UC San Diego

UC San Diego Electronic Theses and Dissertations

Title

Worker Mobility: Agglomeration, Wage Changes, and Entrepreneurship

Permalink

https://escholarship.org/uc/item/1n339736

Author

Zeng, Wendy Su

Publication Date

2021

Peer reviewed|Thesis/dissertation

eScholarship.org Powered by the California Digital Library
University of California

UNIVERSITY OF CALIFORNIA SAN DIEGO

Worker mobility: Agglomeration, wage changes, and entrepreneurship

A dissertation submitted in partial satisfaction of the
requirements for the degree Doctor of Philosophy

in

Economics

by

Wendy Zeng

Committee in charge:

Professor James E. Rauch, Chair
Professor Lawrence Broz
Professor Munseob Lee
Professor Marc-Andreas Muendler
Professor Johannes Wieland

2021

Copyright

Wendy Zeng, 2021

All rights reserved.

The dissertation of Wendy Zeng is approved, and it is acceptable in quality and form for publication on microfilm and electronically.

University of California San Diego

2021

TABLE OF CONTENTS

Dissertation Approval Page iii

Table of Contents iv

List of Figures vi

List of Tables vii

Acknowledgements ix

Vita x

Abstract of the Dissertation xi

Chapter 1 Worker Circulation and Labor Market Size: Evidence from Brazil 1
1.1 Introduction 1
1.1.1 Literature Review 4
1.2 Data 6
1.2.1 Employment Spells 7
1.2.2 Measuring Labor Market and Firm Sizes 8
1.2.3 Occupation vs. Industry 10
1.2.4 Table of Means 12
1.3 Empirical Methodology 14
1.4 Analysis of Worker Circulation 15
1.4.1 Estimation Results 15
1.4.2 Robustness Checks 17
1.5 Theory 19
1.5.1 Example Setup 20
1.5.2 Comparative statics 21
1.6 Conclusion 23
1.7 Figures and Tables 24

Chapter 2 Climbing the Wage Ladder Through External and Internal Mobility: Evidence from Brazil 38
2.1 Introduction 38
2.2 Data 41
2.2.1 Employment Spells 42
2.2.2 Building wage changes 43
2.2.3 Internal vs External Job Changes 43
2.3 Distinguishing involuntary from voluntary job changes 44
2.4 Wage Changes Resulting from Job Changes 47
2.4.1 Involuntary versus voluntary and external versus internal 47
2.4.2 Months between jobs 48

2.5 Conclusion 50
2.6 Tables and Figures 52

Chapter 3 Former Occupations and Entrepreneurship 57
3.1 Introduction 57
3.2 Data 59
3.3 Managerial Talent and Job Creation 61
3.3.1 Managerial occupations and probability of becoming founding owners . 65
3.3.2 Managerial occupations and firm size 70
3.4 Demand- and Supply-Side Occupations, Entrepreneurship, and Firm Size 78
3.5 Robustness and Extensions 82
3.6 Concluding Remarks 85

Bibliography 86

Chapter 1 Appendix 91
A.0.1 Piecewise Wage Function 91
A.0.2 Appendix Figures and Tables 93

LIST OF FIGURES

Figure 1.1. Map of Brazil 24

Figure 2.1. Histograms with the number of months between jobs, truncated after 13 months 52

Figure 3.1. Shares of workers, founding owners, and new firm employment accounted for by managerial and other occupations 64

Figure A.2. Case 1: Worker mobility when $\beta = 1$ 94

Figure A.3. Case 2: Worker mobility when $\beta < 1$ 95

LIST OF TABLES

Table 1.1. Cell-switching for occ-meso, 25-55 24

Table 1.2. Cell-switching for occ-meso, 16-24 25

Table 1.3. Cell-switching for cnae-meso, 25-55 25

Table 1.4. Cell-switching for cnae-meso, 16-24 25

Table 1.5. Regression results, cell-switching for occ-meso 26

Table 1.6. Regression results, cell-switching for industry-meso 26

Table 1.7. Cell and Regional Statistics 27

Table 1.8. Table of Means for Aggregate 27

Table 1.9. Aggregate regression results, skill firm and urban firm count 28

Table 1.10. Aggregate regression results, regional fixed effects 28

Table 1.11. Aggregate regression results, including cell-switchers 29

Table 1.12. Aggregate regression results, including non-surviving firms 29

Table 1.13. Table of Means for Aggregate, 3occ 30

Table 1.14. Aggregate regression results, skill firm and urban firm count, 3occ 30

Table 1.15. Aggregate regression results, regional fixed effects, 3occ 31

Table 1.16. Table of Means for Aggregate 31

Table 1.17. Aggregate regression results, skill firm and urban firm count 32

Table 1.18. Aggregate regression results, regional fixed effects 32

Table 1.19. Table of Means for Aggregate (CNAE) 33

Table 1.20. Aggregate regression results, industry categorization 34

Table 1.21. Aggregate regression results, industry with regional F.E. 34

Table 1.22. Table of Means for Individual, remove cell-switchers 35

Table 1.23. Individual regression results, skill firm and urban firm count 36

Table 1.24. Individual regression results, regional fixed effects 37

Table 2.1. External and Internal Job Changes Within Cell, by Cause of Separation . . . 53

Table 2.2. Nonemployment Length (in Months) for Job Changes in Table 2.1 53

Table 2.3. Log wage changes for job changes in Table 2.1 53

Table 2.4. Log wage change means with compositional controls 54

Table 2.5. Log wage change means, pooled voluntary and involuntary job changes . . . 54

Table 2.6. Log wage change means, accounting for intervening non-employment 55

Table 2.7. Log wage change means, external, with months between jobs 56

Table 2.8. Log wage change means, internal, with months between jobs 56

Table 3.1. Summary statistics for worker sample used in entrepreneurship regressions 66

Table 3.2. Determinants of entrepreneurship 67

Table 3.3. Mean and median sizes of firms with only one trackable founder 71

Table 3.4. Summary statistics for variables used in new firm size regressions 75

Table 3.5. Determinants of new firm size 76

Table 3.6. Additional occupational determinants of entrepreneurship and new firm size 80

Table 3.7. Occupational determinants of entrepreneurship and new firm size in manufacturing 83

Table A.8. Educational Levels 93

ACKNOWLEDGEMENTS

I am thankful to my advisor, Jim, for his incredible patience, advice, and guidance throughout all these years. It is because of his support that I can be where I am today. I am also grateful for the dedicated time, support, and helpful advice from my committee members Marc, Johannes, Munseob, and Larry, and to several other faculty who have given me guidance.

I am grateful for all of my friends and colleagues at the Econ department, and in particular my entering cohort of 2014 for their support and companionship over the years. Thank you especially to Stephanie and Olga for always being there for me, and to Zhenya for all his support this past year. Finally, I am thankful to my family for always believing in me.

Chapter 1 is currently being prepared for submission for publication. Zeng, Wendy. The dissertation author is the principal researcher on this chapter.

Chapter 2 is currently being prepared for submission for publication. Zeng, Wendy. The dissertation author is the principal researcher on this chapter.

Chapter 3 is co-authored with Muendler, Marc-Andreas and Rauch, James E. and is currently being prepared for submission for publication. The dissertation author is a principal researcher on this chapter.

All errors are my own.

VITA

2014	B.A in Mathematics, Cornell University
2014	B.A in Economics, Cornell University
2015	M.A. in Economics, University of California San Diego
2021	Doctor of Philosophy, University of California San Diego

ABSTRACT OF THE DISSERTATION

Worker mobility: Agglomeration, wage changes, and entrepreneurship

by

Wendy Zeng

Doctor of Philosophy in Economics

University of California San Diego, 2021

Professor James E. Rauch, Chair

This dissertation consists of three chapters that discuss worker mobility in the context of agglomeration, wage changes, and entrepreneurship.

In Chapter 1, we use the Brazilian RAIS data to evaluate the impact of labor market size on worker circulation. Following the literature showing that human capital is more occupation- than industry-specific, I measure labor market size by the log of the number of firms offering jobs within an occupation-region cell. I find that a one standard deviation increase in cell size is associated with a 7.8 percentage point increase in the probability that workers change jobs at least once within a cell, and a 7.6 percentage point increase in the probability that workers who change jobs do so more than once.

In Chapter 2, we document empirical statistics for internal worker job movements within the firm in addition to their external movements between firms using Brazilian RAIS data. We also better distinguish between voluntary and involuntary job separations by utilizing RAIS reporting of cause of separation and length of nonemployment between jobs. Including internal worker movements increases total job separations in the data by 47%, voluntary job separations by 40%, and involuntary job separations by 50%. We find that wage gains are comparable when workers do on-the-job search, transferring internally versus quitting a job and switching externally. The wage losses when workers are fired and return either to the same firm or move to another firm are similarly comparable. We interpret these results as evidence that external and internal labor markets are integrated.

In Chapter 3, we consider founders of limited liability firms who previously held jobs in the formal sector of Brazil. Managers account for four percent of former job holders but nearly one quarter of new firm employment at startup. Little of their overrepresentation as founders or the larger size of their startups is explained by their previous wages or other standard human capital variables. Among non-managerial former occupations, we examined those clearly connected to the demand side (sales) and the supply side (technology, purchasing). Only purchasing was comparable to managerial occupations in entrepreneurship and new firm size.

Chapter 1

Worker Circulation and Labor Market Size: Evidence from Brazil

1.1 Introduction

A well-established stylized fact in urban economics is that larger labor markets and urban areas are more productive (Moretti, 2011; Rosenthal and Strange, 2004). However, the literature is still working on the sources underlying the increased productivity. One way in which larger labor markets are more productive is the increased match quality between workers and firms (Dauth et al., 2019; Wheeler, 2008). One channel through which workers can increase their match quality is through the circulation of workers among firms.

Given the evidence that agglomeration promotes better worker-firm match quality and that worker circulation is a channel to increase match quality, it seems natural to conclude that agglomeration promotes worker circulation. That is, we posit that agglomeration promotes higher match quality through increased worker circulation. However, the relationship between agglomeration and worker circulation is theoretically ambiguous. Agglomeration creates more job opportunities and makes it easier to find them. This facilitates the movement of workers between jobs, but also improves the matches between workers and their initial jobs, making it less likely they will leave their initial jobs.

A number of papers investigate whether workers are more likely to change jobs in denser labor markets (Amarante, da Silva and do Monte, 2019; De Blasio and Di Addario, 2005;

Harmon, 2013; Andersson and Thulin, 2013) and whether they are more likely to change industry in denser labor markets (Wheeler, 2008; Bleakley and Lin, 2012). However, the literature has not investigated whether workers in larger or denser labor markets circulate more within their occupations. This difference is important because results at the region level may be driven by lower search costs. Results within occupation, however, are more likely to reflect better matching rather than lower search costs since workers are able to make better use of their existing skills. Evidence supports that workers make better use of their existing skills when they remain in the same occupation, even relative to remaining in the same industry (Kambourov, Manovskii and Manovskii1, 2009; Sullivan, 2010; Ellison, Glaeser and Kerr, 2010). Then, if workers are employed in jobs that use their skills, it is more likely that workers will be better able to find a quality match.

I use the RAIS Brazilian dataset of matched firm-worker data for the period 2002-2014 to study worker circulation. I track worker movement between jobs, within an occupation, as a function of the number of within-occupation firms as well as a function of the total number of firms in a region. A firm is defined to be within an occupation if it employs workers in that occupation.[1] I find that as the number of within-occupation firms increases, holding regional number of firms constant, workers are less likely to stay in their initial jobs and more likely to change jobs within region and occupation. I also find similar results as the regional number of firms increases holding constant within-occupation firm composition, but the impacts of the number of within-occupation firms are quantitatively larger. I also add regional fixed effects, dropping regional number of firms, to control for better communications and transportation infrastructure in larger or denser regions. I find that the impacts of the number of within-occupation firms are almost unchanged. A standard deviation increase in the number of within-occupation firms, holding the number of regional firms constant, increases the percentage points (ppts) of workers switching jobs at least once by 5.2, and increases the ppts of workers

[1] I will refer to the within-occupation firms in a region as the same-skill labor market for a particular occupation, and refer to the total number of firms in a region as the regional labor market.

switching jobs twice by 8.2. A standard deviation increase in regional firm number, maintaining within-occupation firm composition, increases the ppts switching once by 2.0 and twice by 3.6.

To interpret my results, I develop a model of the impact of agglomeration on worker circulation within skill groups. The literature tends to focus on the reduction in search costs caused by agglomeration. Insofar as this occurs at the regional level, due to better communication and transportation infrastructure or reduced distance between firms, it will be captured by the region fixed effects. The relatively large impacts of the number of within-occupation firms, controlling for regional fixed effects, shows there is an impact of agglomeration at the occupational level that is not captured at the regional level. Though this could again be due to search costs, I focus instead on what I call matching costs. In my model, search costs are zero within the set of firms in a region that employ workers in a given skill group. These firms are located in a space of worker characteristics rather than a geographic space. A firm's location is determined by the skill combination of its ideal worker. A worker with given skills will be employed by the firm closest to him in the space of worker characteristics, all else equal. I show that in this model, the standard assumption that the costs to productivity of distance between worker and firm increase at least linearly with distance (Hotelling, 1929; Salop, 1979; Helpman and Krugman, 1985) yields the results that worker mobility increases in the number of firms. I also construct an example, detailed in the appendix, showing that if the productivity costs of distance increase at a decreasing rate, it is possible for worker mobility to decrease with the number of firms.[2] Intuitively, if the productivity costs of distance increase at a decreasing rate, the effect of decreasing the distance between the worker and the next closest firm is dominated by the effect of decreasing the distance between the worker and his current employer.

[2]Note that because search costs and matching costs are conceptually distinct, it is conceivable that agglomeration increases circulation among all jobs by reducing search costs, but reduces circulation among jobs within an occupation by reducing matching costs. Thus agglomeration may promote career-changing, but not worker circulation within an occupation.

1.1.1 Literature Review

This paper connects to several strands of existing literature. First, it is closely related to the empirical literature that studies whether workers are more likely to switch jobs in a larger regional labor market. The story in this literature is that job density affects circulation through lower search costs; more job options makes it easier to find a job. The correlation between these lower search costs and worker circulation is unclear, with different studies providing mixed evidence.

A recent work by Amarante, da Silva and do Monte (2019) analyze Brazilian municipalities using the RAIS data and find that workers in denser regions have higher inter-firm mobility. Harmon (2013) analyzes job tenure among workers in Denmark and investigates the effect of more unemployed workers in the local region, controlling for labor market tightness. He finds a positive correlation between larger labor markets and job tenure (a dummy variable for whether a job lasts more than two years), suggesting lower circulation. De Blasio and Di Addario (2005) run a Poisson maximum-likelihood regression on the number of jobs an individual holds over his entire work life and find mobility to be slightly higher for blue-collar workers but overall insignificant when all worker data is combined. Andersson and Thulin (2013) investigate the likelihood that a worker switches employers within a year depending on employment density. They find that doubling the density leads to a 0.2% increase in the probability of employer switching. These papers study search cost reduction by exploiting the variation in the density of the regional labor market, which I include as a control variable. I build upon this literature by focusing on what I call a reduction in matching costs rather than in search costs, in which having more jobs around allows workers to choose jobs closer suited to their skills. To do this, I exploit the variation of density of the same-skill labor market, which is a smaller and more specific subset of the regional labor market. In addition, I impose a restriction to workers who stay within the same labor market throughout the sample period to address the migration channel. Furthermore, I use a more detailed dataset with a longer time frame and individual worker

tracking that can tell if workers experience multiple job changes or stay at their initial job over 13 years[3].

This paper complements the empirical literature on worker mobility across industries in a larger labor market. Bleakley and Lin (2012) and Wheeler (2008) investigate the rate at which workers in US cities change occupation and industry based on population density. They find that workers switch occupation and industry less in denser areas, although this result is reversed for young workers. While they study movement between industries as it relates to market size, I study movement only within industry (and location) in relation to same-skill market size[4]. Their results complement my findings that workers change more within occupation when the labor markets are larger.

This paper also relates to the literature on job circulation and knowledge spillovers in high-tech clusters. Fallick, Fleischman and Rebitzer (2006) study the relation between IT clusters in different states and the probability that a worker employed for two consecutive months in the high-tech computer industry changes employers within that time frame. They find a "California effect" leading to increased worker mobility compared to other states, attributed to the lack of enforcement of noncompete statutes. Touching on the focus of this paper, they also find that density of employment in the cluster and within the industry have weak effects on month-to-month worker mobility. I hope to show broader insight into this question by including labor markets of all sizes and skills in the analysis.

Finally, this paper builds upon the literature that studies if workers better retain their capital when they remain in the same industry versus remaining in the same occupation. Traiberman (2019) studies reallocation costs and finds large costs of occupational switching compared to industry switching. Workers build specific human capital over time that is destroyed when they must switch occupations. Specifically, Traiberman finds that 57% of the variation in lifetime

[3]The De Blasio and Di Addario (2005) data also includes lifetime number of jobs, but this includes temporary jobs with no method to exclude these jobs. In my data, I attempt to take out temporary jobs by only including jobs with tenure over 3 months.

[4]Note that in their narrative, workers use the larger labor market to better learn their type, while my narrative assumes that workers already know their type since they only move within occupation.

earning changes is accounted for by occupations compared to only 16% by sectors, suggesting that workers better retain their human capital when they remain in the same occupation. Further evidence in the literature also suggests that workers retain more of their skill when they remain in the same occupation as opposed to when they remain in the same industry, although this varies among industries (Kambourov, Manovskii and Manovskii1, 2009; Sullivan, 2010). I use these insights and sort labor markets by skill using occupational categorization, leaving the industry variable as a robustness check. Analysis of the RAIS data also shows that workers are more likely to remain in labor markets partitioned by occupation than labor markets partitioned by industry.

Section 1.2 of the paper describes the data. Section 1.3 describes the methodology, while section 1.4 provides the empirical interpretation of the results. Section 1.5 discusses the model, and section 1.6 concludes.

1.2 Data

The RAIS data consist of linked employer-employee records on socio-economic information from the Brazilian labor ministry. Each year contains observations of that year's reported jobs in the formal labor market and includes variables such as employee ID, age, salary, dates of employment, employer ID, etc.[5] Our data cover the period 2002-2014.

We use the data to identify formal employment spells for each worker. From the RAIS data, we use the variables age, education, tenure, municipality, year, sex, date of birth, cause of leave, worker ID, occupation, industry, month of admission, month of leave, employer ID, monthly average wage, and hours contracted per week. Employer ID is a 14-digit code. The first 8 digits give the firm ID which we use to identify a worker's firm. The remaining 6 digits identify the plant ID, which we drop. We also have data for the yearly minimum wage and can map each municipality to its respective meso and microregion. The tables of means for these variables and others are provided at the end of this section.

[5] We do not observe workers who work in the informal sector, which comprises around 1/3 of Brazil's workforce.

1.2.1 Employment Spells

The goal in the data is to identify employment spells for individual workers. We limit our observations to prime age males and their primary, long-term jobs. To this end, we restrict the data to males ages 25-55 [6] and employment spells in which they work at least 20 hours a week and, if they left the job, worked there at least 3 months. [7]

Our first concern is workers who might report under another worker's ID. Since we base employment/unemployment spells on the observations under a given worker ID, we try to ensure that no more than one worker is reporting their employment history under a single worker ID. We remove workers who report as female at any point, even if the majority of that worker's ID reports as male. We also remove those who report different birth dates or who never report their birth date in any of their observations. We also remove workers if their education variable is not reported for any observation since we would like to control for education but would ideally want a worker's full work history. Simply removing missing education observations would give us an incomplete work history.

Now, we can build employment spells. We first build the employment spells that have terminated, i.e., the worker left the job. RAIS reports a variable, cause of leave, that is positive if a worker left his job that year[8] and zero otherwise. For example, say that a worker holds a job from 2002 until 2004, when he is fired. He then takes another job from 2007 and leaves in 2008. Finally, he begins working at yet another job in 2010, and stays there until the end of the sample period. He will be in the RAIS data for the years 2002-2004, 2007-2008, and 2010-2014. The observations in 2002, 2003, 2007, and 2010-2014 will report zero for cause of leave because he did not separate from his job in those years. The observations in 2004 and 2008 will report

[6] The legal working age in Brazil begins at age 16, while most of the population retires from the working force around age 65.

[7] Amarante et al. (2019) differ from our treatment of the RAIS data. They include women and use a balanced panel from 2003-2013, throwing out workers who do not have observations for each of those years. However, the use of a balanced panel, presumably to get rid of workers who may disappear into the informal sector, creates selection and describes a very specific sample of workers who were able to find employment in all years of the period, while our analysis is relevant for a broader population.

[8] The different values of cause of leave indicate the reason for the job separation.

a positive value for cause of leave, which we obtain by taking only values for which cause of leave is positive. This gives us a unique observation for each job match that the worker has left. Then we take the month of separation and the tenure variable (in months) to get the start and end dates of the job and we have the employment spells that have terminated. If the worker reports 0 for the month of separation even though cause of leave is positive, we assume their month of separation to be in December.

This leaves us to build the employment spells that have not terminated. We take all observations in the year 2014 with zero values of cause of leave to get all of the employment spells that have not terminated. In our example, we get the worker's 2014 observation. To calculate the start dates, we subtract tenure from December 2014.

Now that we have the employment spells, we need to order the employment spells and address spells with overlapping dates. For each worker, we sort their spells by order of earliest admission date and then by longest tenure, for cases with identical admission dates. Going down this ordered list, the worker's last employment date among all the previous jobs is recorded into the observation.

We do not analyze within-firm transfers as separate employment spells. As determined by the ordering above, if a worker is employed at the same firm for consecutive employment spells, we merge the employment spells into a single spell, retaining the earliest job start date as well as the latest job end date among the consecutive employment spells.

1.2.2 Measuring Labor Market and Firm Sizes

We partition labor markets by worker occupation[9] and location. Ideally, the goal is to define a labor market as a market in which, if a worker were to switch from one job to another, the worker would be able to find a new job without having to pay the costs of either moving or retraining. This is also the sense in which agglomeration economies are the product of local labor markets.

[9]The next subsection discusses the choice between occupation and industry

Occupation is aggregated at the 2-digit level, out of 5 digits, based on Brazil's occupation classification system from 1994[10]. The first digit characterizes different areas of work, such as science vs law. The 2-digit level is a fairly broad categorization, such as scientist vs. engineer. The third digit might differentiate between different fields of science, such as a chemist vs a physicist. The 4th and 5th digit together characterize this even further into, for example, different types of physicists i.e. nuclear vs thermal physicist. We choose the 2-digit occupation level which would be similar to grouping all the scientists from the hard sciences, or grouping together doctors and nurses, where the professions, level of education, and training can vary quite significantly. There are 84 different categories in the data at the 2-digit occupation level. We group at the 3-digit level as well for robustness.

Location is aggregated at the meso region level. The meso level is a fairly large region. Figure 1 shows two maps of Brazil. The left figure shows the entire country of Brazil, with the 5 regions outlined in black borders, the states partitioned in dark grey borders, and municipalities in white borders. To get an idea of the size of a meso region, the right figure shows the state Rio Grande de Sul, with one of the meso regions colored in red. In the right figure, the other meso regions are of similar size, drawn with black borders. The meso region is partitioned with lighter grey borders into micro regions, while the municipalities are drawn in the lightest grey. Since we have municipality information for each worker, we map each municipality to the larger meso regions. We also partition location at the micro region level for robustness.

Thus, our labor market cell becomes defined as a two-digit occupation code at the meso level.

We measure the size of the labor market cells by counting the number of jobs (*n_jobs_8*) and the number of firms (*n_firms_8*) existing within our defined universe of employment spells for each occupation-region cell in December 2008, and take the log (*l_njobs* and *l_nfirms*). The number of firms is the number of unique firm IDs reported by workers in a labor market. The

[10] A new occupation code system was introduced in 2002 that saw a gradual implementation until 2010, so for observations that use the new 2002 occupational code, we do a mapping back onto the occupational code system from 1994.

choice of the month December is due to a thirteenth salary, a special December payment made in some sectors for workers reported in RAIS, which incentivizes workers to appear in the RAIS data in December.[11] We measure the size of the regional labor market by counting the number of jobs and firms (*n_urbanjobs* and *n_urbanfirms*) for each meso in December 2008 and take the log (*l_urbanjobs* and *l_urbanfirms*). The log of the ratio of jobs to firms gives the log average firm size in the occupation-region cell and meso (*l_avgskillfirmsize* and *l_avgurbanfirmsize*). If there were no jobs reported in December 2008, the missing value for both number of jobs and number of firms was replaced with a value of 1.[12]

We calculate individual firm sizes and the log, *firm_size* and *l_firmsize*, by counting the number of workers in our universe at a firm for a given year in December. From this we can get a firm's average size, *avg_size* and respective log *l_avgindfirmsize*, over all the years it employs at least one worker in December, as well as the ratio, *firmsizeratio*, of a firm's current size to its last observed size in December 2014, *size_2014*. Summary statistics for labor market sizes are included in the tables of means at the end of this section.

1.2.3 Occupation vs. Industry

RAIS reports both a worker's occupation and industry, so we choose the variable that best captures our notion of a labor market in the sense that workers can change jobs within this categorization without a wage decrease due to skill differential. The literature highlights the significance of occupational tenure on human capital. The RAIS data can also show the labor market partition, whether by occupation or industry, that is best able to retain its workers after a job change.

We can get a sense of how fluidly workers move between labor markets by looking at three categories: the percentage of workers who stay in the same meso location, the percentage

[11] The number of firms is the number of firms with positive employment in an occupation or industry in December 2008.

[12] Because some workers quit their jobs before December 2008 or start their jobs after this date, it is possible for the size of a worker's labor market cell to be zero.

who stay in the same two-digit occupational or industry classification, and the percentage who stay in the same labor market cell (both region and occupation/industry) when workers transition between jobs. We run regressions on the effect of the log size of the labor market on these percentages with municipality and industry fixed effects. We separate workers into two age groups to consider the possibility that younger workers differ in job circulation behavior from experienced workers in their prime age. A younger worker looking for experience and trying out different types of jobs might be willing to switch jobs, location, and even occupation more often than an older, more experienced worker. These older workers are more likely to be settled down and unwilling to uproot from a certain location. They might have accumulated experience in a particular occupation and be less willing to start afresh. Therefore, we might expect older workers to be more likely to stay within the same labor market. We set the age bracket for young workers to be 16-24 and prime age workers to be 25-55.

Tables 1.1-1.4 report the table of means for occupation and industry categorization for prime age and young workers. In Table 1.1, *percent_occ* is the percentage of job changes that stay within the 2-digit occupation for all prime age workers, *percent_muni* the percentage that stay within the meso region, and *percent_muniocc* the percentage that stay within the same meso region and occupation cell. Table 1.2 reports the means for occupation and meso cells for the young, and Table 1.3 and 1.4 report the means for industry and meso cells. The number of observations reports the number of unique ind/occ-location cells that have at least one worker who changed jobs.

Across both groups, the means of staying in cell are greater for occupation than for industry. Prime age workers have a higher mean of staying in cell of 15% for occupation-meso cells compared to 12% for industry-meso cells. The means are also higher for young workers, with 11% for occupation compared to 9% for industry. Occupation-meso cells are on average smaller than industry-meso cells since there are 11,187 occupation-meso cells compared to 7,241 industry-meso cells for prime age workers. All else equal, staying within the occupation-meso cell should be less likely. This makes the larger mean of staying in the occupation-meso cell

more significant, implying that staying within occupation is strongly favored over staying within industry.

Increasing the size of the same-skill labor market is correlated with an economically significant higher likelihood of staying within the same occupation-meso or industry-meso cell. Table 1.5 reports that a unit increase in the log number of same-skill firms increases the job changes within cell by 4.7 ppts, holding average firm size constant. The distribution of firm sizes across occ-meso cells has a standard deviation of 1.97 according to Table 1.7. I compute that a standard deviation increase in size leads to a 9.26 ppt increase in the job changes that remain within cell, nearly two thirds the mean of staying in the occ-meso cell. From Table 1.6, a standard deviation increase in size leads to a 9.96 ppt increase in job changes that remain within cell, which is actually higher than the mean of staying in the industry-meso cell. I conclude that in both measures, the impact of labor market size on probability of staying within cell is large and significant. I will partition the labor markets primarily by occupation, and leave industry partitioning for robustness.

1.2.4 Table of Means

Table 1.8 gives the table of means for aggregate observations at the cell by cohort year level. The cohort year is the start year of each worker's initial observable job in the data. Workers are grouped by the occupation-meso cell of their first job and by cohort year. Those with a cohort year before 2002 are dropped, so that there is a group of workers who entered the labor market in 2002, 2003, and so on until 2014 for each cell. There are 123,314 observations from 11,356 cells and 13 cohort years.

P*ercent_stay* is the percentage of workers in a group that stay in their initial job until the end of the sample period in 2014. To illustrate how I calculate this percentage for workers in cohort 2004 in a given cell, I first take all workers with cohort year 2004 who started their first job in that cell. I drop workers who retired or died by the end of 2014, took a job outside of the occ-meso cell, or started at a firm that did not survive until the end of 2014. This is done

to address the issue of cell-switching and because individuals at firms that don't survive must leave their first job, but not in the spirit of circulation to find a better match. I then count the workers who stay at their initial firm until 2014 and divide by the number of workers. Similarly, *percent_change1* is the number of workers that change jobs at least once during the sample period divided by the number of workers. *Percent_change2* is the number of workers that change jobs at least twice over the sample period, divided by the number of workers that change jobs at least once. A*vg_changes* counts the total number of job changes divided by the number of workers. These calculations are then repeated for all the cells for cohort 2004, and then repeated for all of the cohort years.

Only 112,639 of the cells have non–cell-switchers in the data. Of these, the mean cell has a high percentage of workers who remain in their first firm at 56% and very few job changers at 7%, with the average number of job changes per person at 0.1. Less than half of the cells contain workers that change jobs at least once, but among those, 19% change jobs twice in the mean cell.

Variants of these variables have *_m* for measures that do not drop cell-switchers and *_f* for measures that do not drop workers who started at firms that did not survive until the end of 2014. Including cell-switchers drops the percentage of stayers to 41% and significantly increases the mean percentage of changers to 38%.

Table 1.22 gives the table of means at the individual worker level, dropping cell-switchers and workers who started at non-surviving firms and those with cohort years before 2002 for nearly 6 million observations. We drop workers if their firm ID is missing. Similar to Table 1.8, *stay* is a dummy for whether the worker stayed at his initial job, *change1* for whether the worker changed jobs at least once, *change2* for whether the worker changed firms at least twice given one change, and *changes* is the total number of job changes over the sample period.[13] 52% of the workers stayed at their jobs the entire period while 15% changed jobs at least once, leaving about 33% of the workers leaving their initial job but not finding another job in the formal sector.

[13]Note that these two variables stay and change1 are not the same because a worker can quit his job but not find another job, giving him a value of 0 for both dummies.

Out of the workers who changed jobs at least once, almost a third changed jobs at least twice.

Age is given at the time of the cohort year. Education, numbered 1 to 11, represents different educational attainments detailed in the appendix. The relevant educational levels are 5 (completed middle school), 7 (completed high school), and 9 (completed college). The average worker in our sample has completed a little more than a middle school education. The size of the firm and ratio to its size in 2014 is from the cohort year.

1.3 Empirical Methodology

The different measures of worker circulation include the probability of staying at the same firm, changing jobs at least once, changing jobs more than once, and the number of job changes. To estimate the impact of the same-skill labor market size on circulation, I include the job and firm sizes of the skill-specific labor market:

$$y_{jt} = \beta_1 \log(n_{1j}) + \beta_2 \log(n_{2j}) + \alpha_{1t} + \alpha_{2k} + \alpha_{3m} + \varepsilon_{jt} \tag{1.1}$$

where j denotes the occ-meso cell, t the cohort year, k the occupation, and m the meso region. The left hand side variable y is one of the circulation measure variables. On the right hand side, n_{1j} and n_{2j} are the number of firms and jobs in the skill-specific labor market, respectively. α_1, α_2, and α_3 are cohort, occupational, and regional fixed effects. Standard errors are clustered at the cell by cohort year level.

The number of firms and jobs are highly correlated, so I then instead compute the implied effect of market size by adding $\beta_1 + \beta_2$ which introduces a variable for average same-skill firm size. We can rewrite the equation to make the effect of labor market size more obvious:

$$y_{jt} = \underbrace{(\beta_1 + \beta_2)}_{\text{labor market size}} \log(n_{1j}) + \underbrace{\beta_2}_{\text{avg firm size}} \log(\frac{n_{2j}}{n_{1j}}) + \alpha_{1t} + \alpha_{2k} + \alpha_{3m} + \varepsilon_{jt} \tag{1.2}$$

$\beta_1 + \beta_2$ measures the effect of increasing the number of same-skill firms, keeping constant the average same-skill firm size.

I also estimate the regression adding the regional job and firm size, which requires dropping the regional fixed effects. Again, the implied effect of market size requires adding the two coefficients for job and firm size:

$$y_{jt} = (\beta_1 + \beta_2)\log(n_{1j}) + \beta_2 \log(\frac{n_{2j}}{n_{1j}}) + (\beta_3 + \beta_4) log(n_{1m}) + \beta_4 log(\frac{n_{2m}}{n_{1m}}) + \alpha_{1t} + \alpha_{2k} + \varepsilon_{jt} \tag{1.3}$$

where n_{1m} and n_{2m} are the number of firms and jobs in the regional market. Here, we still measure the effect of increasing the share of the same-skill labor market, holding constant the regional market size. Another interesting analysis is to look at the effect of increasing the regional market size, while maintaining the composition of the same-skill labor market in the regional market. To do so, we can add the coefficients of both the same-skill market size and the regional market size. This sum, $\beta_1 + \beta_2 + \beta_3 + \beta_4$, gives us the effect of increasing the number of firms in the region, controlling for the market share of the firms in the occ-meso cell, the average same-skill firm size, and the average regional firm size. We rewrite the regression equation (3), rearranging the variables and changing to coefficients to make it look simpler:

$$y_{jt} = \gamma_1 \log(n_{1j}) + \gamma_2 log(n_{1m}) + \gamma_3 \log(\frac{n_{2j}}{n_{1j}}) + \gamma_4 log(\frac{n_{2m}}{n_{1m}}) + \alpha_{1t} + \alpha_{2k} + \varepsilon_{jt}. \tag{1.4}$$

1.4 Analysis of Worker Circulation

1.4.1 Estimation Results

The results for the aggregate level regressions are reported in Tables 1.9 and 1.10, where we look at the impact of number of within-occupation firms on the four measures of circulation.

Table 1.9 reports the estimates from the regression equation (4). Increasing the size of the same-skill labor market is correlated with increased circulation for all four measures. A unit increase in the log number of same-skill firms decreases the amount of workers who stay at their jobs by 2.39 ppts. The standard deviation for firm size distribution across occ-meso cells is 1.97, and the standard deviation of average percentage that stay across cell by cohort groups is 0.35. I compute that a one standard deviation increase in same-skill firm size leads to a 4.71 ppt decrease in the amount of workers who stay, or a 13% decrease in standard deviation. Table 1.9 also reports the sum of the coefficients of the same-skill labor market size and regional labor market size as b1+b2. A unit increase in the log number of regional firms decreases the amount of stayers by 2.24 ppts. The standard deviation for firm size distribution across meso regions is 1.3. Then a standard deviation increase in regional firm size, holding share of same-skill firms constant, leads to a 2.91 ppt decrease in amount of stayers, about 8% of the standard deviation.

In summary of Table 1.9, a standard deviation increase of 1.97 in the same-skill market size increases the amount of changers by 5.16 ppts, nearly half a standard deviation increase. For workers who change at least twice, the increase is 8.23 ppts, almost a third of a standard deviation increase. Finally, the increase in average changes per person is 8.18 ppts, nearly half a standard deviation increase. The effects of increasing regional market size holding share constant lead to smaller increases, but still significant.

Table 1.10 reports the estimates with regional fixed effects to account for any regional differences in size or communications and transportation infrastructure. The coefficients are nearly identical, which implies that most of the variation in regional fixed effects is captured by the size of the regional labor market. The coefficient for *percent_stay* does drop by less than a half. However, the coefficient for this measure drops very little at the individual level regressions, and I will investigate more into this drop in the aggregate level regressions in future work. For all four measurements of worker circulation, increasing market size seems to encourage worker circulation, with increasing the size of the skill specific market being more impactful.

1.4.2 Robustness Checks

Including cell-switchers and workers in non-surviving firms

To address the worry that switching out of the occ-meso cells causes the correlation, I include cell-switchers in the regression data to show that the probability of any job change is still increasing with the number of firms. The sign and significance of the coefficients does not change. I also then include workers whose first firm no longer existed in 2014, dropping cell-switchers again, and see that the sign and significance of the coefficients does not change.

3-digit occupation partitioning

I then use different partitioning to create the same-skill labor market cells. First, I partition by 3-digit occupations to create 3-digit occ-meso cells. This partitioning quadruples the number of cells from the original 2-digit occ-meso partitioning to 42,968 cells, and the standard deviation in cell size stays mostly the same at 1.92. The estimates from the regression are shown in Table 1.14. The coefficients decrease a little in magnitude in every regression but remain significant.

Micro region partitioning

I then partition by micro regions instead of meso regions to create occ-micro cells, again increasing the number of cells to 43,445 with a standard deviation in market size of 1.83. The coefficients in the regression, shown in Table 1.17, are actually very similar to those in the original regression, especially for *percent_change1* and *percent_change2*.

2-digit industry partitioning

I also partition by the 2-digit industry code to create industry-meso cells. The number of cells decreases to 7,371 and the standard deviation in cell size is 2.12. The coefficients are very similar as when partitioning by occ-meso, looking at Table 1.20.

Individual level regressions

I run the regressions at the individual level, reported in Tables 1.23 and 1.24. The coefficients on same-skill firm size is higher for change1 and changes.

The average firm size in the skill specific labor market follows the sign of the coefficient for market share, suggesting that a larger average firm size results in workers being more likely to change jobs. The coefficient and standard deviation are smaller so the relation is weaker, and the coefficient for likelihood to stay is insignificant. The average firm size in the regional labor market seems to have the opposite effect, though the effects are still smaller and less significant.

Regarding age, workers entering the labor market at an older age are more likely to stay at their job and also less likely to change jobs at all levels. A standard deviation increase in age of 8.58 years corresponds to being 0.043 more likely to stay at the job and 0.083 less likely to experience a job change. The coefficients in all four regressions are significant.

Ratio, which compares the firm size in 2014 to the size when the worker entered, follows the same trend as age. Workers who enter a firm that is growing is more likely to stay at that firm until the end of the sample period, and also less likely to change jobs once. It is harder to interpret the coefficients for workers who change jobs more than once since we need to also consider the following firm the worker matches with.

1-digit occupation partitioning

To address concerns regarding the high mobility rates across occupational categories and cells, I partition the cells by the 1-digit occupation and meso regions. There were 9 occupational categories and 137 meso partitions, resulting in 1,233 cells with a standard deviation in cell size of 1.44.

Once we defined a broad occupational categorization, the mobility between occupation and cells is reduced, with 45% of job changes staying within the same occupation and 29% staying within cell. In the 2-digit occupation partitioning for prime age workers, an average across cell of 25% of job changes stayed within the same occupation, and 15% in the same cell.

In the 3-digit occupation partitioning, 18% stayed within occupation and 11% in the same cell.

Coefficients in the circulation regression remain significant in the same direction as with the various previous partitioning. Compared with the 2-digit occupation and meso partitioning, an increase in a standard deviation in cell size brings about similar decrease in probability of staying and changing jobs twice given one change. There is a much higher probability of changing jobs at least once, with a 8.74 ppt increase compared to 5.16 ppt increase.

1.5 Theory

To formally investigate what drives the results in the data, I develop a model of the impact of agglomeration on worker circulation within a skill group. The theory embeds a model of worker preferences among different firms in a location model set in characteristics space. In this model, there are a finite number of skill-specific firms in a space. To abstract from search costs and focus on matching costs, distance between these firms measures their distance in characteristics space, and information is perfect.[14] It follows that increasing firm density decreases the distance between neighboring firms. Workers are distributed uniformly over this space. A measure of the closeness of the skill match between a firm and worker pair is determined by the inverse of their distance. Increasing distance raises worker productivity costs which lowers wages. Firms and workers are static and cannot change their location.

In the first period, workers choose jobs at the closest firm in the characteristics space. Then in the second period, a share $s << 1$ of firms receive a productivity shock. For now, we assume that if a firm receives a productivity shock, its closest neighboring firms do not receive a shock. We leave investigation of correlated shocks to future work. Workers matched with negatively shocked firms may switch to neighboring firms, depending on the closeness of the initial match versus the distance to the neighboring firms. On the other hand, workers matched

[14] Imagine that all of these firms are located in a large mall. Packing more firms into the mall has no effect on search costs because they are all listed in the directory, which is costless to peruse. However, worker circulation is affected because the worker considers the distance between his skills and his initial job in this characteristics space in comparison to the distance between his skills and the other firms.

with firms neighboring a positively shocked firm may switch to the positively shocked firm, again depending on the closeness of the initial match versus the distance to the positively shocked firm.

We will compare worker mobility in response to shocks across regions that differ in firm density. Regions may also differ in number of workers, but that will not affect our results. In particular, the probability that a worker switches firms in response to a negative shock to his firm or a positive shock to a neighboring firm is not affected by the number of workers per firm.

In the first period, workers in a region with higher firm density will be closer on average to the firms with which they are matched and therefore have higher wages. This suggests they will be less willing to switch firms when shocks arrive in the second period. On the other hand, the neighboring firms to which they can switch to are also closer in characteristics space. We will show that, in a standard setting, an increase in firm density unambiguously increases the probability that a worker will switch firms, as we find in our empirical results. Our model will also provide intuition for this strong result. In the Appendix, we will show that it is possible to reverse this result by moving away from the standard setting in a particular way. Again, our model will provide intuition for how the result can be reversed.

1.5.1 Example Setup

We represent characteristics space by a circle of unit circumference. There are n symmetric firms, indexed by j, located equidistantly on the circle, so the distance between any two firms equals $\frac{1}{n}$.[15] Workers are uniformly distributed on the circle and maximize their wages. A worker's wage is determined by his productivity at the firm with which he is matched. Let $w_i(j)$ denote the wage of worker i at firm j. We assume

$$w_i(j) = a - c(d_{ij}) + \varepsilon_j,$$

[15] Firms do not make any decisions. We can also think of these firms as production technologies that workers utilize.

where a is the productivity, absent any productivity shock, of a worker who is perfectly matched with a firm; d is the distance between worker i and firm j; c is the productivity cost of distance; and ε is a firm-specific productivity shock that is zero for all firms in the first period. Following Hotelling (1929) and Salop (1979), we assume that productivity declines linearly with distance, and choose cost function such that $c(d) = d$. It can be shown that any function c that increases at least linearly with distance will yield qualitatively the same results. We consider a function c that increases less than linearly with distance in the Appendix.[16]

1.5.2 Comparative statics

A positive shock causes the positively shocked firm to attract workers from its two closest neighbors, and a negative shock causes the negatively shocked firm to lose workers to its two closest neighbors. Although the results are identical for a small enough shock, it is easier to analyze the case where workers only leave one firm, so we consider a negative shock and how the results change with n.

In the first period, all firms are equally productive, so a worker located on one side of a firm will be indifferent between working at that firm and its nearest neighbor on that side when he is equidistant between them. Since the distance between any two firms equals $\frac{1}{n}$, the distance between the indifferent worker and each firm is $\frac{1}{2n}$. Now consider a firm that receives a negative productivity shock in the second period. Let x be the distance from the shocked firm of the worker who is just indifferent between staying and moving to the neighboring firm nearest to him. Given the uniform distribution of workers over the circle, the probability that a worker at a firm that receives a negative productivity shock stays at his firm equals $\frac{x}{1/2n}$, and the probability that he switches firms equals $1 - \frac{x}{1/2n}$. Since the impacts of negative and positive shocks are symmetric, it follows that the probability that any worker in the region switches firms equals $s(1 - \frac{x}{1/2n})$.

[16] We note that many standard functions that increase less than linearly with distance yield the same qualitative results as functions that increase linearly with distance. The function in the appendix that provides a different result is a rather special piecewise function.

To solve for x, note that x and the distance of the indifferent worker to the neighboring firm nearest to him must sum to $\frac{1}{n}$. Thus the distance in characteristics space of the indifferent worker to his best alternative firm is $\frac{1}{n} - x$. Now use the condition that the wage of the indifferent worker at the negatively shocked firm must be equal to his wage at his best alternative firm: $a - x + \varepsilon = a - (\frac{1}{n} - x)$, where $\varepsilon < 0$. This yields the solution $x = \frac{1}{2}(\frac{1}{n} + \varepsilon)$. Plugging in x to the earlier result, the probability that any worker in the region switches firms equals $s(1 - \frac{1/n + \varepsilon}{2/2n}) = -s\varepsilon n$. Since $\varepsilon < 0$, this probability is unambiguously increasing in n.

The intuition for this strong result is that, as the number of firms increases, the size of the productivity shock increases relative to the distance between firms. The share of workers whose best firm changes in response to the shock therefore increases. If the size of the productivity shock were instead proportional to the distance between firms, the result would not hold. However, there is no reason for the size of a productivity shock such as a change in management to be affected by the distance between firms in characteristics space.

What happened to the countervailing effect of higher n, that workers are closer to the firms with which they are initially matched? Consider the expression for the probability that any worker in a region switches firms, $s(1 - \frac{x}{1/2n})$. Holding x constant, this probability decreases as n increases. This countervailing effect is always dominated by the decrease in x. However, this need not be the case if the productivity cost of distance increases less than linearly. In the Appendix, we construct an example where, after a certain distance, the productivity cost of distance increases less steeply, and show that it is possible for an increase in the number of firms to decrease the probability that a worker switches firms. The intuition is that the reduction in average distance of a worker to his initial firm has a much larger effect on productivity than the reduction in distance to the nearest neighboring firm.

1.6 Conclusion

We have provided evidence that one way in which agglomeration improves matching between workers and firms is by facilitating worker circulation among firms within same-skill labor markets. Regional labor market density also increases worker circulation, but the effect is weaker.

Our results imply that the process by which the matching benefits of agglomeration are realized is likely to reinforce another benefit of agglomeration, the spillover of ideas between firms through worker circulation. A challenge for future research is to integrate the matching and spillover benefits of worker circulation into one empirical and theoretical framework.

1.7 Figures and Tables

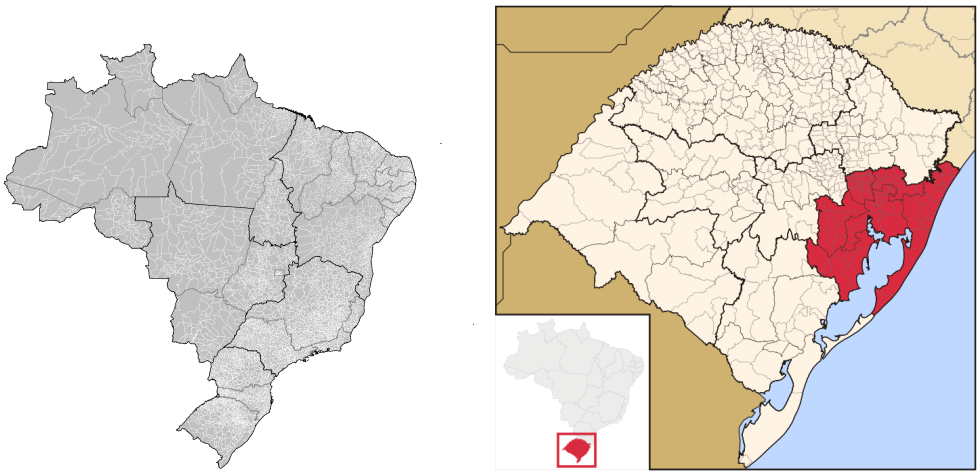


Figure 1.1. Map of Brazil
The left figure shows the states of Brazil in dark outlines and the individual municipalities in the lighter white lines. The right figure shows the state of Rio Grande de Sul with one of the meso regions highlighted in red. Micro regions within the meso regions are outlined in slightly lighter grey, while the municipalities are shown as the smallest partitions.

Table 1.1. Cell-switching for occ-meso, 25-55

	mean	sd	min	max
percent_occ	0.25	0.20	0.00	1.00
percent_muni	0.61	0.21	0.00	1.00
percent_muniocc	0.15	0.13	0.00	1.00
n_jobs_8	1439.89	6451.73	1.00	220093.00
n_firms_8	320.23	1326.54	1.00	47491.00
l_njobs	5.12	2.26	0.00	12.30
l_nfirms	4.02	1.92	0.00	10.77
l_avgskillfirmsize	1.10	0.76	0.00	6.16
N	11182			

Table 1.2. Cell-switching for occ-meso, 16-24

	mean	sd	min	max
percent_occ	0.17	0.17	0.00	1.00
percent_muni	0.64	0.22	0.00	1.00
percent_muniocc	0.11	0.11	0.00	1.00
n_jobs_8	1504.07	6594.36	1.00	220445.00
n_firms_8	333.93	1353.34	1.00	47504.00
l_njobs	5.28	2.15	0.00	12.30
l_nfirms	4.15	1.84	0.00	10.77
l_avgskillfirmsize	1.13	0.75	0.00	6.16
N	10732			

Table 1.3. Cell-switching for cnae-meso, 25-55

	mean	sd	min	max
percent_cnae	0.18	0.16	0.00	1.00
percent_muni	0.63	0.21	0.00	1.00
percent_municnae	0.12	0.13	0.00	1.00
n_jobs_8	1370.41	6194.77	1.00	242794.00
n_firms_8	178.09	899.10	1.00	44948.00
l_njobs	4.68	2.58	0.00	12.40
l_nfirms	3.04	2.07	0.00	10.71
l_avgskillfirmsize	1.64	1.10	0.00	6.94
N	7241			

Table 1.4. Cell-switching for cnae-meso, 16-24

	mean	sd	min	max
percent_cnae	0.13	0.13	0.00	1.00
percent_muni	0.65	0.22	0.00	1.00
percent_municnae	0.09	0.10	0.00	1.00
n_jobs_8	1409.41	6278.23	1.00	242794.00
n_firms_8	183.14	911.34	1.00	44948.00
l_njobs	4.80	2.52	0.00	12.40
l_nfirms	3.12	2.04	0.00	10.71
l_avgskillfirmsize	1.68	1.08	0.00	6.94
N	7040			

Table 1.5. Regression results, cell-switching for occ-meso

	(1)	(2)	(3)	(4)	(5)	(6)
l_njobs	0.0350***			0.0296***		
	(28.83)			(26.36)		
l_nfirms		0.0510***	0.0470***		0.0404***	0.0357***
		(30.14)	(27.90)		(24.40)	(21.51)
l_avgskillfirmsize			0.0194***			0.0220***
			(10.20)			(12.35)
Occ. FE	Yes	Yes	Yes	Yes	Yes	Yes
Meso FE	Yes	Yes	Yes	Yes	Yes	Yes
N	11187	11187	11187	10732	10732	10732
R^2	0.657	0.656	0.662	0.533	0.526	0.535

t statistics in parentheses. (1)-(3) denote ages 25-55, (4)-(6) for ages 16-24

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 1.6. Regression results, cell-switching for industry-meso

	(1)	(2)	(3)	(4)	(5)	(6)
l_njobs	0.0271***			0.0223***		
	(27.97)			(24.97)		
l_nfirms		0.0498***	0.0470***		0.0403***	0.0377***
		(30.97)	(27.98)		(28.30)	(26.49)
l_avgskillfirmsize			0.00895***			0.00803***
			(6.42)			(6.83)
Ind. FE	Yes	Yes	Yes	Yes	Yes	Yes
Meso FE	Yes	Yes	Yes	Yes	Yes	Yes
N	7241	7241	7241	7040	7040	7040
R^2	0.573	0.590	0.593	0.558	0.573	0.577

t statistics in parentheses. (1)-(3) denote ages 25-55, (4)-(6) for ages 16-24

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 1.7. Cell and Regional Statistics

	count	mean	sd	min	max
occ-meso cell	11356	3.96	1.97	0.00	10.77
meso region	137	8.48	1.30	4.23	12.26
micro region	558	6.83	1.39	2.08	11.99
industry-meso cell	7371	3.13	2.12	0.00	10.90
3occ-meso cell	42968	2.46	1.92	0.00	10.62
2occ-micro cell	43445	2.50	1.83	0.00	10.51
1occ-meso cell	1233	6.68	1.44	1.61	11.22

Table 1.8. Table of Means for Aggregate

	count	mean	sd	min	max
percent_stay	112639	0.56	0.35	0.00	1.00
percent_change1	112639	0.07	0.13	0.00	1.00
percent_change2	49733	0.19	0.27	0.00	1.00
avg_changes	112639	0.10	0.19	0.00	4.00
percent_stay_m	118627	0.41	0.34	0.00	1.00
percent_change1_m	118627	0.38	0.28	0.00	1.00
percent_change2_m	96285	0.40	0.31	0.00	1.00
avg_changes_m	118627	0.70	0.69	0.00	10.00
percent_stay_f	119114	0.48	0.36	0.00	1.00
percent_change1_f	119114	0.09	0.14	0.00	1.00
percent_change2_f	58386	0.18	0.25	0.00	1.00
avg_changes_f	119114	0.12	0.20	0.00	4.00
cohort	123314	2007.93	3.73	2002.00	2014.00
n_jobs_8	123314	1677.75	6974.69	1.00	220093.00
n_firms_8	123314	370.19	1433.03	1.00	47491.00
l_njobs	123314	5.32	2.43	0.00	12.30
l_nfirms	123314	4.19	2.04	0.00	10.77
n_urbanjobs	123314	133354.69	303978.56	1.00	2.81e+06
n_urbanfirms	123314	11474.86	23124.44	1.00	210029.00
l_urbanjobs	123314	10.11	3.33	0.00	14.85
l_urbanfirms	123314	7.96	2.70	0.00	12.26
l_avgskillfirmsize	123314	1.13	0.78	0.00	6.16
l_avgurbanfirmsize	123314	2.14	0.71	0.00	3.17
N	123314				

Table 1.9. Aggregate regression results, skill firm and urban firm count

	(1) percent_stay	(2) percent_change1	(3) percent_change2	(4) avg_changes
l_nfirms	-0.0239***	0.0262***	0.0418***	0.0415***
	(-16.65)	(40.36)	(16.88)	(44.07)
l_urbanfirms	0.00153	-0.0109***	-0.0140***	-0.0171***
	(1.02)	(-16.17)	(-5.64)	(-17.38)
l_avgskillfirmsize	0.0188***	0.0104***	0.0102***	0.0163***
	(11.87)	(14.62)	(3.91)	(15.70)
l_avgurbanfirmsize	0.102***	0.00467***	0.00328	0.00789***
	(35.16)	(3.57)	(0.72)	(4.15)
Cohort FE	Yes	Yes	Yes	Yes
Occ. FE	Yes	Yes	Yes	Yes
Implied effect for increasing size while maintaining marketshare				
b1+b2	-0.0224	0.0153	0.0278	0.0245
tstat	-33.87	51.09	25.91	56.40
N	112639	112639	49733	112639
r2	0.469	0.229	0.170	0.268

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 1.10. Aggregate regression results, regional fixed effects

	(1) percent_stay	(2) percent_change1	(3) percent_change2	(4) avg_changes
l_nfirms	-0.0137***	0.0264***	0.0424***	0.0418***
	(-9.93)	(39.96)	(16.73)	(43.68)
l_avgskillfirmsize	0.0109***	0.0105***	0.0110***	0.0163***
	(7.22)	(14.52)	(4.13)	(15.56)
Cohort FE	Yes	Yes	Yes	Yes
Occ. FE	Yes	Yes	Yes	Yes
Meso FE	Yes	Yes	Yes	Yes
N	112639	112639	49733	112639
R^2	0.534	0.237	0.178	0.275

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 1.11. Aggregate regression results, including cell-switchers

	(1) percent_stay_m	(2) percent_change1_m	(3) percent_change2_m	(4) avg_changes_m
l_nfirms	-0.0546***	0.105***	0.0219***	0.209***
	(-60.80)	(136.16)	(17.45)	(106.51)
l_urbanfirms	0.0230***	-0.0757***	-0.00343**	-0.140***
	(29.16)	(-111.43)	(-3.18)	(-81.17)
l_avgskillfirmsize	0.0109***	0.0165***	-0.00279	0.0348***
	(8.40)	(14.76)	(-1.60)	(12.27)
l_avgurbanfirmsize	0.122***	-0.114***	-0.0188***	-0.204***
	(61.88)	(-67.27)	(-7.00)	(-47.17)
Cohort FE	Yes	Yes	Yes	Yes
Occ. FE	Yes	Yes	Yes	Yes
N	118627	118627	96285	118627
R^2	0.616	0.580	0.350	0.537

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 1.12. Aggregate regression results, including non-surviving firms

	(1) percent_stay_f	(2) percent_change1_f	(3) percent_change2_f	(4) avg_changes_f
l_nfirms	-0.0434***	0.0311***	0.0446***	0.0484***
	(-43.37)	(61.91)	(25.74)	(67.48)
l_urbanfirms	0.0223***	-0.0142***	-0.0195***	-0.0215***
	(25.26)	(-32.05)	(-12.93)	(-33.99)
l_avgskillfirmsize	0.0245***	0.0116***	0.0101***	0.0177***
	(17.21)	(16.20)	(4.57)	(17.36)
l_avgurbanfirmsize	0.139***	-0.00132	0.00581	0.000114
	(64.47)	(-1.22)	(1.75)	(0.07)
Cohort FE	Yes	Yes	Yes	Yes
Occ. FE	Yes	Yes	Yes	Yes
N	119114	119114	58386	119114
R^2	0.576	0.269	0.175	0.314

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 1.13. Table of Means for Aggregate, 3occ

	count	mean	sd	min	max
percent_stay	271454	0.58	0.39	0.00	1.00
percent_change1	271454	0.06	0.15	0.00	1.00
percent_change2	74383	0.19	0.30	0.00	1.00
avg_changes	271454	0.08	0.21	0.00	6.50
cohort	324472	2007.82	3.74	2002.00	2014.00
n_jobs_8	324472	623.23	3126.99	1.00	174602.00
n_firms_8	324472	157.40	745.98	1.00	40956.00
l_njobs	324472	4.05	2.49	0.00	12.07
l_nfirms	324472	3.11	2.06	0.00	10.62
n_urbanjobs	324472	157680.85	359887.37	1.00	2.81e+06
n_urbanfirms	324472	13240.73	27277.57	1.00	210029.00
l_urbanjobs	324472	9.46	4.33	0.00	14.85
l_urbanfirms	324472	7.48	3.48	0.00	12.26
l_avgskillfirmsize	324472	0.94	0.85	0.00	8.19
l_avgurbanfirmsize	324472	1.98	0.91	0.00	3.17
N	324472				

Table 1.14. Aggregate regression results, skill firm and urban firm count, 3occ

	(1)	(2)	(3)	(4)
	percent_stay	percent_change1	percent_change2	avg_changes
l_nfirms	-0.0121***	0.0195***	0.0350***	0.0294***
	(-13.28)	(47.60)	(17.55)	(50.65)
l_urbanfirms	-0.00552***	-0.00761***	-0.0115***	-0.0106***
	(-5.70)	(-17.47)	(-5.70)	(-17.24)
l_avgskillfirmsize	0.0259***	0.00388***	0.00744**	0.00560***
	(24.88)	(8.29)	(3.19)	(8.46)
l_avgurbanfirmsize	0.0686***	0.00616***	0.00382	0.0111***
	(30.07)	(6.00)	(0.89)	(7.64)
Cohort FE	Yes	Yes	Yes	Yes
Occ. FE	Yes	Yes	Yes	Yes
b1b2	-0.0176	0.0119	0.0235	0.0188
tstat	-33.17	49.89	23.77	55.53
N	271452	271452	74378	271452
r2	0.387	0.139	0.143	0.166

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 1.15. Aggregate regression results, regional fixed effects, 3occ

	(1) percent_stay	(2) percent_change1	(3) percent_change2	(4) avg_changes
l_nfirms	-0.00797***	0.0204***	0.0366***	0.0307***
	(-8.87)	(48.96)	(18.02)	(52.05)
l_avgskillfirmsize	0.0238***	0.00361***	0.00720**	0.00521***
	(23.37)	(7.66)	(3.05)	(7.82)
Cohort FE	Yes	Yes	Yes	Yes
Occ. FE	Yes	Yes	Yes	Yes
Meso FE	Yes	Yes	Yes	Yes
N	271452	271452	74378	271452
R^2	0.426	0.143	0.150	0.171

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 1.16. Table of Means for Aggregate

	count	mean	sd	min	max
percent_stay	302918	0.58	0.38	0.00	1.00
percent_change1	302918	0.06	0.15	0.00	1.00
percent_change2	84681	0.20	0.30	0.00	1.00
avg_changes	302918	0.08	0.21	0.00	6.00
cohort	354154	2007.87	3.74	2002.00	2014.00
n_jobs_8	354154	573.13	3333.20	1.00	162071.00
n_firms_8	354154	127.87	679.65	1.00	36537.00
l_njobs	354154	4.07	2.38	0.00	12.00
l_nfirms	354154	3.04	1.93	0.00	10.51
n_urbanjobs	354154	41474.81	148893.68	1.00	2.10e+06
n_urbanfirms	354154	3605.84	11442.01	1.00	160916.00
l_urbanjobs	354154	8.20	3.61	0.00	14.56
l_urbanfirms	354154	6.27	2.86	0.00	11.99
l_avgskillfirmsize	354154	1.03	0.82	0.00	7.78
l_avgurbanfirmsize	354154	1.93	0.86	0.00	3.84
N	354154				

Table 1.17. Aggregate regression results, skill firm and urban firm count

	(1) percent_stay	(2) percent_change1	(3) percent_change2	(4) avg_changes
l_nfirms	-0.0168***	0.0238***	0.0424***	0.0363***
	(-18.27)	(56.21)	(22.00)	(61.08)
l_urbanfirms	-0.00667***	-0.00769***	-0.0143***	-0.0116***
	(-7.12)	(-17.87)	(-7.48)	(-19.13)
l_avgskillfirmsize	0.0178***	0.00534***	0.0125***	0.00837***
	(18.06)	(11.81)	(6.05)	(13.17)
l_avgurbanfirmsize	0.0515***	-0.00931***	-0.0201***	-0.0135***
	(30.88)	(-12.16)	(-5.92)	(-12.58)
Cohort FE	Yes	Yes	Yes	Yes
Occ. FE	Yes	Yes	Yes	Yes
b1b2	-0.0235	0.0161	0.0282	0.0248
tstat	-53.50	79.76	33.13	87.29
N	302918	302918	84681	302918
r2	0.380	0.130	0.128	0.153

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 1.18. Aggregate regression results, regional fixed effects

	(1) percent_stay	(2) percent_change1	(3) percent_change2	(4) avg_changes
l_nfirms	-0.0111***	0.0243***	0.0451***	0.0371***
	(-12.22)	(56.53)	(22.66)	(61.25)
l_avgskillfirmsize	0.0143***	0.00550***	0.0113***	0.00859***
	(14.88)	(12.04)	(5.37)	(13.38)
Cohort FE	Yes	Yes	Yes	Yes
Occ. FE	Yes	Yes	Yes	Yes
Micro FE	Yes	Yes	Yes	Yes
N	302918	302918	84679	302918
R^2	0.423	0.139	0.142	0.161

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 1.19. Table of Means for Aggregate (CNAE)

	count	mean	sd	min	max
percent_stay	67963	0.56	0.35	0.00	1.00
percent_change1	67963	0.07	0.12	0.00	1.00
percent_change2	30507	0.18	0.25	0.00	1.00
avg_changes	67963	0.09	0.17	0.00	4.00
percent_stay_m	71916	0.40	0.34	0.00	1.00
percent_change1_m	71916	0.40	0.28	0.00	1.00
percent_change2_m	60089	0.41	0.30	0.00	1.00
avg_changes_m	71916	0.76	0.71	0.00	10.00
percent_stay_f	73016	0.47	0.37	0.00	1.00
percent_change1_f	73016	0.08	0.13	0.00	1.00
percent_change2_f	36381	0.17	0.23	0.00	1.00
avg_changes_f	73016	0.10	0.18	0.00	3.00
cohort	75829	2007.93	3.72	2002.00	2014.00
n_jobs_8	75829	2802.46	11455.84	1.00	321469.00
n_firms_8	75829	258.22	1215.85	1.00	54199.00
l_njobs	75829	5.55	2.72	0.00	12.68
l_nfirms	75829	3.54	2.10	0.00	10.90
n_urbanjobs	75829	141992.45	320124.32	1.00	2.81e+06
n_urbanfirms	75829	12147.85	24330.83	1.00	210029.00
l_urbanjobs	75829	10.00	3.60	0.00	14.85
l_urbanfirms	75829	7.89	2.90	0.00	12.26
l_avgskillfirmsize	75829	2.01	1.28	0.00	6.94
l_avgurbanfirmsize	75829	2.11	0.76	0.00	3.17
N	75829				

Table 1.20. Aggregate regression results, industry categorization

	(1) percent_stay	(2) percent_change1	(3) percent_change2	(4) avg_changes
l_nfirms	-0.0248***	0.0233***	0.0380***	0.0357***
	(-17.75)	(40.33)	(17.06)	(43.68)
l_urbanfirms	0.00993***	-0.0116***	-0.0170***	-0.0171***
	(6.25)	(-17.67)	(-7.26)	(-18.44)
l_avgskillfirmsize	0.0177***	0.00536***	0.00555*	0.00654***
	(13.20)	(9.69)	(2.28)	(8.34)
l_avgurbanfirmsize	0.0829***	0.00428**	0.00204	0.00682**
	(22.63)	(2.84)	(0.38)	(3.18)
Cohort FE	Yes	Yes	Yes	Yes
Ind. FE	Yes	Yes	Yes	Yes
b1b2	-0.0149	0.0117	0.0210	0.0186
tstat	-17.19	32.69	16.66	36.71
N	67963	67963	30507	67963
r2	0.494	0.234	0.172	0.265

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 1.21. Aggregate regression results, industry with regional F.E.

	(1) percent_stay	(2) percent_change1	(3) percent_change2	(4) avg_changes
l_nfirms	-0.0171***	0.0240***	0.0391***	0.0367***
	(-12.80)	(40.83)	(17.20)	(44.15)
l_avgskillfirmsize	0.0126***	0.00565***	0.00861***	0.00722***
	(9.79)	(10.03)	(3.45)	(9.03)
Cohort FE	Yes	Yes	Yes	Yes
Ind. FE	Yes	Yes	Yes	Yes
Meso FE	Yes	Yes	Yes	Yes
N	67963	67963	30507	67963
R^2	0.558	0.242	0.184	0.273

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 1.22. Table of Means for Individual, remove cell-switchers

	count	mean	sd	min	max
stay	5889026	0.52	0.50	0.00	1.00
change1	5889026	0.15	0.36	0.00	1.00
change2	5889026	0.05	0.22	0.00	1.00
changes	5889026	0.24	0.70	0.00	16.00
age	5889026	35.91	8.51	25.00	55.00
edu	5889026	5.61	2.13	1.00	11.00
cohort	5889026	2008.17	3.96	2002.00	2014.00
changed_cell	5889026	0.00	0.00	0.00	0.00
firm_size	5889026	1645.32	8091.64	1.00	110058.00
l_firmsize	5889026	4.49	2.51	0.00	11.61
size_2014	5889026	1779.51	9192.88	1.00	107795.00
firmsizeratio	5889026	1.26	4.05	0.00	3430.00
avg_size	5889026	1584.32	8297.22	1.00	98530.54
l_avgindfirmsize	5889026	4.39	2.46	0.00	11.50
n_jobs_8	5889026	27296.49	47319.77	1.00	220093.00
n_firms_8	5889026	5043.20	9555.40	1.00	47491.00
l_njobs	5889026	8.87	1.84	0.00	12.30
l_nfirms	5889026	7.12	1.81	0.00	10.77
l_avgskillfirmsize	5889026	1.76	0.72	0.00	6.16
n_urbanjobs	5889026	650623.96	925503.79	1043.00	2.81e+06
n_urbanfirms	5889026	50211.94	69054.65	69.00	210029.00
l_urbanjobs	5889026	12.38	1.50	6.95	14.85
l_urbanfirms	5889026	9.90	1.42	4.23	12.26
l_avgurbanfirmsize	5889026	2.47	0.24	1.62	3.17
l_marketshare	5889026	-2.78	1.12	-9.94	-0.90

Table 1.23. Individual regression results, skill firm and urban firm count

	(1) stay	(2) change1	(3) change2	(4) changes
l_nfirms	-0.0260***	0.0396***	0.0388***	0.0670***
	(-7.70)	(14.11)	(6.57)	(10.20)
l_urbanfirms	0.0129***	-0.0230***	-0.0226***	-0.0380***
	(3.47)	(-7.31)	(-3.74)	(-5.43)
l_avgskillfirmsize	0.0150**	0.0219***	0.0476***	0.0448***
	(3.24)	(6.47)	(9.12)	(6.09)
l_avgurbanfirmsize	0.0152*	0.00949	0.00125	0.0188
	(2.09)	(1.96)	(0.13)	(1.67)
age	0.00109***	-0.00350***	-0.00401***	-0.00674***
	(7.68)	(-30.64)	(-28.25)	(-21.73)
firmsizeratio	0.00323***	-0.00120***	0.000704*	-0.00179***
	(5.35)	(-5.16)	(2.49)	(-4.90)
l_firmsize	-0.00147	-0.000853	0.00153*	-0.00130
	(-1.52)	(-1.67)	(2.06)	(-1.34)
Occ. FE	Yes	Yes	Yes	Yes
Cohort FE	Yes	Yes	Yes	Yes
Education FE	Yes	Yes	Yes	Yes
b1b2	-0.0131	0.0166	0.0162	0.0290
tstat	-7.889	10.33	6.045	9.514
N	5889026	5889026	877459	5889026
r2	0.306	0.104	0.105	0.104

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 1.24. Individual regression results, regional fixed effects

	(1) stay	(2) change1	(3) change2	(4) changes
l_nfirms	-0.0242***	0.0426***	0.0445***	0.0737***
	(-7.72)	(15.35)	(8.44)	(11.43)
l_avgskillfirmsize	0.0109*	0.0194***	0.0432***	0.0388***
	(2.27)	(6.38)	(9.31)	(5.92)
age	0.00114***	-0.00354***	-0.00408***	-0.00680***
	(8.01)	(-30.74)	(-27.52)	(-21.82)
firmsizeratio	0.00315***	-0.00118***	0.000695*	-0.00175***
	(5.34)	(-5.17)	(2.44)	(-4.90)
l_firmsize	0.000204	-0.000904	0.00148*	-0.00135
	(0.22)	(-1.86)	(2.07)	(-1.48)
Occ. FE	Yes	Yes	Yes	Yes
Meso FE	Yes	Yes	Yes	Yes
Cohort FE	Yes	Yes	Yes	Yes
Education FE	Yes	Yes	Yes	Yes
N	5889026	5889026	877459	5889026
R^2	0.313	0.107	0.108	0.107

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Chapter 1 is currently being prepared for submission for publication. Zeng, Wendy. The dissertation author is the principal researcher on this chapter.

Chapter 2

Climbing the Wage Ladder Through External and Internal Mobility: Evidence from Brazil

2.1 Introduction

Job changes are an important source of wage gains for workers. This observation is well documented for external job changes between firms. Topel and Ward (1992) find that external changes account for a third of wage increases for young workers. Workers also increase wages by changing jobs within a firm. When changing internally, workers can search for jobs in the external market for bargaining power, and vice versa when changing externally.

Voluntary job changes are more likely than involuntary job changes to result in wage gains. However, the ability to distinguish voluntary from involuntary job changes has been limited in the literature. Current on-the-job search papers categorize all job-to-job moves as voluntary (Haltiwanger et al., 2018; Haltiwanger, Hyatt and McEntarfer, 2018), where these moves are categorized as job-to-job and voluntary if no full quarter of nonemployment occurs in between jobs. A problem with this method is that it captures workers who were fired at the beginning of one quarter and found a new job by the end of the next quarter, allowing almost six months of unemployment to be categorized as a voluntary job-to-job change. Moreover, the literature rarely counts internal job changes. We will show that these internal job changes are

quantitatively important parts of both voluntary and, more surprisingly, involuntary job changes, and a key source of wage gains in the former case.

We provide a more nuanced and accurate description of voluntary vs involuntary job changes, in addition to observing both internal and external job changes. The Brazilian matched employer-employee database RAIS reports all job separations, so we can categorize external vs internal job moves by observing changes in employer code. Workers also report separation cause, including quits, transfers, and fires, which we use to help characterize voluntary and involuntary job changes. We also calculate the months between jobs, a more accurate measure of nonemployment, using exact start and end dates.

With more accurate categorization, we can give an accurate count of the frequency of voluntary and involuntary job changes in both the internal and external categories. We find that internal job changes account for nearly a third of all voluntary job changes. Surprisingly, internal job changes also account for a third of involuntary job changes. Accounting for internal job changes thereby raises the estimated rates of involuntary and voluntary job turnover by roughly 50% and 40%, respectively. This highlights the importance for macro job search models to account for internal job changes.

We compute wage changes for the various job change categories. For the two largest voluntary job change categories, external quits and internal transfers, we find mean wage gains of nine and eight percent, respectively. For each of the two largest involuntary job change categories, external fires and internal fires, we find mean wage losses of three percent. The similarities across internal and external job changes suggest integration of the internal and external job markets, and that workers search for offers in both the internal and external job markets for bargaining power.

Related Literature. Our paper builds upon the literature measuring on-the-job search. Mortensen and Pissarides (1992, 1994) develop a dynamic, stochastic framework to study worker and job flows where exogenous destruction of jobs leads to involuntary separation. Mortensen (1994) adds an on-the-job search feature to this model, allowing workers to leave a previous firm after finding a better offer. These job to job movements are voluntary separations. Empirically,

they use data from the National Longitudinal Survey of Mature Men to describe quit statistics. These quits can further be broken down into job-to-job quits versus job-to-unemployment quits, which provide measures of voluntary and involuntary separation rates for workers moving between firms (Akerlof et al., 1988). While the survey provides detailed descriptions of reasons behind job separation, the sample itself is limited; the cohort includes 5,020 men aged 45-59 who are given thirteen surveys over 24 years.

In recent on-the-job search models, voluntary and involuntary job changes are categorized based on nonemployment between jobs. Workers who do not experience nonemployment between jobs are assumed to be voluntary job-to-job (Haltiwanger et al., 2018; Haltiwanger, Hyatt and McEntarfer, 2018). Due to nonemployment length between jobs often being the only source of information to infer voluntary vs involuntary, the terms job to job flows and voluntary flows are used interchangeably. As far as we know, the on-the-job search literature does not account for workers moving within firms.

Our paper is also closely related to the literature connecting wage gains to job changes. Topel and Ward (1992) find that wage changes when changing jobs account for a third of the wage growth in the early careers of young men. They find that job changes during the first ten years in the labor market account for about two thirds of total job changes in a career. There is some analysis of voluntary mobility. Antel (1986) reports large and significant wage effects accompanying voluntary job changes. Antel (1991), using the NLS data for Young Men, finds that voluntary mobility leads to higher wage gains when workers experience an unemployment period in between jobs. A large literature on unemployment insurance, wage gains, and unemployment length studies how longer unemployment periods allow workers to search for better jobs and increase their post-unemployment wages (Acemoglu and Shimer, 2000; Hall and Milgrom, 2008; Ehrenberg and Oaxaca, 1976).

Regarding internal job changes in this literature, Frederiksen, Halliday and Koch (2016) look at executive job changes within firms and across firm mobility to study short and long-term wage growth. Van der Klaauw and Da Silva (2011) use a Portuguese dataset to investigate

internal job changes, both in hierarchical levels and employer-reported promotions. They find that employers report promotions when the job description and tasks change as well as when the employee receives a substantial wage increase. Neither of these papers investigates the topic of voluntary vs involuntary changes, nor do they compare wage gains in the internal and external markets.

This paper is organized as follows. In Section 2.2, we provide a detailed description of the data. Section 2.3 presents our empirical analysis of nonemployment length between jobs and how we use it to characterize voluntary vs involuntary job changes. Section 2.4 analyzes log wage changes, and Section 2.5 concludes.

2.2 Data

Our data are drawn from the linked employer-employee records RAIS (Relação Anual de Informações Sociais of the Brazilian labor ministry MTE). RAIS is organized by job spell. Each job spell is identified by employee ID, firm-plant ID, and dates of job accession and separation. The first eight digits of the 14-digit firm-plant ID identify the firm to which the plant belongs. Worker age, education, and gender are recorded, as are earnings, industry, and occupation. By Brazilian law, every private or public-sector employer must report this information every year.[1]

The RAIS database was extensively revised in 2002, which we therefore use as our first year of observation. Our last year of observation is 2014, after which Brazil entered a period of negative GDP growth and political instability. From the RAIS data, we construct job changes and wage changes using the variables worker ID, employer ID, monthly average wage, hours contracted per week, month of admission, month of separation, year, and cause of separation. We also use the variables age, education, tenure, municipality, sex, date of birth, occupation, and industry. We have data for the yearly minimum wage and can map each municipality to its respective meso and microregion.

[1] I do not observe workers who are employed in the informal sector, which comprises around 1/3 of Brazil's labor force.

2.2.1 Employment Spells

We identify employment spells for each worker during our sample period. We limit our worker sample to prime age (25-55) males and their primary, long-term jobs, keeping only employment spells with 20 or more hours worked per week and, if the employment spell ended, with a duration of at least 3 months.[2] We remove workers from the sample if they ever report as female, report different birth dates, or are missing education in any observation, and also remove those who never report a birth date across all observations.

Next, we build employment histories for each worker in our sample. We start with employment spells that have terminated, indicated by the variable cause of separation. To illustrate, suppose that a worker has a job from 2002 until 2004, when he is fired. He then takes another job from 2007 and quits in 2008. He then works at another job in 2010, and is still working there by the end of the sample period. He will be in the RAIS data for the years 2002-2004, 2007-2008, and 2010-2014. The observations in 2004 and 2008 will report a positive cause of separation value. Taking each observation with a positive cause of separation value provides all employment spells that ended. The month of separation, year of observation, and tenure variable give the start and end dates of each employment spell. If the month of separation is missing, we assume the month to be December. We then add employment spells that are ongoing. We take all observations in 2014 with zero as the cause of separation value to get all of these spells, and calculate start date using the tenure variable.

We order the employment spells by admission date, and by employment length if two spells have the same admission date. After ordering the spells, we proceed to build job change observations. We define a job change as a spell followed by subsequent spell down the ordered list. We calculate the wage change and nonemployment length for each job change. Nonemployment is calculated by taking the difference between the second job's start date and the worker's last employment date from previous jobs. This value is negative if two spells overlap in dates.

[2]The legal working age in Brazil begins at age 16, while most of the population retires from the working force around age 65.

Regarding overlap, we remove job changes with a negative nonemployment length of three months or more. This is because we do not consider working two jobs simultaneously for over 3 months to be a job change, but more like taking a secondary job.

2.2.2 Building wage changes

The wage data in RAIS are reported in monthly salaries. The units are multiples of the monthly minimum wage to make the data comparable across jobs. We take monthly wage, divide by ratio of hours worked weekly to regular hours in a work week[3], and obtain the monthly wage under regular hours.

We calculate log wage changes by subtracting the current job's starting log wage from the log of the average monthly wage over the entire final year of the previous job.

2.2.3 Internal vs External Job Changes

If a job change results in the worker moving to a different firm, we categorize the change as external. Each firm has a unique firm ID reported in each observation, which allows us to identify external vs internal changes. One of our important contributions is to systematically count internal as well as external job changes.

We focus on job changes within same-skill labor markets. We partition labor markets, which we call cells, by worker occupation and location. The goal is to define a labor market as one in which, if a worker were to switch from one job to another, the worker would be able to find a new job without having to pay the costs of either moving or retraining. We impose a restriction to workers who stay within the same labor market throughout the sample period to address the migration channel. We refer to these workers as ones that change jobs within cell.

We partition the cells by the two-digit occupation code at the meso regional level. This is the same specification as in Chapter 1, which provides a detailed description of the specification levels, so we summarize it briefly again. Occupation is aggregated at the 2-digit level, out of 5

[3]The regular working week in Brazil is 44 hours over a six-day period according to Brazilian legislation

digits, based on Brazil's occupation classification system from 1994. The 2-digit level is fairly broad, and would differentiate between a scientist and an engineer. Location is aggregated at the meso region level. A meso region is fairly large; smaller than a state but larger than a micro region.

Table 2.1 counts the total number of external vs internal job changes within cell in a 1% random sample of the data. There are two points of interest to note. First, 32% of the total number of job changes occur internally. It follows that most of the literature has been missing a major part of job-to-job transitions. Second, although we might have assumed that most internal job changes are transfers, 76% of them are actually fires. Thus, the rate of involuntary job turnover is much greater than what we see using only external job changes.

2.3 Distinguishing involuntary from voluntary job changes

To classify job changes as involuntary versus voluntary, we will use two pieces of information: cause of separation (fires, quits, and transfers) and months between jobs. There are at least two reasons to expect that time between jobs will be shorter when job changes are voluntary. First, employees may receive offers through on-the-job search and change to their new jobs with minimal or zero intervening nonemployment (Pissarides 1994). This applies to both external and internal job changes. There may even be overlap between the new job and the old job (negative intervening nonemployment), as the worker trains in the new job while retaining the old job. When the job change is internal, the firm may encourage this overlap. Second, workers may choose to leave their jobs and search during periods when the demands for their skills are strong, as opposed to workers who are laid off, which is more likely to coincide with periods when the demands for their skills are weak.

The literature (e.g., Antel (1991)) classifies external quits as voluntary and external fires as involuntary. In Table 2.2 we report descriptive statistics for months between jobs for external fires and quits, as well as internal fires and quits and external and internal transfers.

The mean and median differences between external fires and quits are roughly four and five months, respectively, with the median for external quits being only about one month between jobs. Thus the literature and classification on the basis of months between jobs are in agreement that external fires are involuntary and external quits are voluntary. (Note that, because we treat job spells that overlap by three months or more as holding two jobs rather than as job changes, the minimum in Table 2.2 of months between jobs for all categories of job changes is -2.90.)

The literature provides little guidance as to whether internal job changes should be classified as involuntary or voluntary. We first consider internal transfers, the most intuitively familiar category of internal job changes. Table 2.2 shows that the mean for months between jobs for internal transfers is about two months and the median is zero months, clearly marking internal transfers as voluntary job changes. Next we consider the most common but less intuitive category of internal job changes, internal fires. Standard theory tells us that fires indicate bad matches between workers and firms, yet in the case of internal fires these matches are formed again. Given the high frequency of these job changes, it seems likely that they result either from seasonal employment or from layoffs followed by recalls. If we look closely at some of the special features of the Brazilian labor market, they may help explain why internal fires are so common. Under Brazilian labor laws, formal sector employees have the right to a 30-day vacation every 12 months of work. In addition, every year in December, all employees have the right to receive what is called a 13th salary. Employees who have worked all the 12 months receive the full payment, while employees who have not receive a proportional payment. Employees dismissed for *justa causa* (a termination for just cause) lose the right to receive this benefit. These laws could create incentive for employers to terminate their employees, particularly in the case of *justa causa*, to avoid giving employees their month vacation or paying the 13th salary. Table 2.2 shows that the mean and median months between jobs for internal fires are about eight and six months, respectively, compared to eleven and six months for external fires. Figure 2.1, panels (a) and (b), shows histograms for months between jobs for both external and internal fires. We see that very short nonemployment durations are actually more common for external fires, which

have higher mean duration only because very long periods between jobs are more common. On the basis of this information we classify internal fires as involuntary job changes.

The least common internal job change category is internal quits, accounting for only 5.6 percent of internal job changes. We conjecture that these are either failed external searches--the worker quit to find a better job elsewhere, but did not--or seasonal employment, where the seasonal nature is driven on the worker side. Table 2.2 shows that the mean and median months between jobs for internal quits are almost the same as for internal fires. Figure 2.1, panels (c) and (d) shows histograms for months between jobs for both external and internal quits. The histogram for internal quits looks much more like the histogram for external fires in panel (a) than the histogram for external quits in panel (c). We conclude that internal quits are best classified as involuntary job changes.

Finally, we consider external transfers. These are the smallest external job change category, accounting for only 7.3 percent of external job changes. We conjecture that external transfers involve some agreement between firms. Table 2.2 shows that the median months between jobs for external transfers are almost exactly the same (roughly zero) as for internal transfers, but the mean months between jobs for external transfers are roughly the same as for external quits. Internal transfers and external quits are both voluntary job change categories. Figure 2.1, panels (e) and (f), shows that the histogram for external transfers looks much more like the histogram for internal transfers than the histogram for external quits. External transfers should clearly be classified as voluntary job changes.

In summary, we have three voluntary job change categories: external quits, internal transfers, and external transfers, and three involuntary job change categories, external fires, internal fires, and internal quits. 79 percent of job changes are involuntary and 21 percent are voluntary. Accounting for internal job changes increases the number of involuntary job changes by 50 percent and the number of voluntary job changes by 40 percent.

We should note that Table 2.2 also reveals an interesting contrast between external and internal job changes within each cause of separation: the standard deviation of months between

jobs is always greater for external changes. Moving outside of one's firm opens up a greater range of possibilities.

2.4 Wage Changes Resulting from Job Changes

2.4.1 Involuntary versus voluntary and external versus internal

We expect that a worker who changes jobs voluntarily typically does so in expectation of a raise. Haltiwanger, Hyatt and McEntarfer (2018) find that voluntary (by their definition) job-to-job changes disproportionately allocate workers from less productive to more productive firms, where more productive firms offer higher wages.

Table 2.3 provides descriptive statistics for log wage changes for the job changes in Table 2.1, following the format of Table 2.2 for ease of comparison. The two largest categories of voluntary job changes, external quits and internal transfers, show mean wage gains of nine and eight percent, respectively. The two largest categories of involuntary job changes, external fires and internal fires, both show mean wage losses of three percent. The smallest categories of voluntary and involuntary job changes, respectively external transfers and internal quits, show mean wage changes of four and zero percent, respectively. The zero percent wage change for internal quits supports our conjecture that these are instances where workers returned to their old jobs after searching for new jobs but failing.

The similarity in wage changes across external and internal job changes for the main categories of both voluntary and involuntary job changes suggests that external and internal labor markets are integrated, in the sense that workers appear to be searching for external and internal jobs simultaneously and using offers in one market to negotiate wages in the other. At the same time, the descriptive statistics in Table 2.3 show that the external and internal labor markets are very different. Within fires, quits, and transfers, the standard deviations of log wage changes are greater for external than for internal job changes. The same is true for the interquartile ranges, except for transfers where the external and internal interquartile ranges are roughly equal. The

external labor market offers a greater range of opportunities but also poses more downside risk.

In Table 2.4, we examine the robustness of our results in Table 2.3 when controlling for the characteristics of the workers who change jobs, their locations, and the years in which the job changes occur. The first column of Table 2.4 simply replicates the mean log wage changes in Table 2.3, with external fires as the omitted category so its mean log wage change is given by the constant. The coefficients on the indicators for the other job change categories yield the mean log wage changes in Table 2.3 relative to external fires. In the second column we add indicators for each age in years (from 25 to 55) and each education level provided in RAIS (listed in Appendix Table A.8). The coefficients on the job change categories and the constant are virtually unchanged from the first column. In the third column we add indicators for occupation and location at the levels used to define labor market cells, and an indicator for the last year of the initial job spell of the job change. (The data on worker age, education, location, and occupation are also taken from the last year of the initial job spell of his job change.) Again, the various coefficients are virtually unchanged.

In Table 2.5, we examine more formally the hypothesis that external and internal labor markets are integrated. We pool together the job changes we have classified as voluntary, external quits and internal and external transfers, and the job changes we have classified as involuntary, external and internal fires and internal quits, and regress the log wage changes for voluntary and involuntary job changes on a constant, an indicator for internal, and compositional controls. If external and internal labor markets are integrated we expect the coefficients on the internal indicator to be zero. The second and fourth columns of Table 2.5 show that when all compositional controls are included we cannot reject the hypothesis that the coefficient on the internal indicator is zero for either voluntary or involuntary job changes.

2.4.2 Months between jobs

We use months between jobs to help distinguish voluntary from involuntary job changes. Then we find that wage changes are greater for voluntary job changes. A natural question to ask

is, how much, if any, of the larger wage changes for voluntary job changes can be accounted for by fewer (or zero) months between jobs?

In and of themselves, months between jobs might be associated with larger or smaller wage changes. On the one hand, more months between jobs might indicate that workers are being more selective before accepting their new jobs. On the other hand, more months between jobs might indicate that workers are failing to find new jobs and have to settle for less when their savings run out. Antel (1991) finds, using the U.S. National Longitudinal Survey of Young Men, that an unemployment spell between jobs is associated with higher wage gains. If his results apply, it might be that accounting for months between jobs will make the difference in wage changes between voluntary and involuntary job changes even larger.

To investigate, in the first column of Table 2.6 we regress log wage changes for all job changes in our sample on an indicator for an intervening spell of nonemployment and all fixed effects used in the last column of Table 2.4. An intervening spell of nonemployment is associated with a wage loss of more than five percent, as we would expect since we are not controlling for the involuntary versus voluntary nature of the job changes. In the second column of Table 2.6 we add the logarithm of months between jobs when the number of months is positive; this can be interpreted as the interaction between the indicator for an intervening spell of nonemployment and the logarithm of months between jobs since it equals zero when there is no intervening spell and log months otherwise. The coefficients on both the indicator and its interaction with log months are negative and economically and statistically significant. We then add indicators for the five categories of job changes to the specifications in columns (1) and (2), with external fires as the omitted category as in Table 2.4. In the third column of Table 2.6 the coefficients on the indicators for the five categories of job changes are barely changed relative to Table 2.4. In the fourth column of Table 2.6 the impacts on wage changes of all voluntary job changes decrease substantially relative to external fires: compared to Table 2.4, the impacts are about one, one and a half, and two and a half percent smaller for external quits, external transfers, and internal transfers, respectively.

The coefficients on the indicator for intervening spell of nonemployment and its interaction with log months between jobs shrink in absolute value but remain statistically significant between columns (1) and (3) and between columns (2) and (4), respectively, of Table 2.6. These results contradict Antel (1991) and indicate that in Brazil during the period 2002-2014 longer job search is associated with lower wage gains. For this reason controlling for months between jobs does account for a small but economically significant part of the difference in wage changes between voluntary and involuntary job changes.

Tables 2.7-2.8 consider the impacts of months between jobs on wage changes in more depth. It appears from a comparison of the two tables that months between jobs are penalized more for external than for internal job changes. For internal transfers, a spell of nonemployment actually has a positive association with log wage changes. We conjecture that for internal transfers a spell of nonemployment indicates that the worker has received external training (or internal, unpaid training) for a better paid position within the firm.

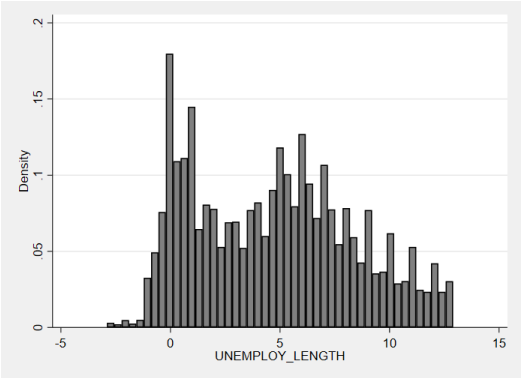
2.5 Conclusion

We used the Brazilian RAIS data to describe both internal and external worker job mobility. We have also provided a more detailed and accurate description of voluntary versus involuntary changes based on reported separation causes and nonemployment length between jobs. Based on these categorizations, we show that including internal job changes increases the total amount of voluntary job changes by 40%, and the total amount of involuntary job changes by 50%. Wage gains for the two main categories of voluntary job changes and wage losses for the two main categories of involuntary job changes are similar. This is evidence that the internal and external labor markets are integrated, meaning that workers search for, consider, and bargain with options in both markets when switching to new jobs.

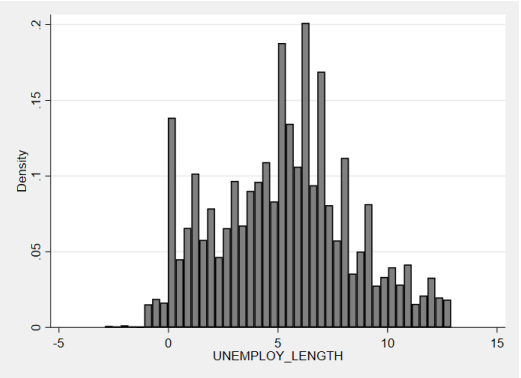
A major implication of our finding that internal job changes account for a significant percentage of total job changes, both voluntary and involuntary, is that including internal changes

would greatly affect current measures of job turnover in the literature that only records external job changes. This could discipline future calibrations of job turnover in the macro job search literature. The next step to our research would be to show the effects of including internal worker circulation rates on a macro calibration of a job search model.

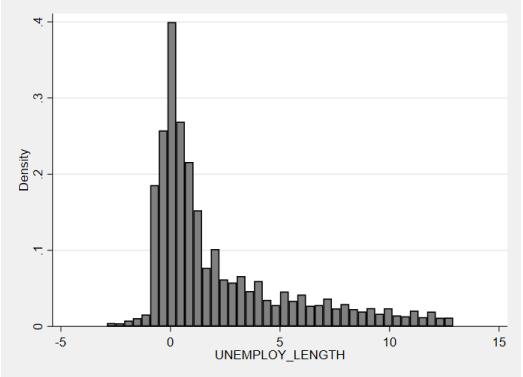
2.6 Tables and Figures



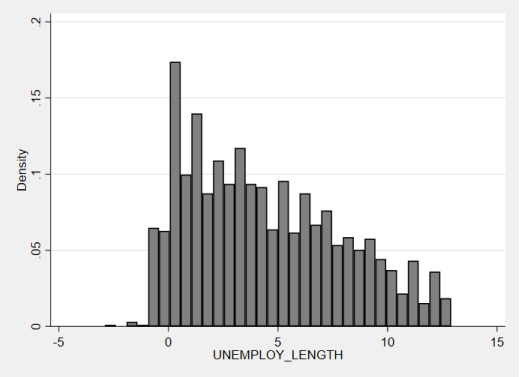
(a) External fires



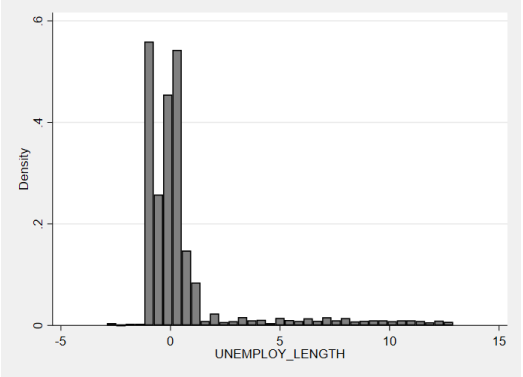
(b) Internal fires



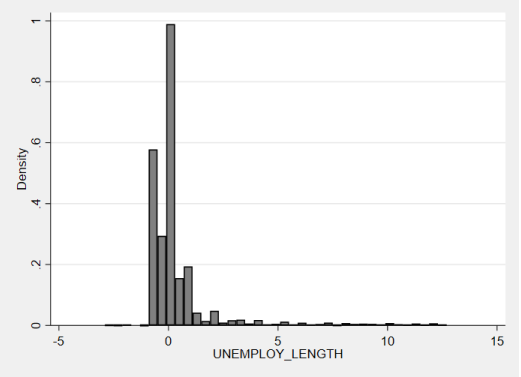
(c) External quits



(d) Internal quits



(e) External transfers



(f) Internal transfers

Figure 2.1. Histograms with the number of months between jobs, truncated after 13 months

Table 2.1. External and Internal Job Changes Within Cell, by Cause of Separation

	external	internal
fires	72468	33448
quits	14158	2473
transfers	6857	8350
total	93483	44271

Table 2.2. Nonemployment Length (in Months) for Job Changes in Table 2.1

	count	mean	sd	min	max	p25	p50	p75
external fires	72468	11.09	14.84	-2.90	146.20	2.50	6.50	13.10
internal fires	33448	8.22	9.04	-2.90	127.10	3.60	6.10	9.10
external quits	14158	6.75	13.64	-2.90	142.10	0.10	1.30	7.10
internal quits	2473	8.96	12.33	-2.90	120.10	2.10	5.30	10.70
external transfers	6857	6.67	16.08	-2.90	131.10	-0.50	0.10	3.70
internal transfers	8350	1.79	7.87	-2.90	127.10	-0.50	0.00	0.50

Table 2.3. Log wage changes for job changes in Table 2.1

	count	mean	sd	min	max	p25	p50	p75
external fires	72468	-0.03	0.45	-4.32	2.97	-0.20	0.00	0.18
internal fires	33448	-0.03	0.34	-3.47	3.35	-0.12	-0.01	0.10
external quits	14158	0.09	0.46	-3.74	3.05	-0.12	0.08	0.32
internal quits	2473	-0.00	0.40	-3.47	2.17	-0.09	0.01	0.14
external transfers	6857	0.04	0.34	-3.66	3.53	-0.06	0.04	0.16
internal transfers	8350	0.08	0.29	-3.35	3.13	-0.04	0.05	0.18

Table 2.4. Log wage change means with compositional controls

	(1) log wage change	(2) log wage change	(3) log wage change
Internal fires	-0.00134	0.000193	-0.000165
	(-0.40)	(0.06)	(-0.05)
External quits	0.120***	0.118***	0.115***
	(23.44)	(22.67)	(22.81)
Internal quits	0.0296**	0.0303**	0.0289**
	(3.20)	(3.28)	(3.20)
External transfers	0.0696***	0.0707***	0.0671***
	(14.31)	(13.99)	(12.92)
Internal transfers	0.113***	0.115***	0.108***
	(24.33)	(23.90)	(20.09)
_cons	-0.0305***	-0.0309***	-0.0299***
	(-11.83)	(-11.68)	(-16.57)
Age FE	No	Yes	Yes
Edu FE	No	Yes	Yes
Occ. FE	No	No	Yes
Meso FE	No	No	Yes
Year FE	No	No	Yes
N	137754	137754	137754
R^2	0.012	0.013	0.031

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 2.5. Log wage change means, pooled voluntary and involuntary job changes

	voluntary log wage change	voluntary log wage change	involuntary log wage change	involuntary log wage change
internal dummy	0.0112*	0.00833	0.00242	0.000798
	(2.10)	(1.52)	(0.74)	(0.25)
_cons	0.0725***	0.0733***	-0.0310***	-0.0305***
	(18.14)	(23.85)	(-12.01)	(-18.17)
Age FE	Yes	Yes	Yes	Yes
Edu FE	Yes	Yes	Yes	Yes
Occ. FE	No	Yes	No	Yes
Meso FE	No	Yes	No	Yes
Year FE	No	Yes	No	Yes
N	29365	29361	108389	108387
R^2	0.003	0.024	0.002	0.023

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 2.6. Log wage change means, accounting for intervening non-employment

	(1)	(2)	(3)	(4)
	log wage change	log wage change	log wage change	log wage change
nonemploy dummy	-0.0545***	-0.0319***	-0.0175***	-0.00704
	(-13.12)	(-7.65)	(-3.86)	(-1.56)
nonemploy interact		-0.0184***		-0.0129***
		(-16.15)		(-11.72)
Internal fires			0.000578	-0.000109
			(0.18)	(-0.03)
External quits			0.114***	0.105***
			(22.55)	(21.15)
Internal quits			0.0294**	0.0281**
			(3.25)	(3.12)
External transfers			0.0609***	0.0506***
			(11.13)	(9.09)
Internal transfers			0.101***	0.0840***
			(18.06)	(15.22)
_cons	0.0406***	0.0460***	-0.0137**	-0.00239
	(10.12)	(11.07)	(-2.92)	(-0.49)
Age FE	Yes	Yes	Yes	Yes
Edu FE	Yes	Yes	Yes	Yes
Occ. FE	Yes	Yes	Yes	Yes
Meso FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
N	137754	137754	137754	137754
R^2	0.023	0.027	0.031	0.033

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 2.7. Log wage change means, external, with months between jobs

	fires log wage change	quits log wage change	transfers log wage change
nonemploy dummy	-0.00134	-0.00375	-0.0119
	(-0.18)	(-0.35)	(-1.46)
nonemploy interact	-0.0171***	-0.0177***	-0.0162***
	(-11.02)	(-6.42)	(-5.71)
_cons	-0.000727	0.105***	0.0521***
	(-0.10)	(11.25)	(9.02)
Age FE	Yes	Yes	Yes
Edu FE	Yes	Yes	Yes
Occ. FE	Yes	Yes	Yes
Meso FE	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
N	72464	14152	6845
R^2	0.028	0.038	0.069

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 2.8. Log wage change means, internal, with months between jobs

	fires log wage change	quits log wage change	transfers log wage change
nonemploy dummy	0.0151	0.0220	0.0185*
	(0.50)	(0.35)	(2.22)
nonemploy interact	-0.00413	-0.00932	0.00522
	(-1.92)	(-1.12)	(1.83)
_cons	-0.0399	-0.00853	0.0750***
	(-1.32)	(-0.14)	(14.92)
Age FE	Yes	Yes	Yes
Edu FE	Yes	Yes	Yes
Occ. FE	Yes	Yes	Yes
Meso FE	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
N	33448	2453	8339
R^2	0.030	0.135	0.052

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Chapter 2 is currently being prepared for submission for publication. Zeng, Wendy. The dissertation author is the principal researcher on this chapter.

Chapter 3

Former Occupations and Entrepreneurship

3.1 Introduction

We expect individuals who rise to the level of management to have an aptitude for managing and high generic ability or human capital. Both of these traits should make these individuals more likely to become entrepreneurs and to found larger firms, conditional on entrepreneurship.

Using a sample of Brazilian workers who left formal sector jobs to start limited liability firms or find new formal sector employment during the period 2002-2014, we find that former managers account for 17 percent of entrepreneurs and nearly one quarter of new firm hiring at startup, despite being only four percent of the sample. Former managers have much greater generic human capital than workers with other former occupations, as indicated most importantly by their earnings in their last jobs but also by their educational attainments. Yet our regression analysis shows that generic human capital variables account for only a small part of the higher propensity of former managers to become entrepreneurs or the larger sizes of their firms at startup. After also controlling for former employer fixed effects, cause of separation from former employer, and many other variables, we ascribe the remainders to the managerial aptitude or "talent" that led them to employ their human capital in management. In our preferred specifications, directors, managers, and supervisors are, respectively, roughly six times, four

times, and two times more likely than non-managerial workers to become founding owners of limited liability firms, and these firms employ 40, 18, and 14 percent more workers at startup. We judge our results to be closer to the model of Lucas (1978), where managerial talent is distinct from generic ability, than to the model of Murphy, Shleifer and Vishny (1991), where generic ability is more highly rewarded in entrepreneurship than in wage employment but there is no distinct talent for management.

After demonstrating that managerial talent has impacts on entrepreneurship and new firm size beyond what can be explained by generic human capital, we investigate whether other occupational skills have comparable impacts. The literature on the determinants of firm success has emphasized both the importance of firm productivity and of developing a customer base (Jovanovic, 1982; Foster, Haltiwanger and Syverson, 2008, 2016; Gourio and Rudanko, 2014). We therefore examine occupations connected to technology and sales, respectively. The work of Amiti and Konings (2007) and Halpern, Koren and Szeidl (2015) showing that access to imported inputs increases productivity suggests that we also examine occupations connected to purchasing. Only purchasing has impacts on both entrepreneurship and new firm size comparable to those of managerial occupations.

The importance of management for firm performance has been well established for existing firms (Bloom and Van Reenen, 2007; Bloom et al., 2013). Fairlie and Robb (2007) are much closer to our work in that they consider business *founders*. They find that "previous work experience in a managerial capacity" is associated with larger (log) sales and a higher probability of having employees in the founding owner's firm, controlling for many human capital attributes (but not the founder's previous earnings). Their data do not allow them to examine the determinants of ownership and do not cover the legal form typically used by large firms.

As we do, Levine and Rubinstein (2017) track former employees to self-employment in limited liability (incorporated) firms. However, they do not know whether these self-employed individuals are founding owners as opposed to joining existing businesses. They are more

concerned with which individuals increase their earnings through self-employment than with which individuals create jobs. Choi et al. (2021) track former employees to "founding teams" that contain a mixture of owners of sole proprietorships and individuals with positive earnings in the first year of operation of incorporated firms and sole proprietorships. Their main interest is in the impact of premature death of a founding team member on subsequent firm performance.

The next section of our paper describes our data, our identification of firm founders, and how we link them to previous employment. Section 3.3 examines the tendency of former managers to be overrepresented in entrepreneurship and to found larger firms. Section 3.4 does the same for non-managerial occupations. Section 3.5 explores the robustness of our results. Section 3.6 concludes.

3.2 Data

Our study combines two Brazilian government data bases. Linked employer-employee data allow us to associate workers to firms. Records of "proprietors" (owners) from the official firm registry allow us to associate owners to firms. Integrating the two data bases allows us to associate owners to workers.

The linked employer-employee data base RAIS (*Relação Anual de Informações Sociais* of the Brazilian labor ministry *MTE*) is organized by job spell. Each job spell is identified by employee ID, firm-plant ID, and dates of job accession and separation. The first eight digits of the 14-digit firm-plant ID identify the firm to which the plant belongs. Worker age, education, and gender are recorded, as are earnings, industry, and occupation. By Brazilian law, every private or public-sector employer must report this information every year.[1] De Negri et al. (1998) compare labor force information in RAIS to that in a main Brazilian household survey (PNAD)

[1]RAIS primarily provides information to a federal wage supplement program (*Abono Salarial*), by which every employee with formal employment during the calendar year receives the equivalent of a monthly minimum wage. RAIS records are then shared across government agencies. An employer's failure to report complete workforce information can, in principle, result in fines proportional to the workforce size, but fines are rarely issued. In practice, employees and employers have strong incentives to ascertain complete RAIS records because payment of the annual public wage supplement is exclusively based on RAIS. The ministry of labor estimates that well above 90 percent of all formally employed individuals in Brazil are covered in RAIS.

and conclude that, when comparable, RAIS delivers qualitatively similar results to those in the national household survey. Menezes-Filho, Muendler and Ramey (2008) apply the Abowd et al. (2001) earnings-estimation methodology to Brazil and show that labor-market outcomes from RAIS broadly resemble those in France and the United States, even after controlling for selection into formal employment, except for unusually high returns to high school and college education and to experience among males.

The ownership records are from the Brazilian tax authority Secretaria da Receita Federal (RF). For each firm with recorded owners the RF data base provides the owners' dates of accession to the firm. Ownership shares and any dates of separation are not provided.[2] We refer to the owners present at firm startup as the *founding owners*. Many firms do not have recorded owners. In particular, owners of "non-incorporated" firms are typically not recorded in the RF data base. These firms tend to be small. We restrict our analysis to firms with limited liability legal form (Sociedade Empresária ou Simples Limitada), which account for 53.9 percent of new private-sector, for-profit firms and 65.1 percent of their December employment during the period 2002-2014.[3]

Starting in 2002, RAIS includes the worker tax ID in addition to the RAIS-specific employee ID. This enables the Brazilian central bank BACEN to match owners to workers. The so-combined data are provided to us by BACEN, which is permitted to share these data with researchers who have independent access to RAIS. 2002 is therefore the first year of our sample period. In 2014 Brazil entered a political crisis. GDP growth remained marginally positive in 2014 but there was a severe recession with negative GDP growth of 3-4 percent in both 2015 and 2016. We therefore make 2014 the last year of our sample period.

We identify new firms as firms with eight-digit IDs that did not appear in RAIS prior to

[2]The RF records on owners by firm are publicly available through an online query interface maintained by the federal Brazilian tax authority.

[3]We exclude from this calculation firms with any state ownership, cooperatives, business groups, and foreign subsidiaries. Of the 46.1 percent of new private-sector, for-profit firms that do not have limited liability legal form, over 95 percent are sole proprietor commercial companies (legal form Empresário (Individual)).

2002 (we search back to 1986, the first year available to us).[4] Founding owners are then defined as owners whose years of accession to a firm are the same or earlier than the first year the firm appears in RAIS. Having identified founding owners, we trace them to the most recent jobs they held prior to (or in) their new firms' first years. We only consider jobs held for at least three months. We use the last month that the proprietor was employed to identify the most recent job. If there is more than one most recent job, we use the highest paying job. We break any remaining ties by keeping the job with the highest firm ID number, or the last alphabetical occupation description.

A concern in U.S. studies such as Levine and Rubinstein (2017) is that owners should be employed by the firms they own so we can be sure they are not passive investors. In our data, only 1.4 percent of founding owners of limited-liability firms who can be traced to last jobs from which they separated are employed by the firms they start. Tax incentives in Brazil strongly favor paying executive compensation as dividends rather than salaries. Moreover, 72 percent of founding owners leave their previous jobs, which they have no reason to do if they are passive investors. We shall see below that when the founding owners do not leave their previous jobs the firms they start tend to be significantly smaller, consistent with these firms suffering from lack of owner attention compared to firms founded by owners who leave their jobs.

3.3 Managerial Talent and Job Creation

For the remainder of this paper we focus on new firms whose founding owners can be traced to previous employment. We hypothesize that workers with greater managerial talent are more likely to have held managerial positions in their last jobs. We can identify three levels of managerial occupations in RAIS: "directors," "managers," and "supervisors." We hypothesize that among workers whose last position was in management, those with more managerial talent rose to the level of manager and those with the most managerial talent rose to the level of director.

Occupational classifications in the RAIS data follow the CBO (Classificação Brasileira

[4]Firms that are sold do not receive new tax IDs, so we do not identify them as new.

de Ocupações). We use the second edition, published in 2002, which codes 2511 occupations at the six-digit level. For the sample of workers used in Figure 3.1 (or Table 3.1) below, 3.2 percent have occupations coded only using the first edition of the CBO, published in 1994. We use the workers who have both 1994 and 2002 codes to find the modal 2002 code for each 1994 code, and use these modal 2002 codes to assign 2002 codes to workers with only 1994 codes.[5].

We include as directors and managers all workers whose last jobs were coded in principal subgroups 12 and 14, respectively. This excludes directors and managers in principal subgroup 13, which is comprised of "directors and managers of personal, social, and cultural enterprises," "directors and managers of health service enterprises," and "directors and managers of educational service institutions." We consider adding these in a robustness check. Supervisors, unlike directors and managers, are not grouped together by the CBO. We included every six-digit occupation title containing the word "supervisor," plus occupations in four-digit groups with "supervisor" in the group title.

We were able to trace 880,558 founding owners to previous employment. As mentioned earlier, 72 percent of these separated from their last jobs.[6] An advantage of working with this sample is that we can control for cause of separation (e.g., fired versus quit) in the regressions below. We compare these founding owners to other workers who separated from their last jobs, but found jobs with different firms instead of founding their own new firms. More specifically, we consider workers whose last jobs were in the firm-years in which founding owners held their last jobs. This allows us to include firm-year fixed effects in regressions with founding owner status as the dependent variable, so that differences in propensity to become entrepreneurs is identified using characteristics of workers originating from the same firms in the same years. Firm-year fixed effects will control for differences in "entrepreneurial culture" between firms.

[5]More specifically, we use the modal one-digit 2002 code to assign workers with only 1994 codes to the 2002 major (one-digit) occupation groups, the modal two-digit 2002 code to assign these workers to the 2002 principal (two-digit) subgroups, and the full mode 2002 code to assign these workers to specific 2002 occupations

[6]563,132 of these have codes for cause of separation, and 71,750 are not recorded as having separated but are also no longer recorded as employed in their previous jobs. We include all founding owners not recorded as separated from their last jobs in the firm size regressions in subsection 3.3.2.

Prior to the regression analysis of the next two subsections, we show in Figure 3.1 the shares of workers, founding owners, and new firm employment accounted for by managerial and other occupations. The sample of founding owners in the second bar of the figure is those who 1) started limited-liability firms that first appeared in RAIS during the period 2002-2014, 2) can be traced to last jobs from which they separated, and 3) have, from those last jobs, occupation codes and other worker characteristics used in the entrepreneurship regressions below. Working backward, the sample of workers in the first bar of the figure is those who 1) were employed in firm-years in which founding owners held their last jobs, 2) separated from those jobs and found new (formal sector) jobs or became founding owners before the end of 2014, and 3) have, from their last jobs, occupation codes and other worker characteristics used in the entrepreneurship regressions below. Working forward, the sample of firms for which employment is counted in the third bar of Figure 3.1 is those with limited liability that first appeared in RAIS during 2002-2014 and had at least one founder in the sample for the second bar of the figure. We use employment in December of the year in which the firm first appeared in RAIS.[7] Letting n equal the number of founders for a given new firm, we attribute $1/nth$ of its December employment to each founder since we have no data on ownership shares. In Figure 3.1 n is restricted to the founders in our sample. We could also include in n founders who did not separate from their last jobs, or founders who did not separate from their last jobs and founders who cannot be traced to previous jobs, but neither of these changes substantially affects the third bar in Figure 3.1.

Figure 3.1 shows that, although only four percent of workers who left their jobs held managerial occupations, they account for 17 percent of founding owners of limited liability firms and nearly one-quarter of the employment of those firms (attributable by previous occupation to founding owners who separated from their last employers) in December of their founding years. The overrepresentation of former managers in founding owners and new firm job creation is greatest for former directors and least for former supervisors.

[7] For every worker with an employment spell reported during a given year, RAIS provides an indicator whether the employment spell lasted through December 31, facilitating a count of the number of employees at that point in time.



Figure 3.1. Shares of workers, founding owners, and new firm employment accounted for by managerial and other occupations

The overrepresentation of former managers in entrepreneurship and new firm employment may, of course, have many causes other than talent for management. To investigate we estimate a series of regressions in which an indicator for firm founder or the logarithm of new firm employment is the dependent variable. We control for a steadily increasing number of alternative explanations of the association of former managers with entrepreneurship and new firm employment. We will use linear regressions to examine the impacts of former managerial occupations on the probability of becoming founding owners in the next subsection, and on the logarithm of new firm size in following subsection. We will call the regressions in the next subsection "entrepreneurship regressions" and the regressions in the following subsection "firm size regressions."

3.3.1 Managerial occupations and probability of becoming founding owners

Table 3.1 provides descriptive statistics for the worker sample to be used in the entrepreneurship regressions.[8] As expected, every group of workers with a managerial occupation has higher mean wages, tenure, age, and post-high school education than non-managerial workers. Moreover, directors have higher means than managers and managers have higher means than supervisors for all of these characteristics. The variable "lag" equals the difference in calendar years between the year the worker starts his new job or founds his firm and the year in which he separates from his previous job.[9] 78 percent of non-managerial workers, 81 percent of former directors, and 84 percent of former managers or supervisors have lag values equal to zero or one, so most workers who are going to return to the formal sector do so quickly. Directors, managers, and supervisors are all less likely to have been fired from their last jobs than non-managerial workers, but managers and supervisors are also less likely than non-managerial workers to have quit. Investigation of other causes of separation shows that workers in managerial occupations are far more likely than workers in non-managerial occupations to have retired from their last jobs.

Table 3.2 reports our entrepreneurship regressions. The first column includes no covariates other than indicators for the three managerial occupation levels. The coefficients indicate that former directors, managers, and supervisors are 8.0, 5.7, and 1.3 percentage points more likely to become founding owners than workers who had other former occupations. These are very large effects relative to the mean founding owner probability of 1.1 percent, as we would have expected from Figure 3.1.[10]

[8] More precisely, the worker sample is for all regressions that include firm-year fixed effects. There are 42,110 workers who become founding owners who separate from firm-years from which no other workers separate. These observations are automatically dropped from the fixed effect regressions.

[9] During the intervening years the worker may be out of the labor force, unemployed, employed in the informal sector, or owner of an informal sector firm.

[10] If founders of non-limited liability firms are included as former workers who did not become entrepreneurs, they slightly lower the coefficients on the managerial occupation indicators. 93 percent of our founding owners founded limited liability firms.

Table 3.1. Summary statistics for worker sample used in entrepreneurship regressions

	(1) all	(2) non-managers	(3) directors	(4) managers	(5) supervisors
Founding owner	0.011	0.010	0.088	0.064	0.022
Deflated wage (reals)	507.387	473.027	3585.887	1575.744	833.536
	271.029	264.398	1636.095	916.075	500.888
Tenure (months)	31.929	31.086	53.704	50.942	50.325
	13.100	12.900	26.400	23.300	23.900
Age (years)	32.407	32.254	41.837	35.647	35.509
	30.000	30.000	41.000	34.000	34.000
Female = 1	0.361	0.361	0.281	0.325	0.379
Less than high school	0.440	0.452	0.081	0.135	0.208
High school graduate	0.348	0.347	0.202	0.323	0.391
Some college	0.056	0.054	0.075	0.104	0.115
College graduate	0.153	0.144	0.629	0.434	0.283
Graduate degree	0.003	0.003	0.013	0.005	0.002
Fired	0.673	0.679	0.553	0.454	0.595
Quit	0.171	0.172	0.188	0.143	0.129
Other cause of separation	0.156	0.148	0.258	0.403	0.276
Lag (years)	1.200	1.212	1.013	0.895	0.933
Lag share = 0 or 1	0.782	0.779	0.811	0.843	0.843
occ group 0	0.001	0.001	0.000	0.000	0.000
occ group 1	0.032	0.011	1.000	1.000	0.000
occ group 2	0.088	0.091	0.000	0.000	0.025
occ group 3	0.085	0.088	0.000	0.000	0.044
occ group 4	0.350	0.352	0.000	0.000	0.577
occ group 5	0.142	0.144	0.000	0.000	0.156
occ group 6	0.070	0.072	0.000	0.000	0.041
occ group 7	0.179	0.185	0.000	0.000	0.076
occ group 9	0.023	0.023	0.000	0.000	0.034
Observations	51636208	49420552	83230	1005648	1126778

Founding owner equals 1 if a former worker becomes a founding owner of a limited liability firm and equals 0 if the former worker remains a worker. Former workers who do not reappear in the formal sector, who become founding owners of firms with other legal forms, or who become non-founding owners are not included in the sample. Less than high school includes former workers who are illiterate, primary school dropouts or graduates, middle school dropouts or graduates, or high school dropouts.

Table 3.2. Determinants of entrepreneurship

	(1)	(2)	(3)	(4)	(5)	(6)
Director	0.0799***	0.0727***	0.0720***	0.0585***	0.0524***	0.0527***
	(0.000370)	(0.000354)	(0.000353)	(0.000401)	(0.000401)	(0.000401)
Manager	0.0569***	0.0468***	0.0470***	0.0321***	0.0298***	0.0306***
	(0.000107)	(0.000107)	(0.000107)	(0.000223)	(0.000222)	(0.000223)
Supervisor	0.0128***	0.0142***	0.0145***	0.0178***	0.0111***	0.0116***
	(0.000102)	(0.0000980)	(0.0000977)	(0.0000987)	(0.000100)	(0.000100)
Log deflated wage					0.00295***	0.00307***
					(0.0000264)	(0.0000264)
Log tenure					0.00222***	0.00255***
					(0.0000166)	(0.0000172)
Log age					0.0131***	0.0137***
					(0.0000553)	(0.0000557)
Female = 1					−0.00309***	−0.00301***
					(0.0000341)	(0.0000341)
occ group 0				0.0142***	0.00969***	0.00979***
				(0.00166)	(0.00166)	(0.00166)
occ group 1				0.0224***	0.0137***	0.0140***
				(0.000221)	(0.000223)	(0.000223)
occ group 2				0.0189***	0.00984***	0.0100***
				(0.000118)	(0.000125)	(0.000125)
occ group 3				0.0113***	0.00669***	0.00700***
				(0.000114)	(0.000115)	(0.000115)
occ group 4				0.000104	−0.00145***	−0.00117***
				(0.000108)	(0.000109)	(0.000109)
occ group 5				0.0157***	0.0145***	0.0147***
				(0.000110)	(0.000110)	(0.000110)
occ group 6				−0.00239***	0.000579***	0.000599***
				(0.000138)	(0.000139)	(0.000139)
occ group 7				−0.00174***	−0.00191***	−0.00189***
				(0.000107)	(0.000107)	(0.000107)
occ group 9				0.00249***	0.00179***	0.00193***
				(0.000139)	(0.000139)	(0.000139)
Constant	0.0101***	0.00968***	0.00796***	0.00261***	−0.0714***	−0.0748***
	(0.0000152)	(0.0000139)	(0.0000244)	(0.000101)	(0.000281)	(0.000284)
ymean	0.0116	0.0110	0.0110	0.0110	0.0110	0.0110
r2	0.00652	0.145	0.150	0.153	0.156	0.157
N	51678318	51636208	51636208	51636208	51636208	51636208

Standard errors in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Dependent variable is indicator for whether former worker becomes a founding owner. Columns (2)-(6) include origin firm x year fixed effects. Columns (3)-(6) include lag indicators for 1 through 20 years (0 lag is omitted category). Columns (5)-(6) include indicators for 10 education categories (illiterate is omitted category). Column (6) includes cause of separation fixed effects. Real wages are provided in 1994 levels, using the inflation rate from December of each worker's last job. Tenure is provided in months and age in years. 42110 founding owners are dropped from the sample in columns (2)-(6) because no workers leave the owner's last firm in the same year.

In the second column of Table 3.2 we add firm-year fixed effects, so that we identify the effect of former occupation on entrepreneurship using the differences between workers who

left the same firm in the same year. This addresses the concern that managers might come from especially "entrepreneurial" firms. The coefficients on the director and manager indicators decrease somewhat and the coefficient on the supervisor indicator increases. In the third column we include 20 indicators for the number of years from 1 to 20 between when a worker left his job and when he either started another (formal sector) job or founded a limited liability firm. The coefficients on these indicators (not shown) are nearly all positive, increase monotonically to a peak at eight years, and decline monotonically thereafter. A plausible explanation for the positive coefficients is that if a worker disappears from the formal sector for years only to reappear, there is a substantial chance that he was running an informal sector firm that he chose to formalize.[11] This explanation is consistent with the negative association of lag years with firm size that we find below. The coefficients on the managerial occupation indicators are virtually unchanged. In the fourth column we add indicators for the nine major CBO occupation groups.[12] Directors and managers belong to group 1, "Senior members of the government, leaders of organizations of public interest and companies and managers." Supervisors belong mainly to group 4, "Workers in administrative services." The coefficients on the director and manager indicators fall substantially, both by just under two-thirds of the coefficient on the indicator for occupation group 1, suggesting that some of the association of former directors and managers with entrepreneurship can be ascribed to generic "leadership." We postpone our main discussion of the coefficients on the major occupation group indicators themselves to Section 3.4, after we add indicators for more specific non-managerial former occupations.

In the fifth column of Table 3.2 we add various measures of human capital. Perhaps the most important is the log wage of the worker at his last job. This wage includes all compensation that is taxable income or subject to Brazilian social security contributions; see Hirakawa, Muendler and Rauch (2010, Appendix A) for complete details. It should be a good proxy for

[11] The inverted-U traced by these coefficients is a puzzle, however.

[12] There are actually ten major groups, but groups 7 and 8 have the same title, "Workers producing industrial goods and services." We let group 8 be the omitted category.

generic worker ability.[13] We also add indicators for the eleven educational attainment categories reported in RAIS (coefficients not shown), log age, log months of tenure in last job, and gender, which can affect a worker's propensity to become a founding owner in ways other than through his generic ability. The signs of all coefficients are as expected.[14] The coefficients on the managerial occupation indicators fall 10, 7, and 38 percent, respectively, for directors, managers, and supervisors. The coefficient on the indicator for the major CBO occupation group to which directors and managers belong falls 39 percent. The coefficients on the indicators for the other high human capital occupation groups, 2 ("professionals of sciences and the arts") and 3 ("middle-level technicians"), are similarly reduced.

In the last column of Table 3.2 we add fixed effects for all 29 causes of separation of workers from their last jobs. The concern here is that managers are less likely to have been fired from their last jobs than other workers, so their greater propensity to become founding owners could reflect selection. Managers are also more likely to have retired from their last jobs than other workers, and retirement could encourage entrepreneurship by providing a pension. Despite these concerns, inclusion of cause of separation fixed effects has virtually no effect on the coefficients on the managerial occupation indicators. It may be that by only considering workers who remain in the formal sector as employees or founding owners we have taken care of most potential selection effects.

The last column of Table 3.2 is our preferred specification. Having been a director, manager, or supervisor respectively increases the probability of becoming a founding owner by 5.3, 3.1, and 1.2 percentage points. These effects imply that directors, managers, and supervisors are, respectively, roughly six times, four times, and two times more likely than non-managerial workers to become founding owners.

Another way to quantify the impacts of having held a managerial occupation is to

[13] Choi et al. (2021, p. 2) follow a similar procedure. They write, "Leveraging the longitudinal structure of the matched employer-employee data, we use each founding team member's most recent earnings prior to joining the startup as a proxy for their human capital."

[14] We omitted the lowest educational attainment category, and all coefficients for the other education categories are positive and significant, with the highest coefficients for some college and college graduate.

compare them to the increase in the probability of becoming a founding owner attributable to standard human capital variables. We will compute the latter by multiplying the coefficients in the last column of Table 3.2 on log age, education, female, log tenure, and log wage by the differences between the means of these variables for former managers (the most common former managerial occupation among founders) and former holders of non-managerial occupations. This computation yields a value of 0.0099, 90 percent of the mean of the dependent variable but less than one-third of the coefficient on manager in the last column of Table 3.2. We interpret this result to show that the managerial talent of former managers has more than three times the impact on the probability they will become entrepreneurs than does their greater generic ability relative to former holders of non-managerial occupations.

3.3.2 Managerial occupations and firm size

The results of the previous subsection are consistent with the hypothesis that former managers are more likely to start their own firms because their managerial talent will make their firms more productive, but are also consistent with former managers having personal traits such as high self-esteem (Levine and Rubinstein, 2017) that increase their propensities to become firm owners but do not necessarily make their firms more productive. In this subsection we examine the association of former managerial occupations with new firm size as measured by employment. This should increase with firm productivity, and is also of interest because it measures new job creation.

New firms may have more than one traceable founder. We will consider firms for which all founders are traceable so that we know the former occupation, demographic characteristics, and other relevant information for every founder. We do not require founders to have separated from their former employers because we are not comparing them to individuals who separated from their former employers and remained workers. Our criteria yield a sample of 537,741 new firms, of which 71.2 percent have one traceable founder, 26.0 percent have two, 2.3 percent have three, and 0.5 percent have four or more. 1.0 percent of the firms in our sample also have one

or more Brazilian legal entities (presumably firms) as founding owners, of which 17.9 percent are the origin firms of one or more of the individual founding owners. In the latter case it seems highly likely that the new firms are divestitures or corporate spinoffs from the origin firms, so we can expect them to be much larger than typical new firms. This may also be true when the founding firms are not origin firms of the founding individuals, but in this case it seems equally plausible that the founding individuals recruited the founding firms.

With these considerations in mind we compute mean and median new firm sizes for the 382,844 firms in our sample with only one traceable founder, so that we can classify them unambiguously by the (former) occupation of the single individual founder. We also allow firm founding owners to override individual founding owners, so that new firms with firm founding owners are not included in the computations for new firms with individual founding owners. Table 3.3 shows that, as we could anticipate from Figure 3.1, new firms founded by former managers are larger on average than new firms founded by workers who held non-managerial positions, and that this size difference increases with managerial rank. However, median new firm size mostly does not differ across managerial versus non-managerial founders. Evidently superior managerial talent of former managers mainly reveals itself in the right tail rather than throughout the distribution. Table 3.3 also shows that new firms with at least one existing firm founder are far larger than new firms with only individual founders, especially if the individual founder came from one of these existing firms.

Table 3.3. Mean and median sizes of firms with only one trackable founder

	No. of workers		
Owner	Mean	Median	Obs
All	4.5	2	382844
Non-management	3.9	2	306301
Directors	8.6	3	6107
Managers	5.5	2	48431
Supervisors	5.1	2	17976
Existing firm	20.1	4	3424
Existing origin firm	37.7	8	605

We now turn to regressions to further explore the determinants of new firm size. In light of the skewness of the distribution of new firm sizes indicated by Table 3.3, we use the log of firm size as our dependent variable. Since we do not know the ownership shares when there is more than one founding owner, we use the averages of the individual owner characteristics as our explanatory variables. For example, if there are two individual founders and only one is a supervisor, the value of the supervisor indicator is one-half. All regressions were repeated using new firms with only one individual founder as in Table 3.3, but there were no qualitative differences in the results (these regressions are available on request).

Table 3.4 reports means (and medians for non-categorical variables) of the regressors used to explain new firm size.[15] A comparison with Table 3.1 shows that the mean wages, tenure, age, and post-high school education are higher for founding owners than for all workers who left the origin firms. Means for cause of separation are not directly comparable to Table 3.1 because founding owners who have no cause of separation are included – together the indicators for causes of separation, retains previous job, and no cause of separation but job disappears are mutually exclusive and exhaustive.[16] The mean of the firm share of founding owners who started firms in the same year or year subsequent to the calendar year in which they left their last jobs is 66.9 percent.[17] A somewhat surprising fact is that the mean of the firm share of founding owners who started firms in the same industries as their previous jobs is only 26.6 percent, even though same industry is defined at the two-digit level, a fairly high level of aggregation.[18] This fact is

[15] After we add "meso" region fixed effects to the firm size regressions we lose firms founded in 2014 because region identifiers are missing for that year. Since these fixed effects are included in our preferred specification we show the summary statistics for the remaining 517,628 firms instead of the full sample of 537,741 firms.

[16] The most plausible explanation for the last category is that these founding owners left their jobs and their causes of separation were incorrectly coded as 0. Another explanation is that the former employers of these founding owners exited. We find this was the case for 19.8 percent of these founding owners compared to 4.2 percent of the employers of founding owners with causes of separation.

[17] A founder who retains his last job is coded as lag = 0. A founder with no cause of separation but his job disappears is assigned a lag equal to the difference between the calendar year in which he started his firm and the calendar year in which his job was last observed, as is also the case for founders with causes of separation.

[18] We identify firm industry by the mode industry assigned to its employees in its founding year. We use the CNAE (Classificação Nacional de Atividade Econômica) industry classification. The indicator for same industry takes the value one when the two-digit industry assigned to the founding owner at his last job equals the mode two-digit industry of the employees of his new firm.

consistent with a number of findings that suggest that workers move relatively easily between industries (Kambourov, Manovskii and Manovskii1, 2009; Traiberman, 2019).

Table 3.5 reports our regressions for log new firm size, where firm size is measured by number of employees in December of the founding year. As in Table 3.2, the first column includes no covariates other than indicators for the three levels of managerial occupations. The coefficients on these indicators imply that firms founded by former directors, managers, and supervisors are respectively 83, 32, and 19 percent larger than other new firms. These are somewhat smaller than the percentage differences in means in Table 3.3, not because of differences with the coefficients estimated using new firms with only one individual founder but because the skewness of the distribution of new firm sizes causes the differences between the means of log sizes to be smaller than between the logs of mean sizes.

In the second column of Table 3.5 we add fixed effects for year of firm founding, 59 two-digit industries, and 137 "meso" regions in which new firms are located. We identify firm locality by the mode locality assigned to its employees in its founding year. The coefficients on the director and supervisor indicators shrink somewhat. In the third column of Table 3.5 we add indicators for two, three, and four or more individual founders, for at least one firm founder, and for at least one origin firm founder. As expected from Table 3.3, having at least one firm founder and especially having at least one origin firm founder are associated with much larger new firm sizes. Having four or more individual founders is associated with a 20 percent larger new firm. Having two (three) individual founders is associated with a slightly smaller (larger) new firm size. (The association of two founders with a slightly smaller new firm is a puzzle and does not reappear when we restrict our sample to manufacturing firms in Section 3.5.) Including these five indicators leads to a substantial reduction in the coefficient on the director indicator, mainly because former directors are much more likely to found new firms in partnership with existing firms.

Indicators for one-digit occupation categories are added in the fourth column of Table 3.5. The coefficients on the director and manager indicators fall substantially, both by just over

40 percent of the coefficient on the indicator for occupation group 1. This is similar to what we saw in Table 3.2 and again suggests that some of the impact of directors and managers can be ascribed to generic "leadership." In the fifth column of Table 3.5 we add the same measures of human capital as in Table 3.2: log wage in last job, indicators for the eleven educational attainment categories reported in RAIS, log age, log months of tenure in last job, and gender. Wage and tenure are associated with larger new firm size, age and female with smaller firm size. The education coefficients (not reported) show a weak tendency for higher education to be associated with smaller firm size. The coefficients on the managerial occupation indicators fall slightly.

In the last column of Table 3.5 we add other variables pertaining to firm founders. The most important is an indicator for whether the founder's last job was in the same two-digit industry as the firm he starts. New firms in the same two-digit industries as their founders' last jobs are 19 percent larger, all else equal. This result is consistent with Azoulay et al. (2020), who find that previous work experience of founding owners in the same two-digit industry as the firms they start is associated with various measures of firm success. We also add a complete set of indicators for cause of separation from last job and remains employed in last job, with last job disappears but no cause of separation as the omitted category. We report the coefficient on the indicator for retains last job, which is of special interest and is economically and statistically significant. New firms whose founding owners retain their last jobs are eight percent smaller, all else equal. This negative effect could occur because the new owners are not devoting their full attention to these firms or because they do not expect these firms to generate enough income for them to leave their jobs. (Only 1.0 percent of new firms employ their founding owners, so the negative effect is not due to non-employment of founding owners at new firms whose founding owners retain their last jobs.[19]) We also include indicators for the number of years between the calendar year in which the founder was last observed in his previous job and the calendar year in

[19]If we add to the last column of Table 3.5 an indicator for at least one founding owner is an employee, all regression coefficients are virtually unchanged and the coefficient on the indicator is slightly smaller than implied by increasing mean new firm size by one employee.

Table 3.4. Summary statistics for variables used in new firm size regressions

Number of workers	4.470
Median	2.000
Deflated wage (reals)	821.102
Median	403.930
Tenure (months)	52.416
Median	28.900
Age (years)	36.427
Median	35.000
Director	0.018
Manager	0.124
Supervisor	0.047
Two founders	0.259
Three founders	0.023
Four+ founders	0.005
Firm founder	0.010
Origin firm founder	0.002
Female founder	0.357
Less than high school	0.235
High school graduate	0.361
Some college	0.088
College graduate	0.311
Graduate degree	0.006
Fired	0.507
Quit	0.099
Other cause of separation	0.015
Holds previous job	0.290
Lag (years)	1.477
Lag share = 0 or 1	0.669
Prop same industry	0.266
Multiple founder same pis	0.001
occ group 1	0.164
occ group 2	0.178
occ group 3	0.122
occ group 4	0.216
occ group 5	0.155
occ group 6	0.007
occ group 7	0.115
occ group 8	0.017
occ group 9	0.025
Observations	517628

Each observation of a founding individual owner characteristic is the mean over the characteristics of the founding individual owners when there is more than one. The table reports the means over all observations, except where the medians are indicated. Real wages are provided in 1994 levels using the inflation rate from December of each founding owner's last job. Less than high school includes founding owners who are illiterate, primary school dropouts or graduates, middle school dropouts or graduates, or high school dropouts.

Table 3.5. Determinants of new firm size

	(1)	(2)	(3)	(4)	(5)	(6)
Director	0.603***	0.556***	0.440***	0.395***	0.383***	0.335***
	(0.00997)	(0.00986)	(0.0101)	(0.0134)	(0.0134)	(0.0134)
Manager	0.275***	0.280***	0.271***	0.223***	0.220***	0.162***
	(0.00411)	(0.00404)	(0.00404)	(0.00976)	(0.00979)	(0.00985)
Supervisor	0.175***	0.155***	0.155***	0.149***	0.140***	0.132***
	(0.00648)	(0.00635)	(0.00633)	(0.00643)	(0.00654)	(0.00652)
Two founders			−0.0234***	−0.0228***	−0.0222***	−0.0227***
			(0.00286)	(0.00285)	(0.00285)	(0.00284)
Three founders			0.0364***	0.0350***	0.0329***	0.0301***
			(0.00831)	(0.00830)	(0.00830)	(0.00828)
Four+ founders			0.185***	0.182***	0.177***	0.168***
			(0.0180)	(0.0180)	(0.0180)	(0.0179)
Firm founder			0.641***	0.632***	0.622***	0.610***
			(0.0141)	(0.0141)	(0.0141)	(0.0141)
Origin firm founder			0.306***	0.305***	0.305***	0.275***
			(0.0326)	(0.0326)	(0.0326)	(0.0324)
Log deflated wage					0.00833***	0.0206***
					(0.00177)	(0.00180)
Log tenure					0.00527***	0.00345*
					(0.00130)	(0.00135)
Log age					−0.0243***	−0.00470
					(0.00557)	(0.00564)
Female founder					−0.0345***	−0.0257***
					(0.00302)	(0.00301)
Prop same industry						0.171***
						(0.00313)
Holds previous job						−0.0878***
						(0.00541)
Multifounder same pis						0.0665
						(0.0538)
occ group 1				0.108***	0.112***	0.129***
				(0.0136)	(0.0138)	(0.0138)
occ group 2				0.0922***	0.0958***	0.0925***
				(0.0107)	(0.0112)	(0.0112)
occ group 3				0.0362***	0.0436***	0.0492***
				(0.0109)	(0.0110)	(0.0109)
occ group 4				0.115***	0.129***	0.131***
				(0.0105)	(0.0107)	(0.0106)
occ group 5				0.0135	0.0263*	0.00537
				(0.0107)	(0.0107)	(0.0107)
occ group 6				0.0718***	0.0628***	0.0906***
				(0.0186)	(0.0188)	(0.0187)
occ group 7				0.0219*	0.0234*	0.0181
				(0.0110)	(0.0110)	(0.0109)
occ group 9				−0.0544***	−0.0555***	−0.0686***
				(0.0132)	(0.0132)	(0.0132)
Constant	0.772***	1.233***	1.170***	1.078***	1.236***	1.223***
	(0.00142)	(0.288)	(0.287)	(0.287)	(0.291)	(0.290)
y	0.825	0.833	0.833	0.833	0.833	0.833
r2	0.0149	0.0979	0.103	0.105	0.106	0.113
N	537741	517628	517628	517628	517628	517628

Standard errors in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Dependent variable is log of December employment in the firm's founding year. Each observation of a founding individual owner characteristic is the mean over the characteristics of the founding individual owners when there is more than one. Columns (2)-(6) include fixed effects for the firm start year, two-digit industry, and meso region. Columns (5)-(6) include indicators for 10 education categories (illiterate is omitted category). Column (6) includes lag indicators for 1 through 11 years (0 lag is omitted category) and indicators for 29 causes of separation, with separation = 0 and job disappears the omitted category. Real wages are provided in 1994 levels, using the inflation rate from December of each worker's last job. Tenure is provided in months and age in years. There is no indicator for occupation group 0 because no firms in our sample have a founding owner in that group.

which he founded his firm, with zero years as the omitted category. The coefficients on all of these indicators (not shown) are negative, with the greatest absolute values in years two through four, each of these three lags being associated with an 11 percent smaller new firm. Finally, we include an indicator for when two or more founding owners have the same identification number in RAIS. Table 3.4 shows that this rarely happens, and the estimated coefficient is statistically insignificant.[20]

Column (6) is our preferred specification. Inclusion of the additional variables causes the coefficients on the director and manager occupation indicators to shrink substantially. The coefficients on the three managerial occupation levels now imply that firms founded by former directors, managers, and supervisors are respectively 40, 18, and 14 percent larger than other new firms.

As we did at the end of the previous subsection, we can also quantify the impacts of having held a managerial occupation by comparing them to the impacts of standard human capital variables. We will compute the increase in firm size attributable to these variables by first computing the means of log age, education, female, log tenure, and log wage for the founders of the new firms in the Table 3.5 sample who are former managers or former holders of non-managerial occupations, and then multiplying the differences between these means by the corresponding coefficients in the last column of Table 3.5. This computation yields a value of 0.014, less than one tenth of the coefficient on the manager indicator in the last column of Table 3.5. We interpret this result to show that the greater managerial talent of former managers has more than ten times the impact on new firm size than does their greater generic ability relative to former holders of non-managerial occupations. This is a much larger relative impact than we found for the probability of entrepreneurship. The reason is not that the human capital coefficients are relatively smaller but that the differences in means between former managers and former holders of non-managerial occupations are smaller. This occurs because we are now

[20]This appears to occur when one or more founding owners has no employment record in RAIS but is related to another founding owner, and probably indicates a husband-and-wife founding partnership. However, it seems unlikely that this indicator captures a majority or even a large minority of husband-and-wife founding partnerships.

comparing former managers who became founding owners to former holders of non-managerial occupations who also became founding owners, instead of comparing all former managers to all former holders of non-managerial occupations.

Our results support the view that generic ability is more highly rewarded in entrepreneurship than in wage employment, so that individuals with greater generic ability select into entrepreneurship and found larger firms (Murphy, Shleifer and Vishny, 1991). At the same time our results indicate that managerial talent is distinct from generic human capital, as in Lucas (1978), and is in fact a more powerful factor in leading individuals to become entrepreneurs and in making their firms more productive, as reflected in their sizes. Former managers have more human capital by every conventional measure, but the effects of this human capital on their propensities to become entrepreneurs and the sizes of their new firms are small relative to the effects of managerial talent as proxied by their former occupation.

3.4 Demand- and Supply-Side Occupations, Entrepreneurship, and Firm Size

We now consider whether the skills associated with some non-managerial occupations have impacts on entrepreneurship and new firm size comparable to those of managerial talent. We employ the same strategy of estimating the effects of former occupations controlling for measures of generic human capital, especially previous earnings. As explained in our Introduction, we focus on non-managerial occupations connected to firm productivity and developing a customer base. For convenience we will refer to these as supply- and demand-side occupations, respectively.

On the supply side we consider four types of occupations: engineers (*engenheiro*), technologists (*tecnólogo*), technicians (*técnico*), and purchasing (occupation titles with the root *compra*). Engineers are drawn entirely from major occupation group 2 (professionals) and technologists are drawn almost entirely from that group, whereas technicians are drawn

almost entirely from major occupation group 3 (mid-level technicians) and purchasers are drawn entirely from that group. On the demand side we considered occupation titles with the root *vend*. These are about evenly divided between major occupation group 3 and major occupation group 5, "service workers, salespeople in stores and markets." In the latter group are several occupations that would seem to employ skills with little potential for founding successful limited liability firms, such as street vendors. We therefore created three subsets: wholesalers (*vendedors em comercio atacadista*), retailers (*vendedors de comercio varejista*), and sales, where the latter are drawn entirely from major occupation group 3 and include "autonomous commercial representatives," which does not have the root *vend*.

Table 3.6 adds indicators for the occupation types we have created to our preferred specifications of the regressions for entrepreneurship and firm size (columns (6) of Tables 3.2 and 3.5). Purchasing immediately stands out for coefficients that are comparable to the coefficient for supervisors in the entrepreneurship regression and the coefficient for managers in the firm size regression. No other non-managerial occupation type has coefficients in both regressions that are as large as those for supervisors.

A distinguishing characteristic of purchasing is that it is necessarily a business-to-business (B2B) activity. Choi et al. (2021, p. 20) write, "a greater share of the organizational capital is likely embedded in founding teams of B2B businesses due to the importance of managing relationships." A possible explanation of the benefits of a purchasing occupation for entrepreneurship and new firm size, therefore, is that the activity entails building relationships with suppliers that are useful in the new business.

The indicator for major occupation group 1 ("Senior members of the government, leaders of organizations of public interest and companies and managers") has coefficients in Table 3.6 that are very close to those for the indicator for supervisor. A natural interpretation is that former holders of these occupations have leadership ability that increases their effectiveness as entrepreneurs independently from the impact of talent in managing a for-profit enterprise. The coefficients in Table 3.6 on the indicator for major occupation group 2 are just below those on

Table 3.6. Additional occupational determinants of entrepreneurship and new firm size

	(1) Founding owner	(2) Log firm size
Director	0.0561***	0.337***
	(0.000400)	(0.0134)
Manager	0.0367***	0.161***
	(0.000223)	(0.00985)
Supervisor	0.0147***	0.125***
	(0.000100)	(0.00660)
Engineer	0.0112***	0.0580***
	(0.000211)	(0.0104)
Technologist	−0.00213	−0.169
	(0.00131)	(0.0864)
Technician	−0.0000588	−0.0108
	(0.000120)	(0.00811)
Purchasing	0.0167***	0.193***
	(0.000428)	(0.0225)
Sales	0.000823**	0.0451**
	(0.000252)	(0.0170)
Wholesale	0.00739***	−0.00178
	(0.000339)	(0.0192)
Retail	0.0586***	−0.0508***
	(0.000199)	(0.0134)
Vend	0.00857***	−0.0141
	(0.000174)	(0.0125)
Log deflated wage	0.00218***	0.0199***
	(0.0000265)	(0.00181)
Log tenure	0.00261***	0.00370**
	(0.0000172)	(0.00135)
Log age	0.0142***	−0.00627
	(0.0000555)	(0.00564)
Female founder	−0.00321***	−0.0233***
	(0.0000340)	(0.00303)
occ group 0	0.00909***	
	(0.00165)	
occ group 1	0.0128***	0.128***
	(0.000223)	(0.0138)
occ group 2	0.00956***	0.0849***
	(0.000127)	(0.0112)
occ group 3	0.00600***	0.0460***
	(0.000145)	(0.0122)
occ group 4	0.000804***	0.129***
	(0.000108)	(0.0106)

Table 3.6 (Cont.): Additional occupational determinants of entrepreneurship and new firm size

occ group 5	0.00178***	0.0335**
	(0.000113)	(0.0112)
occ group 6	−0.00104***	0.0916***
	(0.000138)	(0.0187)
occ group 7	−0.00193***	0.0186
	(0.000107)	(0.0109)
occ group 9	0.00177***	−0.0692***
	(0.000139)	(0.0132)
Two founders		−0.0227***
		(0.00284)
Three founders		0.0304***
		(0.00827)
Four + founders		0.169***
		(0.0179)
Brazil firm founder		0.608***
		(0.0141)
Origin firm is founder		0.277***
		(0.0324)
Prop same industry		0.173***
		(0.00314)
Holds previous job		−0.0881***
		(0.00541)
Multiple founder same pis		0.0669
		(0.0538)
Constant	−0.0707***	1.217***
	(0.000284)	(0.290)
ymean	0.0110	0.833
r2	0.162	0.114
N	51636208	517628

Standard errors in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$
Column (1) includes all controls in Column (6) of Table 2. Column (2) includes all controls in Column (6) of Table 5.

the indicator for supervisor, yielding an impact on the probability of becoming an entrepreneur just below the mean probability and a nine percent increase in new firm size. This result for professionals is consistent with the findings of (Levine and Rubinstein, 2017) that "nonroutine cognitive abilities" predict selection of workers into incorporated self-employment.

3.5 Robustness and Extensions

Bloom and Van Reenen (2007, p. 1371) report finding "a strong and positive correlation between firm size and management practices" using a sample of 732 medium-sized manufacturing firms in the United States, the United Kingdom, France, and Germany. Their paper is typical of much of the literature in its focus on manufacturing firms, so our results for new firms in Brazil will relate better to this literature if we restrict new firms to manufacturing. We would also like to know whether our results are robust to this restriction. In Table 3.7 we repeat the regressions in Table 3.6, letting the dependent variable in the entrepreneurship regression equal one only if the former worker founds a manufacturing firm (CNAE with first two digits 10-33) and restricting the sample for the new firm size regression to manufacturing firms. We add to the entrepreneurship regression an indicator for whether the industry assigned to the worker at his last job was in the manufacturing sector. There is little qualitative change in our results for the impacts of former managerial and former non-managerial occupations. In the entrepreneurship regression the coefficients on the former managerial occupations continue to be ranked director, manager, supervisor, and all of them are larger relative to the mean of the dependent variable than they were for the full sample. In the firm size regression the coefficient on director is substantially larger and the coefficients on manager and supervisor essentially exchange places relative to the full sample regression. Purchasing is still the only former non-managerial occupation with larger coefficients in both regressions than the smallest coefficients for former managerial occupations.

In all our regressions we have entered our measures of generic human capital linearly, but the literature does not necessarily predict that propensity to become an entrepreneur in particular increases monotonically with worker ability. Poschke (2013), for example, predicts a U-shaped relationship between entrepreneurship and ability. We experimented with consecutively adding to both regressions in Table 3.6 the second, third, and fourth powers of the previous log wage, probably our best proxy for generic worker ability. These experiments left the coefficients on the former occupation indicators virtually unchanged. Notably, in the quadratic

Table 3.7. Occupational determinants of entrepreneurship and new firm size in manufacturing

	(1) Founding owner	(2) Log firm size
Director	0.0155***	0.509***
	(0.000178)	(0.0496)
Manager	0.01000***	0.124***
	(0.0000995)	(0.0373)
Supervisor	0.00402***	0.166***
	(0.0000448)	(0.0198)
Engineer	0.00123***	0.0356
	(0.0000941)	(0.0395)
Technologist	−0.000643	−0.0864
	(0.000583)	(0.255)
Technician	−0.000248***	0.0429
	(0.0000534)	(0.0282)
Purchasing	0.00618***	0.285***
	(0.000191)	(0.0707)
Sales	0.000700***	0.0560
	(0.000113)	(0.0635)
Wholesale	0.00100***	0.115
	(0.000152)	(0.0709)
Retail	0.0113***	0.0544
	(0.0000890)	(0.0496)
Vend	0.00265***	−0.0160
	(0.0000778)	(0.0460)
Manufacturing	0.00767***	
	(0.0000553)	
Constant	−0.0238***	0.631
	(0.000128)	(0.730)
ymean	0.00208	1.111
r2	0.124	0.111
N	51636208	57890

Standard errors in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$
Column (1) includes all controls in Column (6) of Table 2. Column (2) includes all controls in Column (6) of Table 5.

specifications for previous log wage the coefficient on the linear term remains positive and statistically significant in both the entrepreneurship and firm size regressions, and the coefficient on the squared term is negative and statistically significant in the entrepreneurship regression and statistically insignificant in the firm size regression.

We have interpreted the economically large impacts of former managerial occupations on propensity to become a founding owner and on new firm size as evidence for the importance of managerial "talent" or aptitude. One could argue that a better term would be "experience." Recall, however, that we control for log months of tenure in the worker's previous job. An increase in tenure from the minimum of three months to the sample mean in the entrepreneurship regression of Table 3.6 increases the probability of becoming a founding owner by 0.6 percent compared to the mean probability of 1.1 percent, and an increase in tenure from the minimum to the sample mean in the log firm size regression of Table 3.6 increases firm size by about one percent. There are many reasons why tenure in any previous job would, conditional on leaving that job and independent of generic ability, have a positive impact on probability of becoming an entrepreneur and on new firm size. The founders of successful firms interviewed by Bhide (2000, p. 32) "typically imitated someone else's ideas that they often encountered in the course of a previous job." Presumably the longer was a worker's tenure in his previous job, the more ideas he encountered and the better he learned them. Muendler and Rauch (2012) provide evidence that workers who become entrepreneurs recruit their co-workers to their new firms. The longer their tenure the more co-workers they will have gotten to know.

The question is whether tenure in managerial occupations is more important for entrepreneurial productivity than tenure in other occupations. We investigated this question by adding interactions between log months of tenure and the eleven indicators for former managerial, demand- and supply-side occupations to the regressions in Table 3.6. The coefficients on these interactions were positive and statistically significant for ten of the eleven occupation groups in the entrepreneurship regression, but no more quantitatively important for the managerial occupations than the other occupations. In the firm size regression the coefficients on the interactions were positive, economically and statistically significant for directors; negative, economically and statistically significant for supervisors; positive, economically small and marginally statistically significant for managers and technicians; and statistically insignificant for the other occupation groups. We conclude that the evidence does not support the view that experience makes a greater

contribution to entrepreneurial productivity for former managers than for other former workers.

Larger firms tend to have more layers of hierarchy (Caliendo, Monte and Rossi-Hansberg, 2015). Larger firms are therefore more likely to employ supervisors, managers, and especially directors. If experience at a large firm helps a worker to make the firm he founds larger, this could account for some of the association of former managerial occupations with new firm size. To control for this possibility we added the log of origin firm size to the right-hand side of the firm size regression in Table 3.6. Its coefficient was statistically insignificant and the coefficients on the occupation indicators were virtually unchanged.

3.6 Concluding Remarks

We have investigated the impacts of managerial talent and occupational skills of entrepreneurs on new job creation. Our focus on initial firm size is supported by Sterk, Sedláček and Pugsley (2021), who find that ex ante heterogeneity is the main determinant of subsequent firm differences. Nevertheless, in future research it would be straightforward to extend our investigation to the impacts of managerial talent and occupational skills of entrepreneurs on firm survival or employment growth.

Our investigation leaves open the question of the extent to which managerial talent can be quantified by the good management practices identified by Bloom and Van Reenen (2007). This will inform the degree to which managerial "talent" can be taught, and thus the emphasis that policy should place on such training.

Chapter 3 is co-authored with Muendler, Marc-Andreas and Rauch, James E. and is currently being prepared for submission for publication. The dissertation author is a principal researcher on this chapter.

Bibliography

Abowd, John M., Francis Kramarz, David N. Margolis, and Kenneth R. Troske. 2001. "The Relative Importance of Employer and Employee Effects on Compensation: A Comparison of France and the United States." *Journal of the Japanese and International Economies*, 15(4): 419–436.

Acemoglu, Daron, and Robert Shimer. 2000. "Productivity gains from unemployment insurance." *European economic review*, 44(7): 1195–1224.

Akerlof, George A, Andrew K Rose, Janet L Yellen, Laurence Ball, and Robert E Hall. 1988. "Job switching and job satisfaction in the US labor market." *Brookings papers on economic activity*, 1988(2): 495–594.

Amarante, Patrícia Araújo, Magno Vamberto Batista da Silva, and Paulo Aguiar do Monte. 2019. "Does the spatial density of employment stimulate inter-firm worker mobility? An analysis of Brazilian municipalities." *Annals of Regional Science*.

Amiti, Mary, and Jozef Konings. 2007. "Trade Liberalization, Intermediate Inputs and Productivity: Evidence from Indonesia." *American Economic Review*, 97(5): 1611–1638.

Andersson, Martin, and Per Thulin. 2013. "Does spatial employment density spur inter-firm job switching?" *Annals of Regional Science*, 51(1): 245–272.

Antel, John J. 1986. "Human capital investment specialization and the wage effects of voluntary labor mobility." *The Review of Economics and Statistics*, 477–483.

Antel, John J. 1991. "The wage effects of voluntary labor mobility with and without intervening unemployment." *ILR Review*, 44(2): 299–306.

Azoulay, Pierre, Benjamin F Jones, J Daniel Kim, and Javier Miranda. 2020. "Age and High-Growth Entrepreneurship." *American Economic Review: Insights*, 2(1): 65–82.

Bhide, Amar. 2000. *The Origin and Evolution of New Businesses.* Oxford and New York:Oxford University Press.

Bleakley, Hoyt, and Jeffrey Lin. 2012. "Thick-market effects and churning in the labor market: Evidence from US cities." *Journal of Urban Economics*, 72(2-3): 87–103.

Bloom, Nicholas, and John Van Reenen. 2007. "Measuring and Explaining Management Practices Across Firms and Countries." *Quarterly Journal of Economics*, 122(4): 1351–1408.

Bloom, Nicholas, Benn Eifert, Aprajit Mahajan, David McKenzie, and John Roberts. 2013. "Does Management Matter? Evidence from India." *Quarterly Journal of Economics*, 128(1): 1–51.

Caliendo, Lorenzo, Ferdinando Monte, and Esteban Rossi-Hansberg. 2015. "The Anatomy of French Production Hierarchies." *Journal of Political Economy*, 123(4): 809–852.

Choi, Joonkyu, Nathan Goldschlag, John C Haltiwanger, and J. Daniel Kim. 2021. "Founding Teams and Startup Performance." National Bureau of Economic Research Working Paper 28417.

Dauth, Wolfgang, Sebastian Findeisen, Enrico Moretti, and Jens Südekum. 2019. "Matching In Cities." *National Bureau of Economic Research*.

De Blasio, Guido, and Sabrina Di Addario. 2005. "Do workers benefit from industrial agglomeration?" *Journal of Regional Science*, 45(4): 797–827.

De Negri, João Alberto, Paulo Furtado, Natália Ribeiro de Souza, and Jorge Saba Arbache. 1998. "Mercado Formal de Trabalho: Comparação entre os microdados da RAIS e da PNAD." *IPEA Texto para Discussão*, 840. Instituto de Pesquisa Econômica Aplicada, Rio de Janeiro.

Ehrenberg, Ronald G, and Ronald L Oaxaca. 1976. "Unemployment insurance, duration of unemployment, and subsequent wage gain." *The American Economic Review*, 66(5): 754–766.

Ellison, Glenn, Edward L. Glaeser, and William R. Kerr. 2010. "What causes industry agglomeration? Evidence from coagglomeration patterns." *American Economic Review*, 100(3): 1195–1213.

Fairlie, Robert W, and Alicia Robb. 2007. "Families, Human Capital, and Small Business: Evidence from the Characteristics of Business Owners Survey." *ILR Review*, 60(2): 225–245.

Fallick, Bruce, Charles A Fleischman, and James B Rebitzer. 2006. "Job-Hopping in Silicon Valley: Some Evidence Concerning the Microfoundations of a High-Technology Cluster." *Review of Economics and Statistics*, 88(3): 472–481.

Foster, Lucia, John Haltiwanger, and Chad Syverson. 2008. "Reallocation, Firm Turnover, and Efficiency: Selection on Productivity or Profitability?" *American Economic Review*,

98(1): 394–425.

Foster, Lucia, John Haltiwanger, and Chad Syverson. 2016. "The Slow Growth of New Plants: Learning about Demand?" *Economica*, 83(329): 91–129.

Frederiksen, Anders, Timothy Halliday, and Alexander K Koch. 2016. "Within-and cross-firm mobility and earnings growth." *ILR Review*, 69(2): 320–353.

Gourio, Francois, and Leena Rudanko. 2014. "Customer Capital." *Review of Economic Studies*, 81(3): 1102–36.

Hall, Robert E, and Paul R Milgrom. 2008. "The limited influence of unemployment on the wage bargain." *American economic review*, 98(4): 1653–74.

Halpern, Laszlo, Miklos Koren, and Adam Szeidl. 2015. "Imported Inputs and Productivity." *American Economic Review*, 105(12): 3660–3703.

Haltiwanger, John C, Henry R Hyatt, Lisa B Kahn, and Erika McEntarfer. 2018. "Cyclical job ladders by firm size and firm wage." *American Economic Journal: Macroeconomics*, 10(2): 52–85.

Haltiwanger, John, Henry Hyatt, and Erika McEntarfer. 2018. "Who moves up the job ladder?" *Journal of Labor Economics*, 36(S1): S301–S336.

Harmon, Nikolaj A. 2013. "Are Workers Better Matched in Large Labor Markets?" *Job market paper, Princeton University*.

Helpman, Elhanan, and Paul R Krugman. 1985. *Market structure and foreign trade: Increasing returns, imperfect competition, and the international economy.* MIT press.

Hirakawa, Oana, Marc-Andreas Muendler, and James E Rauch. 2010. "Employee Spinoffs and Other Entrants: Stylized Facts from Brazil." National Bureau of Economic Research Working Paper 15638.

Hotelling, Harold. 1929. "Stability in Competition Author (s): Harold Hotelling Published by : Wiley on behalf of the Royal Economic Society Stable URL : http://www.jstor.org/stable/2224214." *The Economic Journal*, 39(153): 41–57.

Jovanovic, Boyan. 1982. "Selection and the Evolution of Industry." *Econometrica*, 50(3): 649–670.

Kambourov, Gueorgui, Iourii Manovskii, and Iourii Manovskii1. 2009. "OCCUPATIONAL SPECIFICITY OF HUMAN CAPITAL*." *Source: International Economic Review*, 505468(1): 63–115.

Levine, Ross, and Yona Rubinstein. 2017. "Smart and Illicit: Who Becomes an Entrepreneur and Do They Earn More?" *Quarterly Journal of Economics*, 132(2): 963–1018.

Lucas, Robert E. Jr. 1978. "On the Size Distribution of Business Firms." *Bell Journal of Economics*, 9(2): 508–23.

Menezes-Filho, Naércio Aquino, Marc-Andreas Muendler, and Garey Ramey. 2008. "The Structure of Worker Compensation in Brazil, with a Comparison to France and the United States." *Review of Economics and Statistics*, 90(2): 324–346.

Moretti, Enrico. 2011. *Local labor markets.* Vol. 4, Elsevier B.V.

Mortensen, Dale, and Christopher Pissarides. 1992. "The cyclical behavior of job creation and job destruction." Discussion Paper.

Mortensen, Dale T. 1994. "The cyclical behavior of job and worker flows." *Journal of Economic dynamics and control*, 18(6): 1121–1142.

Mortensen, Dale T, and Christopher A Pissarides. 1994. "Job creation and job destruction in the theory of unemployment." *The Review of Economic Studies*, 61(3): 397–415.

Muendler, Marc-Andreas, and James E Rauch. 2012. "Mobilizing Social Capital Through Employee Spinoffs." National Bureau of Economic Research Working Paper 18459.

Murphy, Kevin M, Andrei Shleifer, and Robert W Vishny. 1991. "The Allocation of Talent: Implications for Growth." *Quarterly Journal of Economics*, 106(2): 503–530.

Poschke, Markus. 2013. "Who Becomes an Entrepreneur? Labor Market Prospects and Occupational Choice." *Journal of Economic Dynamics and Control*, 37(3): 693–710.

Rosenthal, Stuart S., and William C. Strange. 2004. *Chapter 49 Evidence on the nature and sources of agglomeration economies.* Vol. 4, Elsevier Inc.

Salop, Steven C. 1979. "Monopolistic Competition with Outside Goods." *The Bell Journal of Economics*, 10(1): 141.

Sterk, Vincent, Petr Sedláček, and Benjamin Pugsley. 2021. "The Nature of Firm Growth." *American Economic Review*, 111(2): 547–79.

Sullivan, Paul. 2010. "Empirical evidence on occupation and industry specific human capital." *Labour Economics*.

Topel, Robert H, and Michael P Ward. 1992. "Job mobility and the careers of young men." *The Quarterly Journal of Economics*, 107(2): 439–479.

Traiberman, Sharon. 2019. "Occupations and import competition: Evidence from Denmark." *American Economic Review*, 109(12): 4260–4301.

Van der Klaauw, Bas, and António Dias Da Silva. 2011. "Wage dynamics and promotions inside and between firms." *Journal of Population Economics*, 24(4): 1513–1548.

Wheeler, Christopher H. 2008. "Local market scale and the pattern of job changes among young men." *Regional Science and Urban Economics*, 38(2): 101–118.

Chapter 1 Appendix

A.0.1 Piecewise Wage Function

We consider a function c that increases less than linearly with distance and specify a piecewise cost function

$$c(d) = \begin{cases} d & d \leq \bar{d} \\ \bar{d} + \beta(d - \bar{d}) & d > \bar{d} \end{cases} \quad \text{where } \bar{d} \in (0, \frac{1}{2n}),$$

where $\bar{d} \in (0, \frac{1}{2n})$, $\beta > 0$, and keeping all else the same as before.[21] This allows us to get the results that worker circulation is decreasing in n with a small enough β. Since all else is the same, we still have that the probability that any worker in the region switches firms equals $s(1 - \frac{x}{1/2n})$. We solve for x under the assumption that $x \leq \bar{d}$. This assumption will be possible because we can make x sufficiently small through shrinking β.[22] The wage conditions for the indifferent worker give us the equation: $a - x + \varepsilon = a - \bar{d} - \beta(\frac{1}{n} - x - \bar{d})$. This yields the solution $x = \frac{\beta/n + (1-\beta)\bar{d} + \varepsilon}{1+\beta}$. Plugging this value in, we have that the probability that any worker in the region switches firms is $s(1 - \frac{\beta/n + (1-\beta)\bar{d} + \varepsilon}{(1+\beta)/2n}) = s[1 - \frac{2}{1+\beta}[\beta + ((1-\beta)\bar{d} + \varepsilon)n]]$. The derivative of this is negative in n if $(1-\beta)\bar{d} > -\varepsilon$, so with small enough β, this probability is decreasing in n.

Graphically, we can compare the results for both the case presented in the main paper where $\beta = 1$ and when we shrink β. Figure A.2 illustrates the case with $\beta = 1$. In the first graph, the left y-axis represents firm 1 and the wage function firm 1 offers, while the right y-axis

[21] Note that setting $\beta = 1$ gives us the example shown main paper.

[22] This can be seen algebraically, noting that $\varepsilon < 0$, but it is easier to see graphically, which we demonstrate at the end.

represents firm 2 and the wage function from firm 2. Firm 1 initially offers a wage function that intercepts the y-axis at a. After the shock, firm 1 now offers a lower wage function with an intercept of $a+\varepsilon$. The location on the x-axis of the intercept of the wage functions from firm 1 and firm 2 give us the location of the marginal worker who is indifferent between working at either firm. Without a shock, this location is equidistant between the two axes, or the two firms. This location shifts to the left after the shock that lowers firm 1's wage function, and the distance of the shift represents the amount of workers who leave firm 1.

The second graph in Figure A.2 shows what happens when we increase the number of firms. The right y-axis has shifted to the left to reflect the fact that the distance between the two firms is smaller, and the wage function for firm 2 shifts left as well. We see that the amount of workers who leave firm 1 does not decrease, although firm 1 now employs fewer workers to start with. This trend continues in the third graph. By now, it is clear that the amount of workers who leave firm 1 do not decrease with every iteration as the number of firms increases and the distance between firm 1 and 2 decrease. Since firm 1 starts off with fewer and fewer workers as n increases, this means that the percentage of firm 1's workers that stay after the shock is decreasing in n.

Figure A.3 illustrates the case with a very small β. The setup is very similar to that from figure A.2. The difference now is that the amount of workers who leave firm 1 after the negative shock shrinks as the distance between the two firms decreases. The amount of workers who stay with firm 1 after the negative shock changes very little as the distance decreases. If you assume that this decrease in the amount of stayers is negligible, it is clear that as n increases and firm 1's initial amount of workers decreases, the percentage of stayers as n increases is increasing. In other words, the denominator shrinks faster than the numerator.

A.0.2 Appendix Figures and Tables

Table A.8. Educational Levels

Level	Educational Attainment
1	Illiterate
2	Some primary
3	Completed primary
4	Some middle school
5	Completed middle school
6	Some high school
7	Completed high school
8	Some college
9	Completed college
10	Masters Degree
11	Ph.D.

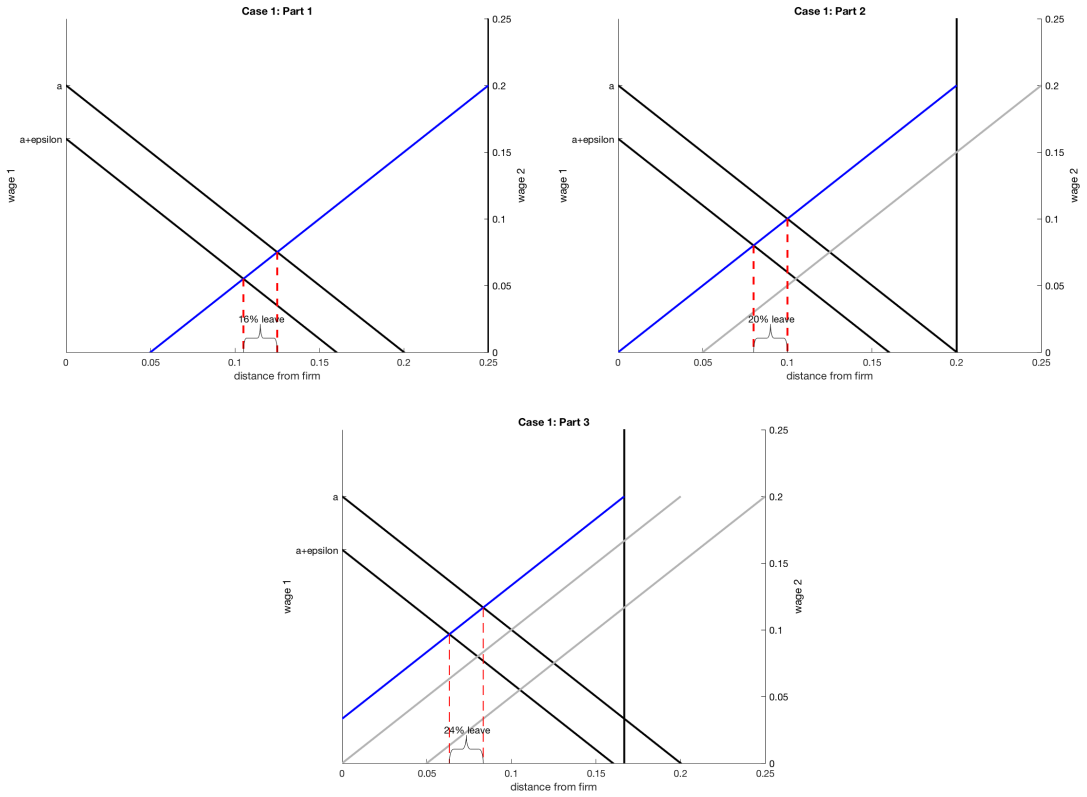


Figure A.2. Case 1: Worker mobility when $\beta = 1$.
The sequence of graphs illustrate worker mobility as the number of firms increases. The two vertical bars in each graph represent two neighboring firms. There are 4 firms in the first graph, so the distance between each firm is 0.25. The upper black line gives the wage of a worker at firm 1, decreasing with distance from the firm. The blue line similarly gives the wage of a worker at firm 2. The right dotted red line where the two wage lines intersect marks the location of the indifferent worker. The lower black line gives the new wage when firm 1 undergoes a negative shock. The left dotted red line marks the new location of the indifferent worker. The distance between the dotted red lines marks the percentage of firm 1's workers who switch to firm 2 after realization of the shock. In the second graph, the number of firms increases to 5, so the distance between the two firms shrinks. The blue line gives the wage corresponding to firm 2's new location, with the former blue line in the first graph in grey for comparison. We can see that the percentage of firm 1's workers who switch firms increases. This statistic further increases in the third graph with 6 firms, showing that worker mobility is increasing in number of firms.

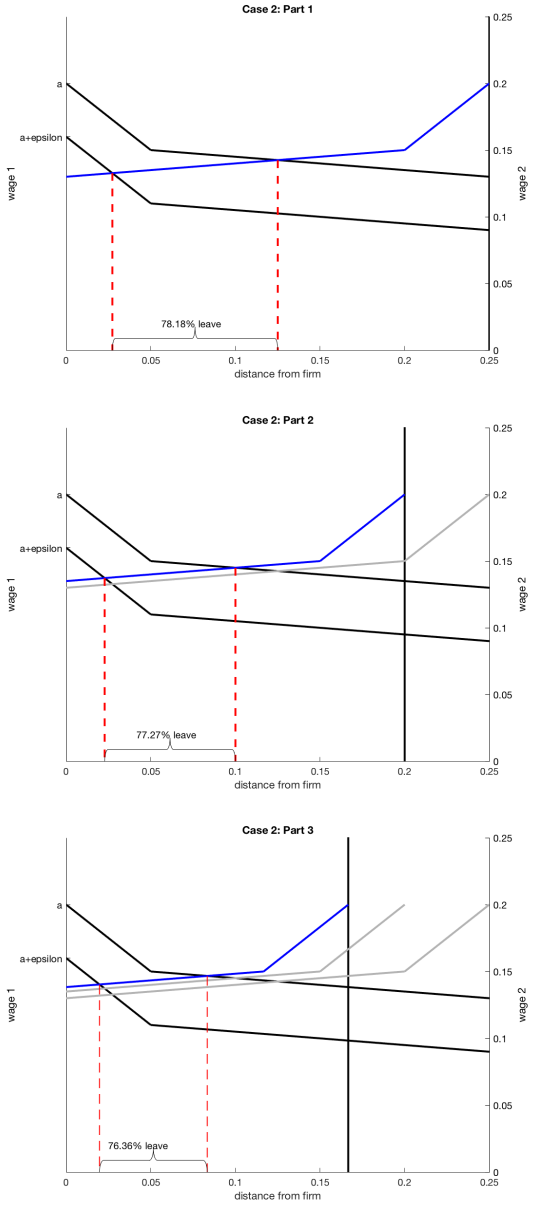


Figure A.3. Case 2: Worker mobility when $\beta < 1$.
The setup is the same as that in Figure A.2 except for the β. However we now have that increasing the number of firms leads to fewer workers switching to firm 2.