the current occupational status. Then, given an initial distribution $\Psi(h, z, o_{-1}, n_{-1})$ and a sequence of aggregate productivity $\{A_t\}_{t=0}^{\infty}$, an equilibrium for this economy can be defined as a sequence of value functions $\{V_t^i\}_{t=0}^{\infty}$, prices for labor efficiency units $\{\rho_t(A)\}_{t=0}^{\infty}$, wage distributions $\{w_t(h,A)\}_{t=0}^{\infty}$, labor market tightness $\{\theta_h\}_{t=0}^{\infty}$, decision rules $\{o_t^i(h, z_t, o_{t-1}, n_{t-1})\}_{t=0}^{\infty}$, $\{n_{sub,t}(z_t, n_{t-1}; A)\}_{t=0}^{\infty}$, $\{n_t(z_t, n_{t-1}; A)\}_{t=0}^{\infty}$, as well as a distribution of individuals $\{\Psi(h, z_t, o_{t-1}, n_{t-1})\}_{t=0}^{\infty}$, that solve:

1. Given $\{\rho_t(A), w_t(h,A), \{f(\theta_h)\}\}_{t=0}^{\infty}$, the decision rules $\{o_t^i(h, z_t, o_{t-1}, n_{t-1})\}_{t=0}^{\infty}$, $\{n_{sub,t}(z_t, n_{t-1}; A)\}_{t=0}^{\infty}$, and $\{n_t(z_t, n_{t-1}; A)\}_{t=0}^{\infty}$, solve equations (1.1), (1.2), (1.3), (1.4), (1.5), and (1.6).

2. Given $\{\rho_t(A), w_t(h,A)\}_{t=0}^{\infty}$ and the zero profit condition, the job filling rates $\{q^h(\theta_h)\}_{t=0}^{\infty}$ solve equations (1.8) and (1.9).

3. The wage distribution $\{w_t(h,A)\}_{t=0}^{\infty}$ satisfies the Nash Bargaining solution in

equation (1.11).

4. Given the sequence $\{\rho_t(A), w_t(h, A), \theta_h\}_{t=0}^{\infty}$, the sequence of distributions of individuals $\{\Psi(h, \theta_t, o_{t-1}, n_{t-1})\}_{t=0}^{\infty}$ is consistent with the decision rules and the transition matrix T:

$$\Psi(x_{t+1}) = T\Psi(x_t) \tag{1.12}$$

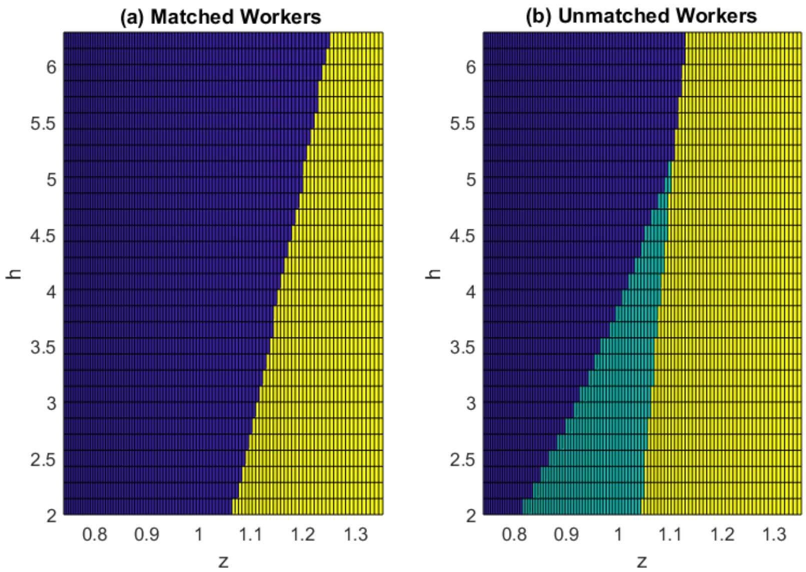
5. Price sequence $\{\rho_t(A)\}_{t=0}^{\infty}$ satisfies the intermediate goods market clearing condition (1.10).

Appendix A.3 presents the set of equilibrium conditions, with 18 equations and 18 unknowns, used to solve for the stationary equilibrium and the transition dynamics of the model. Appendix A.4 presents the algorithm used to solve for the stationary equilibrium and appendix A.5 the algorithm used to solve for the transition dynamics.

1.4.6 Intuition behind the mechanisms

This section presents a simplified partial equilibrium analysis to develop intuition about how the proposed mechanisms work, which at the same time correspond to three model-implied predictions that are empirically tested in the next section. For the analysis, I compute the occupational decision rules of the individual over the space (h, z), for a given set of prices $(\rho(A), w_h(A))$ and job finding probability $f(\theta_h)$, and it is assumed that the initial productivity z is equal to the signal q.[19]

[19] For the analysis, some arbitrary prices and job finding probability are used, which are not necessarily those from the stationary equilibrium.



Panel(a): blue color corresponds to the option of staying as an employed worker and yellow color to the choice of being an entrepreneur. Panel (b): blue color corresponds to remaining unemployed, green to being subsistence self-employed and the yellow to the choice of being an entrepreneur.

Figure 1.6: Occupational Decision Rules: $f(\theta) = 0.28$

1.4.6.1 "Labor Force Composition" channel

Figure 1.6 shows the occupational choice of individuals over the space formed by labor skills h and entrepreneurial ability z. Panel (a) presents the optimal choices for matched workers and panel (b) shows the optimal choices for unmatched workers. In panel (a), the blue color represents choosing to remain as a wage worker and the yellow color represents becoming an unmatched worker. In panel (b), the blue color represents the choice of being an unemployed worker, green the option of subsistence self-employment, and the yellow the option of being an entrepreneur.

Panel (a) shows that only those matched workers with a relatively high entrepreneurial ability z choose to quit their jobs to become unmatched workers. Similarly, in panel (b) individuals deciding to be entrepreneurs are those who have a relatively high entrepreneurial ability, while those with higher relative labor skills prefer to keep searching for a job. The positive slope of the division between the unemployed and entrepreneur choice spaces reflects how the value of the outside option to entrepreneurship is increasing with labor skills.

From analyzing panels (a) and (b) jointly, we see that the parameter space for which an unmatched worker chooses to become an entrepreneur is larger than for a matched worker. The reason for this can be decomposed in two parts. First, matched workers can choose to be a wage worker in the current period, which is not an available option for unmatched workers. This implies a lower outside option to self-employment for unmatched workers in the current period, generating an expansion of the entrepreneurial choice space for them. Second, if the probability of finding a job is lower than 1, the value of labor skills relative to entrepreneurial ability decreases for unmatched workers. This means a steeper division line between the decision rule spaces of unmatched workers, reinforcing the expansion of the space where they choose to become entrepreneurs. The opposite happens with matched workers, for whom the division line becomes less steep. Suppose an employed worker chooses to start a business and the business fails. In that case, it is going to take more time to find another job if the job finding probability is low, which reduces the expected value of becoming an entrepreneur. The imperfect probability of finding a job affects high-skill workers disproportionately more because the difference between their wages and what they receive in unemployment is bigger than for low-skill workers. The intuition for this difference is developed further below in the explanation for the "labor market tightness" channel. In the decision rule space, this means a less steep division line. This is what I call the "fear to fail" effect. Only workers who get a high draw of entrepreneurial ability, with a low failure probability, decide to become entrepreneurs. Therefore, labor market frictions generate a wedge between the outside options to entrepreneurship of matched and unmatched workers with the same labor skills, increasing the space on which unmatched workers become entrepreneurs. Thus, we should expect unemployed workers to be more prone to transit into self-employment than matched workers.

Also, there is a substantial chance that unemployed workers decide to become subsistence self-employed. This area is characterized by combinations with low z and low h. Under this technology there is no fixed operational cost, so even individuals with low entrepreneurial ability might find this option appealing. However, conditional the prices used here, this is only true for unmatched workers. All matched workers who decide to quit become entrepreneurs. This property arises from the fact that the yellow area from panel (a) is a subset of the yellow area of panel (b). Because every matched worker who quits must then solve the unmatched worker problem, we know that they are quitting to become entrepreneurs. No matched worker decides to quit to become an unemployed worker or subsistence self-employed. This should reinforce the idea that unemployed workers are more prone to start businesses than employed workers.

Therefore, the larger set of combinations (h, z) over which unemployed workers choose to start a business either as subsistence self-employed or as entrepreneurs with respect to employed workers produce the first model-implied prediction:

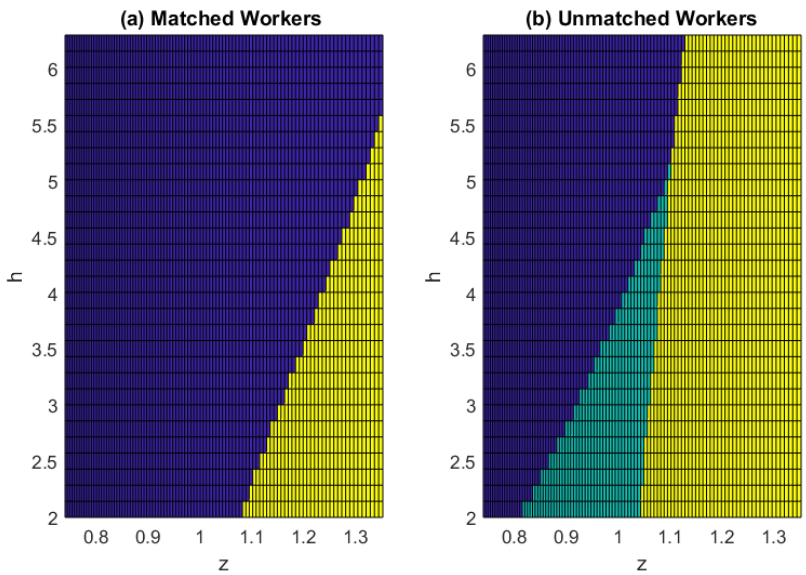
Prediction 1: *Unemployed workers are more likely to start businesses than employed workers.*

This prediction is the mechanism by which the "labor force composition" channel works. If the unemployment rate increases during recessions, then we should expect an increase in the number of individuals deciding to start businesses. In addition, if the distribution of those individuals over the space (z, h) accumulates a significant amount of people for whom the subsistence alternative is the optimal decision, then we should expect that most of that increase will take the form of subsistence self-employment. Also, because of the fall in the number of employed workers, we should expect a decrease in transitions from employment into entrepreneurship, which on average are high quality transitions.

The previous analysis may not hold in a general equilibrium context in which prices $(w_h(A), \rho(A))$ and the job finding probability $(f(\theta_h))$ adjust over the cycle.

1.4.6.2 "Labor Market Tightness" channel

Here, I present a comparative statistics exercise to explore how the job finding probability dynamics shape the occupational choices of matched and unmatched workers. A lower job finding probability increases the wedge between the outside options of matched and unmatched workers. This means that a fall in the job finding probability increases the space of combinations (h, z) over which an unemployed workers choose to start a business, while this space shrinks for employed workers. We can see this by comparing Figure 1.7, which shows the occupational decision rule with a job finding probability $f(\theta_h) = 0.1$, with the baseline case from Figure



Panel(a): blue color corresponds to the option of staying as an employed worker and yellow color to the choice of being an entrepreneur. Panel (b): blue color corresponds to remaining unemployed, green to being subsistence self-employed and the yellow to the choice of being an entrepreneur.

Figure 1.7: Occupational Decision Rules: $f(\theta) = 0.10$

1.6 that uses $f(\theta_h) = 0.28$.

The comparison between both Figures 1.6 and 1.7 shows that the decrease in the job finding probability discourages wage workers from starting businesses. This happens purely because if the business fails, it will be harder to find a job again, which decreases the value of the fallback option. In other words, there is a stronger "fear to fail" effect. Importantly, this "fear to fail" effect disproportionately affects people with higher labor skills because their future outside option is associated with a higher wage, and thus the cost of being an unemployed worker increases. In other words, the increase in the cost of failure is relatively bigger for highly skilled workers because their expected wage loss increases when the job finding probability falls. This implies that it is more likely that highly skilled workers switch their occupational decision from starting a business to staying as an employed worker because of the "fear to fail" effect.

Regarding unmatched workers, the lower job finding probability increases the incentive to enter into self-employment. This happens because the expected value of being unemployed is lower which means a lower outside option to self-employment.

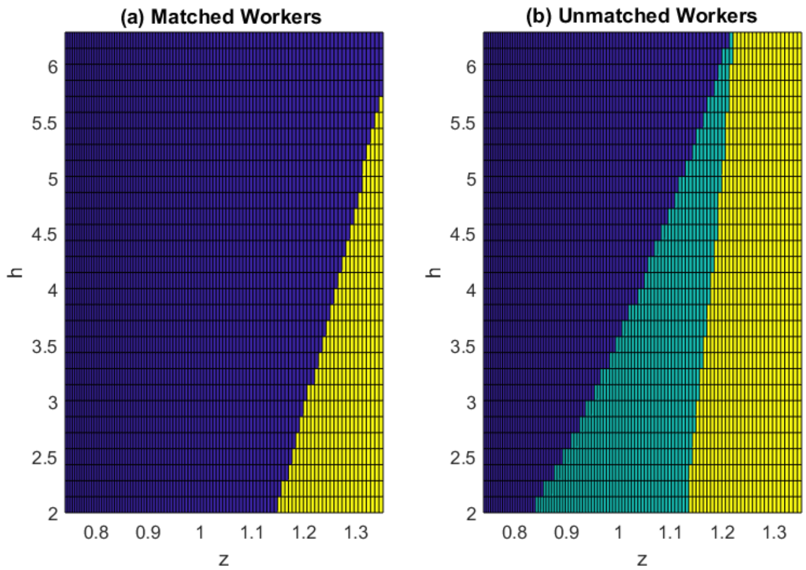
Therefore, a decline in the job finding probability should discourage entry from employment, specially for high skilled workers, because of the higher "fear to fail" effect, and it should encourage the entry from unemployment. This correspond to the second testable prediction derived from the model.

Prediction 2: *A decline in the job finding probability discourages entry from employment, specially for high skilled workers, and encourage entry from employment.*

This second prediction corresponds to the mechanism through which the "labor market tightness" works.

1.4.6.3 "Profitability" channel

Keeping everything else constant, a decrease in aggregate productivity reduces business profits. Figure 1.8 shows that the space over which both matched and unmatched workers decide to be an entrepreneur shrinks after a 7% decrease in aggregate productivity. Only those individuals with high relative entrepreneurial ability still want to be entrepreneurs. This is what I call the "profitability" channel. Through this channel we should expect that recessions lead to a decline in business



Panel(a): blue color corresponds to the option of staying as an employed worker and yellow color to the choice of being an entrepreneur. Panel (b): blue color corresponds to remaining unemployed, green to being subsistence self-employed and the yellow to the choice of being an entrepreneur.

Figure 1.8: Occupational Decision Rules: $A = 0.93A^{\star}$

entry of individuals with lower entrepreneurial ability from both unemployment and employment.

Both the "labor market tightness" and "profitability" channels discourage the entry of employed workers. However, there are two important differences. First, the "labor market tightness" discourages disproportionately high skilled workers from starting businesses, while the effect from the "profitability" channel is the same for all types of workers. Second, in the data, the aggregate demand recovers faster than the job finding probability. For the Great Recession, the job finding probability recovered its pre-recessionary level just in October 2014. Therefore, a pure aggregate demand channel struggles to generate a persistent decline in the entry of new businesses as we saw in the aftermath of the Great Recession, which can be matched better by the “labor market tightness” channel. I will revisit these ideas in the quantitative analysis section.

1.4.6.4 Initial size and growth dynamics

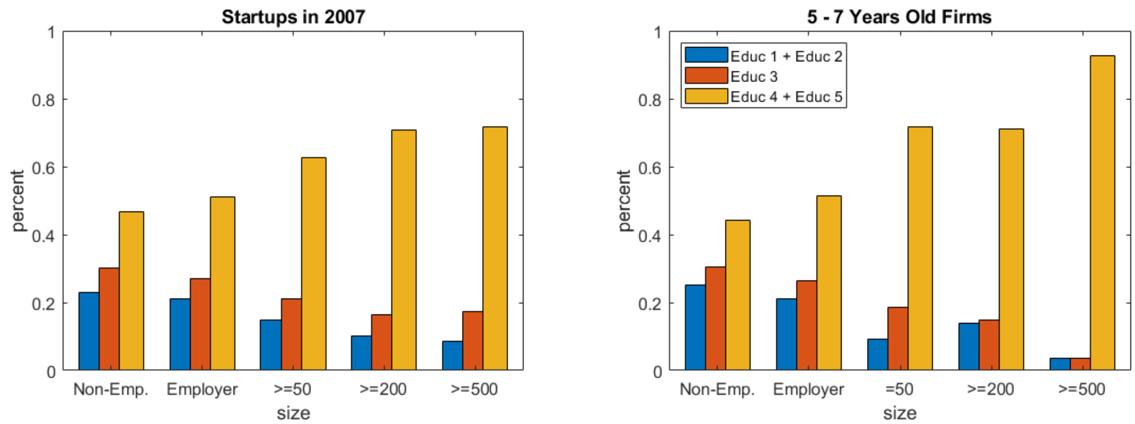
Now, we turn to the question of how the labor force status and educational attainment are related to the growth potential of startups.

Regarding the educational attainment of business owners, the previous analysis showed that an imperfect probability of finding a job produces a flatter division line in the decision rule space of matched workers, discouraging disproportionately the entry of high-skill workers (see Figure 1.6). This implies that high-skill workers only decide to start a business when the entrepreneurial idea is good enough to compensate for both the forsaken wage from quitting the current job and the potential wage loss in the event of a business failure, which depends on how much time will be needed before finding a new job. Then, high-skill workers start businesses with a higher average initial productivity z, which means a larger average initial size of startups. Regarding their growth potential, the higher threshold for z also implies that business owned by high skill workers will never become smaller because they will decide to exit before that. In addition, the permanent heterogeneity arising from the positive correlation between average productivity and labor skills ($\mu_{zh} > 0$) makes businesses owned by high-skill individuals to grow larger. Therefore, we should expect high-skill workers to start businesses with a larger average initial size and a higher potential to grow.

Prediction 3: *High-skill workers start businesses with a larger average initial size and higher potential to grow.*

Using data for the universe of nonemployer and employer firms in the U.S. from the Survey of Business Owners 2007 (SBO 2007), Figure 1.9 shows the firm size distribution by the founders' education level. The left panel shows the initial size distribution of startups in 2007, and the right panel shows the size distribution of 5-7-years-old firms in 2007. We see that highly educated individuals start most of the startups with a large average initial size. For 5-7-year-old firms, we see that the differences in size found at the year of birth persist, with most large firms being owned by highly educated owners. This evidence is, in principle, consistent with prediction 3. However, in the model, if there is no permanent heterogeneity in the productivity z across labor skills, we would see just a few high skilled owners surviving in the long run. This might impede the model to reproduce the fact that high-skill individuals own most large businesses in the long run. The positive correlation between labor skills h and entrepreneurial ability z helps the model to capture this empirical regularity. A formal empirical analysis for prediction 3 will be performed in the quantitative section.

Regarding the differences in business performance between those started by previously employed and unemployed workers, the publicly available data doesn't provide the necessary information to test this prediction. However, previous works have found empirical support for this. In particular, Galindo Da Fonseca (2019), using Canadian administrative data, shows that unemployed workers are more likely to become self-employed than wage workers, but they start smaller firms that are



Notes: The sample only includes firms that are owned by the founder. "Educ 1 + Educ 2" is high-school or less (left bin). "Educ 3" is incomplete college (middle bin). "Educ 4 + Educ 5" is complete college or graduate studies (right bin). Author calculations with the Survey of Business Owners 2007 (SBO PUMS).

Figure 1.9: Firm size distribution by education of business founders

more likely to exit.

1.4.6.5 Putting everything together

The first two model-implied predictions presented above give rise to a selection at entry mechanism that shapes the entry and composition of business founders over the cycle consistently with the suggestive evidence presented in Section 1.3. Through the "labor force composition" and the "labor market tightness" channels, the labor market dynamics should generate a decline in the entry of employer businesses in downturns and a shift in the composition of startups toward more businesses started from unemployment and fewer high-skill workers.

The missing generation of startups should account for a large part of the decline in the aggregate job creation during downturns, but also there should be an effect arising from the change in the composition of business founders. In particular, if the cohort of new businesses contains fewer high-skill founders, we should see a slower growth and a weaker contribution of new and young businesses to the recovery of the aggregate employment creation. The next section will address this using the calibrated model to take the general equilibrium effects into account in the analysis.

1.5 Quantitative Analysis

This section uses the proposed framework to study how the labor market dynamics affect the entry decisions and the composition of business founders over the cycle and quantify how the entry and compositional dynamics shape the recovery of aggregate job creation in the aftermath of an economic downturn.

The analysis is divided into three parts. First, I perform an empirical analysis using individual- and firm-level data for the U.S. to test the three model-implied predictions presented in Section 1.4. Then, in the second part, I present the calibration and stationary equilibrium results. Finally, in the third part, I performed two perfect foresight transition dynamics exercises. First, I feed the model with an exogenous aggregate productivity sequence that triggers a path of the unemployment rate that mimics the one exhibited by the U.S. during and after the Great Recession. This exercise aims to assess the ability of the model to reproduce the labor market and firm dynamics observed in the data, and quantify the effect of the labor market dynamics on the entry, composition of business founders, and aggregate job creation. Then, to understand the channels by which the labor market dynamics affect the entry and composition of business founders, I compute the impulse response functions for a one-time unexpected negative shock to aggregate productivity under perfect foresight, and I perform a counterfactual exercise keeping fixed the value of the fallback option in the event of a business failure.

1.5.1 Empirical Analysis

This section provides empirical support for the three model-implied predictions derived in Section 1.4. For the analysis, I use individual-level data from the Survey of Income and Participation Program (SIPP) and firm-level data from the Survey of Business Owners 2007 (SBO).

1.5.1.1 Survey of Income and Participation Program

The SIPP provides a continuous series of national panels, with sample size ranging from approximately 14,000 to 52,000 interviewed households. I use the panels beginning in 1996, 2001, 2004, and 2008, providing monthly data between 1996 and 2013. The SIPP has a very rich and complex structure that includes data about occupation, education, earnings, demographics, assets, labor market history, businesses, and family, among other variables. With the labor market history, we can keep track of the whole path of occupations for all individuals in the sample. When individuals own a business, we know its legal form of organization (incorporated/unincorporated). The SIPP follows up to two wage jobs and two businesses simultaneously, with starting and ending dates for each spell, and provides the hours worked at each job or business. This allows me to account for multiple jobs and businesses, making possible a more rigorous identification of the main occupation, which is key for identifying transitions between labor force status.

For the analysis, I construct monthly transition among four labor market states: employment (E), unemployment (U), subsistence self-employment (S), and entrepreneurship (F). To take the model to the data I have to proxy S and F. From the CPS, we know that around 40-50% of incorporated businesses are employer firms, while only 10-15% of unincorporated firms have one or more employees. Consistent with this fact, and following the identification assumption from Levine and Rubinstein (2018), I approximate S with the universe of unicorporated businesses plus the incorporated businesses with owner working less than 35 hours and F with the universe of incorporated businesses with owners working 35 or more hours.

1.5.1.2 Prediction 1: Transition Probabilities

To test whether the probability of transitioning to self-employment is higher from unemployment than from employment, I estimate the following Multinomial Logit Regression model, following Levine & Rubinstein (2018):

$$Ln(P_{Jit}/P_{Oit}) \quad = \quad \beta_{JO} + \beta_{JOX} X_i + \epsilon_{JOit} \tag{1.13}$$

where the dependent variable $Ln(P_{Jit}/P_{Oit})$ is the log-odds ratio of person i being subsistence self-employed (J=S) or an entrepreneur (J=F) rather than occupation O at time t. Two different models are estimated. First, I estimate the model using only the sample of employed workers at time $t-1$ ($O = E$), and next using only the sample of unemployed workers at $t-1$ ($O=U$). X_i: is a categorical variable for education (five categories).

	Multinomial Logit Model		Linear Probability Model	
Previous Status	S	F	S	F
E	0.142	0.039	0.144	0.038
	(0.002)	(0.001)	(0.002)	(0.001)
U	0.604	0.055	0.543	0.055
	(0.012)	(0.004)	(0.009)	(0.004)
($\tilde{\beta}_1$ - $\tilde{\beta}_2$)	***	***	***	***
(Pseudo) R2	0.8297	0.8297	0.0004	0.0004
N obs	9,696,770	9,696,770	5,262,966	5,262,966
Individual FE	no	no	yes	yes
State and Time FE	yes	yes	yes	yes

($\tilde{\beta}_1$ - $\tilde{\beta}_2$): Difference test. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively. Results are expressed as percentages.

Table 1.1: Margins (Predicted Probabilities)

The results are presented in Table 1.1. The probability of starting a business is higher for unemployed than for employed workers, giving empirical support to prediction 1. This difference is much larger for subsistence self-employment. Figure 1.10 presents the probability of transitioning into S and F by the level of educational attainment. We can see that the probability of starting a business as "S" for employed workers is U-shaped. This is consistent with the findings from Poschke (2013), who using the CPS shows that there is a U-shaped relationship between the probability of entrepreneurship and both a person's schooling and wage when employed. My finding goes a little further, by distinguishing between two types of self-employment, and by showing that only ES transitions are U-shaped, while EF transitions increase monotonically with the level of education. A second interesting

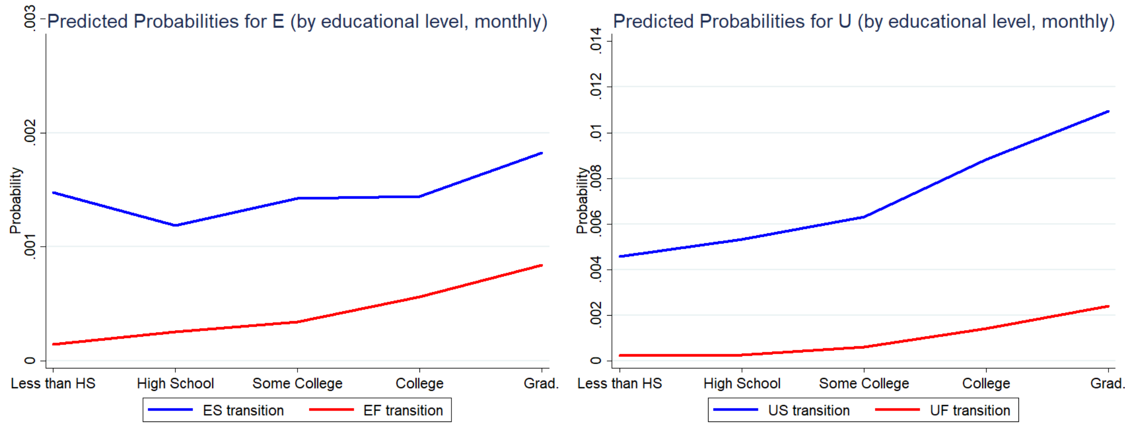


Figure 1.10: Transition Probabilities by Educational Attainment

finding is that the probability of starting a business as "F" is steeper in education for employed than for unemployed workers.

1.5.1.3 Prediction 2: Cyclicality of Transition Rates

To study the dynamic effects of the job finding probability on transition between labor market states, I estimate a set of linear probability models using the job finding probability as explanatory variable. The reduced form is given by:

$$E_{JOist} = \beta_{JO} + \beta_{JOf} f(\theta_t) + \beta_{JOX} X_{it} + \beta_{JOfX} f(\theta_t) * X_{it} + \epsilon_{JOist} \quad (1.14)$$

where E_{JOist} is a binary indicator that equals 1 if person i in the state of residence s is observed transiting from occupation O to occupation J at time t, and 0 otherwise. $f(\theta_t)$ is the probability of finding a job in period t. X_{it} is a categorical variable for education (five categories).

Table 1.2 presents the marginal effects for the job finding probability at monthly

	(1)				(2)			
	ES	US	EF	UF	ES	US	EF	UF
$f(\theta)$	0.2467***	0.7808	0.1449***	0.2782	0.2541***	0.7532	0.1515***	0.2967
	(0.0579)	(0.5076)	(0.0314)	(0.1807)	(0.0583)	(0.5125)	(0.0318)	(0.1840)
Mean	0.1444	0.6993	0.0420	0.0741	0.1444	0.6993	0.0420	0.0741
BC Controls	No	No	No	No	Yes	Yes	Yes	Yes
R-squared (pseudo)	0.0062	0.0105	0.0188	0.0494	0.0063	0.0108	0.0191	0.0494
N obs	3,932,492	248,673	3,932,492	237,562	3,903,661	245,694	3,903,661	234,720

Notes: Results are expressed as percentages. All regressions control for state, and time fixed effects. Model (2) includes change in the state unemployment rate as control. Heteroskedasticity robust standard errors are in parenthesis.

Table 1.2: Transition Probabilities (Marginal Effects)

frequency. The second group of specifications include the change in the state unemployment rate as a control for the aggregate demand conditions. The transition probabilities from employment to both subsistence self-employment and entrepreneurship are positively correlated with the probability of finding a job. To give an economic interpretation to the magnitude of these effects, I include the mean values for each transition. In the case of the transition probability from employment to entrepreneurship, the mean value is 0.042%. If the job finding probability goes up in 0.1, then a coefficient of 0.1449% means that the transition probability from employment to entrepreneurship increases in 0.0145%, which represent an increase of 35% with respect to the mean value.

I also explore whether the sensitivity of the transition rates with respect to the job finding probability depends on the level of human capital. The results are presented in Table 1.3. The probability of transitioning from employment to entrepreneurship is positively correlated with the job finding-rate for all of the five educational levels, but the effect becomes stronger as the level of education increases. This gives empirical support to prediction 2. If we believe that high skill workers

$f(\theta)$	(1) ES	(1) US	(1) EF	(1) UF	(2) ES	(2) US	(2) EF	(2) UF
$EDUC = 1$	0.2708***	0.6109	0.0409	0.0005	0.2663***	0.6356	0.0413	0.0077
	(0.0987)	(0.6053)	(0.0316)	(0.1305)	(0.0992)	(0.6136)	(0.0319)	(0.1307)
$EDUC = 2$	0.2618***	0.3316	0.0585**	0.2821**	0.2726***	0.3572	0.0656**	0.2959**
	(0.0635)	(0.5549)	(0.0306)	(0.1404)	(0.0637)	(0.5622)	(0.0318)	(0.1444)
$EDUC = 3$	0.1698***	1.1477**	0.1200***	0.3796*	0.1757***	1.1127**	0.1250***	0.3955**
	(0.0680)	(0.5608)	(0.0336)	(0.2049)	(0.0684)	(0.5645)	(0.0339)	(0.2083)
$EDUC = 4$	0.2866***	1.1336	0.2491***	0.2474	0.2914***	0.9175	0.2573***	0.2957
	(0.0759)	(0.9226)	(0.0526)	(0.4444)	(0.0766)	(0.9250)	(0.0532)	(0.4545)
$EDUC = 5$	0.3676***	0.9858	0.3571***	0.6856	0.3881***	0.9391	0.3698***	0.6932
	(0.1131)	(1.4388)	(0.0795)	(0.7961)	(0.1140)	(1.4688)	(0.0803)	(0.8072)
Mean	0.1444	0.6993	0.0420	0.0741	0.1444	0.6993	0.0420	0.0741
BC Controls	No	No	No	No	Yes	Yes	Yes	Yes
R-squared (pseudo)	0.0062	0.0105	0.0188	0.0494	0.0063	0.0108	0.0191	0.0494
N obs	3,932,492	248,673	3,932,492	237,562	3,903,661	245,694	3,903,661	234,720

Notes: Results are expressed as percentages. All regressions control for state, and time fixed effects. Model (2) includes change in the state unemployment rate as control. Heteroskedasticity robust standard errors are in parenthesis.

Table 1.3: Transition Rates (Marginal Effects)

are those who in average start business with a higher potential to grow, then this finding support the idea that the cohorts of businesses born in downturns contain fewer potentially high-growth entrepreneurs. The probability of transitioning from unemployment to subsistence self-employment is negatively correlated with the job finding-rate, but not significantly.

1.5.1.4 Prediction 3: Growth Potential of Startups by Founders' Education

Finally, I test empirically whether high-skill workers start businesses with a higher potential to grow (prediction 3). For the analysis, I use the Survey of Business Owners 2007 (SBO PUMS 2007), a representative survey covering the universe of 26 million employer and non-employer businesses in the U.S. The SBO provides information about both firm and business owners' characteristics. In particular, the SBO includes information about age, size, and industry of the firms, and age and level of education of the owners.

Table 1.4 shows the result for a regression analysis using the log of employment of the firm as dependent variable and the level of education of the business owners as explanatory variable. I control for the age of the business owner, and whether the business is owned by the original founder. In terms of the firm, I use age, state and industry fixed-effects. The first two columns show the results for the total universe of businesses and for only employer firms, respectively. The last two columns show the results for the same analysis, but restricting the sample only to the universe of businesses for which the current owner is also the founder. Across all the specifications we see a positive correlation between the firm size and the educational attainment of the business owner, except for individuals with graduate studies. These results are consistent with previous finding in the empirical literature.[20]

1.5.2 Parametrization and Stationary Equilibrium

A first set of parameters is calibrated externally. The model frequency is assumed to be monthly, so β is set to 0.996. The standard deviation of the distribution of labor skills σ_h is set to 0.327, which allows the best possible match with the educational attainment distribution for five categories from the CPS. The separation rate s is set to 0.0149, which is the average monthly rate of employment-to-unemployment flows in the CPS for the period 2005-2006. The workers' Nash bargaining power η is set to 0.24. The literature has used a wide range of values for this parameter, from Shimer (2005) using 0.72 to Hagedorn & Manovskii (2008) using 0.052. Recall

[20] See Brown, Earle, Kim & Lee (2019).

Dependent variable: log(employment)				
Educ	Full Universe	Only Employers	Full Universe	Only Employers
High School	0.0981	0.0001	0.0714***	0.002
	(0.0033)	(0.0065)	(0.0036)	(0.0080)
Incomplete College	0.1370***	0.0545***	0.0849***	0.0181***
	(0.0035)	(0.0068)	(0.0038)	(0.0083)
College	0.2721***	0.2554***	0.1697***	0.1476***
	(0.0034)	(0.0065)	(0.0037)	(0.0080)
Graduate	0.2913***	0.2116***	0.1969***	0.0909***
	(0.0036)	(0.0071)	(0.0039)	(0.0086)
Constant	-0.2766***	1.4635***	-0.2055***	1.4937***
	(0.0091)	(0.0220)	(0.0092)	(0.0269)
R2	0.1830	0.1266	0.1285	0.1077
N obs	2,134,142	1,238,702	1,388,554	689,046
Only founded	no	no	yes	yes
Owner's age ¿=25	yes	yes	yes	yes
Age Fixed-Effects?	yes	yes	yes	yes
State Fixed-Effects?	yes	yes	yes	yes
Industry Fixed-Effects?	yes	yes	yes	yes

Notes: Reference category corresponds to less than high-school.

Table 1.4: Growth Potential of Startups by Founders' Education

that the latter work shows that low workers' bargaining power and a high unemployment benefit generates wage rigidity closer to what we see in the data. With respect to the elasticity of the matching function, I set ψ to 0.2, which is the value from Hall (2005). This parameter corresponds to one minus the elasticity of the job finding probability $f(\theta_h)$ with respect to the labor market tightness θ_h, so a lower value of ψ means a higher elasticity of $f(\theta_h)$ with respect to θ_h. The parameter Y^{ss}, which reflects the unemployment benefit or the value of leisure, is set to 0.4. This is the standard value used in the literature. In the model, $Y^{ss} * w_h^{ss}$ is not the only component of the opportunity cost of being a worker because individuals also have the option to start a business, making business earnings also a fundamental determinant for the worker's reservation utility used in the Nash Bargaining solution.

Parameter	Value	Description	Source
β	0.996	Discount factor	CHW (2007)
σ_h	0.327	Std. dev. h (log distr.)	Educ. distr. (CPS 2016)
s	0.0149	Separation probability (monthly)	EU flows (CPS 2005-2006)
η	0.24	Workers' bargaining power	HW (2008) 0.052 - Shimer (2005) 0.72
ψ	0.24	Matching function elasticity	Hall (2005)
Y^{ss}	0.4	UI / Leisure	Standard
$A^{\star}$	1	Aggregate productivity (SS)	Normalization

Table 1.5: Externally Calibrated Parameters

Finally, the aggregate productivity A is normalized to 1 in equilibrium. Table 1.5 summarizes the seven externally calibrated parameters and the sources.

The remaining 17 parameters are calibrated internally. Table 1.6 presents the results. I discipline the model to match selected key features of the labor market and firm dynamics in the U.S. Even though every targeted moment is determined simultaneously by all parameters, I discuss each of them in relation to the parameter that, intuitively, yields the most identification power. Data moments from CPS are monthly averages over the period 2005-2006 unless otherwise specified. Data moments from BDS are annual average over the period 2001-2007.

The mean of the labor skills distribution (μ_h) determines the size of the effective labor force in terms of efficiency units, so I calibrate it to match the average size of employer businesses. The distribution of labor skills is determined by its lower and upper bounds (h_l, h_h), which I use to set the ratio of wages between high and low skilled workers.

To discipline the flows between labor market states, I use the parameters of the entrepreneurial ability stochastic process, job finding probability, and search efficiency. The span control parameter α pins down the value of owning a business relative to the wages earned as an employed worker, which determines the entrepreneurship rate in the economy. The exit rates primarily respond to the parameters related to entrepreneurial ability, and business owners can exit endogenously either into unemployment or to employment. I set the persistence of the entrepreneurial ability (ρ_z) to discipline the total exit rate from subsistence self-employment and entrepreneurship.[21] Then, I discipline the unemployment rate using the standard deviation of the entrepreneurial ability (σ_z), which pin downs the exit into unemployment. The exit into employment is determined mostly by the search efficiency parameters (φ^S, φ^F). These parameters also pin down the entry rate into both technologies since exit has to be equal to entry in the stationary equilibrium. The parameter (A_{sub}) is set to discipline the subsistence self-employment rate. The cost of posting vacancies for each labor market segment (c_h) is set to match an equal job finding probability across all the labor market segments, which determines the flows from unemployment into employment and the probabilities of receiving job offers for business owners. The flows from employment into unemployment are pinned down by the exogenous separation rate s. This calibration strategy disciplines completely the flows between labor market states and the labor force composition: employment rate, unemployment rate, subsistence self-employment, and entrepreneurship.

To discipline the firm dynamics structure in the model, I use the fixed operational cost (ϕ), the Pareto exponent (ξ), the standard deviation of the initial

[21]If the draw of z is bad enough, a business owner will decide to close the business to become an unemployed worker ($z < \underline{z}$) If a business owner receives a job offer, and the realization of z makes owning a business a better option than unemployment but worse than taking the job ($\underline{z} < z < \overline{z}$), then the owner accept the offer and become an employed worker. The probability of receiving a job offer depends on the search efficiency of each technology.

productivity (σ_s) to match the 5-years survival rate of employer businesses, the 1-year survival probability of entrepreneurs, and the relative size of entrants with respect to incumbent firms. In the model, employer businesses are defined as those hiring more than 1.5 units of the labor efficiency good, and the remaining businesses are classified as non-employer firms. The correlation between the mean of the entrepreneurial ability and labor skills (μ_{zh}) is used to introduced permanent heterogeneity in the size of the businesses to match the distribution by education from the SBO, where more educated individuals own larger firms on average. A second source of permanent heterogeneity is introduced with the span control parameter of the subsistence technology (α_{sub}), which allows the model to account better for the skewed firm size distribution by chosen it to match the share of total employment in firms using less than 50 units of labor.[22] The cost shifter (κ) and the elasticity (γ) from the convex hiring cost function help to discipline the share of total employment of young firms (1-5 years old) and the employment share of firms hiring more than 500 units of labor.

1.5.3 Cross-Sectional Implications

Now, I present the main cross-sectional implications of the calibrated model at the stationary equilibrium.

[22] Gavazza, Mongey & Violante (2018) also use a similar kind of permanent heterogeneity. They argue that it helps to decouple age and size, which tend to be too strongly correlated in standard firm dynamics model with mean reverting productivity.

Parameter	Value	Description	Target	Source	Data	Model
Panel A: Labor market dynamics						
α	0.75	Span control parameter	Entrepreneurship rate	CPS	3.8%	3.6%
A_{sub}	0.48	Sub. tech. scale param.	Subsistence SE rate	CPS	6.4%	6.4%
σ_z	0.05	Standard deviation of z (log distr.)	Unemployment Rate	CPS	4.4%	4.4%
ρ_z	0.96	Persistence of z	Entrepreneurship exit rate	CPS	5.1%	3.1%
φ^S	0.45	Search efficiency sub. tech.	SE flows	CPS	6.2%	6.5%
φ^F	0.21	Search efficiency entrepreneurs	FE flows	CPS	4.2%	3.9%
m	0.496	Matching efficiency	Labor Market Tightness (θ_h)	HM 2011	0.63	0.63
c_h	Distr.	Cost of vacancies	Job Finding Rate	CPS	0.28	0.28
(h_l, h_h)	(2, 6.3)	Bounds for labor skills h	$w_{h=5}/w_{h=1}$	CPS	2.82	2.40
Panel B: Firm dynamics						
ϕ	2.50	Fixed operational cost	Employer businesses survival rate at age 5	BDS	0.49	0.41
ξ	3.2	Pareto exponent	Relative size of employer entrants	BDS	0.51	0.58
μ_{zh}	distr.	z and h dependence	Avg. size by educ of founder	SBO 2007	distr.	distr.
σ_s	0.10	Std. Dev. initial productivity	Entrepreneurs survival rate at age 1	SIPP	0.68	0.59
κ	1.5	Hiring cost scale param.	Cohort employment share at age 1-5	BDS	13.2%	20.3%
γ	2.0	Convexity of hiring costs	Employment share $n > 500$	BDS	0.47	0.40
α_{sub}	0.10	Sub. tech. span control param.	Employment share $n < 50$	BDS	0.31	0.35
μ_h	3.80	Mean of h (Labor skills)	Average employer firm size	BDS	23	22.6

Panel A: CPS data corresponds to monthly averages for period 2005-2006, except for mean and bounds of labor skills that use data from 2017.
Panel B: BDS data correspond to annual averages over the period 2001-2007. SBO data correspond to SBO PUMS 2007. SIPP data corresponds to monthly averages over period 1996-2013.

Table 1.6: Internally Calibrated Parameters

1.5.3.1 Labor Skills and Occupational Distributions

The labor force and educational composition shape the effects of the "labor force composition" and "labor market tightness" channels in the model. Therefore, the model needs to fit the composition of individuals in terms of these two characteristics in a good way in order to quantify the effect of the labor market dynamics on the entry and composition of business founders properly.

The results in Table 1.6 show that the model presents a good fit in terms of the labor force composition. In particular, the unemployment rate in the model, which drives the "labor force composition" channel, matches the target of 4.4%.

The distribution of labor skills pins down the wage distribution in the model. Figure 1.11 presents the comparison between the model and the data. The good fit of the model matching this moment is also an important feature. It determines the

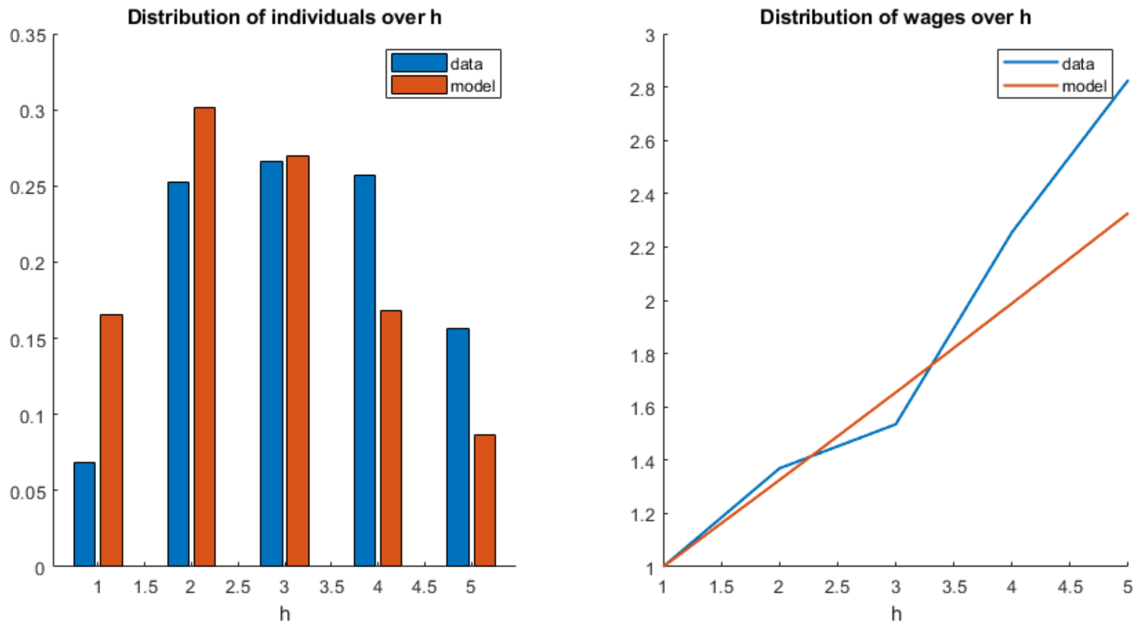


Figure 1.11: Labor skills distribution and equilibrium wages

distribution of outside options for potential business owners, which generates the selection at entry mechanism by education. This feature is key for the effects of the "labor market tightness" channel on entry decisions.

1.5.3.2 Type of business and composition by founders' characteristics

Table 1.7 reports some empirical compositional moments and their model-generated counterparts. The composition by type of business, which is a targeted moment, replicates the fact that subsistence self-employment is almost twice the entrepreneurship rate in the data (proxied by unincorporated and incorporated businesses, respectively). This difference increases when we look just at startups, which is a non-targeted moment that the model captures well.

The composition of entrants in terms of the previous labor force status is also

	CPS		Model	
	Sub. Tech.	Entrep.	Sub. Tech.	Entrep.
TYPE OF BUSINESS				
All Businesses				
Share as % of labor force	6.4%	3.6%	6.4%	3.6%
Share as % of all businesses	69.6%	31.4%	69.6%	31.4%
Startups				
Share as % of labor force	0.42%	0.04%	0.50%	0.04%
Share as % of entry	91.3%	8.7%	80.5%	19.5%
COMPOSITION BY PREVIOUS LABOR FORCE STATUS				
Startups				
E	76.5%	86.4%	73.6%	86.6%
U	23.5%	13.6%	26.4%	13.4%
COMPOSITION BY HUMAN CAPITAL				
All Businesses				
<= High School	42.2%	27.5%	36.9%	29.1%
Incomplete College	27.4%	26.4%	34.2%	31.1%
>= College	30.4%	46.1%	29.0%	39.8%
Startups				
<= High School	47.0%	31.1%	44.0%	28.0%
Incomplete College	26.8%	25.1%	30.3%	33.0%
>= College	26.2%	43.8%	25.7%	39.0%

Notes: All values from CPS are weighted. Basic Monthly CPS 2005-2006.

Table 1.7: Compositional distribution by type of business and characteristics of business founders

well replicated by the model, which is not explicitly targeted. Subsistence self-employment exhibits a larger share of entry from unemployment than entrepreneurship. The model can also satisfactorily replicate the observed educational composition of business owners in the CPS, for both existing and new businesses. We can see that the entrepreneurial businesses are formed by a larger share of highly educated owners than subsistence self-employment, with a similar pattern for both existing and new businesses. The model generates these features due to the selection at entry mechanism that makes individuals with different characteristics to choose different kinds of businesses.

Shares (in %)	BDS and NES Data		Model
	Number	Share	Share
Non-employers	21,708,021	78.2	69.9
Employer firms	6,049,655	21.8	30.1

Notes: NES and BDS data in 2007.

Table 1.8: Non-employers and employer businesses

1.5.3.3 Firm age and size distribution

Table 1.8 shows the composition of businesses in terms of non-employer and employer businesses. We can see that the model replicates well the larger share of non-employer businesses observed in the data, which is a non-targeted moment. All the non-employer businesses in the model are subsistence self-employed, while employer businesses are both subsistence self-employed and entrepreneurs.

Table 1.9 presents the age and size distribution of employer businesses only. The contribution of entrants to aggregate job creations is somewhat overestimated in the model with respect to the BDS data. The entry rate of employer businesses in the model is mainly driven by the entry rate of entrepreneurs, which is pinned down by the exit rate in the stationary equilibrium. So, the higher entry and exit rates for incorporated businesses in the CPS than in the BDS are passed on into the model. However, the lower 1-year survival rate of entrepreneurial startups helps to correct the problem, although not completely.

Shares (in %)	New	Young	Old	1-49	50-499	500+
BDS DATA						
Employer firms	10.6	31.7	57.7	95.6	4.0	0.4
Employment	3.0	13.2	83.8	31.0	21.5	47.5
MODEL						
Employer firms	16.2	38.6	45.2	96.2	3.5	0.3
Employment	6.2	20.3	73.5	36.2	23.5	40.3

New firms: less than 1 year old; Young firms: 1-5 years old.

Table 1.9: Firm age and size distribution

1.5.4 Great Recession: Entry, Composition, and Job Creation Dynamics

This section presents a perfect foresight transition dynamics exercise aimed, first, to validate the aggregate dynamics of the model, and, second, to quantify the effect of the labor market dynamics on entry decision, the composition of business founders, and the recovery of aggregate job creation during and after the Great Recession.

The exercise is as follows. The economy is assumed to be at the stationary equilibrium at $t = 1$. Then, in $t = 2$, the model is fed with an unexpected exogenous aggregate productivity sequence that triggers a path of the unemployment rate that mimics the one exhibited by the U.S. over January 2008 - December 2017. From $t = 3$ onwards, individuals have perfect foresight about the future path of aggregate productivity and, therefore, about the future path of the distribution of individuals, the price of the labor good, wages, and the job finding probability.

For the transition, I impose ad hoc wage rigidity as follows:

$$w_{h,t} \;=\; \gamma w_h^{Nash} + (1-\gamma) w_{h,t}^{Nash}$$

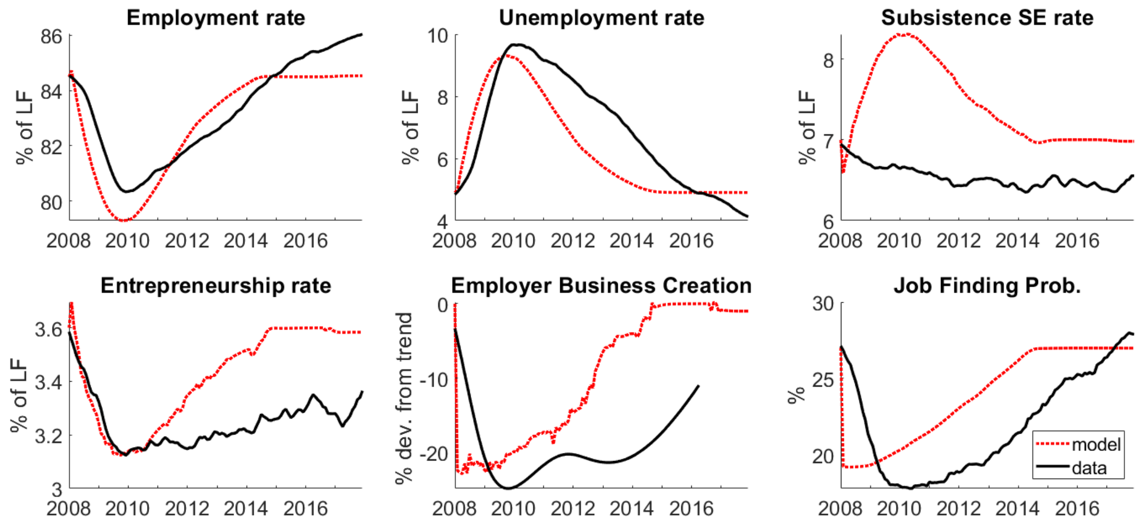
Here, w_h^{Nash} corresponds to the steady state wages. From Hagedorn & Manovskii (2008), the elasticity of wages with respect to aggregate productivity is 0.449. I discipline the value of gamma in the model to match this elasticity, which implies $\gamma = 0.93$. This feature is needed because though wages are set by Nash bargaining, their dynamics are much more volatile than what we see in the data.[23]

1.5.4.1 Aggregate Dynamics

Figure 1.12 presents the empirical and model-implied dynamics for the components of the labor force, the job finding probability, and employer business entry, which are not explicitly targeted in the transition except for the unemployment rate path. The data for employer business creation corresponds to the BDS, and the empirical counterpart for all the other variables is constructed using CPS data.

The model replicates well the persistent decline of the job finding probability during the Great Recession. The model is also able to satisfactorily replicate the procyclical trajectories of the employment and entrepreneurship rates. The decrease in the employment rate from $t = 3$ onwards is explained by the lower job finding rate that reduces the flows from unemployment to employment. This also explains the increase in the unemployment rate from 5% to 9%. This increase in the unemploy-

[23] The excessive volatility of wages in models with labor market frictions is known in the literature as the "Shimer puzzle".



Notes: Actual and model-predicted time-paths. Labor market data corresponds to CPS and employer business creation to BDS. BDS data is transformed into monthly frequency by using cubic interpolation.

Figure 1.12: Actual and model-predicted time-paths

ment rate would have been even higher in the model if the flows from unemployment to subsistence self-employment did not increase during this period. The decline in the entrepreneurship rate is driven by the lower entry from employment and fewer people choosing to stay as entrepreneurs. The model predicts a contracyclical behavior for subsistence self-employment driven by the increase in the transitions from unemployment into subsistence self-employment, which is not observed in the data. The analysis performed with CPS data in Section 1.3 showed that during 2009 and 2010 the U.S. economy actually exhibited a higher entry into self-employment as the model predicts. However, in the data, we also observe an increase in the exit rate that is not matched by the model. Such difference makes the subsistence self-employment rate constant in the data but increasing in the model. Appendix A.6 includes the complete set of 16 transition rate paths.

Finally, employer business creation is countercyclical consistently with its empirical counterpart from the BDS. Even though the model is able to increase persistence, the decline of employer business entry is still more persistent in the data than in the model. I will explain further how the "labor force composition" and the "labor market tightness" channels shape this decline in Section 1.5.5.

1.5.4.2 Entry of Business Founders

Figure 1.13 presents the distribution of entrants over their initial productivity for two periods in the transition, December 2007 and December 2009. The first row distinguishes between subsistence self-employed and entrepreneurs, and the second row between nonemployer and employer businesses.

There are two things to note here. First, the similarity between the panels in the first and second rows is because subsistence self-employed start businesses mostly as nonemployers, while all the entrepreneurs start businesses as employers. Second, entry into subsistence self-employment (new nonemployer firms) increases in recessions, mostly driven by low-quality business founders, while entry into entrepreneurship (new employers firms) decreases, especially for high-quality business founders. These dynamics are generated by the higher entry of low-skilled business founders from unemployment and lower entry of highly educated business founders from employment. Next, I turn the analysis to examine the entry composition in terms of these two business founders' characteristics.

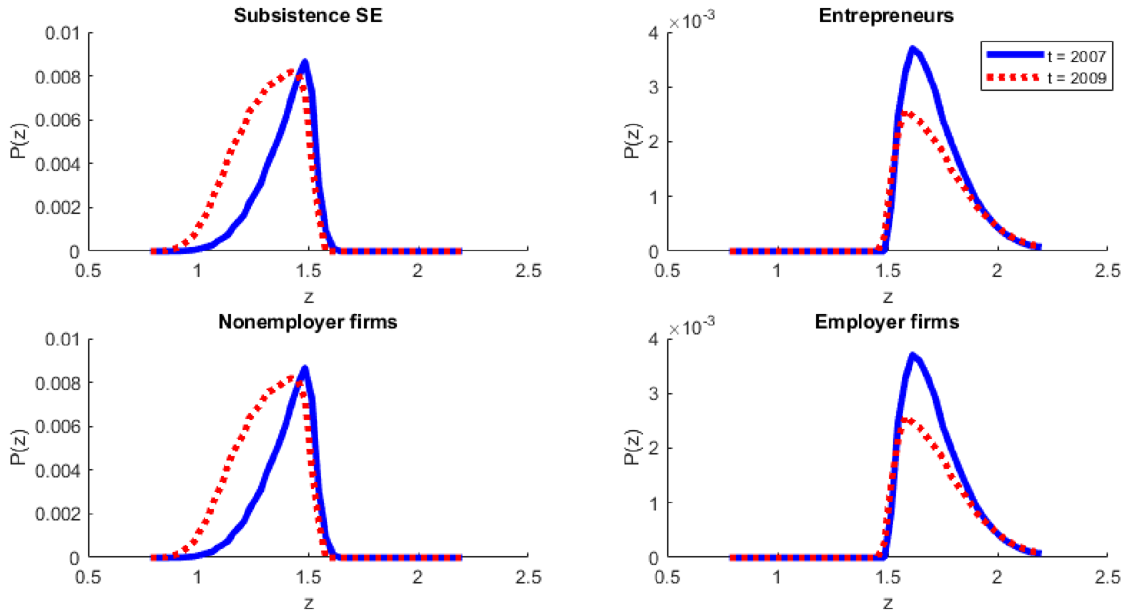


Figure 1.13: Initial productivity distribution

1.5.4.3 Composition of business founders

Figure 1.14 presents the composition of employer startups in terms of the previous labor force status and educational level of the founders in the model. The jump in the unemployment rate in 2009 and 2010 is passed on into the composition of business founders through the "labor force composition" channel, and the decline in the job finding probability through the "labor market tightness" channel. Both channels increase the entry from unemployment, making the share of new employer businesses starting from unemployment to increase from 16% to almost 30%. The decline in the job finding probability through the "labor market tightness" channel reduces the share of high-skill founders in the composition of employer startups from 42% to 31%.

Therefore, the composition of founders of employer firms shifts toward fewer

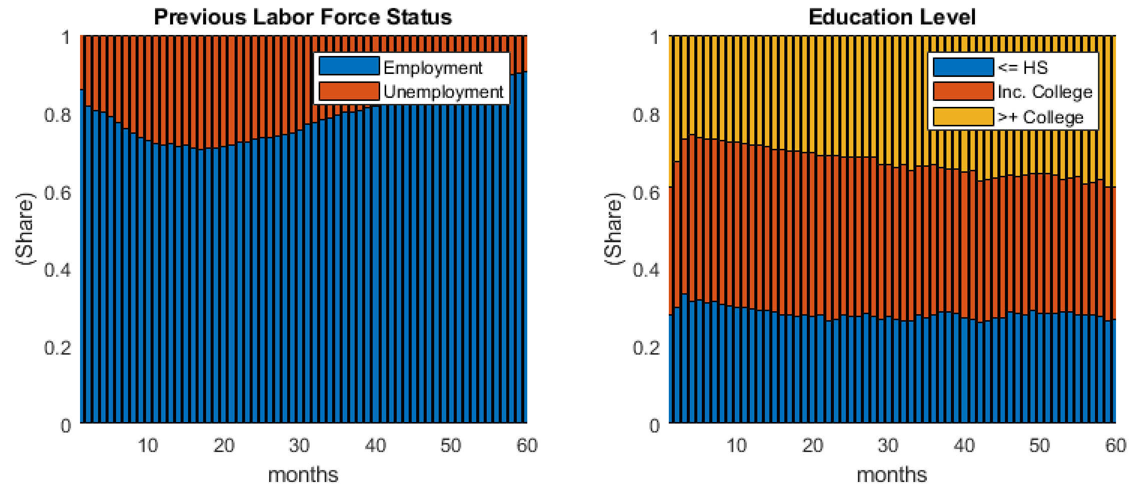


Figure 1.14: Composition of employer startups by founders' characteristics

highly educated individuals and more people coming from unemployment.

1.5.4.4 Initial productivity and growth potential of entrants

Figure 1.15 presents the model-predicted initial productivity distribution of employer startups by the education level of the founders for two periods in the transition, the initial stationary equilibrium in December 2007 and 24 periods later in December 2009.

First, we can observe that the right side of the distribution shrinks disproportionately for highly educated individuals. Second, the mass of entry falls for every level of education, but the decline increases as we increase the level of education of the founders. Because initial productivity pins down the initial average size of startups, the first result implies that new businesses will start relatively smaller, especially those owned by highly skilled individuals. The second result implies that there will be fewer high skilled individuals deciding to start businesses, which make

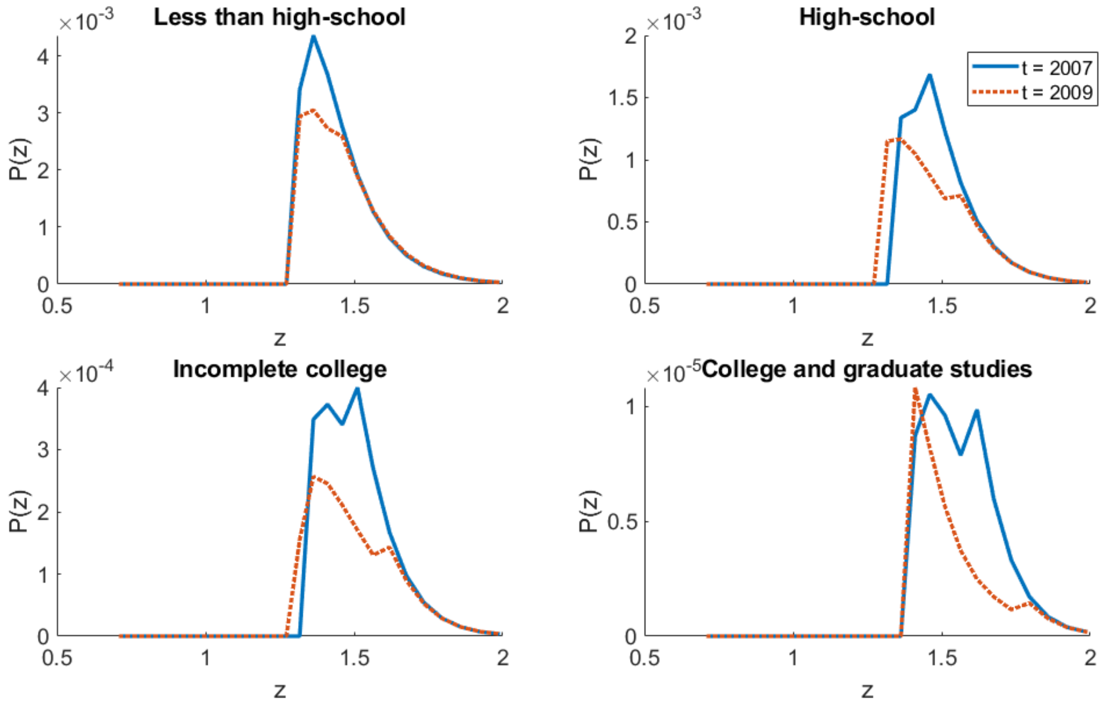


Figure 1.15: Distribution of entrants over initial productivity

the cohort to contain fewer high-growth entrepreneurs.

Therefore, employer businesses born in periods with a high unemployment rate and low job finding probability will start smaller and with a lower potential to grow because of the fewer highly educated individuals starting businesses. These two features will reduce the contribution of entrants and young firms to the recovery of the aggregate job creation in the aftermath of an economic crisis.

1.5.4.5 Job Creation Dynamics

In Figure 1.16, the left panel shows the model-predicted job creation and its empirical counterpart from the BDS. The model replicates well the path of the job creation, with a faster fall at the beginning of the transition, consistent with the sharp decline in the job finding probability. The magnitude and persistence of the fall in the aggregate job creation respond to the lower labor demand by existing firms, decline in entry, smaller initial size of startups, and their lower potential to grow.

The right panel presents a counterfactual exercise to quantify the role of the compositional change in the recovery of job creation. First, I fix the educational composition of entrants, and then, on top of that, I fix the previous labor force status composition. The results show that by fixing the composition of entrants, the initial fall is somewhat reduced because now the initial size of startups doesn't decline, but this improvement is small. So, the contribution of the change in founders' composition to the decline in the aggregate job creation is minimal at the beginning of the recession. However, over time, the compositional shift starts to matter because young businesses grow less, so their contribution to job creation is smaller. Therefore, slower growth of young businesses in the aftermath of the recession prevents aggregate job creation from a faster recovery.

Regarding the relative contribution between previous labor force status and educational attainment composition to the slower recovery of job creation, we can see that the educational composition change accounts for most of the effect. The effect arising from the shift in the previous labor force status, which corresponds to the "labor force composition" channel, is small. In other words, the decline in the growth potential of startups is primarily explained by the shift in the educational composition of business founders toward fewer highly educated individuals.

We can summarize the findings of this section as follows: (i) entry due to

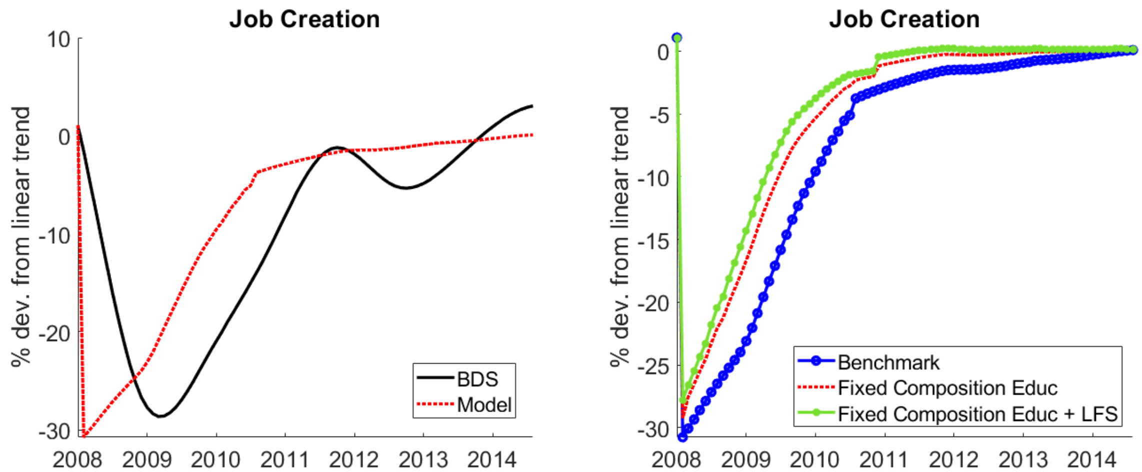


Figure 1.16: Aggregate Job Creation

stopgap motives into subsistence self-employment increases during downturns while entry into entrepreneurship declines, and (ii) composition of entrepreneurs shifts toward fewer highly educated individuals, making the new cohorts to contain fewer potentially high-growth startups. Both features hinder job creation recovery, keeping the labor market depressed longer, and the entrepreneurship entry persistently low.

1.5.5 Understanding the Mechanisms

The previous analysis showed that the "labor market tightness" channel accounts for most of the slower recovery of aggregate job creation due to its negative effect on startups' growth potential. However, the total effect depends on both the growth potential and the number of new businesses started. This section investigates further the role of the "labor force composition" and the "labor market tightness" channels in driving the aggregate dynamics.

1.5.5.1 "Labor Force Composition" channel

Even though the "labor force composition" channel doesn't play an important role in shaping the growth potential of startups, it is an important driver for entry into subsistence technology. Figure 1.17, left panel, shows the transitions from unemployment to subsistence self-employment for the transitional dynamics exercise from Section 1.5.4. It also includes a counterfactual exercise, in which additional flows from unemployment into subsistence self-employment are not allowed throughout the entire transition. In the model, the subsistence technology absorbs a significant fraction of the increase in unemployment during recessions. The right panel shows the unemployment rate for both cases. The difference between the benchmark and the counterfactual exercise shows that the unemployment rate would have increased by 1.5 p.p. more at the peak of the Great Recession without the subsistence alternative.

In the data, the entry into subsistence self-employment (proxied as unincorporated businesses) increases during recessions but the exit rate also follows a similar pattern, making the subsistence self-employment rate to follow a non-increasing path. This suggests that unemployed individuals use self-employment as a stopgap activity for short periods of time.

1.5.5.2 "Labor Market Tightness" Channel

To disentangle the role of the "labor market tightness" channel from the "profitability" channel, I perform a counterfactual exercise in which the "labor market

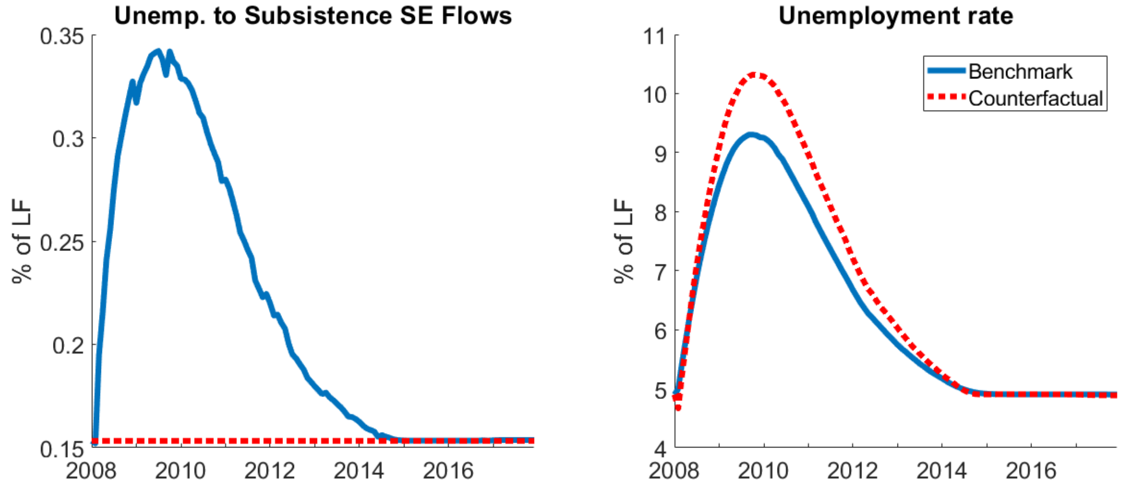


Figure 1.17: Transitions from Unemp. to Subsistence SE and Unemployment rate

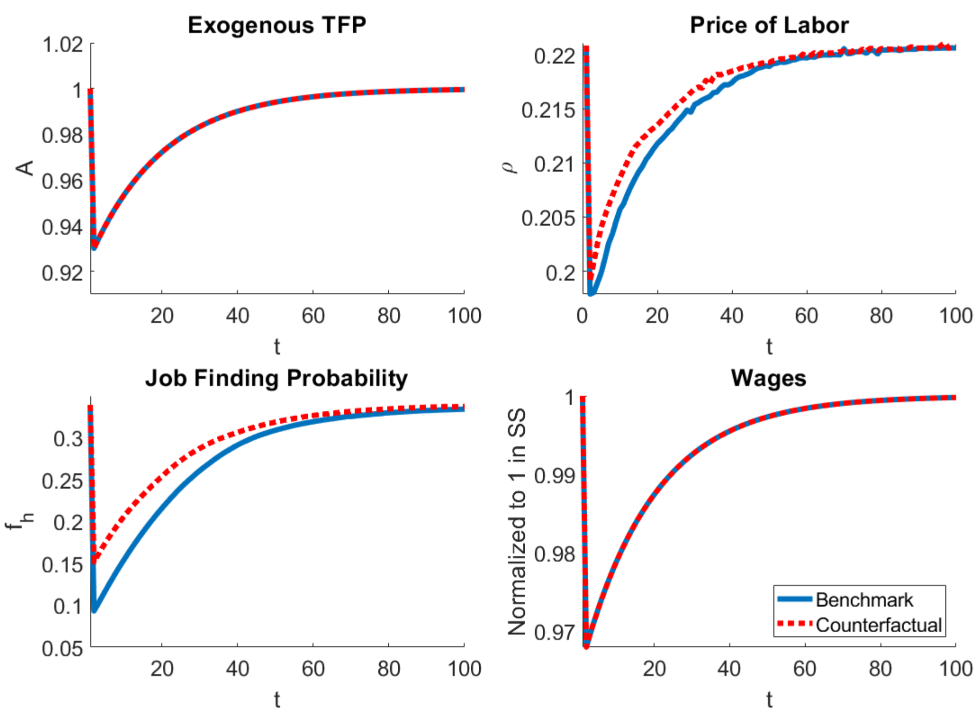


Figure 1.18: Aggregate productivity and job finding probability

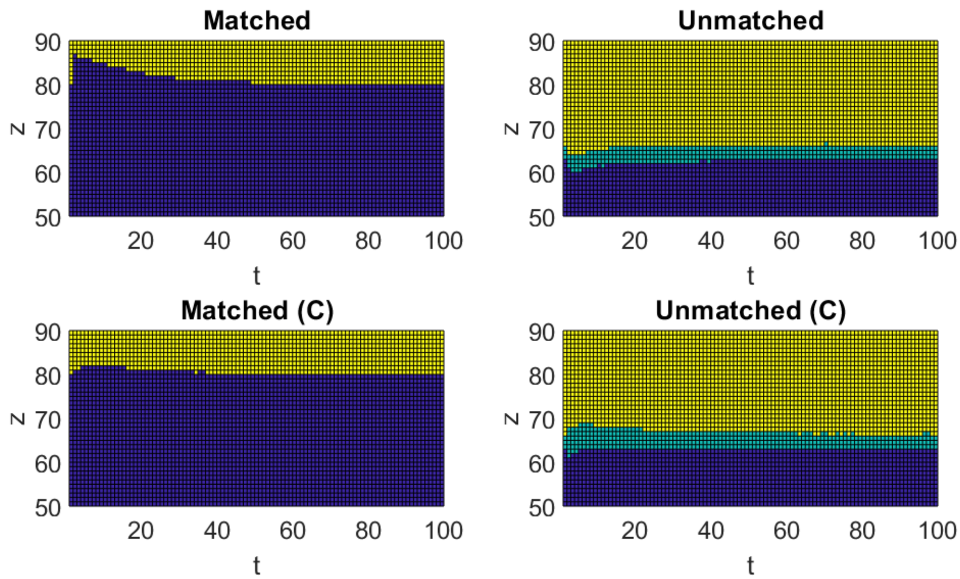
tightness" channel is muted. To do so, I keep fixed the job finding probability at the steady state level that individuals see for their occupational choice problem throughout the transition. By doing this, I keep constant the value of the fallback option in the event of business failure for the potential entrants, canceling the "fear to fail" effect.

The transition dynamics for this exercise is as follows. The economy is assumed to be at the stationary equilibrium at $t = 1$. Then, at $t = 2$ the economy is shocked by an unexpected 7% decrease in the aggregate productivity. After that, individuals have perfect foresight about the future path of aggregate productivity and, therefore, about the future path of the distribution of individuals, the price of the labor good, wages, and the job finding probability. The same kind of wage rigidity from the first exercise is used here. After the one-time initial shock, aggregate productivity follows a mean reverting AR(1) process, given by:

$$A' = (1 - \rho_A) * A^\star + \rho_A A \tag{1.15}$$

for which I use $\rho_A = 0.95$.

Figure 1.18 shows the exogenous path of aggregate productivity used to perform the exercise, price of labor efficiency units, wages, and job finding probability. Wages follow the path of the exogenous TFP because of the exogenous wage rigidity imposed throughout the transition. The price of labor efficiency units falls slightly less in the counterfactual because of the smaller fall in the labor demand. However, its path is also rigid because of the exogenous wage rigidity. The job finding prob-



Panel(a): blue color corresponds to the option of staying as an employed worker and yellow color to the choice of being an entrepreneur. Panel (b): blue color corresponds to remaining unemployed, green to being subsistence self-employed and the yellow to the choice of being an entrepreneur.

Figure 1.19: Paths of Decision Rules

ability falls less in the counterfactual. This happens because, without the "fear to fail" effect, the entry decisions decline less, especially for highly educated individuals, making the number of new businesses to fall by less and start bigger. Moreover, because now the cohort of new businesses contains more highly educated workers, it grows faster, fostering the job creation recovery. I will explain each one of these steps next.

Figure 1.19 presents the occupational decision rules paths for both a matched and unmatched highly educated individuals. The first row corresponds to the benchmark transition, and the second row to the counterfactual. In the benchmark transition, we see that the cutoff $\bar{z}$ above which a highly educated matched individual decides to be an entrepreneur moves up, discouraging potential entrants with good entrepreneurial ideas from starting businesses. In the counterfactual, in the absence of the "fear to fail" effect, this shift is smaller. The remaining effect corresponds to the "profitability" channel. If we perform the same counterfactual analysis for an individual with lower educational attainment, the "profitability" effect would be larger because the "fear to fail" for low-skill workers is smaller.

Figure 1.20 presents the results for both the benchmark (upper panels) and the counterfactual (lower panels) exercises in terms of the initial productivity distribution of new businesses, distinguishing between the subsistence self-employment and entrepreneurship. In the counterfactual scenario, the "profitability" channel generates a rightward shift in the left part of the quality distribution of new entrepreneurial businesses, with almost no change in the right side, difference that corresponds to the "fear to fail" effect. This means that the "fear to fail effect" makes the cohort of startups to start with a smaller average initial size and a lower potential to grow.

In Figure 1.21, the left panel presents the counterfactual analysis for the entry rate into entrepreneurship, which can also be thought of as the entry rate of employer businesses. The "labor market tightness" channel account almost for one-third of the total decline in the entry rate. The right panel presents the results for the aggregate job creation. At the beginning of the transition, the "labor market tightness" channel has a very small effect, but its role becomes more important as time goes by. In the model, as in the data, most of the contribution of startups to the aggregate job creation comes from their growth, so the missing generation of new business doesn't diminish the aggregate job creation contemporaneously as in the

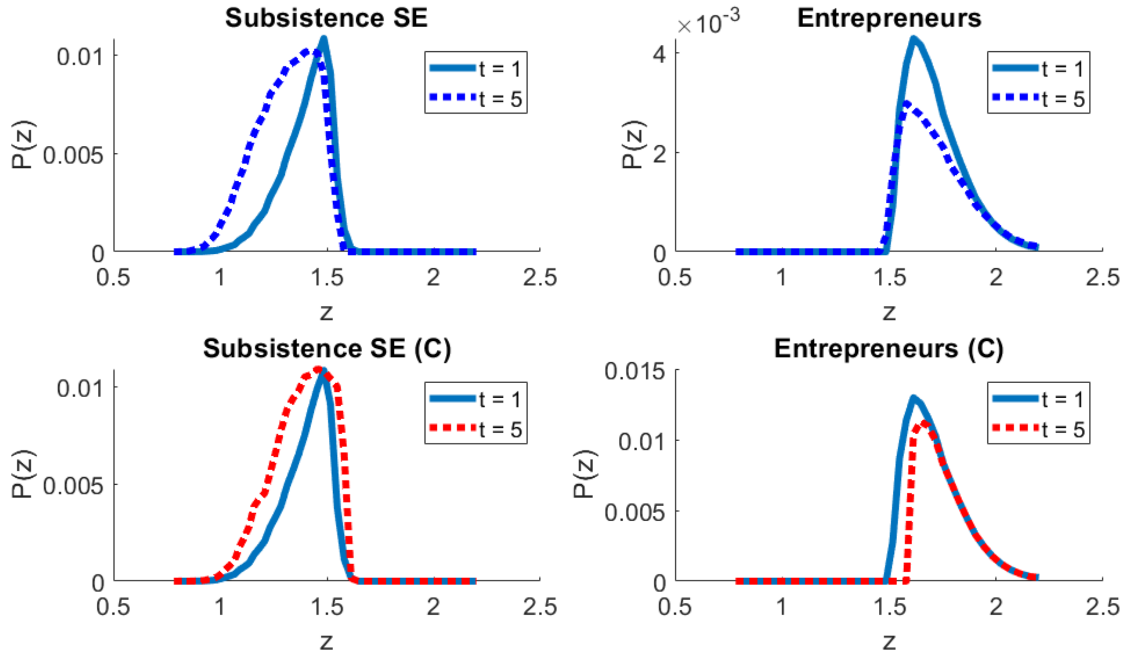


Figure 1.20: Initial productivity distribution of new businesses

following periods. In the model, the negative effect of lower entry is reinforced by the composition change of business founders toward fewer highly educated individuals, which reduces the entry of high-growth startups disproportionately. Therefore, both the decline in entry and the change in composition hinder aggregate job creation recovery in the aftermath of an economic crisis.

Finally, Figure 1.22 shows the paths of the labor force components. Because fewer individuals are discouraged from starting businesses, we see a smaller fall in the entrepreneurship rate. The smaller decline in entry plus the bigger size of startups helps to avoid a further decline in the job finding probability, making the unemployment rate jump less. Then, because the unemployment rate and the job funding probability respond less, the subsistence self-employment also increase slightly less.

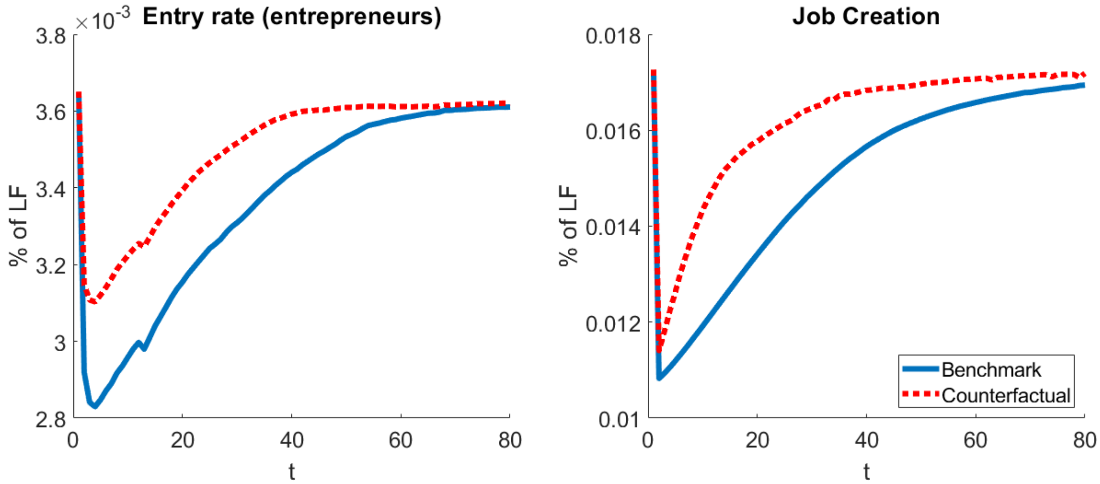


Figure 1.21: Entry rate of entrepreneurs and job creation

The entrepreneurship rate, measured as a share of the labor force, falls from 0.034 to 0.028 in the benchmark model, and from 0.034 to 0.030 in the counterfactual. This means that the "labor market tightness" channel (and "fear to fail" effect) contributes to a 33% to the total decline in the entrepreneurship rate. Regarding the unemployment rate, in the benchmark model, it jumps from 5% to 8% while in the counterfactual it goes from 5% to 7%. Therefore, the "labor market tightness" channel accounts again around a 30% of the total increase. Finally, from the benchmark analysis, we can see that without the 3% increase in the subsistence self-employment rate, the unemployment rate would have gone from 5% to 11%, which highlights the role of the subsistence self-employment activity as a shock smoother during downturns.

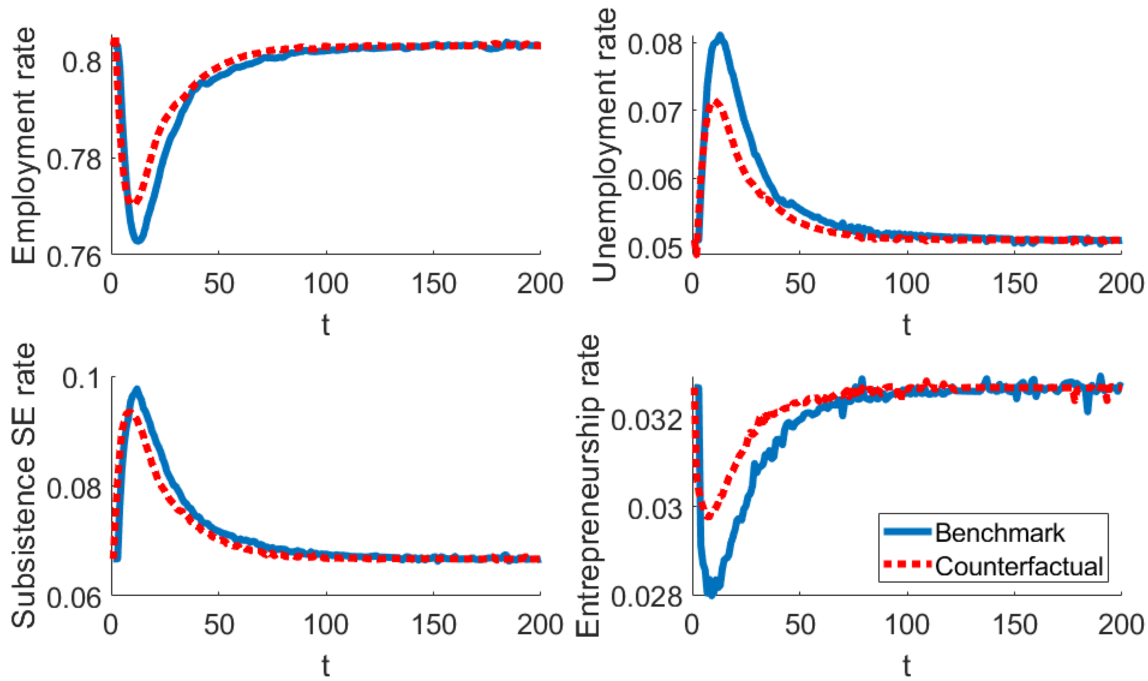


Figure 1.22: Path of Labor Force Components

1.6 Conclusions

This paper studies how labor market dynamics affect the decision to start a business and the growth potential of startups over the cycle, and the final effects on the aggregate job creation. The results show that entry due to stopgap motives increases in recessions, but entry into entrepreneurship, which is more related to employer businesses, falls. The type of individuals deciding to start businesses also change over the cycle. In particular, in recessions, the composition of business founders shifts toward more previously unemployed and fewer highly educated individuals. Because highly educated individuals are precisely those individuals more likely to start a business with a high potential to grow, the change in business founder composition also shifts toward businesses with a lower potential to grow.

Therefore, during downturns, there is a decline in the entry of employer firms and a shift in the business composition toward startups with a lower growth potential. The structural model shows that labor market dynamics during recessions account for a 33% of the decline in the entrepreneurship rate, and a 30% of the increase in the unemployment rate at the peak of the recession period. It also highlights that the initial decline in the aggregate job creation in downturns is mainly driven by the decline in the entry of employer businesses, but as the economy recovers, the lower growth of young businesses becomes more important. In this framework, firm and worker dynamics interact in equilibrium to amplify the effects and persistence of an aggregate productivity/demand shock: a lower job creation of startups declines further the job finding probability, deterring, even more, and more persistently, the entry of startups, especially high-growth businesses. This mechanism generates a slower recovery in the entry of employer businesses and aggregate job creation, consistent with the labor market dynamics in the aftermath of the Great Recession in the U.S.

The next steps of this project are aimed to deal with some potential concerns that might matter for the analysis. First, the possible existence of a downward trend in the job finding probability post-2000 might lead to overestimating the role of the cyclical dynamics of the unemployment rate and job finding in driving the entry and composition of startups. To deal with this concern, I plan to check the results from the reduced form-analysis by incorporating a linear time trend into the regression model used to study the cyclical relationship between the job finding probability and the business creation by employed workers (prediction 2). For the macroeconomic analysis, I plan to redo the first transition dynamic exercise by feeding the model now with a sequence of aggregate productivity shocks aimed to reproduce the deviations of the job finding probability with respect to its linear trend instead of the dynamics of the unemployment rate in levels. If the job finding probability has a downward linear trend, then this modified exercise will show us how much of the total effect on entry rate, the composition of business founders, and aggregate job creation corresponds to the cyclical component. Then, the difference with respect to the original exercise would correspond to the trend component.

The second concern relates to financial frictions. The positive correlation between business founders' educational attainment and the growth potential of their businesses estimated with SBO data is robust to the inclusion of the amount of starting capital as an explanatory variable. However, the negative relationship between the job finding probability and business creation by employed workers estimated SIPP data, especially for the highly educated, does not include controls related to a financial component. The structural model gives light about the pure effect without financial frictions, but this omission might influence the results from the regression analysis. To check whether the results hold when we bring financial constraints into the analysis, I plan to include assets as an additional explanatory variable in the regression analysis, which is information also available in a topic module of the SIPP.

A third and final concern relates to the relationship between skills and entrepreneurship. In the paper, I rely on the positive empirical correlation between the educational attainment of business founders and the potential to grow of their businesses to motivate the mechanism by which high-skill workers start businesses with higher growth potential. In the model, this mechanism is related to two features. First, an endogenous selection mechanism arises because high-skill workers start businesses only when the entrepreneurial idea is good enough to compensate for the forsaken wage. Second, an exogenous positive correlation between labor skills and the permanent component of entrepreneurial ability. However, other idiosyncratic characteristics besides education can also be considered part of the individuals' skill set, which might matter for their businesses' future performance. For future considerations, I intend to go deeper into this analysis. I plan to explore questions like: Are these are sector-specific skills? Do individuals open a firm in the same sector where they were working as employees? Do the experience and age of business founders also matter for the future performance of their businesses?

Chapter 2: Earnings Effect of Startup Employment (coauthored with Nathalie Gonzalez and Álvaro Silva)

2.1 Introduction

Startups and young businesses are the main contributors to aggregate employment creation. However, most startups exit and, among those who survive, just a small fraction grow large and end up boosting job creation.[1] At the individual-level, the higher failure probability of startups implies a higher risk associated with the decision of founders to start businesses and also with the decision of individuals to join them as employees. Startup exits can create costly spells of unemployment for their founders and employees, and even if the startup does not exit, the likely outcome of never growing large could leave them trapped in a small firm, getting stuck in a lower rung of the job ladder. Therefore, either in the form of exit or by remaining small, a startup failure can generate scarring effects on the future career path of the founders and employees.

In this paper, we study and quantify the short- and medium-term effects of employment spells in startups on future earnings of employees.[2] A better understanding

[1] Haltiwanger et al. (2013) and Decker et al. (2014) provide empirical evidence on the importance of young firms in aggregate job creation and their up or out dynamics

[2] On the side of business founders, Garcıa-Trujillo (2020) argues that deteriorated labor market

of these effects is important for at least two reasons. First, business formation depends on the incentives of potential entrepreneurs to start new businesses but also on the individuals' incentives to join them as employees. How large and persistent the difference in earnings is will determine how reluctant individuals will be to accept job offers from relatively risky businesses such as startups.[3] Then, if startups cannot build the required organization capital by hiring individuals with the desired characteristics, their probability of survival and potential to grow will diminish.[4] A second important reason relates to the welfare of the individuals. If an employment spell at a failed startup makes individuals to fall off the job ladder, either by unemployment or by ending up in a worse job, then the resulting accumulated long-term effects might be large along the career path of the individuals.[5]

To perform the empirical analysis, we use the Chilean Unemployment Insurance data. This data allows us to keep track of the individuals' career paths in terms of the firms they work for, their earnings, and periods of non-employment. To carry out the analysis, we build a balanced panel with all the individuals between 18 and 50 years old who made at least one job-to-job or nonemployment-to-job transition after 2012, and who are observed at least 60 months after the first transition date (5 years). A key element of our analysis is the definition of startups. We classify

conditions during economic downturns discourage individuals from starting businesses with potential to grow because it becomes harder to find another job when the businesses fail, which increases the cost of a business failure.

[3] Jarosh (2015) and Pinheiro & Visschers (2015) argue that workers value both job wages and job security because job loss depresses workers' future wages and employment, generating large cumulative earnings losses. This feature makes individuals reluctant to accept job offers from risky firms paying similar wages than other safer firms.

[4] Choi, Goldschlag, Haltiwanger & Kim (2021) argue that the success of a startup derives from the organizational capital that is created at the stage of formation which is inalienable from the founding team itself.

[5] See Carrington & Fallick (2017) for a complete review of theories of costly job displacement.

a firm as a startup if the firm is in its first year of operation and if less than 30% of its employees come from the same previous employer. This additional requirement helps us to avoid the wrong classification of a large company that opens a new branch but with a new firm identifier as a startup.

We start our empirical work estimating the earnings differential over the five years after the job transition between workers who transition to a startup and workers who transition to an established firm. To do so, we regress the logged earnings on an indicator variable that takes value 1 for the individuals that move to startups and 0 for those who moved to established firms. In our first specification, we only include calendar year-month and transition year-month fixed-effects, but we do not control for any worker characteristics. For this specification, we find that those who moved toward startups earned over the next 5 years on average 20.7% less than those who transitioned to existing firms. However, this result includes a sorting component because workers that transition to a startup are potentially different from workers that transition to an established firm.

To isolate the pure effect of the treatment, i.e., having an employment spell in a startup, we follow closely the approach developed by Burton, Dahl & Sorenson (2018), to adjust for worker heterogeneity. Rather than estimating the earning equation with linear adjustments for the effects of the worker characteristics, we instead match on those characteristics and include a fixed effect for each matched group. This non-parametric approach has the advantage of not assuming that each of the relevant dimensions has additive effects on the expected wage as most of the previous literature does (e.g., Brown & Medoff (2003)). In particular, we construct each group using age, gender, country of birth, and year-month of transition as the relevant individuals' observable variables. Then, we look for the two individuals with the closest previous earnings for each treated worker within each group and select them as controls. By doing this, we construct a ''triplet" formed by a treated individual (worker who transitioned to a startup) and two controls (workers who transitioned to an existing firm). This is our baseline specification. The results show that for our matched sample, the 5-year earning effect of working at a startup is -13.8%, which means that 6.9 percentage points of the unmatched difference correspond to sorting. To disentangle how much of this effect is associated with a lower average wage and how much with more frequent or longer periods of unemployment, we estimate a new regression model dropping the periods in which individuals do not receive wages. We find that the differential effect attributed to only a decline in the average wages is -10.5%, leaving 3.3 percentage points attributable to more periods of unemployment.

In the next part of our empirical analysis, we shed light on whether the 5-year earning differential from transitioning to a startup varies across firm size. We do this in light of Babina, Ma, Moser, Ouimet & Zarutskie (2019) who, using the data from the US Census Bureau's Longitudinal Employer-Household Dynamics (LEHD) and from the Longitudinal Business Database (LBD), find the existence of a positive size-pay premium. This means that transitioning to a large startup could potentially alleviate the startup earnings penalty. We also find a size-pay premium for medium-term earnings. On top of that, we find that the startup wage penalty decreases as the startup size increases. Further reversing the negative effect of transitioning to a startup for large-size startups.

Next, we explore whether the contemporaneous effects are different than the 5-year effects. To do so, we compute the differences in earnings during the first 12 months after a transition and estimate the same coefficients. We find that the negative effect goes from -13.8% in the original matched sample to -11.6%, which suggests that the negative differential increases as the individuals' careers progress.

Finally, we study potential cyclical differences in the average earnings 5-years after a job transition. We re-estimate our baseline specification for two periods of time. First, we restrict the panel to those individuals who transitioned to a new job between 2009-10, which corresponds to the period in which Chile was hit by the Great Recession. Second, we restrict the panel to include all individuals that transitioned to a new job between 2012-13, which can be thought of as a normal time. The results for the matched sample show an earning differential of -13.9% for the period 2012-13, larger in magnitude than the -10.0% estimated for the Great Recession. This smaller negative effect during a recession might be explained by the idea that fewer firms are created during bad times, with only those with a higher probability of success being able to start (Ates & Saffie (2021)). This might suggest that the higher quality of startups created during a recession more than offset the larger negative effect associated with falling off the job ladder during recessions (Haltiwanger, Hyatt, Kahn & McEntarfer (2018)).

The remaining structure of this document is as follows. Section 2 summarizes the related literature and contributions. Section 3 explains the data construction. Section 4 presents the empirical results. Section 5 concludes.

2.2 Literature Review

Our work contributes to three strands of the literature. First, it relates to the work that studies relative earnings of startup employees with respect to workers in established firms. Brown & Medoff (2003) is one of the first papers to study the relationship between firm age and wages. Using a small sample of 500 individuals from the household Survey of Consumers conducted by the Survey Research Center at the University of Michigan, they find that older firms pay higher wages than younger firms. However, after controlling for worker characteristics, the relationship becomes insignificant or negative. More recently, Burton et al. (2018) and Babina et al. (2019), controlling not only by worker characteristics but also by firm heterogeneity, find a small but positive young-firm wage premium. Most of the work done so far has focused the analysis on studying the contemporaneous earnings differentials between workers at startups and workers at older firms without paying attention to the medium- and long-term effect on the future earning trajectories. One exception, and closely related to our work, is Sorenson, Dahl, Canales & Burton (2021) who, using administrative data from Denmark, estimates the earnings differentials over the next 10 years after the initial transition. In a similar fashion, we extend the analysis to account for both the contemporaneous and medium-term effects, but we go further by disentangling the effect between changes in wages and frequency of unemployment spells and by exploring a business cycle component. Most importantly, all the previous work has used administrative data of developed countries for the analysis, so, to the best of our knowledge, our paper is the first one to study such effects using administrative data for a country like Chile. Chile has a labor market with higher regulations and informality, and lower social security, what makes falling off the job ladder more costly. Given that this has a direct impact on the outside options of working for a startup, individuals might be more reluctant to accept job offers from startups in this context because of potentially larger and more persistent effects on their careers.

Second, this paper also is related to the work studying scarring effects of unemployment spells on job and earnings prospects. Davis & Von Wachter (2011) using Social Security records for the United States find that at the time of displacement, real earnings fall sharply, and even 20 years after displacement, annual earnings are 10 to 20 percent below pre-displacement earnings. They also document that the present-value earnings losses associated with job displacement are highly sensitive to labor market conditions at the time of displacement, with those happening in recessions being nearly twice as large as those occurring during an expansions. Krolikowski (2017) and Jarosh (2015) argue that existing models used to study unemployment fluctuations have difficulty generating this observed magnitude and persistence of post- displacement earnings losses. Searching for an explanation, they propose search models with job ladders in which workers coming from unemployment are matched to more risky businesses, i.e., with a higher separation probability. This makes individuals spend more time climbing up the ladder after a displacement, matching the magnitude and persistence of the earning losses from the data. Our work shares a similar notion, in the sense that startups are more risky than existing businesses, offering jobs with higher separation rates. We find that the medium-term earning losses associated with a transition toward a startup instead of to an existing firm are associated with a decline in the average wages of the future jobs but also with more unemployment periods.

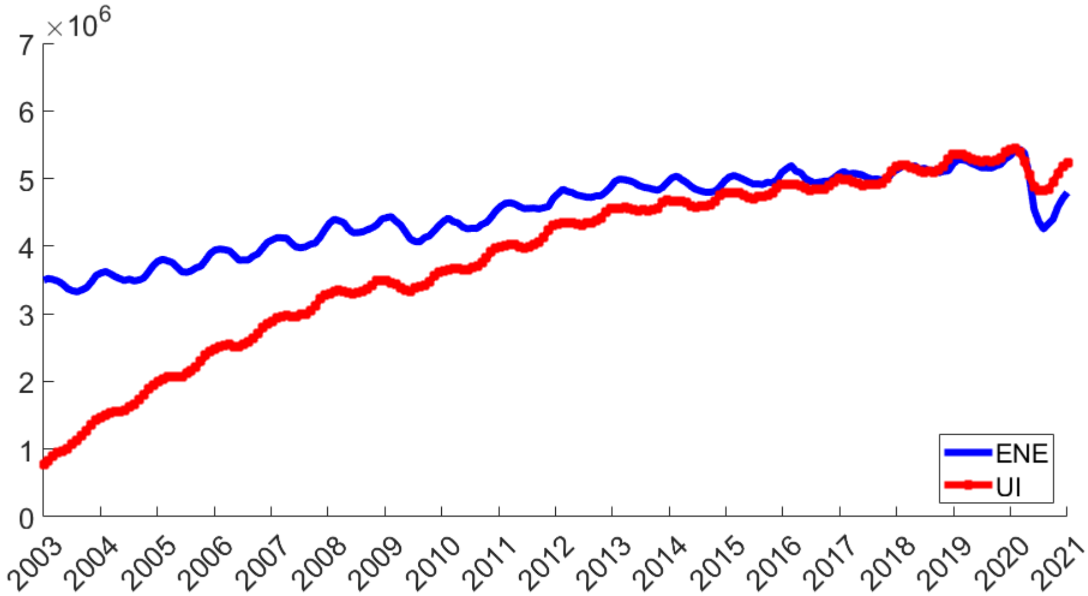
Third, our work also contributes to the literature focusing on joint worker and firm dynamics, in particular the branch that studies the characteristics of the individuals working for startups. In this line, Ouimet & Zarutskie (2014) using administrative data from the U.S. find that young firms disproportionately employ and hire young workers. Engbom (2019) show that older individuals are more reluctant to make job-to-job transitions because they have already reached higher rungs in the job ladder, making a separation that causes falling off the ladder more costly for them. Dinlersoz, Hyatt & Janicki (2019) argue that young firms tend to hire younger workers and provide them with lower earnings compared to more established firms. They formalize this idea in a model with an entrepreneurial sector, for which mainly individuals with low assets accept job offers. Haltiwanger et al. (2018) find that during the Great Recession, the firm wage ladder collapsed, with net worker reallocation to higher-wage firms falling to zero. They also argue that the earnings consequences from this lack of upward progression are sizable. We contribute to this literature by showing that individuals hired by startups in Chile are more likely to be males and, on average, have lower previous earnings than the individuals hired by established firms. In contrast to what the literature has documented for the U.S., we do not find that startups in Chile hire younger workers more intensively.

2.3 Data

Our data comes from unemployment insurance administrative records (UI). The unemployment insurance covers all workers over 18 years old who are employed in private sector salaried jobs. Workers under 18 years old, the self-employed, and public sector employees are excluded from this system. Participation in the unemployment insurance is compulsory for everyone starting a new job after 2002 and voluntary for everyone else. The idea behind the voluntary enrollment of existing work relations was to increase the support to the introduction of this new labor cost. We exclude from our analysis data before 2009 since at that point we can closely match the number of private salaried workers reported by the national statistical institute as Figure 2.1 shows. Therefore, the UI data corresponds to a matched employer-employee dataset, similar to the LEHD in the United States.

2.3.1 Startup identification

A key variable for our empirical analysis is the identification of the age of the firm. To define a firm as a startup, we take advantage of the fact that we have the universe of private employers in our data and define the month of birth of a firm as the first month the firm makes a contribution to the unemployment insurance on behalf of an employee. We define startups as firms with 12 months of age or less. In addition, to avoid mislabelling an established firm as a startup we require that the share of employees that come from the same previous employer to be less than



Note: The ENE data comes from household surveys. We spliced two different surveys using the growth rate and leveraging the fact that there is one year for which we observe both surveys.

Figure 2.1: Private Salaried Employment

30%.[6]. Note that this definition of a startup is not the same across papers in the literature. For example, Sorenson et al. (2021) labels firms as startups during their first 4 years of operation. Babina et al. (2019) classify a firm as young during the first three years of operation. [7]

2.3.2 Earnings

Our dependant variable of interest is the individual's earnings. We have information on monthly earnings from all the jobs held at any time period. We build the total monthly earnings of a worker by adding up the income from all the jobs held in a given month. Earnings refer to taxable salaries. We do not have information

[6]This applies for firms with starting size of more than 3 workers. Firms that start with 2 workers or less are going to have at least 50% of their workers coming from the same previous employer firm.

[7]Note that some of the quantitative difference between our results and previous literature could be coming from this definition. The up or out dynamics of young firms imply that transitioning to a one year old firm is different from transitioning to a 4 year old firm.

on equity payments. However, unlike the US, in Chile equity payments are not a common practice.[8]

2.3.3 Identifying job transitions

To identify job transitions, we follow the literature in restricting the analysis to the main job of an individual. We define the main job as the job with the highest earnings. We drop the information on secondary jobs when building the transitions panel but we take them into account in total earnings as described above. By dropping the non-primary jobs we lose 3% of observations. This is an indication of a low incidence of workers with multiple jobs.

2.4 Empirical Analysis

2.4.1 Data Construction

A central element to our analysis comes from the data structure. We build a balanced panel of individuals starting in the period of their first transition to a new job after 2012 and include 5 years of observations after that first transition. We start with 2012 because, by that time, the Chilean economy had recovered from the Great Recession, and for our baseline analysis, we want to abstract from cyclical effects. In section 2.4.8 we look into the differential effect of working for a startup over the cycle. We restrict our sample to individuals between the ages of 18-50 at the time of their first transition to abstract from retirement decisions. Since we are

[8] In the future we plan to study if this is a relevant margin by using tax data.

going to work with a balanced panel, we also exclude people that exit the formal labor market before 60 months after their first transition. An individual leaves the formal labor market the last time an employee makes a contribution to the UI on his behalf. Since startups are more likely to close than established firms (Decker et al. (2014)), an important component of the effect of transitioning to a startup is its impact on job security. If an individual has a gap in his UI contributions, we assume that he was out of a job.[9] Note that the Chilean labor market has an informality sector of 30% of employment, so some of the zeros in our balanced panel can be periods with informal employment, but given the nature of our data we are not able to observe that income.[10]

2.4.2 Summary Statistics

Table 2.1 shows the descriptive statistics of workers in our balanced panel. An individual enters our sample the first time she transitions to a new job after 2012. This means that for every individual we have information on the earnings in her previous job. As we described in the identification strategy section, we use this information to build the counterfactual when identifying the causal effect of transitioning to a startup. Notice that the way in which we define the starting month means that workers enter the panel in different calendar months.

Table 2.1 shows that 13% of the individuals in our panel made their first transition after 2012 to a startup, while 87% made their first transition after 2012

[9]We replace his earnings with the first percentile of earnings in the panel to avoid distorting the coefficients given that we are using the log of earnings as our variable of interest. We ran the specifications in levels replacing missing earnings with zeros and got similar elasticities.

[10]The source of value for informal employment is the Chilean Institute of National Statistics

to an established firm. The difference between the average monthly earning of those who moved to an existing firm and those who moved to a startup is 19%. In our panel, workers whose first transition is to a startup are almost a year older than the workers whose first transition after 2012 is to an established firm. This result contrast with previous findings in the literature. Sorenson et al. (2021) shows that, on average, workers that transition to startups in Denmark tend to be younger than workers that transition to established firms. In line with our expectations, the share of females who transition to a startup is lower than the share of females who transition to an established firm. This is in line with the literature that shows that females are more risk averse than men Charness & Gneezy (2012). Finally, we see that workers who transition to established firms tend to have lower previous earnings. Interestingly, workers who move to a startup have previous earnings 11% lower than workers that transition to an established firm. Notice that the difference is not as large as the difference in monthly earnings. The difference between the 19% and the 11% is suggestive evidence of a negative effect of transitioning to a startup.

These descriptive statistics indicate that workers who transition to startups are different from those who transition to established firms. In other words, selection plays a central role in the earnings differentials between individuals transitioning to a startup and individuals who transition to an existing firm. In the next section, we describe our strategy to overcome this selection that would bias the estimation of the difference we would obtain by using a simple mean difference.

	Movers to all Firms	Movers to Existing Firms	Movers to Startups
Monthly earnings	445,537	456,753	369,475
	(552,233)	(561,077)	(481,200)
Age	34.8	34.7	35.6
	(8.53)	(8.52)	(8.56)
Female	0.35	0.35	0.32
	(0.48)	(0.48)	(0.47)
Previous earnings	327,960	332,775	295,307
	(387,863)	(393,186)	(347,883)
Number of Individuals	1,161,620	1,012,346	149,274
Observations	69,697,200	60,740,760	8,956,440

Mean value is in the main line and standard deviation in parenthesis. The balanced panel contains 60 months after the first transition to a new job after 2012 for individuals between 18 and 54 years old. Monthly and previous earnings are in current Chilean pesos.

Table 2.1: Descriptive Statistics of Balanced Panel

2.4.3 Identification Strategy-Earnings Differential

The main goal of this paper is to identify the impact of transitioning to a startup on the medium-term earnings of a worker. In other words, we want to know what is the medium-term effect of working for a startup versus working for an established firm. The main challenge to estimating the causal effect of working for a startup is to identify the relevant control group. The ideal control group would be a set of workers identical to the workers that transitioned to a startup but that instead of transitioning to a startup transitioned to an established firm. We build the control group for each of the workers that transition to a startup- our treated group- using the methodology proposed by Burton et al. (2018). The idea is to find the two most similar workers that transitioned to an established firm to each of the workers that transitioned to a startup and compare the average earnings among them. This is a matching strategy in the same spirit of a propensity score matching. However, in contrast with the traditional propensity score matching, Burton et al. (2018)'s approach is non-parametric. More precisely, we build subsamples of individuals of the same age, gender, date of the first transition after 2012, and country of origin. Within those subsamples, we choose the two workers with the most similar previous earnings and generate an identifier at the group level. Each treated worker and the two nearest neighbors constitute a triplet in our panel. To make sure that individuals in a triplet are similar enough, we restrict the control workers to those individuals with previous earnings with less than 10% of difference to the treated worker.[11]

The advantage of this non-parametric approach is that it does not impose or assumes a linear relationship between earnings and worker characteristics. Note that to the extent that previous earnings have information about non-observable ability and preferences, matching on previous earnings among individuals with the same personal characteristics allows us to account for those non-observable.

2.4.4 Estimating Equations

Unmatched specification

To quantify the extent to which selection plays a role in the raw differences

[11] Note that we are doing matching with replacement. This means that workers in the control group could potentially be part of multiple triplets. Given the sample size this is not an issue, 3% of the control group are in more than one triplet.

we observe, we start by calculating the average difference in medium-term earnings between workers that transitioned to a startup and workers that transitioned to an established firm. Equation 2.1 describes the way in which we calculate the raw difference using fixed-effect for the date of the first transition after 2012 and date fixed effects. This is what we called the unmatched specification in the tables below.

$$ln(Earnings)_{it} = \beta_1 + \beta_2 startup_i + \gamma_\tau + \eta_t + \varepsilon_{i,t} \tag{2.1}$$

Where $startup_i$ is an indicator variable that takes the value of one if the worker enters the panel after transitioning to a startup and zero if the worker enters the sample after transitioning to an established firm. γ_τ is the set of time of transition fixed effects, and eta_t is the set of time fixed effect.

Matched Specification-Triplets

There are two differences between the unmatched and the matched specifications. The first one is the sample of workers included in the estimation. For the unmatched specification, we use all individuals in the balanced panel, whereas for the matched sample we include all the workers in the treated group and the two closest workers in the control group. The second difference is the triplet fixed effect. The triplet fixed effect accounts for sorting to the extent that among individuals in the same triplet the one that ends up working for a startup is random. In the final section of the paper, we discuss some of the threats to this identification assumption and the steps we plan to implement to address them. Note that since workers in a triplet made their first transition during the same month, the triplet fixed effect absorbs the transition date fixed effect.[12]

$$ln(Earnings)_{it} = \alpha_1 + \alpha_2 startup_i + \nu_g + \psi_t + \epsilon_{i,t} \tag{2.2}$$

Where ν_g is the coefficient of a dummy that takes a value of one for all the individuals in triplet g.

Matched Specification- (wages only)

An interesting question we can answer in this setting is, to what extent are the observed differences in average earnings coming from lower wages versus more periods of unemployment. We address this question by estimating the differences in earnings conditioning on being employed, or what is equivalent in our setting, conditional on having a contribution to the UI. In practical terms, this means that we exclude all the observations with zero earnings and run the specification described by equation 2.2.

2.4.5 Baseline Results

Table 2.2 summarizes the results of the three estimating equations described above. The first column shows the difference in earnings of workers that made their first transition after 2012 to a startup versus an established firm. The average difference in earnings is -20.7%. This difference is a combination of the effect of working for a startup and selection. By selection we mean the difference in earnings

[12]Note that the number of observations in column 2 of table 2 are less than 3 times the observations of workers that transitioned to a startup. This is because of the restriction we imposed on the difference of the previous earnings of the workers in the same triplet to be smaller than 10%

that comes from workers of startups having different characteristics.

Column 2 reports the results we obtain when using the matching triplets approach described above. The assumption behind the matching triplets on worker characteristics is that workers of the same age, gender, country of birth, similar previous earnings, and that made their first transition after 2012 at the same time are comparable groups. If this is true, then the difference in earnings of those who transition to a startup vs those who transition to an established firm is the effect of transitioning to a startup. The difference between the coefficient of startups in column 1 and the coefficient of startups in column 2 is the degree to which selection explains the total effect.

The effects on earnings of working for a startup can be reflected in lower monthly wages or in fewer periods of employment. To identify the importance of each of these mechanisms, we run the same specification as in column 2, but excluding the periods of non-employment. The coefficient of startups in this column tells us the difference in earnings coming only from monthly wages. As expected, the effect on wages is smaller than the total effect. This means that workers that transition to a startup tend to have more unemployment spells. The difference in monthly wages is significant statistically and economically. On average, working for a startup reduces average monthly wages in 10.5% even 5 years after the transition.

VARIABLES	Unmatched log(earnings)	Matched log(earnings)	Matched (wages only) log(earnings)
Startup	-0.207***	-0.138***	-0.105***
	(0.00050)	(0.00050)	(0.00033)
Transition year-month FE?	yes	no	no
Calendar year-month FE?	yes	yes	yes
Triplet FE?	no	yes	yes
Number of observations	69,697,200	24,884,760	18,004,190

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 2.2: Medium Term Earning Effects of Working at a Startup

2.4.6 Taking into account the size of the startup

We know that startups tend to start small Decker et al. (2014) and that small firms have different characteristics than large firms. Previous evidence for developed countries points to the fact that startups of different size have different effects on the earnings path of workers. Babina et al. (2019) show that large firms have a positive wage premium when compared to small firms. Sorenson et al. (2021) find that the effect on earnings of transitioning to a startup also depends on the startup size. To explore if this is a relevant margin in the Chilean context, we check if the size of the startup affects the medium-term earnings penalty. We do this by adding size categories to our baseline specification. We include 8 bin sizes: 1, 2-9, 10-19,

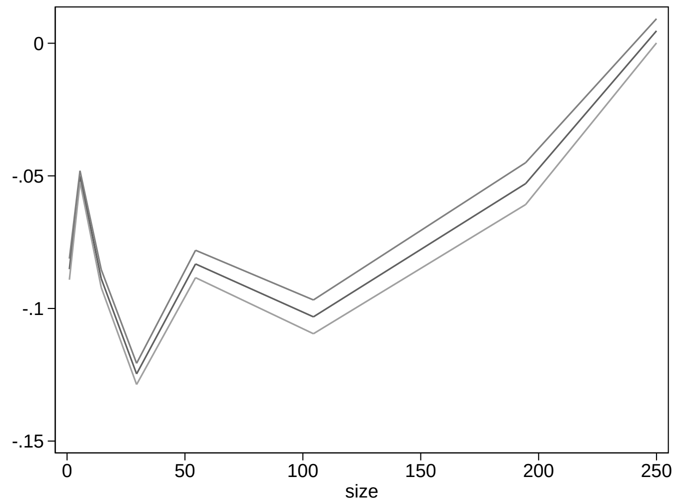


Figure 2.2: Effect of transitioning to a Startup by Size

20-39, 40-69, 70-139, 140-249, 250+.

$$ln(Earnings)_{it} = \rho_1 + \rho_2 startup_i + \sum_{s=1}^{8}(\delta_s Size_s + \lambda_s Size_s * startup_i) + \gamma_\tau + \eta_t + \varepsilon_{i,t} \tag{2.3}$$

Where $Size_s$ is a dummy that takes the value of one if the firm to which the individual makes the first transition after 2012 is in bin size s. Figure 2.2 shows the λ_s coefficients over the mean value of the size bin. Over most of the size distribution the effect of negative, except for large startups. Working for a startup of 250 or more workers has a positive premium in earnings when compared to an established firm. A table with the full set of coefficients is included in Appendix B.1.

2.4.7 Contemporaneous Effects

In this section, we study the effects of working for a startup but only on the earnings of the first year after the job transition. The earnings of the first year are more likely linked to the employment spell at the startup, and if startups pay a wage premium to attract workers, we should expect a positive effect of working for a startup. For the case of the US, Babina et al. (2019) find that after controlling for selection, the contemporaneous effects of working for a startup are slightly positive. Table 2.3 shows that in the Chilean case even the contemporary effects are negative, although smaller than the effects over the 5-years horizon analyzed before. This means the effect of working for a startup goes beyond the first year after the transition. This can be interpreted as evidence of scarring effect of startup employment. Notice that to get a smaller coefficient in the contemporaneous analysis, the penalty after the first year not only has to remain, but it has to be larger. This could take the form of lower wages or fewer employment periods.

VARIABLES	Unmatched log(earnings)	Matched log(earnings)	Matched (wages only) log(earnings)
Startup	-0.170*** (0.00085)	-0.116*** (0.00079)	-0.114*** (0.00059)
Transition year-month FE?	yes	no	no
Calendar year-month FE?	yes	yes	yes
Triplet FE?	no	yes	yes
Number of observations	19,594,008	7,089,432	5,528,329

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 2.3: Contemporaneous Effects: 12 Months After the Job Transition

2.4.8 Cyclical Effects

The nature of startups is likely to change over the cycle. Ates & Saffie (2021) find that in Chile there is a fewer but better effect on firm creation. The intuition is that during recessions there is a reduction in the credit supply and only the most promising firms are able to have access to financing. To explore if there is a difference on the impact of working for a startup over the business cycle we perform two different exercises. First, we redefine the balanced panel using transitions that occurred between 2009-2010 and estimate the baseline regression in equation 2.2. Second, we redefine the panel by only including transitions that occurred between 2012-2013. Notice that in both cases we are still following each individual in the panel for 60 months.

2009-10			
VARIABLES	Unmatched log(earnings)	Matched log(earnings)	Matched (wages only) log(earnings)
Startup	-0.141***	-0.0997***	-0.0709***
	(0.00051)	(0.000498)	(0.000362)
Transition year-month FE?	yes	no	no
Calendar year-month FE?	yes	yes	yes
Triplet FE?	no	yes	yes
Number of observations	57,538,260	21,622,140	15,198,314

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 2.4: Cycle Effects: Recession

2012-13			
VARIABLES	Unmatched log(earnings)	Matched log(earnings)	Matched (wages only) log(earnings)
Startup	-0.211***	-0.139***	-0.106***
	(0.00054)	(0.00053)	(0.00036)
Transition year-month FE?	yes	no	no
Calendar year-month FE?	yes	yes	yes
Triplet FE?	no	yes	yes
Number of observations	60,141,480	21,767,820	15,552,338

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 2.5: Cycle Effects: Boom

Tables 2.4 and 2.5 show that individuals who transition to a startup in a recession experience a lower cost of startup employment than workers who transition to a startup in normal times. This result is in line with the fewer but better mechanism, by which startups created in recessions have a higher probability of survival and growth potential than those created in normal times. This result can also be interpreted as the increase in the quality of the startup pool more than offsetting the negative effect arising from the higher cost of a startup failure over the future earnings of their employees.

2.5 Conclusions

Startups offer less job security than established firms, with higher exit rates and a higher probability of never growing large. Exits create costly spells of unemployment for their former employees, and smaller firms have wage penalties over large firms. Therefore, either in the form of exit or by remaining small, a startup failure might generate scarring effects on the future career path of individuals. In this paper, we study and quantify the short- and medium-term effects of employment spells in startups on future earnings of employees. To perform the empirical analysis, we use the Chilean Unemployment Insurance data.

We find, first, that those who moved toward startups earned over the next five years on average 20.7% less than those who transitioned to established firms. However, this result includes a sorting component because workers transitioning to startups are different from workers that transition to an established firm. After controlling for worker heterogeneity, that difference is reduced to -13.8%, which implies that an important part of the observed difference comes from sorting. We further decompose this effect and find that 11 percentage points of the overall 5-years effect, comes from lower average wages. The remaining 3.3% arises from more frequent or longer unemployment spells in the future career path. Then, we explore whether these differences vary across firm size. We find a size-pay premium and, on top of that, the startup wage penalty decreases as the startup size increase. The negative effect even reverses when the startup is large. We also find that the contemporary differential, measured as the effect over the first year after the transition to a startup, is -11.6%, smaller than the estimated medium-term effect. This suggests the existence of a scarring effect on earnings of working for a startup. Finally, we also find heterogeneous effects over the business cycle. Transitioning to a startup has less negative medium-term effects during recessions. This is consistent with a "fewer but better" effect documented by previous literature for business creation during recessions in the Chilean context.

In the next iteration of this paper, we plan to tie some of the loose ends highlighted throughout the text. First, we are going to explore the sensitivity of our results to the definition of a startup. Specifically, we want to check if the estimated effect of working for a startup changes if we label firms three years or younger as startups. Second, we want to take a deeper look at the worker heterogeneity side by investigating how the medium-term startup earning penalty depends on the level of human capital (proxied by previous wages), age, and previous employment status. We found a sizable effect on earnings from transitioning to a startup. However, we have not explored how these effects vary across worker characteristics. Third, we also plan to go further in the business cycle analysis to identify the underlying worker and firm dynamics driving the variation in the earning effects over the cycle. Fourth, we also plan to provide additional evidence using an instrumental variable approach in order to rule out issues related to preferences of individuals for small firms (e.g., flexibility and other non-pecuniary effects (Hurst & Pugsley (2011), Hurst & Pugsley (2017)).

Chapter 3: Business Formation in the U.S.: Individual-level Analysis using SIPP data

3.1 Introduction

The literature on firm dynamics in the United States has proliferated enormously in recent years thanks to the incorporation of firm age into business administrative data. This has improved our knowledge about business formation and business dynamism.[1] However, despite its richness, U.S. Census Bureau's firm-level data does not include information about the business owners. This implies that researchers studying business formation from an individual-level perspective have to rely on indirect strategies to identify business owners in the U.S. administrative firm-level data or, alternatively, use survey data. A few recent exceptions have been able to link U.S. administrative data with IRS data that provides both the Social Security Number and the Employer Identification Number for individuals who are business owners.

A recent work that helps to illustrate the scope of using linkages between tax data and U.S. Census Bureau's data is Azoulay, Jones, Kim & Miranda (2020). They use the Schedule K-1 forms to link businesses from the Census Bureau's Longitudinal

[1] See Haltiwanger et al. (2013).

Business Database (LBD) to their owners' earnings history in the W-2 records as well as to their demographic characteristics in the Census Numident file. By doing this, they were able to find that the mean founder age for the 0.1% fastest growing startups is 45. This is feasible because the K-1 forms work as a bridge between both worlds, firms and individuals, by providing both the Employer Identification Number (EIN) and the Social Security Number (SSN).[2]

However, the approach of using IRS data as a bridge between U.S. Census Bureau's firm- and individual-level data has two important limitations. First, IRS data only can be accessed by Census Bureau's economists. K-1 data is not available through Research Data Centers (RDC) for external researchers. Second, the K-1 dataset identifies owners only of "pass-through" businesses that are not sole proprietors. This means that business ownership data are available for partnership and S-corporations only, which account for approximately half of all new businesses. Therefore, to include sole proprietors and C-corporations into the analysis, researchers still have to rely on indirect identification strategies.[3]

In this paper, considering the limitations to identify business owners in the U.S. administrative data, I use the Survey of Participation Program (SIPP) to provide a rich characterization of the business formation process in the U.S. from an individual-level perspective. The goal of this paper is twofold. First, I develop a

[2] Other examples of work linking IRS data with Census Bureau's data are Goldschlag, Kim & McCue (2017), Nelson (2016), and García-Pérez, Goetz, Haltiwanger & Sandusky (2013).

[3] Azoulay et al. (2020) use the Census Bureau's Annual Survey of Entrepreneurs (ASE) as a bridge to link the LBD with W-2 records for a sample of C-corporation business owners. C-corporations account for approximately 20% of all new businesses. Choi (2017) and Choi et al. (2021) identify business owners using the indirect strategy proposed by Kerr & Kerr (2017), by which the founders are approximated by the three top earners reported in the Longitudinal Employer-Household Dynamics (LEHD).

detailed description of business and owners in the U.S. in terms of their characteristics, to later perform an analysis of survival probabilities and earnings effect of startup spells. Second, I intend to provide a good sense of the appealing of the SIPP to analyze business formation at an early stage. Although the SIPP is much smaller than typical administrative data sets, its advantage is that it is easily accessible, follows individuals for up to four years, records their employer and occupational changes and spell durations, and provides detailed information on their demographic and job characteristics, and earnings. It also provides information that allows us to characterize the early stage of a new business, like start and end dates, legal form of organization, profits, size (categories), and hours spent in the business. The SIPP has been widely used in the literature, specially by the line of research studying wages dynamics and labor mobility.[4] However, the SIPP has been less used to study entrepreneurship.

I start the empirical work by providing a rich characterization of the business formation process in the United States using data from the SIPP. To do so, I perform a descriptive statistics analysis in terms of the number and size of new and established businesses, and also in terms of the previous labor force status, educational attainment, and previous earnings of founders. On the side of the businesses, the results show differences in the composition of new and established businesses in terms of their legal form of organization and size. Approximately a 80% of all new ventures correspond to unincorporated businesses. On the side of the individu-

[4] See Barattieri, Basu & Gottschalk (2014), Erosa, Fuster & Kambourov (2016), Moscarini & Postel-Vinay (2016), Fujita & Moscarini (2017), Carrillo-Tudela, Wiczer & Visschers (2019).

als, the results show that individuals starting incorporated businesses are relatively more educated and less likely to come from unemployment than those starting unincorporated businesses. Also, unincorporated business owners have approximately 40% less previous earnings than incorporated business owners.

Next, I compute the probability of survival for a set of time lengths, and across educational attainment of business owners and previous labor market status. The analysis shows that incorporated businesses are more likely to survive than unincorporated businesses, and the survival probability increases when the owner is a highly educated individual. The results also suggest that the previous labor market status of the business founder does not affect the probability of survival, conditional on the type of business they started.

Finally, I study the earnings differential arising from startups spells. To do so, I use a three-block spell structure in which the first block corresponds to the previous occupation, the second block to the spell as business owner, and the third block is the new job or unemployment. If the businesses does not exit, then there is no third block. This structure allow us to compute the earnings differentials between the labor earnings before and after the self-employment spell. The results show an important source of heterogeneity related to the business cycle, with the difference in earnings being strongly procyclical. If we think of the decision to start a business as an occupational choice, in the spirit of Garcıa-Trujillo (2020), then a spell as a business owner can be also be interpreted as accepting a job offer from a firm located in a particular rung in a firm wage ladder. In such framework, starting a business might also work as a mechanism that individuals use to climb up the ladder in booms. This seems to be an interesting line of work for future research.

The remaining structure of this document is as follows. Section 2 explains the data construction. Section 3 presents the empirical results. Section 4 concludes.

3.2 Data

To perform most part of the empirical analysis, I use the Survey of Income and Participation Program (SIPP). The SIPP survey design is a continuous series of national panels, with a sample size ranging from approximately 14,000 to 52,000 interviewed households. The duration of each panel ranges from 2.5 years to 4 years. The survey used a 4-month recall period, with approximately the same number of interviews being conducted in each month or "wave" of the 4-month period. The interviews occurred in waves where each household responded to questions regarding the past four months. As a result, each household was interviewed every four months for the duration of the panel.

The SIPP provides useful individual- and firm-level information. For individuals, we can keep track of their labor history and earnings. When they start a business, we know when they keep a paid job or if the business becomes the main activity.[5] We also have demographic and financial information. Wealth and assets are not included in the analysis of this work, but they are available. For the businesses, the information is not so rich as for individuals but still is very useful for the analysis. It provides the start and end dates of the businesses that a person starts and unique identifiers for each one of them. Also, we have information about the

[5]The SIPP keeps track of up to two jobs and two businesses simultaneously

industry, profits, size (categories), and legal form of organization.

For the analysis, I use the SIPP panels 1996, 2001, 2004, and 2008. The resulting dataset corresponds to a monthly frequency panel with 10.7 million observations. To construct the data, I follow closely the cleaning procedure from Fujita & Moscarini (2017). First, I construct and clean the data using the extracts from the Center for Economic and Policy Research (CEPR). Then, to define the primary job and business, I use the start and end dates of the spells and the number of hours. The activity with more hours, conditional on being within the range of time in which the job or business is operating, is considered as the main activity. In the data, we see back and forth transitions for businesses with the same identifier between being unincorporated and incorporated. These changes are not related to transitions, so to avoid distortion, I define the legal form of organization as the type with more periods within the spell. Also, there are gaps in the panel. If a transition occurs during a gap, I identify the transition positively only if the gap is smaller than 5 months.

3.3 Empirical Analysis

In this section, I present first a set of descriptive statistics about businesses and owners using the SIPP data. I also complement the analysis in some cases with data from the Current Population Survey (CPS). Then, I present several results related to business dynamics. I compute the probability of survival conditional on several individual characteristics, the impact of self-employment spells on the future labor earnings of individuals, how these effects vary over the business cycle, and how they vary across educational attainment of the individuals.

3.3.1 Characterizing businesses and onwers in the U.S.

Table 3.1 presents a characterization of the businesses and startups using SIPP data for 2005. I also compute the same statistics with CPS data, when it is possible, as a benchmark. First, we see that more than one-third of the total number businesses in the SIPP sample are unincorporated. This difference increases in the sample of startups, defined as businesses in the first month of operation, with more than 80% of the new businesses being unincorporated. The table also reports information about the size distribution of firms, with limitations, and from the CPS we see that the average size of incorporated businesses is approximately 4 while for unincorporated businesses is 0.79. Regarding the size of startup, they are in average smaller but still incorporated startups are created with an initial size 3 or 4 times the size of unincorporated startups.

Year 2005	CPS		SIPP	
	Uninc.	Inc.	Uninc.	Inc.
NUMBER OF BUSINESSES				
All Businesses	10,496,226	5,264,196	10,855,594	5,088,705
	66.6%	33.4%	68.1%	31.9%
Startups	610,730	138,322	326,900	79,315
	81.5%	18.5%	80.5%	19.5%
EMPLOYEES (% of firms)**				
All Businesses				
N>=1	13.2%	38.8%	NA	NA
N>=25	2.3%	10.1%	1.3%	5.4%
P50	0	0	under 25	under 25
Mean	0.79	3.96	NA	NA
Startups				
N>=1	7.5%	30.3%	NA	NA
N>=25	3.9%	9.5%	2.9%	12.4%
p50	0	0	under 25	under 25
Mean	0.66	2.37	NA	NA

Notes: All values are weighted.
** Basic Monthly CPS 2014-2018 and SIPP 2012.

Table 3.1: Business Characteristics

With respect to the individuals deciding to start businesses, Table 3.2 show that incorporated businesses are disproportionately more started from employment as previous labor market status, and they are mainly started by highly educated individuals. In contrast, unincorporated businesses are started by individuals with low educational attainment and an importation fraction by unemployed individuals. Regarding wages, we see that consistently in both the SIPP and CPS, incorporated businesses generate almost twice the average profits of unincorporated businesses.

Year 2005	CPS		SIPP	
	Uninc.	Inc.	Uninc.	Inc.
PREVIOUS OCCUPATION				
Startups				
E	35.8%	59.3%	48.2%	83.0%
NE	64.2%	40.7%	51.8%	17.0%
HUMAN CAPITAL				
All Businesses				
<= High School	42.2%	27.5%	34.9%	21.7%
Incomplete College	27.4%	26.4%	36.2%	34.1%
>= College	30.4%	46.1%	29.0%	44.2%
Startups				
<= High School	47.0%	31.1%	37.0%	27.0%
Incomplete College	26.8%	25.1%	38.3%	32.7%
>= College	26.2%	43.8%	24.6%	40.3%
PREVIOUS WAGE* (annual, thousands)				
Mean	$39.3	$66.2	$27.0	$47.9

Notes: All values are weighted. * CPS ASEC.

Table 3.2: Founder Characteristics

In terms of the age, we see that businesses as well as startups are concentrated in business owners with an average age between 35 and 54 years old, with a mean close to 45. It is worthy to note that the distribution of age is tilted to the right for incorporated businesses, meaning that in average incorporated are started and owned by more experienced people.

3.3.2 Construction of three-block spells $\{XYZ\}$

In the next part of the analysis, I will present the results regarding the probability of survival, the destination after a business closure, and effects on earnings of a startup spell. To compute each of these variables, we first need to construct the transitions between employment, unemployment, unincorporated self-employment,

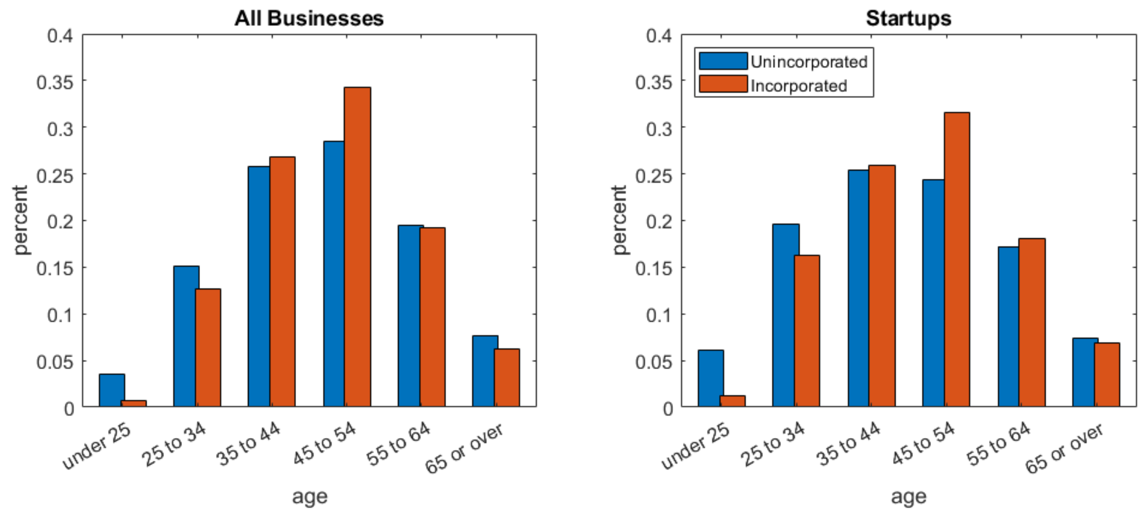


Figure 3.1: Age of Business Owners

and incorporated self-employment. In particular, to estimate the effects on earning of having had a startup employment we need to construct a three-block spell identifier. This is needed because we want to compare the earnings before starting a business with the earnings after the business ends. We can do this because the panel structure of the SIPP allows us to keep track of individuals for up to four years.

The procedure to construct the three-block spells is as follows. I name the initial spell as "X", which can be either employment (E) or unemployment (U). Then, the second spell, which I name "Y", is required to be either an unincorporated (S) or an incorporated business (F). The last third spell, named "Z", has to be E or U. Then, the three-block spell $\{XYZ\}$ is formed by any combination of $X = \{E, U\}$, $Y = \{S, F\}$, and $Z = \{E, U\}$.

We have two important cases for the combinations in $\{XYZ\}$. First, we can have someone who started a business and the business never closed. That kind of spell is going to be named as $XYend$. The second possible outcome is that someone that opens a business, then closes it and goes back to search for a job (in this case $Z = U$), or closes it because she accepts a job offer from another employer firm (in this case $Z = E$).

The three-block spell $XYend$ is used to construct the probability of survival. Conditional on all the individuals that remain in the sample after N months, I compute how many of the have a $XYend$ type of spell. The result is the probability of survival for N months. Exploiting the panel dimension of the SIPP, I do this for N=4, N=12, N=24, and N=36.

The three-block spell XYZ is used to construct two objects. First, it allows to find what is the destination after a business closure. We want to see whether people go to unemployment to employment after a business spell, but most importantly, whether there are differences in the trajectories when $Y = S$ and when $Y = F$. The second object is the earnings differential. We want to know whether the labor earnings for the final block Z are different from the previous earnings in the block X after a transition $Y = \{S, F\}$.

3.3.3 Survival probability

We start this part of the analysis by constructing the probability of survival for three lengths of time: N=12, N=24, and N=36. Figure 3.2 shows how this probability depends on the legal form of organization and on the previous labor market status of the business founders. We can observe that regardless of whether individuals start businesses from employment or from unemployment, incorporated businesses have a higher survival probability than unincorporated businesses. Also, it seems that there is no difference in the probability of survival between those businesses started from unemployment and those started from employment.

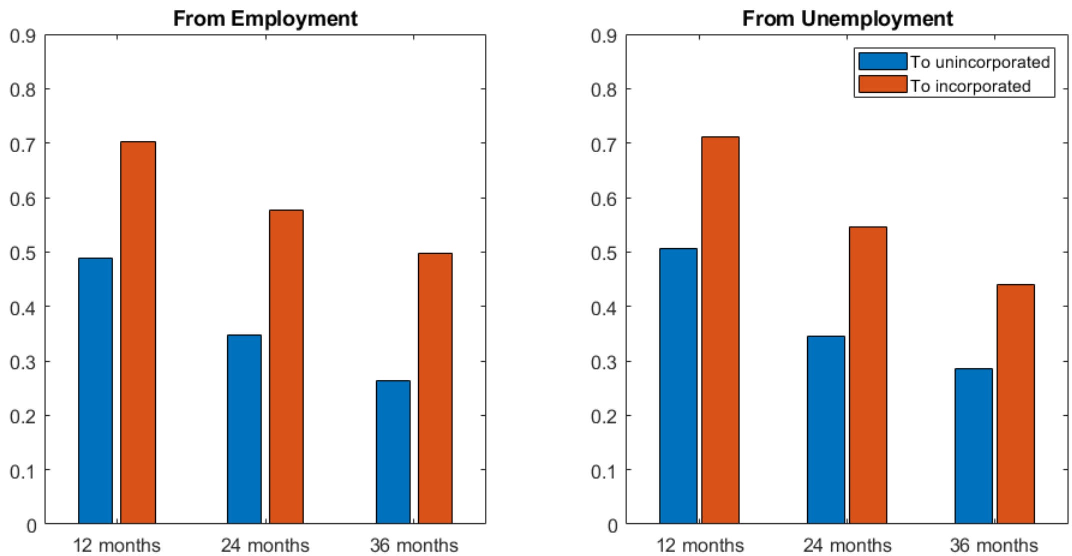


Figure 3.2: Survival Probability by Previous Labor Market Status

Figure 3.3 presents the survival probabilities for the following length of time: N=4, N=12, N=24, and N=36. The analysis is disaggregated by level of educational attainment. The category "Educ==1" corresponds to individuals with complete high school or less educational attainment. "Educ==2" corresponds to individuals with incomplete college. "Educ==3" corresponds to individuals with college or higher level of education. We observe that the probability of survival decreases with time, but the probability of survival of incorporated businesses is monotonically higher than for unincorporated businesses, and the difference seems to even increase over time. Comparing across educational levels, we observe that as time progresses the probability of survival for highly educated decreases less than for other individuals. But the same is not true for the case of unincorporated businesses. This result is important, it points out the different performance of the business owned by highly educated individuals. Garcıa-Trujillo (2020) using the Survey of Business Owners (SBO) also shows that highly educated individuals are relatively more prone to start high growth startups and also that most of the high growth startups are started by highly educated individuals. Therefore, putting together both results, it seems that highly educated individuals are the individuals who start most of the businesses with a high probability of survival and high potential to grow.

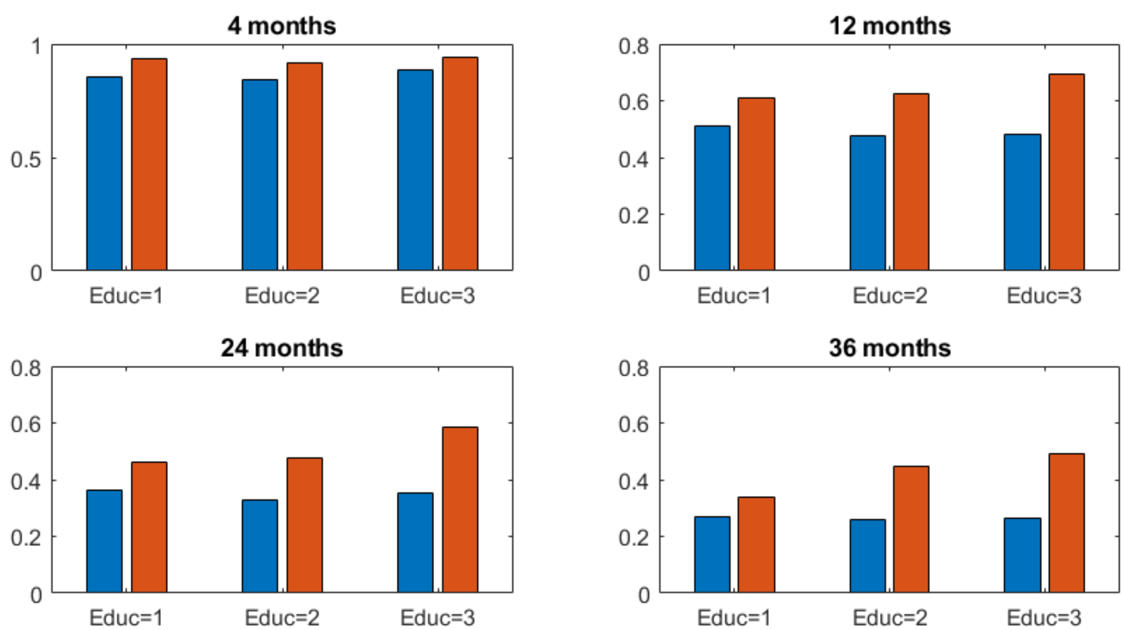


Figure 3.3: Survival Probability by Educational Attainment

Figure 3.4 shows the probability for each possible destination, into employment and unemployment, after a business closure. We observe that unincorporated business owners are more relatively likely to transitioning toward unemployment than incorporated ones, but still the most likely outcome is to find a job before closing the business.

3.3.4 Destination after business closure

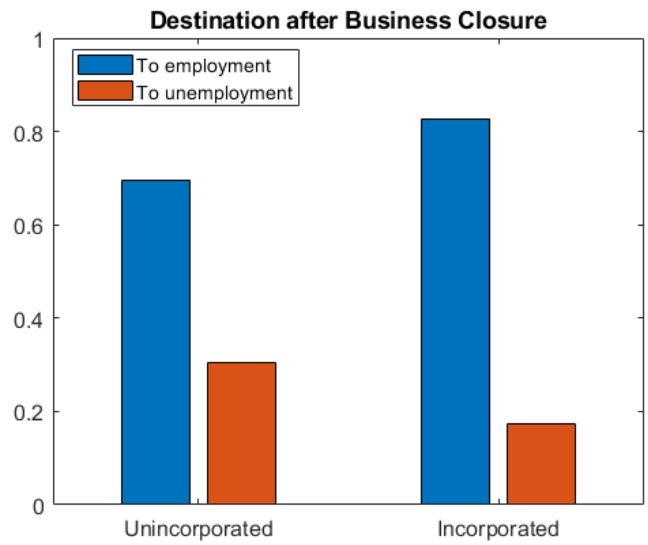


Figure 3.4: Destination after Business Closure

3.3.5 Earnings Change after a Startup Spell

Now we use the three-block spells $\{XYZ\}$ to construct the earning differentials between earnings prior to the business spell and after it. Previous earnings of X corresponds to the average three months before the transition, and earnings in Z are the average earnings three months after the business closure. In the exercises presented below, I'm not controlling for selection, except for the worker characteristics used for the analysis. Therefore, the results cannot be interpreted as causal relationships. Therefore, the goal is just to characterize the earnings dynamics without claiming causality.

Figure 3.5 shows the distribution of the earning differentials for three type of transitions, job-to-job (EE), job-to-unincorporated-to-job (ESE), and job-to-incorporated-to-job (EFE) transitions. These transitions are limited to those for which the medium spell Y lasts at least four months, i.e., the business survives as the main activity of the individual for at least four months. We can observe that the distribution for EE transitions is concentrated in the neighborhood of zero, but the distributions for the other two types of transitions that include a middle spell of self-employment are more disperse. Also, the distribution for the change of earnings for the EFE transition seems to be slightly tilted to the right with respect to the ESE transition, but also slightly more disperse. Next, we turn to disaggregate the analysis by educational attainment to see whether these aggregate distributions hide some kind of heterogeneity.

3.3.6 Business cycle

Now we separate the analysis into two different groups of periods to explore whether the earnings differential varies over the business cycle. Figure 3.6 present the results for the ESE type of transition. The group named Recessions correspond to the periods 2000-02 and 2008-10, and the group Boom to the periods 1996-98, 2003-05, 2011-12. We do not observe much difference between both recessions and booms on the earning differentials for those who had a middle spell of unincorporated business owner.

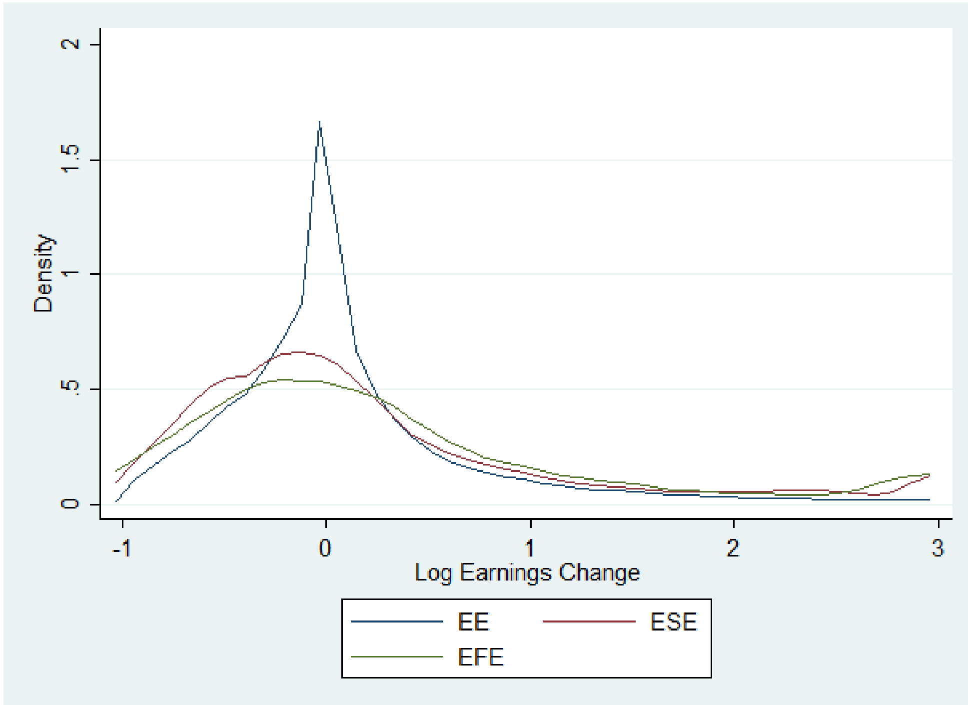


Figure 3.5: Labor Earnings Difference by type Transition

Figure 3.7 presents the same analysis but now for the EFE type of transition. The difference between the distribution of earnings change in recessions and boom is much larger than for the ESE type of transition. We observe that for EFE transitions the distribution of boom tilts to the right while the distribution of recessions slightly to the left. Therefore, it seems that there is an important cyclical component in the earrings change between before and after a self-employment spell when the business is incorporated. This result seems to be in line with what the literature on job ladders and business cycles has found. Haltiwanger et al. (2018) find strong evidence of a firm wage ladder that is highly procyclical. If we think of the decision to start a business as an occupational choice, in the spirit of Garcıa-Trujillo (2020), then a spell as a business owner is equivalent to accepting a job offer from a firm located in a particular rung in the firm wage ladder. Therefore, starting a business in booms can also work as a mechanism that individuals use to climb up the ladder. This seems to be an interesting line of work for future research.

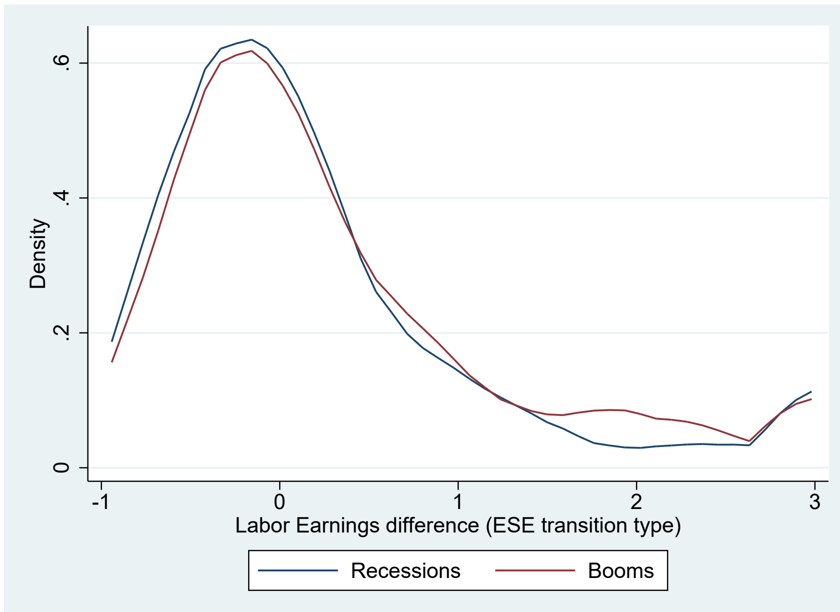


Figure 3.6: Labor Earnings Difference for ESE Transition

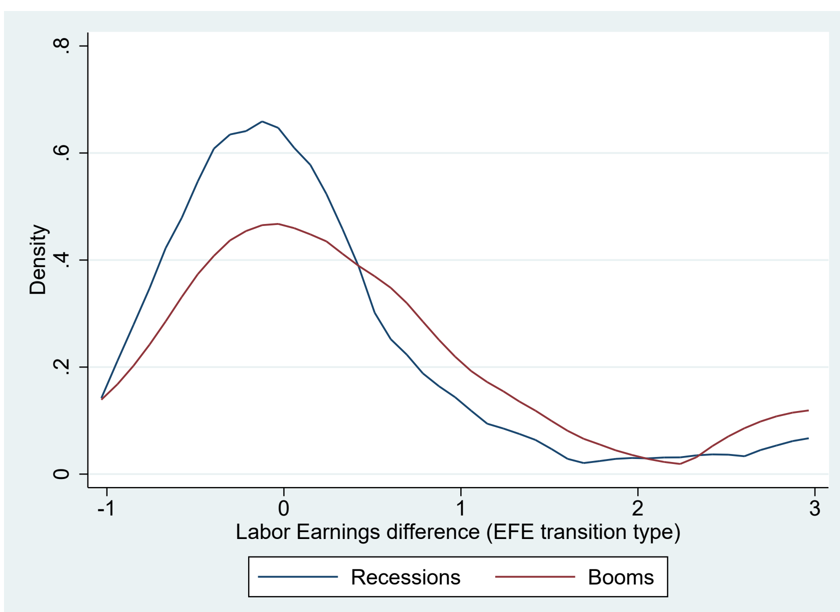


Figure 3.7: Labor Earnings Difference for EFE Transition

3.4 Conclusions

The first goal of this paper is to provide a detailed characterization of the business formation in the United States using data from the SIPP. It shows, on the side of the businesses, differences in the composition of new and established businesses in terms of their legal form of organization and size. On the side of the individuals, the results show that individuals starting incorporated businesses are relatively more educated and less likely to come from unemployment than those starting unincorporated businesses. Also, unincorporated business owners have approximately 40% less previous earnings than incorporated business owners. The analysis done for the probability of survival shows that incorporated businesses are more likely to survive over time than unincorporated businesses, and that probability increases if the owner is a highly educated individual. In terms of the earnings differential analysis, the results show an important source of heterogeneity related to the business cycle. The difference between the labor earnings before and after a self-employment spell is strongly procyclical.

The second goal of this work is to illustrate the potential of a household survey as the SIPP as a source of information to study the early stage of business formation. Its panel property allows to keep track of individuals for up to four years, providing enough time to observe the characteristics of the individuals as workers, their decisions to start businesses, and the events of business closures. Although the SIPP is much smaller than typical administrative data sets, its advantage is that it is easily accessible and provides detailed information at the individual- and firm-level. This might help researchers working in studying business formation at the individual-level to find an alternative to the challenge of identifying business owners in the U.S. Census Bureau's data.

Appendix A: Appendices for Chapter 1

A.1 Autocorrelation: Initial and Future Size of Startups

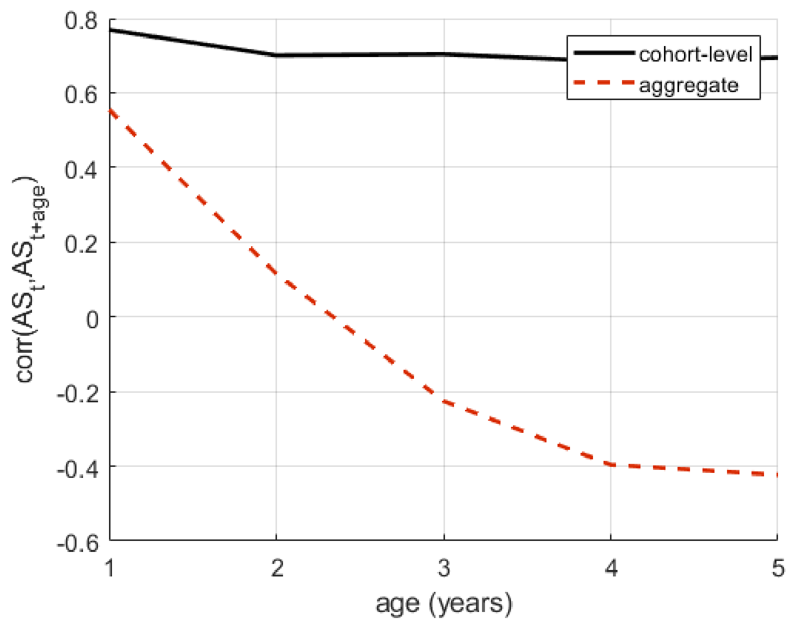


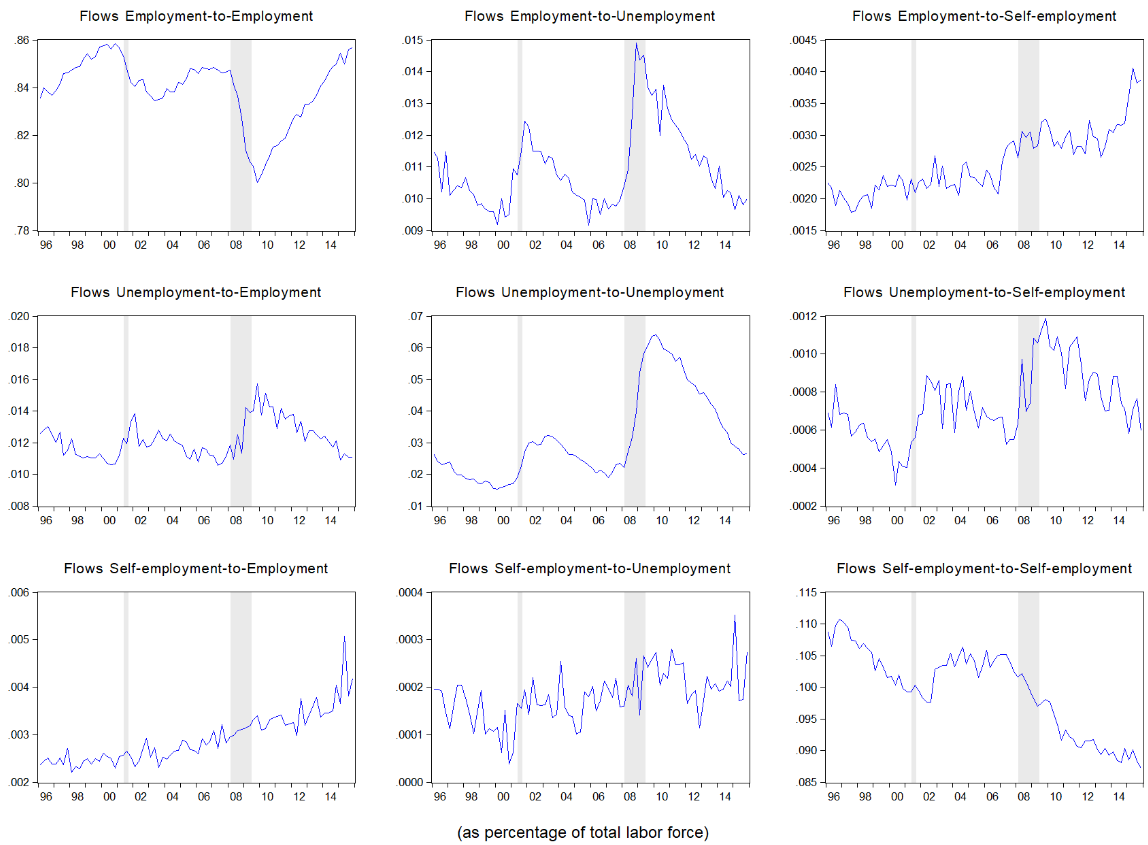
Figure A.1: Autocorrelation: Initial and Future Size of Startups

years window	6-10	11-15	16-20
0-5	0.4196	0.5450	0.4408

Source: author's calculation with BDS data.

Table A.1: Correlation 5-year window: Average size

A.2 Empirical Transition Rates



Transition rates are computed as the number of transitions over the total labor force. All series are seasonally adjusted.

Figure A.2: Empirical Transition Rates

A.3 Equilibrium Conditions

$$
\begin{aligned}
Endogenous\,(18) &= \{W_t, U_t, V_t^E, V_t^U, V_t^S, V_t^F, Opol\_W_t, Opol\_U_t, n_{sub,t}, n_t, w_{ht}, \\
&\quad \theta_{ht}, f_{ht}, q_{ht}, V_{ht}, J_{ht}, \rho_t, \Psi_t\} \\
Exogenous\,(1) &= \{A_t\}
\end{aligned}
$$

$$
\begin{aligned}
W(h,z,n_{-1};A) &= \max[V^W(h,z,n_{-1};A), U(h,z,n_{-1};A)] && \text{(A.1)} \\
DR &: Opol\_W && \text{(A.2)}
\end{aligned}
$$

$$
\begin{aligned}
U(h,z,n_{-1};A) &= \max[V^U(h,z,0;A), V^S(h,z,n_{-1};A), V^F(h,z,n_{-1};A)] && \text{(A.3)} \\
DR &: Opol\_U && \text{(A.4)}
\end{aligned}
$$

$$
\begin{aligned}
V^E(h,z,0;A) &= w_h(A) + \beta[(1-s)\mathbf{E}_{z'/z} W(h,z',n;A') && \text{(A.5)} \\
&+ s\mathbf{E}_{z'/z} U(h,z',o,n;A')]
\end{aligned}
$$

$$
\begin{aligned}
V^U(h,z,0;A) &= Y^{ss} + \beta[f(\theta_h)\mathbf{E}_{z'/z} W(h,z',o,n;A') && \text{(A.6)} \\
&+ (1-f(\theta_h))\mathbf{E}_{z'/z} U(h,z',o,n;A')]
\end{aligned}
$$

$$
\begin{aligned}
V^S(h,z,o_{-1},n_{-1};A) &= A_{sub}A^{\gamma} z n_{sub}^{\alpha} - \rho(A) n_{sub} + \beta \mathbf{E}_{z'/z} U(h,z',o,n;A')] && \text{(A.7)} \\
n_{sub}(z;A) &= \underset{n_{sub}}{\operatorname{argmax}} \left\{ A_{sub}A^{\nu} z n^{\alpha_{sub}} - \rho(A) n_{sub} - g(\chi_{sub}) \right\} && \text{(A.8)}
\end{aligned}
$$

$$
\begin{aligned}
V^F(h,z,o_{-1},n_{-1};A) &= zAn^{\alpha} - \rho(A)n - \phi + \beta \mathbf{E}_{z'/z} U(h,z',o;A')] && \text{(A.9)} \\
n(z,n_{-1};A) &= \underset{n}{\operatorname{argmax}} \left\{ zAn^{\alpha} - \rho(A)n - \phi - g(\chi) \right\} && \text{(A.10)}
\end{aligned}
$$

$$
\begin{aligned}
f(\theta_h) &= \frac{M_h}{S_h} = m\theta_h^{1-\psi} && \text{(A.11)} \\
q^h(\theta_h) &= \frac{M_h}{v_h} = m\theta_h^{-\psi} && \text{(A.12)}
\end{aligned}
$$

$$
\begin{aligned}
V_h(A) &= -c + \beta[q(\theta_h)J_h(A') + (1 - q(\theta_h))V_h(A')] && \text{(A.13)} \\
V_h(A) &= 0 && \text{(A.14)}
\end{aligned}
$$

$$
J_h(A) = \rho(A)h - w_h(A) + \beta[(1-s)J_h(A') + sV_h(A')] \quad \text{(A.15)}
$$

$$
\int h\, d\Psi^W(h, z_t, A) = \int n(h, z_t, A)\, d\Psi^E(z_t, A) \quad \forall t \quad \text{(A.16)}
$$

$$
\begin{aligned}
w_h(A) &= \eta[\rho(A)h + \beta(1-s)\mathbf{E}_{z'/z}J_h(A')] && \text{(A.17)} \\
&+ (1-\eta)[\tilde{U}(h; A) - \beta\tilde{W}_c(h; A')] \\
\tilde{W}_c(h; A') &= (1-s))\mathbf{E}_{z'/z}\tilde{W}(h; A') + s\mathbf{E}_{z'/z}\tilde{U}(h; A')
\end{aligned}
$$

$$
\Psi(x_{t+1}) = T\Psi(x_t) \quad \text{(A.18)}
$$

A.4 Algorithm to solve Stationary Equilibrium

Assumptions:

- Economy is in stationary equilibrium associated with $A = 1$

Algorithm:

1. Set A=1.

2. **Outer loop on $\boldsymbol{\rho(A)}$.** Make a guess for the labor efficiency unit price $\rho(A)$.

3. **Inner loop on $\boldsymbol{w(h, A)}$.** Make a guess for the distribution of wages $w(h, A)$.

   Given the guesses $\rho(A)$ and $w(h, A)$, compute:

   - Labor Market elements: $\theta(h, A), f_h(\theta(h, A)), q_h(\theta(h, A))$ (eq. 6)
   - Do VFI over (eqs. 1-5, 7-8) to compute:
     - Value Functions: $W(h, z, o_{-1}, A)$, $U(h, z, o_{-1}, A)$ and $J(h, A)$
     - Occupational Decision Rules: $o(h, z, o_{-1}, A)$
   - Using Value Functions, compute the implicit wage from Nash Bargaining Solution (eq. 10)
   - Compute $diff_w = w_{implict}(h, A) - w_{old}(h, A)$
   - **Update $\boldsymbol{w(h, A)}$**.If $diff_w > tol_w$, update $w(h, A)$ using a weighted average of previous guess ($w_{old}(h, A)$) and model implied values ($w_{implict}(h, A)$).
   - Go back to step beginning of step 3 and repeat until convergence.

4. **Compute Time Invariant Distribution.** Given the decision rule $o(h, z, o_{-1}, A)$ and the transition matrix for the idiosyncratic productivity Pz, compute the time invariant firm distribution $\Psi(h, z, o)$.

5. **Update guessed $\boldsymbol{\rho(A)}$**:

   - Compute the excess of demand/supply for labor efficiency units (eq. 9)
   - Update bounds of $\rho(A)$ for bisection.
   - Compute $diff_\rho = ub_\rho - lb_\rho$
   - If $diff_\rho > tol_\rho$, go back to step 2 and repeat until convergence.

A.5 Algorithm to solve Transition Dynamics

Assumptions:

- Economy is in stationary equilibrium at $t = 0$, associated with $A = 1$
- Economy is in stationary equilibrium at $t = T$, associated with $A = 1$
- There is a full exogenously given sequence of A_t for t=1...T.

Algorithm:

1. Compute stationary equilibrium with A=1. It corresponds to the initial and final equilibrium. Fix T. Set the exogenously given sequence $\{A_t\}_{t=1}^{T}$, with a 7% decrease in $t = 2$ and a recovery following the following AR(1) process: $A_t = (1 - \rho_A) * \bar{A} + \rho_A * A_{t-1}$.
2. **Outer loop on $\boldsymbol{\rho(A_t)}$.** Make a guess for the path of equilibrium object $\{\rho(A_t)\}_{t=1}^{T}$.
3. **Inner loop on $\boldsymbol{w(h, A_t)}$.** Make a guess for the path of equilibrium object $\{w(h, A_t)\}_{t=1}^{T}$.
4. **Backward Iteration Loop.** Given $\rho_(A_t), w(h, A_t)$ compute for each $t = T...1$:
    - Labor Market elements: $\theta(h, A_t), f_h(\theta(h, A_t)), q_h(\theta(h, A_t))$ (eq. 6)
    - Value Functions: $W(h, z_t, o_{t-1}, A_t)$, $U(h, z_t, o_{t-1}, A_t)$ and $J(h, A_t)$ (eqs. 1-5, 7-8)

- Occupational Decision Rules: $o(h, z_t, o_{t-1}, A_t)$ (from W and U)
- Implicit wage from Nash Bargaining Solution (eq. 10)
- Compute $diff_w = w_{implict}(h, A_t) - w(h, A_t)$
- **Update $\boldsymbol{w(h, A_t)}$** using a weighted average of previous guess and model implied values.
- While $diff_w > tol_w$, iterate until convergence.

5. **Forward Iteration Loop.** Given the initial distribution $\Psi(h, z_1, o_1)$, the sequence of prices $\{\rho(A_t), w(h, A_t)\}_{t=1}^{T}$ and the decision rules $\{o(h, z_t, o_{t-1}, A_t)\}_{t=1}^{T}$, compute the firms distribution forward $\{\Psi(h, z_t, o_t)\}_{t=1}^{T}$.

6. **Update guessed $\boldsymbol{\rho(A_t)}$**:
   - Compute the excess of demand/supply for labor efficiency units (eq. 9)
   - Update bounds of $\rho(A)$ for bisection.
   - Compute $diff_\rho = ub_\rho - lb_\rho$
   - If $diff_\rho > tol_\rho$, go back to step 2 and repeat until convergence.

A.6 Model Implied Transition Rates

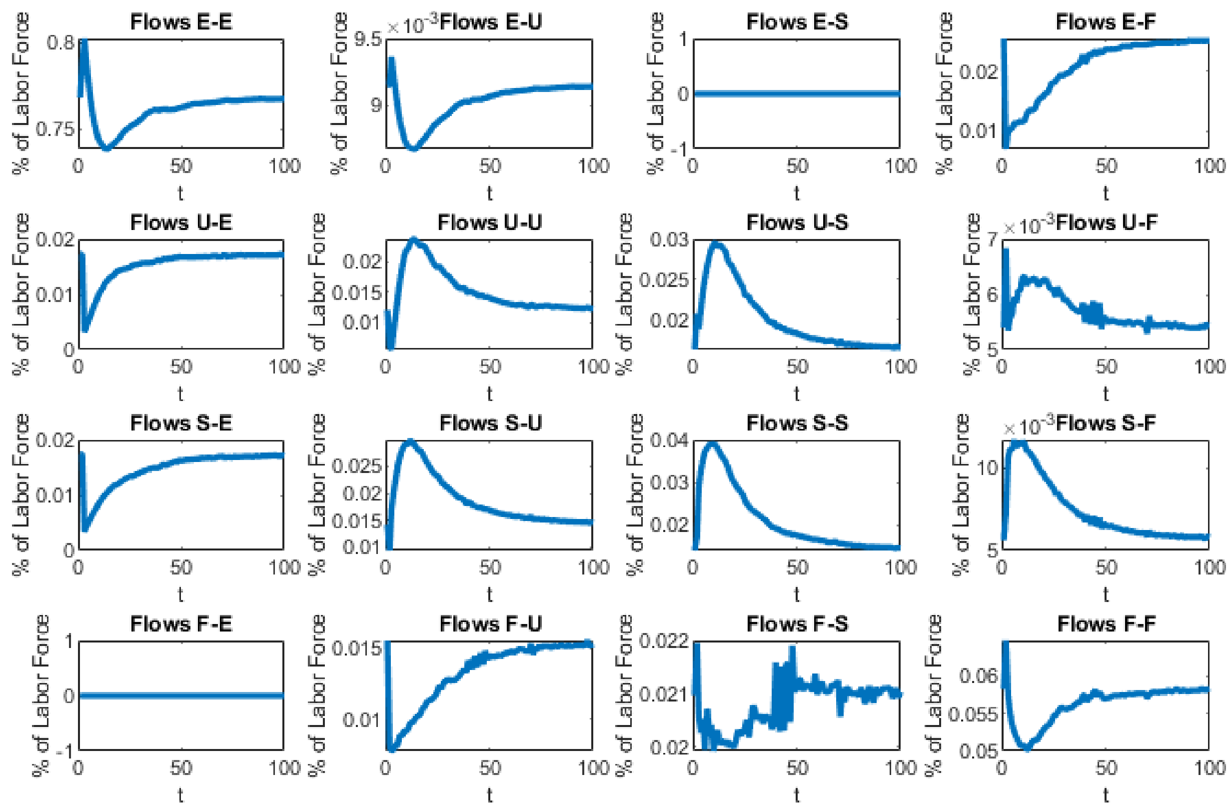


Figure A.3: Model Implied Transition Rates (as % of Labor Force)

Appendix B: Appendices for Chapter 2

B.1 Estimation Results: equation 3

VARIABLES	log(earnings)
2.size_bin	-0.0244***
	(0.00195)
3.size_bin	-0.0264***
	(0.00210)
4.size_bin	-0.0302***
	(0.00212)
5.size_bin	0.0136***
	(0.00222)
6.size_bin	0.0656***
	(0.00218)
7.size_bin	0.140***
	(0.00228)
8.size_bin	0.168***
	(0.00186)
1b.size_bin#1.D1_startup	-0.0852***
	(0.00203)
2.size_bin#1.D1_startup	-0.0502***
	(0.00108)
3.size_bin#1.D1_startup	-0.0887***
	(0.00173)
4.size_bin#1.D1_startup	-0.125***
	(0.00204)
5.size_bin#1.D1_startup	-0.0832***
	(0.00263)
6.size_bin#1.D1_startup	-0.103***
	(0.00325)
7.size_bin#1.D1_startup	-0.0529***
	(0.00403)
8.size_bin#1.D1_startup	0.00467**
	(0.00233)

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Bibliography

Ates, S. T. & Saffie, F. E. (2020). Fewer but Better: Sudden Stops, Firm Entry, and Financial Selection. *American Economic Journal: Macroeconomics, forthcoming.*

Ates, S. T. & Saffie, F. E. (2021). Fewer but Better: Sudden Stops, Firm Entry, and Financial Selection. *American Economic Journal: Macroeconomics.*

Audoly, R. (2020). Firm Dynamics and Random Search over the Business Cycle.

Azoulay, P., Jones, B. F., Kim, J. D., & Miranda, J. (2020). Age and High-Growth Entrepreneurship. *American Economic Review: Insights*, *2*(1), 65–82.

Babina, T., Ma, W., Moser, C., Ouimet, P., & Zarutskie, R. (2019). Pay, Employment, and Dynamics of Young Firms. SSRN Scholarly Paper ID 3425596, Social Science Research Network, Rochester, NY.

Barattieri, A., Basu, S., & Gottschalk, P. (2014). Some Evidence on the Importance of Sticky Wages. *American Economic Journal: Macroeconomics*, *6*(1), 70–101.

Bils, M., Chang, Y., & Kim, S.-B. (2009). Comparative Advantage and Unemployment. SSRN Scholarly Paper ID 1413594, Social Science Research Network, Rochester, NY.

Brown, C. & Medoff, J. L. (2003). Firm Age and Wages. *Journal of Labor Economics*, *21*(3), 677–697. Publisher: The University of Chicago Press.

Brown, J. D., Earle, J. S., Kim, M. J., & Lee, K. M. (2019). Start-ups, job creation, and founder characteristics. *Industrial and Corporate Change*, *28*(6), 1637–1672. Publisher: Oxford Academic.

Burton, M. D., Dahl, M. S., & Sorenson, O. (2018). Do Start-ups Pay Less? *ILR Review*, *71*(5), 1179–1200. Publisher: SAGE Publications Inc.

Carrillo-Tudela, C., Wiczer, D., & Visschers, L. (2019). Cyclical Earnings and Employment Transitions. Technical Report 1548, Society for Economic Dynamics. Publication Title: 2019 Meeting Papers.

Carrington, W. J. & Fallick, B. (2017). Why Do Earnings Fall with Job Displacement? *Industrial Relations: A Journal of Economy and Society*, *56*(4), 688–722. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/irel.12192.

Charness, G. & Gneezy, U. (2012). Strong evidence for gender differences in risk taking. *Journal of Economic Behavior and Organization*, *83*(1), 50–58. Gender Differences in Risk Aversion and Competition.

Choi, J. (2017). Entrepreneurial Risk Taking, Young Firm Dynamics and Aggregate Implications.

Choi, J., Goldschlag, N., Haltiwanger, J. C., & Kim, J. D. (2021). Founding Teams and Startup Performance. Technical Report w28417, National Bureau of Economic Research.

Clementi, G. L. & Palazzo, B. (2016). Entry, Exit, Firm Dynamics, and Aggregate Fluctuations. *American Economic Journal: Macroeconomics*, *8*(3), 1–41.

Davis, S. J. & Von Wachter, T. (2011). Recessions and the Costs of Job Loss. *Brookings Papers on Economic Activity*, 1–72. Publisher: Brookings Institution Press.

Decker, R., Haltiwanger, J., Jarmin, R., & Miranda, J. (2014). The Role of Entrepreneurship in US Job Creation and Economic Dynamism. *Journal of Economic Perspectives*, *28*(3), 3–24.

Decker, R. A., Haltiwanger, J., Jarmin, R. S., & Miranda, J. (2020). Changing Business Dynamism and Productivity: Shocks versus Responsiveness. *American Economic Review*, *110*(12), 3952–3990.

Dinlersoz, E. M., Hyatt, H. R., & Janicki, H. P. (2019). Who works for whom? Worker sorting in a model of entrepreneurship with heterogeneous labor markets. *Review of Economic Dynamics*, *34*, 244–266.

Elsby, M. W. L. & Michaels, R. (2013). Marginal Jobs, Heterogeneous Firms, and Unemployment Flows. *American Economic Journal: Macroeconomics*, *5*(1), 1–48.

Engbom, N. (2019). Firm and Worker Dynamics in an Aging Labor Market.

Erosa, A., Fuster, L., & Kambourov, G. (2016). Towards a Micro-Founded Theory of Aggregate Labour Supply. *The Review of Economic Studies*, *83*(3), 1001–1039.

Fairlie, R. & Fossen, F. (2019). Defining Opportunity versus Necessity Entrepreneurship: Two Components of Business Creation. Technical Report w26377, National Bureau of Economic Research, Cambridge, MA.

Fossen, F. M. (2020). Self-employment over the business cycle in the USA: a decomposition. *Small Business Economics*.

Fujita, S. & Moscarini, G. (2017). Recall and Unemployment. *American Economic Review*, *107*(12), 3875–3916.

Gaillard, A. & Kankanamge, S. (2019). Entrepreneurship, Labor Market Mobility and the Role of Entrepreneurial Insurance. Publication Title: TSE Working Papers.

Galindo Da Fonseca, J. A. (2019). Unemployment, Entrepreneurship and Firm Outcomes. Technical Report 04-2019, Centre interuniversitaire de recherche en économie quantitative, CIREQ. Publication Title: Cahiers de recherche.

García-Pérez, M., Goetz, C., Haltiwanger, J., & Sandusky, K. (2013). Don't Quit Your Day Job: Using Wage and Salary Earnings to Support a New Business.

Garcıa-Trujillo, G. (2020). Startups, Labor Market Frictions, and Business Cycles.

Gavazza, A., Mongey, S., & Violante, G. L. (2018). Aggregate Recruiting Intensity. *American Economic Review*, *108*(8), 2088–2127.

Goldschlag, N., Kim, J. D., & McCue, K. (2017). Just Passing Through: Characterizing U.S. Pass-Through Business Owners. Technical Report 17-69, Center for Economic Studies, U.S. Census Bureau. Publication Title: Working Papers.

Hagedorn, M. & Manovskii, I. (2008). The Cyclical Behavior of Equilibrium Unemployment and Vacancies Revisited. *American Economic Review*, *98*(4), 1692–1706.

Hagedorn, M., Manovskii, I., & Stetsenko, S. (2016). Taxation and unemployment in models with heterogeneous workers. *Review of Economic Dynamics*, *19*, 161–189.

Hall, R. E. (2005). Employment Fluctuations with Equilibrium Wage Stickiness. *American Economic Review*, *95*(1), 50–65.

Haltiwanger, J., Jarmin, R. S., & Miranda, J. (2013). Who Creates Jobs? Small versus Large versus Young. *Review of Economics and Statistics*, *95*(2), 347–361.

Haltiwanger, J. C., Hyatt, H. R., Kahn, L. B., & McEntarfer, E. (2018). Cyclical Job Ladders by Firm Size and Firm Wage. *American Economic Journal: Macroeconomics*, *10*(2), 52–85.

Hombert, J., Schoar, A., Sraer, D., & Thesmar, D. (2020). Can Unemployment Insurance Spur Entrepreneurial Activity? Evidence from France. *The Journal of Finance*, *75*(3), 1247–1285. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/jofi.12880.

Hopenhayn, H. A. (1992). Entry, Exit, and firm Dynamics in Long Run Equilibrium. *Econometrica*, *60*(5), 1127–1150. Publisher: [Wiley, Econometric Society].

Hurst, E. & Pugsley, B. W. (2011). What Do Small Businesses Do?

Hurst, E. & Pugsley, B. W. (2017). *3. Wealth, Tastes, and Entrepreneurial Choice*. University of Chicago Press. Pages: 111-152 Publication Title: Measuring Entrepreneurial Businesses Section: Measuring Entrepreneurial Businesses.

Jarosh, G. (2015). Searching for Job Security and the Consequences of Job Loss.

Kerr, S. & Kerr, W. (2017). Immigrant Entrepreneurship.

Kozeniauskas, N. (2018). What's Driving the Decline in Entrepreneurship?

Krolikowski, P. (2017). Job Ladders and Earnings of Displaced Workers. *American Economic Journal: Macroeconomics*, *9*(2), 1–31.

Levine, R. & Rubinstein, Y. (2018). Selection into Entrepreneurship and Self-Employment. Technical Report w25350, National Bureau of Economic Research, Cambridge, MA.

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

Moreira, S. (2016). Firm Dynamics, Persistent Effects of Entry Conditions, and Business Cycles. SSRN Scholarly Paper ID 3037178, Social Science Research Network, Rochester, NY.

Moscarini, G. & Postel-Vinay, F. (2013). Stochastic Search Equilibrium. *The Review of Economic Studies*, *80*(4 (285)), 1545–1581. Publisher: [Oxford University Press, The Review of Economic Studies, Ltd.].

Moscarini, G. & Postel-Vinay, F. (2016). Wage Posting and Business Cycles. *American Economic Review*, *106*(5), 208–213.

Mueller, A. I. (2017). Separations, Sorting, and Cyclical Unemployment. *American Economic Review*, *107*(7), 2081–2107.

Nakajima, M. (2012). Business Cycles in the Equilibrium Model of Labor Market Search and Self-Insurance*. *International Economic Review*, *53*(2), 399–432. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1468-2354.2012.00686.x.

Nelson, S. C. (2016). Paying Themselves: S Corporation Owners and Trends in S Corporation Income, 1980-2013, 29.

Ouimet, P. & Zarutskie, R. (2014). Who works for startups? The relation between firm age, employee age, and growth. *Journal of Financial Economics*, *112*(3), 386–407.

Pinheiro, R. & Visschers, L. (2015). Unemployment risk and wage differentials. *Journal of Economic Theory*, *157*, 397–424.

Poschke, M. (2013). Who becomes an entrepreneur? Labor market prospects and occupational choice. *Journal of Economic Dynamics and Control*, *37*(3), 693–710.

Poschke, M. (2019). Wage Employment, Unemployment and Self-Employment Across Countries. SSRN Scholarly Paper ID 3401135, Social Science Research Network, Rochester, NY.

Pugsley, B., Sedláček, P., & Sterk, V. (2020). The Nature of Firm Growth. *American Economic Review, forthcoming.*

Salgado, S. (2020). Technical Change and Entrepreneurship. SSRN Scholarly Paper ID 3616568, Social Science Research Network, Rochester, NY.

Schaal, E. (2017). Uncertainty and Unemployment. *Econometrica*, *85*(6), 1675–1721. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.3982/ECTA10557.

Sedláček, P. (2020). Lost generations of firms and aggregate labor market dynamics. *Journal of Monetary Economics*, *111*, 16–31.

Sedláček, P. & Sterk, V. (2017). The Growth Potential of Startups over the Business Cycle. *American Economic Review*, *107*(10), 3182–3210.

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

Shimer, R. (2006). On-the-job search and strategic bargaining. *European Economic Review*, *50*(4), 811–830.

Siemer, M. (2014). Firm Entry and Employment Dynamics in the Great Recession.

Smirnyagin, V. (2020). Compositional Nature of Firm Growth and Aggregate Fluctuations.

Sorenson, O., Dahl, M. S., Canales, R., & Burton, M. D. (2021). Do Startup Employees Earn More in the Long Run? *Organization Science*. Publisher: INFORMS.

Vardishvili, I. (2020). Entry Decision, the Option to Delay Entry, and Business Cycles, 100.