Essays on the Allocation of Talent

Ana Figueiredo

TESI DOCTORAL UPF / ANY 2018

DIRECTOR DE LA TESI

Jan Eeckhout i Isaac Baley

Departament d'Economia i Empresa

Universitat Pompeu Fabra Barcelona

Aos meus Pais.

Acknowledgements

Over the past years, my Grandmother Celeste kept saying "I want to live until you finish your PhD". Her wish was fulfilled, and I have many to thank for making it possible. First, I am thankful to Jan Eeckhout for his guidance, encouragement and support throughout these years. I am also indebted to Isaac Baley for guiding me since his arrival at UPF. The same is true for Edouard Schaal, who helped me during the Job Market with his comments on my papers and reference letter. I also owe many thanks to the faculty at UPF and CREi for their feedback on my work, and to Marta and Laura for their amazing assistance in everything. Outside UPF, I am grateful to Mireia Giné, for her encouragement along the way and endless help in preparing for the Job Market.

I have many more to thank for their friendship. Cristina, Darren, Francisco, Helena, Lorenzo, Marco, María, Miguel, Niklas, Ricardo, Shree, Thomas and Tomás, thank you for all the great moments we had together. These past years would not have been the same without you: I had so much fun! While I was in Barcelona, many people back home also had their part in this journey. *À Irmandade dos Inúteis, à Marta e ao Miguel, obrigada pela amizade e por sempre me fazerem sentir como se nunca tivesse partido. À minha Tia, Avós, Irmã e Sogros, obrigada por todo o carinho sempre que voltava a casa. E, aos meus Pais, obrigada por tudo! Nenhuma palavra desta tese poderia ter sido escrita sem vocês.*

Finally, I am immensely grateful to my partner in crime, Guido. *Gostaria de poder dizer que fui para ti apenas metade daquilo que tu foste para mim durante estes anos. Obrigada pela paciência, por nunca me deixares desistir, e por sempre me arrancares um sorriso. Mais que tudo, obrigada por estares sempre ao meu lado!* "I really don't think it gets any better than this, Vanilla smile and a gorgeous strawberry kiss, Birds sing we swing, Clouds drift by and everything is like a dream, It's everything I wished. Never guessed it got this good, Wondered if it ever would, Really didn't think it could (...) Say it will always be like this, The two of us together, It will always be like this, Forever and ever and ever" (Mint Car, The Cure).

Abstract

This thesis studies the mechanisms behind talent misallocation, how it varies over the business cycle and its implications for wage cyclicality. The first chapter shows that uncertainty about education returns has an important role in the propagation of inequality across generations. I find that a theory of local learning about an uncertain skill premium explains the negative correlation between college enrollment and the share of college graduates when the skill premium is low, and that it accounts for more than half of the enrollment gap between children with low-skill parents and children with high-skill parents. The second chapter examines the dynamics of skill mismatch over the cycle. I provide new evidence that in recessions highly mismatched jobs are destroyed, but also created, and explain this pattern through the lens of a learning model where skill mismatch is unobserved. The last chapter contributes to the ongoing debate about wage cyclicality. I show that excess wage cyclicality of job switchers goes beyond skill mismatch cyclicality, and that skill mismatch amplifies wage cyclicality.

Resumen

Esta tesis doctoral investiga los mecanismos que están detrás de la mala asignación de talento, cómo varía ésta a lo largo del ciclo económico y sus implicaciones para el ciclo salarial. El primer capítulo muestra que la incertidumbre sobre el retorno de la educación tiene un papel importante en la desigualdad. Muestro que una teoría de aprendizaje social sobre el sueldo de los graduados universitarios explica la correlación negativa entre la inscripción universitaria y el porcentaje de graduados universitarios cuando la brecha salarial es baja. También muestro que información imperfecta combinada con el aprendizaje social explica más de la mitad de la brecha que existe entre la inscripción universitaria de los hijos de padres con bajo nivel educativos y hijos de padres con alto nivel educativos. El segundo capítulo examina la dinámica en el ciclo económico del desfase entre las habilidades del trabajador y las habilidades requeridas por el empleo. Presentó evidencia de que en las recesiones los trabajos con un desajuste mayor son tanto destruidos como creados. Explico este patrón a través de un modelo de aprendizaje bayesiano donde no se observa el desajuste de habilidades. El último capítulo contribuye al debate sobre la ciclicidad salarial. Muestro que el exceso de ciclicidad en el sueldo de los trabajadores que cambian de trabajo va más allá de la ciclicidad del desajuste de habilidades, y que el desajuste de habilidades amplifica la ciclicidad salarial.

Preface

In a frictionless economy, individuals choose the education level or occupation where they obtain the highest return for their talent. However, barriers to human capital investment, discrimination in the labor market as well as information and search frictions prevent this from happening. The resulting (mis)allocation of talent has significant effects on the growth rate of an economy (Murphy et al., 1991; Jovanovic, 2014; Hsieh et al., 2018). This doctoral thesis contributes to the effort of understanding the sources of talent misallocation, how it fluctuates over the cycle and its implications for wage cyclicality.

In the first chapter, I explore the role of information frictions and local information transmission in the propagation of inequality across generations. Using school district level data from Michigan, I first document that when the college premium is low, a higher share of college graduates living in a school district is associated with lower college enrollment of students graduating from that district. While this pattern is hard to reconcile through models with local spillovers in the production of human capital, I show that it is consistent with a model featuring imperfect information and local learning. The main novelty is that at the investment stage children are uncertain about the skill-premium, and learn about it in a Bayesian way by observing signals of wages earned by high-skill individuals living in the same location. The local nature of learning implies that the place where children grow up determines the pool of outcomes observed and, therefore, shapes their perceptions about the skill premium. In this environment, more exposure to highly educated neighbors brings more information about the skill premium. However, this only translates into more investment in human capital if the observed wages generate the perception of a higher skill premium. Calibrating the model to match micro data from Detroit, I find that this novel mechanism explains more than half of the dispersion of enrollment across school districts and more than half of the college enrollment gap between children of parents with a college degree and children from parents with a lower education level. Implementing a disclosure policy that corrects inaccurate perceptions about the skill premium closes this gap substantially.

The second chapter, co-authored with Isaac Baley and Robert Ulbricht, studies the dynamics of skill mismatch between workers and occupations over the business cycle. Over the business cycle, economies face a large amount of reallocation: firms enter and exit and workers change jobs. How do business cycles affect the allocation of workers to jobs? While some argue that recessions cleanse the labor market of the worst worker-firm pairs, others view recessions as times when match quality decreases. This is known as the sullying effect of recessions. Using a worker-level panel data from the 1979 National Longitudinal Survey of Youth combined with occupational level data and aggregate data on U.S. unemployment, we provide new evidence that mismatch is procyclical: in recessions, workers skills are more aligned with job skill requirements. Interestingly, we uncover important differences along the flow of job creation and the flow of job destruction. Our results suggest that during recessions highly mismatched jobs are destroyed, consistent with the cleansing effect of recessions, but also created, as suggested by the sullying hypothesis. We then revisit the cyclical behavior of job tenure. To explain the documented patterns, we build a model of learning about unobserved skill mismatch. The novel feature is that recessions are characterized by lower aggregate productivity but also by a larger fraction of matches with high uncertainty. We show that negative productivity shocks destroy (perceived) high mismatched worker-job pairs, but at the same large information frictions potentially create more worker-firm matches with undetected high levels of skill mismatch. As a source of countercyclical uncertainty, we explore the role of occupational switching and document suggestive evidence pointing towards this channel.

In the last chapter, I revisit the issue of wage cyclicality. There is an extensive literature that relies on panel data documenting that wages of new hires are more cyclical than the ones of workers in ongoing job relationships. However, this literature has not yet been able to assess whether these changes in wages over the cycle capture wage cyclicality or instead confounding variation in the wages of new hires due to workers moving to better jobs during expansions. To address this issue, I use the skill mismatch measure developed by Guvenen et al. (2018) as a control in the wage regression. In line with earlier studies, I find that wages of newly hired workers are more cyclical than those workers in ongoing relationships. Further, separating new hires who change jobs between employers and new hires coming from a jobless spell, we find that wage cyclicality of newly hired workers is driven by the former as in Gertler et al. (2016). While they interpret this evidence as capturing changes in match quality over the cycle, I show that wage dynamics of job switchers goes beyond skill mismatch fluctuations. In addition to this, I find that mismatch amplifies wage cyclicality. These results bring important insights to the ongoing debate about what wage setting protocol is consistent with the observed behavior of wages and the role of wage rigidity in search and matching models. In particular, they seem to support Pissarides (2009)'s argument that a good explanation for the unemployment volatility puzzle needs to be consistent with flexible wages.

Contents

List of Figures

List of Tables

Chapter 1

Information Frictions in Education and Inequality

Most of what we know we learn from other people. (Lucas, 1988)

1.1 Introduction

College enrollment in the U.S. exhibits stark socioeconomic differences which contribute to the persistence of inequality across generations. In 2011, the fraction of students who enrolled into college was 83% for children of college educated parents and only 54% for children of parents with a lower education level. While these differences could be explained by differences in the family context, there is now robust evidence showing that the place where children grow up plays an important role (Chetty and Hendren, 2017). Potential channels include the local funding of schools (Bénabou, 1996a,b; Férnandez and Rogerson, 1996; Durlauf, 1996), or human capital spillovers in the production of human capital (Bénabou, 1993; Cavalcanti and Giannitsarou, 2013; Bowles et al., 2014). In this chapter, I propose a novel explanation featuring imperfect information about the skill premium and information transmission at the neighborhood level. This mechanism is motivated by empirical evidence showing that *(i)* in the U.S. individuals are uncertain about the skill premium (Bleemer and Zafar, 2016)[1]; *(ii)* individual perceptions about earnings determine education decisions (Jensen, 2010; Kaufmann, 2014; Hastings et al., 2016; Belfield et al., 2016), and *(iii)* poor students are the most affected by informational barriers (Hoxby and Avery, 2014; Rauh and Boneva, 2017).[2]

The analysis proceeds in three steps. First, I uncover a new empirical fact. Using data from Michigan, I find that a higher share of college graduates living in a school district is associated with lower college enrollment by high-school students from that district, when earnings of those college graduates are sufficiently low. Next, to explain this finding, I develop a theory of local learning about an uncertain skill premium. The local nature of learning implies that the place where children grow up determines the pool of outcomes observed and, therefore, shapes their perceptions about the skill premium. The key and novel insight of the model is that in locations where college graduate earnings are low, but the share of college graduates is high, high-school students have precise information that the value of education is low, hence are less likely to enroll in college. Finally, I calibrate this model, and show that the interplay of imperfect information with local learning explains more than half of the dispersion of enrollment across school districts, and more than half of the enrollment gap between children with low and highly educated parents.

To document the main empirical finding, I use school district level data from Michigan over the period 2008-2014 and exploit variation in the share of college graduates across school districts within a city. My empirical strategy takes into account time varying shocks affecting all cities as well as city level characteristics that might be trending over time (for instance, gentrification or deterioration of housing quality). Further, I show that the observed pattern is not masking the

[1] Bleemer and Zafar (2016) used the Survey of Consumer Expectation, a representative survey of US household heads, to ask about the perceived skill premium. They find that 75% underestimates the skill premium, and that there is a wide dispersion in the perceived premium that goes beyond the fundamental dispersion in skill premium across US MSA's.

[2] Hoxby and Avery (2014) find that low-income students with high-school achievement do not apply to any selective college or university, behavior that contrasts with that of high-income students with similar achievement. Rauh and Boneva (2017) show that students with low socio-economic status perceive both the pecuniary and non-pecuniary returns to college to be lower, when compared to high socio-economic status students.

effects of better schools, credit constraints, differences in students ability or the possibility of sorting across school districts. I address these alternative channels in the following way. First, I control for cohort characteristics (share of females in the 12^{th} grade and average ACT score, a good proxy for ability). Second, I control for socio-economic characteristics of the neighborhood (racial composition, median household income, unemployment rate, and school district size). Third, I include school resources controls: expenditures, local revenues and teacher per student. Fourth, following Oster (2016)'s method, I show that the point estimates from the OLS estimation remain almost unchanged if I adjust them to account for potential selection on unobservables. Overall, the evidence that the relationship between neighborhoods' skill-mix and college enrollment is heterogeneous along the earnings dimension is robust.

Why do individuals with a college degree have negative or positive externalities in others' decision to enroll in college depending on their earnings? This finding is hard to reconcile with existing models of human capital formation with local externalities as these models predict the relationship between a neighborhood skill-mix and college enrollment to be positive and independent of earnings. Thus, in the second part of the paper, I develop a theoretical framework that formally illustrates and quantifies the role of information frictions and local information transmission in explaining the observed pattern.

In the model, parents decide where to locate within a city and children decide whether to invest in education, and become high-skill workers, or not to invest in education, and be low-skill workers. As standard in the literature, the cost of undertaking this investment depends on the child's innate ability and two characteristics of the place where she lives: school quality and the location's skill-mix proxied by the share of high-skill neighbors. The key novelty is that children are uncertain about the returns to the investment in education, and learn about it in a Bayesian fashion by observing signals from wage realizations of individuals living in the same location. Wages differ among high-skill individuals because I consider they are a linear combination of a common and an idiosyncratic term; in turn the skill-composition is different across neighborhoods because of the location decision of parents that depends on exogenous amenities, school quality, and taste heterogeneity.

By Bayesian learning, children's beliefs about the high-skill wage are a weighted average of their prior belief and the public signal observed. Under the assumption that the precision of the signal is proportional to the population size of a location, the share of high-skill neighbors plays two roles. First, it reduces uncertainty about the skill premium, making children more likely to invest in education. Second, it determines the weight children put on the observed signal. This means that the higher is the share of high-skill neighbors, the more precise is the signal and therefore, the more children rely on the information disclosed by their neighbors. In this environment, what happens when children observe a low signal about the high-skill wage? A low signal leads to the perception of a lower skill premium, so a higher share of high-skill neighbors makes children more certain that the skill premium is low, hence higher exposure to high-skill neighbors with low wages translates into lower investment in education. In contrast, a high signal leads to the perception of a higher skill premium. Thus, as the share of high-skill neighbors increases, children have more information that the value of education is high, and therefore are more likely to invest in education. Consistent with the empirical findings, locations with a higher share of high-skill neighbors only have more children investing in education if the earnings of these high-skill neighbors, i.e. the signal observed, are sufficiently high. More exposure to highly educated neighbors brings about more information, but additional information only translates into greater investment in education if it leads to the perception of a higher skill premium.

To evaluate the quantitative implications of imperfect information and the local information transmission mechanism, I discipline the parameters of the model by a set of moments that describe the wage distribution by educational attainment, the distribution of households, and college enrollment across neighborhoods in the city of Detroit in 2013. First, I show that the model provides a good fit for the data. Armed with the calibrated economy, I then ask the following question: by how much would the college enrollment rate change in the absence of the information disclosed by high-skill neighbors? I find that the learning mechanism plays an important role. If individuals did not observe any public signal from high-skill neighbors, the fraction of students enrolling into college would drop from 38% to 21%. This result lies on the fact that learning from high-skill neighbors decreases uncertainty about the skill-premium by 31% and increases its expected value by 2.3%, on average.

In the model, differences in college enrollment across locations arise through three different channels: *(i)* information externalities due to local learning; *(ii)* local spillovers in the cost of human capital formation, and *(iii)* school quality. I decompose the contribution of each channel, and find that local learning is, by far, the most important in explaining inequality across school districts, in particular it accounts for 57% of the observed dispersion in college enrollment across school districts. Furthermore, it has important implications for intergenerational mobility. Due differences in the choice location of parents, the probability of becoming a high-skill worker for a children of low-skill parents is lower than the one for children born to high-skill families: 38% vs. 46%. Learning externalities explain 53% of this difference.

These results highlight the importance of imperfect information and local information transmission for the intergenerational propagation of inequality. Therefore, they have important policy implications for policy-makers interested in addressing opportunity equality, as policies that reduce information frictions differ substantially from policies aimed at tackling liquidity constraints or school quality. In particular, they underline the role of relocation policies such as the Moving to Opportunity that move disadvantageous families to better neighborhoods, and disclosure policies that inform individuals about the skill premium distribution (Hoxby and Turner, 2015; Bleemer and Zafar, 2016; Hastings et al., 2017) as a way to improve outcomes for children born to parents with low levels of education. I simulate such policies in the model. First, I find that moving children with low-skill parents from locations in the first quartile of the college graduate distribution to those in the last quartile increases their probability of enrolling in college from 25% to 49%. More than half of this effect is explained by the information role of neighborhoods. Second, I find that a disclosure policy, which informs children about the distribution of the high-skill wage, increases the college enrollment rate in 20 percentage points. More important, implementing only this policy, while leaving the other sources of inequality—human capital spillovers and school qualities—across neighborhoods at work, reduces significantly inequalities across locations and children from different backgrounds: *(i)* the standard deviation of the college enrollment distribution across neighborhoods reduces in 60%, while the enrollment gap between children with low educated parents and those with highly educated parents reduces in 62%. Given the low cost of these information campaigns, the policy case for implementing them is clear, specially when the success of other policies, such as subsidies or students loans, depends on whether children have have full information on education returns and costs.

1.1.1 Related Literature

This chapter relates to a number of existing literatures. First, it is primarily related to the theoretical literature that studies the role of residential location in determining intergenerational mobility and persistent inequality across generations. This literature has focused on two main channels. One is the *local financing of public schools* (Bénabou, 1996b,a; Férnandez and Rogerson, 1996; Durlauf, 1996). Because schools are funded through property taxes, wealthier families segregate into homogenous communities and poor children attend schools with lower resources. The other channel is *human capital spillovers*. These spillovers have been modeled in different ways. Akerlof (1997) and Akerlof and Kranton (2000, 2002), relate spillovers to the idea of *identity*. In locations where few parents are well educated, obtaining a high level of education may render the feeling of being alienated from those with whom one wants to share an identity. Bénabou (1993), Bowles et al. (2014), Cavalcanti and Giannitsarou (2013) and Kim and Loury (2013) consider instead that either the skill acquisition technology or the cost of human capital formation depend on the human capital of the individual's social network or neighborhood, without specifying a particular mechanism. Lastly, Mookherjee et al. (2010) suggest that location affects parents' aspirations and, thus, children's occupational choice. These theories, however, cannot account for the heterogeneous relationship between college graduates and college enrollment in a location along the earnings dimension. To explain this finding, this chapter introduces uncertainty about the skill premium and local information transmission into an otherwise standard model of human capital formation. To the best of my knowledge, this is the first model of human capital accumulation to take these features into account. I show that the interplay between imperfect information and local information transmission is important for the persistence of inequality across generations. On the one hand, I show it can reconcile my empirical findings. On the other hand, I show that this new channel explains a substantial portion of differences in enrollment across neighborhoods.

Second, this chapter builds on a theoretical and empirical literature that studies environments with information frictions and social learning, and shows how these features affect agents' decision making in different contexts such as technology adoption (Munshi, 2004), fertility decisions (Munshi and Mayaux, 2006), retirement savings (Duflo and Saez, 2003), female labor participation (Férnandez, 2013) and firms' investment decisions (Fajgelbaum et al., 2016). Closely related to this chapter is Fogli and Veldkamp (2011). They focus on explaining the rise of women's labor force participation in a few locations that gradually spread to nearby areas, as information about the costs of working was transmitted locally. My model introduces a similar learning environment in a model of human capital investment with local interactions under uncertainty. In doing so, it shows that imperfect information paired with local information transmission is an important channel through each neighborhoods affect education decisions and the intergenerational propagation of inequality.

Third, the documented facts speak to an important and vast empirical literature aimed at studying the impact of neighborhoods' socioeconomic environment on educational attainment of the young generation. This literature is reviewed in Durlauf (2004) and Topa and Zenou (2015). Despite being key to understanding the implications of the neighborhoods' skill-mix, the existing literature does little to investigate heterogeneity in the effect of neighborhoods' composition on students' educational attainment. The exception is Gibbons et al. (2013) who finds no heterogeneous effects on test scores of students between age 11 and 14 across different location characteristics such as number of students or population density. This chapter suggests that there are also important heterogeneities along the earnings dimension. Furthermore, while most of this literature (Oreopolous, 2003; Kling et al., 2007; Sanbonmatsu et al., 2008; Gibbons et al., 2013; Chetty et al., 2016, among others) treats neighborhoods as a "black box" in terms of the specific causal channels, I am able to shed some light over the role of different mechanisms through which the characteristics of a neighborhood affect educational attainment. In particular, my results suggest that information externalities at the neighborhood level are important: they are able to explain observed regularities and they are quantitatively important when compared to other channels in the literature.

Finally, this chapter is related to a growing literature that studies the role of imperfect information on educational choices. Recent studies show that individuals are uncertain about schooling returns, and that perceptions about the value of education and information constraints have significant impacts on different educational decisions (e.g. Jensen (2010), Attanasio and Kaufmann (2014), Kaufmann (2014), Bleemer and Zafar (2016), Hoxby and Turner (2014) and Belfield et al. (2016) look these effects on the choice to obtain further education, while Stinebrickner and Stinebrickner (2014) and Wiswall and Zafar (2015) focus on the the students' choice of major and Delevande and Zafar (2014) on the university choice).[3] I incorporate these features into a model of human capital accumulation with local externalities, and show that residential location is an important determinant of perceptions about the education value, and that correcting these inaccurate perceptions through a disclosure policy has important impacts in leveling the playing field among children from different backgrounds

The remaining of this chapter proceeds as follows. Section 1.2 presents novel evidence regarding the relationship between neighborhood's characteristics and educational outcomes, and Section 1.3 explains the empirical findings through a model of educational choice with uncertainty and local information transmission. In Section 1.4, I assess the quantitative importance of the proposed mechanism.

[3]The role played by perceptions and information on the decision to pursue further education is transversal to developed and developing countries. In the context of developing countries, Jensen (2010) finds that an intervention in Dominican Republic which informs 8^{th} grade students about actual returns increases school attendance. Also, Attanasio and Kaufmann (2014) and Kaufmann (2014) show that expected returns and risk perceptions are key determinants of education decisions in Mexico. In the context of a high income country, Bleemer and Zafar (2016) have similar results: using survey data for households in the US, they find that a higher perception about the college relative returns, increases the probability of parents sending their child to university. Also in the US, Hoxby and Turner (2015) designed an intervention aimed to improve information of disadvantaged students at the college application stage and find that it made them more likely to submit applications and attend college. Using a unique survey of secondary students in the UK, Belfield et al. (2016) show that perceptions about the returns and the consumption value of education play a role in education decisions.

Section 1.5 concludes.

1.2 Neighborhoods and Education: An Heterogeneous Relationship

In this section, I use school district level data from Michigan and document that when the college premium is low, a higher share of college graduates living in a school district is associated with lower college enrollment of students graduating from that district.

1.2.1 Data and Descriptive Statistics

The empirical exercise relies on a school district panel with annual frequency that runs from 2008 to 2014. This panel combines school district information along three dimensions *(i)* college enrollment, *(ii)* socio-economic characteristics of the school district, and *(iii)* school quality using the following sources:

College enrollment data comes from the Michigan Department of Education (CEPI). I measure college enrollment in a school district as the share of high-school students graduating from a public high-school in that district that enroll in a 4-year college within 6 months after graduation. This data also provides information on total number of students and cohort characteristics, namely students' gender and race per grade and the average American College Testing (ACT) score at the school district level.[4]

Socio-demographic data comes from the Education Demographic and Geographic Estimates of the National Center for Education Statistics (NCES - EDGE). This data has rich information on the socio-economic characteristics of school districts such as racial composition, family median income, unemployment rate and

[4]The ACT is a standardized test that measures high school students' skills to complete college-level work in four different areas (english, math, reading, and science) and is used as a college entrance exam in the United States. There is one ACT score (1 to 36) for each test and a composite ACT score, which is an average of the four tests. In my sample I have information on the latter. More information here: http://www.act.org/.

total population. Using this dataset, allows me to observe median annual earnings by education level, and the education level of individuals over 25 years old.

School district financing data comes from the Common Core of Data of the National Center for Education Statistics (NCES - CCD). Besides detailed information on expenditures and revenues, broken down by source (state, federal and local), this data has information on K-12 enrollment and the number of teachers in public schools per school district.

The sample for the empirical analysis is an unbalanced panel of school districts in Michigan urban areas covering the period from 2008 to 2014 and all public schools within a school district boundaries. Note that in Michigan a significant portion of the student body attends public schools, which mitigates concerns that the sample used is not representative of the whole student population in Michigan.[5] Further, coverage is close to universal, reaching 86% of urban school districts per year, on average.[6]

I summarize the main variables for the analysis in Table 1.A1. We observe that, on average, 33% of high-school students enroll college within 6 months after graduation and 23% of residents with 25 years old or more are college graduates. Descriptive statistics show that median earnings and the share of college graduates vary widely across districts, as do expenditures and revenues per student and student achievement, measured by the average ACT score and college enrollment rate. In addition, Table 1.A2 presents correlations among the main variables. As expected, college graduates' earnings, the average ACT score and the share of college graduates living in the school district are highly and positively correlated with the share of high-school graduates that enroll in a 4-year college within 6 months of graduation. Local revenue per pupil is also positively correlated with college enrollment. However, note that expenditures per pupil show no correlation with this variable. Interestingly, expenditures per pupil exhibits a small, but

[5]For instance, in 2013 around 83% of total students were enrolled in a public school, the vast majority in a local neighborhood school (only 6.5% of those enrolled in public K-12 schools were in a charter or magnet school).

[6]Due to data availability, in the year 2008 I only have data for 131 school districts, which compares to an average of 546.4 for the following years. I show that the results do not rely on including 2008 in the analysis.

negative, correlation with ACT score. The observed pattern seems to suggest that school resources play a small role in student achievement, as measured by college enrollment and the score in the ACT.

1.2.2 Empirical Framework

To formally examine the relationship between the share of college graduates living in a school district and college enrollment, and the presence of heterogeneities along the earnings dimension, I estimate the following equation:

$$\begin{aligned} Enrollment_{ijt} = \beta_0 \; + \; \beta_1 \mathit{College}_{ijt} \; + \; \beta_2 College_{ijt} \times Y_{ijt} \; + \; \beta_3 Y_{ijt} \\ + \; \delta X_{ijt} \; + \; \gamma_j \; + \; \gamma_t \; + \; \rho_j t \; + \; \varepsilon_{ijt} \end{aligned} \tag{1.1}$$

where $Enrollment_{ijt}$ is the share of high-school students graduating from a public schools that enroll in a 4-year college within 6 months of graduation in school district i within city j at year t, $College_{ijt}$ is the share of individuals over 25 years old with a college degree living in school district i within city j at year t, and Y_{ijt} corresponds to the median annual earnings of individuals with a college degree living in school district i within city j at year t. X_{ijt} is a set of school district controls, γ_j and γ_t are city and year fixed effects, and ρ_j is a city-specific time trend. ε_{ijt} is the error term, that captures all unobserved determinants of college enrollment of school district i within city j at year t. I allow for arbitrary within-district correlation of the errors by clustering standard errors at the school district level. Under the standard exogeneity restrictions, the effect of the share of college graduates living in the school district on the college enrollment of high-school graduates is identified by β_1 and β_2,

$$\frac{\partial Enrollment_{ijt}}{\partial \mathit{College}_{ijt}} = \beta_1 + \beta_2 \times Y_{ijt} \tag{1.2}$$

If $\beta_1 > 0$ and $\beta_2 = 0$, the effect of the share of college graduates is constant across different levels of earnings, in line with standard models of human capital formation with local externalities. In contrast, if $\beta_1 < 0$ and $\beta_2 > 0$ or $\beta_1 > 0$ and $\beta_2 < 0$, there is an earnings threshold above which the effect of the share of college graduates living in school district on the college enrollment is positive, and below which is negative.

Identification To identify β_1 and β_2, I exploit variation in the share of residents with a college degree and their median earnings across school districts within a city over time. To illustrate that there are indeed important differences in the magnitude of the main variables of interest between a city's school districts: within a city, the share of college graduates living in a school district varies, on average, between 9% and 37%, and the median annual earnings of individuals with a college degree range, on average, from around 30 000 to 82 000 dollars. The empirical framework exploits these variations to study to what extent the skill-mix of a school district correlates with within-city differences in college enrollment, and whether there are differences in this correlation along the earnings dimension. Note that I do not exploit within school district variation because the variables of interest have little variation within school districts over time when compared to across school districts within a city. For instance, in the panel used for the empirical analysis, the share of college graduates in a school district ranges, on average, from 18% to 22%,

Identification in the OLS framework relies on the assumption that the skill-mix of a school district is exogenous to share of high-school students enrolling in college. However, there are many potential confounders at the school district level that could correlate with both the skill-mix of a school district and college enrollment. school districts with a higher share of college graduates might also be school districts where the ability of high-school students is lower, and lower ability is likely bad for college enrollment. school districts with a higher share of college graduates might also be places where high-school students attend public schools with better resources or have higher family income. For instance, Bayer et al. (2004) and Bayer et al. (2007) find that individuals with a college degree are willing to pay $13.03 more per month than high-school graduates to live in a neighborhood with a higher school quality, as measured by average test scores. They are also willing to pay more for locations with higher population density, average income and a higher share of black residents.

I partially address potential omitted variable bias by including the vector of controls X_{ijt}. First, I control for the characteristics of the cohort that graduated from high-school in a given year by including *(i)* the share of females in the 12th grade[7], and *(ii)* the average ACT score of the graduating class. The latter is particularly important as it allows me to control for the fact that highly educated parents may have children with higher ability, hence more likely to enroll in college. Because in 2007 Michigan implemented a mandatory ACT policy, which requires and pays for college entrance exams for all public school eleventh graders, the average ACT score is a good proxy for the ability of high-school graduates in public schools. Second, X_{ijt} includes school quality measures, namely expenditure, local revenue and teachers per pupil. Thus, the coefficients on $College_{ijt}$ and the interaction term are not capturing the effect of better schools in locations with highly educated adults as suggested by models that explore local funding of education as a mechanism that links neighborhoods to educational outcomes (Bénabou, 1996a,b; Férnandez and Rogerson, 1996; Durlauf, 1996). Third, Equation (1.1) also controls for socioeconomic characteristics at the school district level such as the the share of black and white residents, the median annual family income, the unemployment rate and the median earnings of high-school graduates. Finally, I also control for location attributes by including population density.

OLS estimation of the relationship between college graduates and college enrollment using specification (1.1) also controls for *(i)* unobserved factors that may influence enrollment and are associated with the city to which the school district belongs to; *(ii)* for time varying shocks affecting all school districts; and *(iii)* for unobserved city level characteristics that might be trending over time such as gentrification dynamics or deterioration in housing quality. As a robustness check, I also run a specification that controls for time shocks affecting all school districts within a city. Given this, the key assumption for causal interpretation of β_1 and β_2 is then that unobserved determinants of college enrollment are mean-independent of the share of college graduates and their earnings, conditional on the controls included. I discuss the plausibility of this interpretation further in Section 1.2.5.

[7]It has been widely documented that nowadays females are more likely to enroll in college than males.

1.2.3 Results

I start by estimating the relationship of college enrollment with the neighborhood's skill-mix that with the standard specification in the literature, i.e. a version of specification (1.1) that does not include the interaction term, $College_{ijt} \times Y_{ijt}$. Panel A in Table 1.1 provides points estimates of the coefficients of interest. Column 1 shows a positive and statistically significant relationship between the share of college graduates living in the school district and college enrollment by high-school graduates attending public schools in the same district.

As one can see in column 2, even though ACT scores seem correlated with the share of college graduates and earnings, the coefficient of interest $\hat{\beta}_1$ remains positive, large and significant. Column 3 and 4 present, respectively, results from specifications controlling for socioeconomic conditions of the school district and school resources, in order to account for neighborhood traits that can be correlated with both college enrollment rate and the share of college graduates. Column 5 includes city-year fixed effects, so as to control for shocks affecting all school district in a given city and year. Finally, to address possible concerns over heterogeneous trends, column 6 includes city-specific linear trends. [8] Across all these different specifications, the sign, the magnitude and significance of $\hat{\beta}_1$ are barely affected. Also, note that the fact that the coefficient estimate of $College_{ijt}$ remains unchanged when I introduce school resources variables suggests that the role played by highly educated neighbors goes beyond the school resources, in contrast to what is suggested by models of local public funding proposed by developed by Bénabou (1996b), Bénabou (1996a), Férnandez and Rogerson (1996) and Durlauf (1996). Table 1.A3 reports the coefficients on school quality measures and average ACT score, and shows that while average ACT test score is statistically significant, school quality measures are not.

According to the estimate in column 6, an increase in the share of college graduates living in the school district by 10 percentage points, is associated with an increase of college enrollment at the school district by 3.66 percentage points. This result is in line with the findings in Chetty and Hendren (2017), who find that

[8] As I include additional controls, I loose some observations due to missing variables. My results are robust to restricting the sample to school districts with the full set of controls.

moving to an area with higher college attendance rates at a younger age increases a child's probability of attending college.

Heterogeneity by Earnings The results presented so far show that the relationship between the skills of older neighbors and college enrollment is positive regardless of other socioeconomic characteristics of school districts. However, this result might mask important heterogeneities. I now investigate whether this relationship is heterogeneous along the earning dimensions by replicating all specifications in Panel A, Table 1.1, including now the interaction term, $College_{ijt} \times Y_{ijt}$ as well as Y_{ijt} by itself. I also include as a control the median annual earnings of high-school graduates, which allows me to control for differences in the skill premium across school districts. Panel B in Table 1.1 presents OLS estimates of Equation (1.1).

Column 1 shows that the coefficients of $College_{ijt}$ and the interaction are significant at the 99% confidence level, with $\hat{\beta}_1 < 0$ and $\hat{\beta}_2 > 0$. This result uncovers a threshold in the earnings distribution below which a higher share of college of college graduates living in the school district is associated with a decrease in the college enrollment rate. Panel A in Figure 1.A1 illustrates the average marginal effect of college graduates on college enrollment along the earnings dimension. One can see that as earnings increase, the correlation between the share of college graduates and college enrollment increases. More important, it shows that for low values of college graduates' earnings, this correlation is negative, while at high values is positive.

This pattern remains almost unchanged if I control for the share of females and the average ACT score (column 2), socioeconomic and location characteristics (column 3), school quality measures (column 4), as well as if I take into account time shocks that affect all school districts within a city (column 5) and city-specific linear trends (column 6), with the latter being the preferred specification. As before, when I include school quality controls, not only the coefficient estimates of school quality are small and statistically insignificant, but the magnitude of the coefficients of interest — $\hat{\beta}_1$ and $\hat{\beta}_2$ — are very similar to the ones reported in column 3. This evidence, points against school quality as channel through which neighborhoods affect and educational outcomes.

Table 1.1: College Enrollment and College Graduates

Dependent Variable: Share of High-School Graduates that Enroll in a 4-year College

	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: No Heterogeneity						
College Graduates	0.777***	0.437***	0.387***	0.375***	0.367***	0.366***
	(0.029)	(0.035)	(0.050)	(0.053)	(0.054)	(0.053)
Observations	1841	1839	1827	1818	1818	1818
Adjusted R^2	0.703	0.786	0.798	0.798	0.803	0.801
Panel B: Heterogeneity by Earnings						
College Graduates	-5.989***	-5.508***	-4.763***	-4.783***	-4.771***	-4.708***
	(1.468)	(1.252)	(1.122)	(1.110)	(1.111)	(1.097)
College Graduates × Earnings	0.619***	0.550***	0.478***	0.479***	0.477***	0.471***
	(0.135)	(0.115)	(0.104)	(0.103)	(0.103)	(0.102)
Observations	1841	1839	1827	1818	1818	1818
Adjusted R^2	0.737	0.795	0.804	0.805	0.810	0.807

Notes: The table reports coefficients from an OLS regression with robust standard errors clustered at the school district level reported in parentheses. Column 1 includes only city and year fixed effects. Column 2 to 6 control for characteristics of the graduating class (the share of females among the high-school graduates and the average ACT score). Column 3 to 6 also include socioeconomic controls, which include the share of black and white residents, the unemployment rate, the median family income, school district size. Column 5 includes city-year fixed effects and column 6 city fixed effects and a city-specific time trend. The sample includes all school districts within all MSAs in Michigan over the period 2008 and 2014. ***, ** and * represent statistical significance at 1%, 5% and 10% levels, respectively. Source: CEPI, NCES-EDGE and NCES-CCD.

1.2.4 Robustness Checks

Next, I provide evidence that my results are robust to *(i)* omitted variable bias, *(ii)* a quadratic specification in earnings, *(iii)* adding lagged enrollment as a control, *(iv)* a reduced sample to include only school districts with few non-resident students and the years after the Great Recession, and *(v)* different measures of earnings of college graduates.

Omitted Variable Bias As discussed in Section 1.2.2, there could be many confounders at the school district level correlating with both the share of college graduates, their labor market income and college enrollment. Although we address potential omitted variable bias by (i) cohort variables, (ii) socioeconomic characteristics of the school district and (iii) school quality measures in all our regressions as well as city effects, we cannot fully rule out the existence of unobservable determinants of college enrollment which correlate with initial income inequality even conditional on these controls. Therefore, I now follow the approach of Oster (2016) to evaluate the robustness of my estimates to potential omitted variable bias.

Under the assumption that selection on the observables is proportional to selection on the unobservables by a factor δ[9], Oster (2016)'s bias-adjusted coefficient is

$$\beta_i^* = \hat{\beta}_i - \delta(\tilde{\beta}_i - \hat{\beta}_i)\frac{R^{max} - \hat{R}}{\hat{R} - \tilde{R}}, i = 1, 2 \tag{1.3}$$

where $\hat{\beta}_i$ and $\hat{R}$ are the estimated coefficients and R^2 of column 6 in Panel B of Table 1.1 and $\tilde{\beta}_i$ and $\tilde{R}$ come from OLS estimation of Equation (1.1) with no controls (i.e. not including city and year fixed effects, a city-specific trend and the vector X). δ captures the explanatory power of unobserved variables as a proportion of the explanatory power of observed variables and R^{max} denotes the R^2 of a hypothetical OLS regression if one could control for all relevant (observed and unobserved) variables. To identify β_i^*, I use $\delta = 1$ and $R^{max} = 1$, which yields the identified set for the coefficient estimates $[\hat{\beta}_i, \beta_i^*]$.[10] The identified set for β_1 is $[-4.697, -3.424]$ and for β_2 is $[0.472, 0.331]$. Because both exclude zero, my results can be interpreted to be robust to omitted variable bias under the assumption that selection on the observables is proportional to selection on the unobservables by a factor δ as argued by Oster (2016). Figure 1.A2 plots the average marginal effect with $\hat{\beta}_i$ and β_i^*, and it shows that the heterogeneity in the relationship between college graduates and enrollment along the earnings dimension remains unchanged.

Quadratic Specification Equation (1.1) assumes that the correlation between college graduates and college enrollment is linear in earnings. However, if it is instead quadratic in earnings, approximating it with a linear specification could be driving the negative sign of $\hat{\beta}_1$. Given this, I estimate a version of Equation (1.1) where I consider the effect of college graduates on enrollment to be quadratic in

[9] This assumptions means that $\delta \cdot \frac{cov(x,w_1)}{var(w_1)} = \frac{cov(x,w_2)}{var(w_2)}$, where x is the independent variable of interest, w_1 are observable controls and w_2 are unobservable controls.

[10] According to Oster (2016), to determined the identified set one should set $\delta = 1$ and $R^{max} = min\{2.2\hat{R}, 1\}$.

earnings: $\frac{\partial Enrollment_{ijt}}{\partial College_{ijt}} = \beta_1 + \beta_2 \times Y^2_{ijt}$. Column 1 in Table 1.A5 reports the estimated coefficients, and Panel B in Figure 1.A1 displays the average marginal effect of college graduates on college enrollment along the earnings dimension under this specification. This figure is very similar to the left panel, thus the assumption that the effect of college graduates on enrollment is linear in earnings is suitable.

Lagged Enrollment school districts where college graduates have low earnings could also be the school districts where college enrollment has been low over the years. To account for this issue, I include college enrollment in the previous year as a control. Column 2 in Table 1.A5 and Panel C in Figure 1.A1 show that the main result still holds: there is a threshold in the earnings distribution below which the association between enrollment and college graduates is negative, and above which is positive

Pre-determined Controls Equation (1.1) includes the vector of controls X_{ijt} to take into account the fact that individuals with different education levels may locate in systematically different neighborhoods, whose characteristics might lead to differences in college enrollment. These controls are contemporaneous to the variables of interest, $College_{ijt}$ and Y_{ijt}. While their inclusion might partially control for omitted factors, these variables can themselves be affected by the variables of interest: for instance, it is likely that the unemployment rate or the median family income at the school district level are determined by its skill composition. To assess whether my results are robust to the bad controls problem, I estimate the a version of Equation (1.1) including instead a vector of 2009 school district level controls, X_{ij2009}. Column 3 in Table 1.A5 shows that my findings remain unchanged.[11]

School Choice The empirical analysis in the previous section assumes that high-school graduates live in the school district where they go to school. However,

[11] I estimate the following regression: $Enrollment_{ijt} = \beta_0 + \beta_1 College_{ijt} + \beta_2 College_{ijt} \times Y_{ijt} + \beta_3 Y_{ijt} + \delta X_{ij2009} + \gamma_j + \gamma_t + \rho_j t + \varepsilon_{ijt}$. I use the sample from 2010 to 2014 in this estimation, therefore the results from this estimation should be compared with the ones that exclude the great recession period in Table 1.A5, column 4.

in Michigan there is a school choice program, established in 1996, under which families can opt to move their children out of the schools they would attend by residency to neighboring districts.[12] Between 2008 and 2014, only 19% of the 12th students were non-resident students. Nevertheless, to check the robustness of the results to the inclusion/exclusion of students that do not live in the school district they attend, I re-estimate Equation (1.1) focusing only on school districts which have a low share of non-resident students attending the 12th grade (I fix the share threshold at 10%). Column 4 in Table 1.A5 shows that my findings hold when we exclude from the analysis school districts with a higher share of non-resident students attending 12th grade. More important, the coefficients estimates are relatively similar.[13]

Great Recession The sample used in the empirical analysis covers the period between 2008 and 2014, which includes the period of the Great Recession. According to the National Bureau of Economic Research, the Great Recession began in December 2007 and ended in June 2009. To assess the sensitivity of my results to this period, I re-estimate Equation (1.1) restricting the sample to the years after this period, 2010-2014. I find that the results are robust to the Great Recession (columns 5 and 6 in Table 1.A5): the sign of the estimated coefficients is the same as the one in column 6 in Table 1.1 and their magnitude is relatively unchanged.

Urban and Rural school districts So far, I have focused on school districts that are located within MSA's boundaries. Column7 in Table 1.A5 shows that the sign and significance of coefficients remains unchanged when I also take into account school districts in rural areas, albeit their magnitude are smaller.

Earnings of High-skill Neighbors I replicate the estimation of column 6 in Table 1.1 using median annual earnings of individuals with a post-graduate degree

[12]According to section 105/105C of the Michigan State School Aid Act, all students in Michigan must be allowed to choose to leave their home districts, and when students move districts, the state aid funding travels with them to the destination district. Nevertheless, school districts are allowed to choose whether to accept students from other districts (http://www.michigan.gov/mde/0,4615,7-140-6530_30334-106922--,00.html).

[13]Data on non-resident students per grade and school district comes from the Michigan Department of Education.

and the average between this measure and the median annual earnings of individuals with a college degree so as to capture the earnings of individuals with a college education or higher. Columns 8 and 9 in Table 1.A5 show that using these two alternatives proxies leaves the magnitude, sign and significance of the coefficients of interest relatively unchanged.

All in all, I find robust evidence that there is a threshold in the earnings distribution below which a higher share of college graduates living in the school district is associated with lower college enrollment in that district; and above which this relationship is positive.

1.2.5 Mechanisms

The previous sections reported the results of several descriptive exercises that investigated how a school district skill-mix is related with college enrollment of high-school students living and graduating from that district. In particular, I uncover a new pattern showing that the correlation between the share college graduates and college enrollment at the school district level is only positive if their labor market earnings are sufficiently high.

Why do individuals with a college degree may have negative or positive externalities in the decision to enroll in college depending on their level of earnings? I now argue that a potential mechanism behind the documented evidence is the transmission of information about the returns to education at the local level. First of all, the observed pattern is hard to to reconcile with existing models of human capital formation with local spillovers (Bénabou, 1993; Bowles et al., 2014; Cavalcanti and Giannitsarou, 2013; Kim and Loury, 2013) because these models predict the relationship between the neighborhood's skill-mix and college enrollment to be *(i)* positive and *(ii)* independent of the level of earnings. Second, the controls included in Equation (1.1) exclude the following alternative explanations:

Credit Constraints In school districts where the college premium is low, students could be credit constrained, which is likely to reduce the college enrollment rate. Under this scenario, a negative relationship between the share of college graduates and college enrollment would arise when the earnings of college graduates are low. However, specification (1.1) controls for the median household income, so I compare districts where families have on average the same resources but earnings of college graduates vary.

Ability Parents with a college degree but low earnings, may also have low ability children, a feature that is expected to be negatively associated with college enrollment. Thus, this would potentially generate a negative correlation between the share of college graduates and college enrollment. I address this alternative channel by including the average ACT test score of high-school students in the 12^{th} grade as a I control, which I previously mentioned, is a good proxy for the ability of the students.

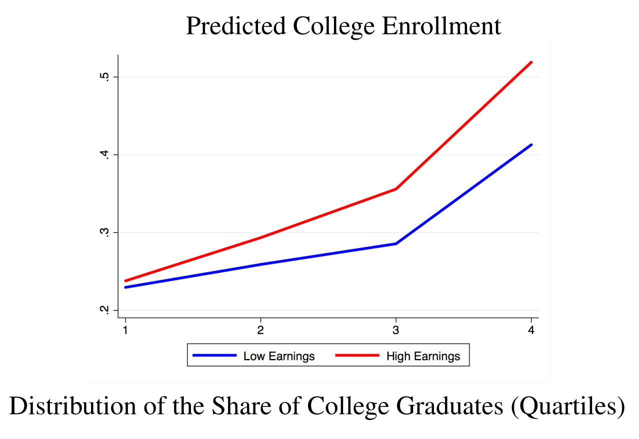
School Resources One may also think that the empirical findings are driven by differences in the resources of public schools across school districts: children living in school districts with where college graduates earn low earnings attend schools with lower resources, hence are less likely to attend college. Nonetheless, I control for school expenditures at the school district level, local funding of the school district, and teachers per student.

To build intuition for the proposed mechanism, under a context of uncertainty, if the assessment of the education value depends on the distributions of educational levels and incomes observed in a neighborhood, then in locations where the earnings of college graduates are low, and the share of college graduates is high, high-school students have a lot of information that the value of education is low, and therefore are less likely to enroll in college. In contrast, in neighborhoods where the exposure to college graduates is high and their earnings are also high, high-school students have a large amount of information suggesting that the returns to education are high.

Figure 1.1 plots the predicted college enrollment for different levels of the neighborhood's skill-mix. The red line corresponds to school districts where college graduates have low earnings, and the blue line represents school districts where college graduates have high earnings. Two important facts can be drawn from that picture. First, when the share of college graduates is low, the difference in college enrollment between locations where college graduates have high earnings and locations where their earnings are low is not significantly different from

0. Second, as the share of college graduates increases, the difference between both groups of school districts widens substantially. This plot strongly supports the local information transmission as a mechanism behind the observed pattern. When students have little information, i.e. live in a district with a low share of college graduates, students' beliefs about the skill premium will not differ between neighborhoods with high and low earnings, and therefore enrollment is similar. As high-school graduates have more labor market information, i.e. are exposed to a higher share of college graduates, they rely more on the information on the information at the neighborhood level. As a result, in places where college graduates earn more, perceptions about college earnings are higher, which translates into a higher college enrollments. Next, I formally illustrate how the empirical findings described in this section are consistent with a theory of local learning about an uncertain skill premium.

Figure 1.1: Local Information Transmission in the Data



Notes: This figure plots predicted college enrollment for each quartile of the distribution of the share of college graduates. The blue and the red line represent, respectively, school districts in the last and first quartile of the distribution of college graduates' earnings. Source: CEPI, NCES-EDGE and author's calculations (2008-2014).

1.3 Education Choice with Information Frictions and Local Learning

Section 1.2 documents that when the earnings of college graduates are low, a higher share of college graduates living in a school district is associated with lower college enrollment of high-school students graduating from that district. This is a surprising result as existing models of human capital formation with human capital spillovers predict this relationship to be positive and independent of earnings. This section outlines a model that provides a natural role for imperfect information and local information transmission in explaining the documented pattern.

Motivated by empirical evidence showing that individuals lack information about education returns and that neighborhoods play a role as an information source, the model makes two key assumptions. First, when deciding whether to become a high-skill worker or not, children do not know the skill premium. Second, children learn about it by observing wage realizations of their direct neighbors. The neighborhood's skill-mix, which is driven by exogenous amenities and dispersion forces (in the form of an inelastic supply of houses in each neighborhood and taste heterogeneity), shapes children's perception about the skill premium and, therefore, the education choice.

Consistent with the facts described in Section 1.2, the model shows that in an environment with imperfect information and local learning, there is a wage threshold below which a higher share of high-skill neighbors living in the neighborhood translate into lower investment in education. To clearly illustrate the mechanism, I make several assumptions that make the model simpler. Section 1.3.6 discusses their implications, and shows that they do not affect the model's key prediction.

1.3.1 Environment

Population There are M households living in a city. Each household is composed by a parent and a child. Parents are of two types, high-skill (H) and low-skill (L), $k \in \{H, L\}$.[14] Each parent provides, inelastically, one unit of labor in

[14] As in Diamond (2016), I use a two skill group model because the largest group divide in wages across education is seen between college and non-college graduates, as found by Katz and

the city, for which she is compensated with a wage. The city is closed, hence the population of high-skill and low-skill parents in the city, M_H and M_L respectively, are exogenously given.

City The city is composed by a set of J discrete neighborhoods, indexed by $j \in \{1, ..., J\} \equiv \mathcal{J}$. Neighborhoods differ in their attractiveness. This can be due to geographical characteristics (weather, coastal access, etc), but also due to man-made features (school quality, retail environment, distances to places of employment, recreation, noisy streets, etc.). I call amenities to all these features that influence a location attractiveness besides rental prices. As in Busso et al. (2013), each neighborhood is characterized a fixed bundle of amenities A_j composed of two skill-specific attributes, $A_j = \{A_{jH}, A_{jL}\}$[15], and school quality q_j. Both A_j and q_j attributes of each location are taken by individuals as exogenously given.[16] All local residents have access to these amenities. Even though the city's high-skill and low-skill populations are exogenous, the quantity of high and low-skill parents living in a given neighborhood j, M_{jH} and M_{jL} respectively, are endogenously determined equilibrium outcomes. The city has sufficient capacity that everyone can reside on it, but I consider that each location j is endowed with an inelastic supply of identical houses H_j as in Bayer et al. (2007) and Ferreyra (2009). Houses are owned by a zero measure of absentee landlords, who rent it to households. Families live in only one house.

Preferences All individuals have preferences over an homogeneous consumption good c and amenities. The consumption good is a tradable numeraire good

Murphy (1992) and Goldin and Katz (2008).

[15]This aims to capture the idea that different types of individuals tend to prefer different types of amenities as in Glaeser et al. (2016) and Diamond (2016). Glaeser et al. (2016) assume that the income share of amenities is higher for skilled than unskilled individuals. Diamond (2016) allows for the utility value of the cities' amenities to differ between high and low skill groups. There is empirical evidence that supports this specification. Bayer et al. (2004) and Bayer et al. (2007) document that individuals with different education levels have a different willingness-to-pay for different location attributes: for instance, when compared to high-school graduates, college graduates are slightly more willing to pay to live in locations that are further away from the workplace and characterized by a higher population density.

[16]Even though this may strike as a strong assumption, in Section 1.3.6 I argue that introducing endogenous amenities (considering, for instance, that a component of neighborhood's attractiveness depends on its skill-mix) would not change the main prediction of the model.

with price normalized to one. For simplicity, I consider that only parents consume. I assume that all individuals have constant absolute risk aversion (CARA) utility over consumption with risk-aversion parameter γ.[17] The utility for an individual i of type $k \in \{H, L\}$ living in neighborhood j is given by

$$U(c^k_{i,j}, \Phi^k_{i,j}) = \frac{-\exp(-\gamma(c^k_{i,j}))}{\Phi^k_{i,j}} \tag{1.4}$$

where $c^k_{i,j}$ is consumption of individual i of with skill-type $k \in \{H, L\}$ living in neighborhood j. $\Phi^k_{i,j}$ maps the attractiveness of neighborhood j to the individual i's utility value for her.

Wages Parents pay for consumption and one unit of housing out of their labor income. I consider wages to be exogenous. Let $w^H \equiv \log(\omega^H)$ and $w^L \equiv \log(\omega^L)$, I assume that $w^H_i = w^H + \epsilon^H_i$, with $\epsilon^H_i \sim \mathcal{N}(0, \sigma^2_{\epsilon^H})$, and that $w^L_i = w^L + \epsilon^L_i$, with $\epsilon^L_i \sim \mathcal{N}(0, \sigma^2_{\epsilon^L})$. $w^H > w^L$. Following empirical evidence showing that wage dispersion is substantially higher among highly educated workers (Lee et al., 2017), I normalize $\sigma^2_{\epsilon^L}$ to 0. Section 1.3.6 discusses the implications if instead $\sigma^2_{\epsilon^L} > 0$.

Timing and Decisions The timing of decisions in the model is the following. Parents draw a wage from the wage distribution corresponding to their skill level and then choose where to locate within the city. Children are born with identical beliefs about the high-skill wage, receive information from high-skill neighbors and update these beliefs. Based on these beliefs, children decide to invest or not in education by comparing the cost of skill acquisition with their perceptions about the skill premium.

1.3.2 Parents' Location Choice

At the very beginning of the period, before their children decide whether to invest or not in education, a k-type parent draws a wage from the k-type wage distri-

[17] In Section 1.3.6, I show that the model's main prediction is qualitatively robust to this assumption.

bution. Then, parents simultaneously choose a neighborhood j to live in such that they maximize their utility taking as given labor income. The location choice is affected by two factors. First, an utility shock associated with living in each neighborhood in the city. This can be interpreted as the idiosyncratic utility cost or benefit of living in a given neighborhood. Second, parents compare the attractiveness of living in different neighborhoods. Taking this into consideration, a parent chooses to live in neighborhood j if either he likes location j for idiosyncratic reasons or because amenities are much better in j. For tractability, I proxy *altruism* by assuming that, when choosing where to locate, parents take into account the school quality of the neighborhood.[18] Parents i with skill level $k \in \{H, L\}$ solves the following program:

$$\underset{j}{\text{Max}} \quad U(c^k_{i,j}, \Phi^k_{i,j}) = \frac{-\exp(-\gamma \cdot c^k_{i,j})}{\Phi^k_{i,j}} \quad \text{subject to} \quad c^k_{i,j} + r_j = w^k_i \tag{1.5}$$

where w^k_i is the wage of parent i with skill level k and r_j is the rent payed to live in neighborhood j. I consider that

$$\Phi^k_{i,j} = q_j \cdot A_{j,k} \cdot \varepsilon_{i,j} \tag{1.6}$$

Individual's i idiosyncratic taste for neighborhood j is denoted by $\varepsilon_{i,j}$. I model this heterogeneity following McFadden (1973).[19] For each parent i, I consider that the idiosyncratic taste for neighborhood j is drawn from a Fréchet distribution (also called the Type II extreme value distribution):

$$Pr(\varepsilon_{i,j} \leq x) = e^{-x^{-\theta}}, \text{for x} > 0, \text{iid}, \theta > 0 \tag{1.7}$$

where the parameter θ reflects the amount of variation in the distribution and is

[18] Alternatively, I could assume that parents have warm-glow preferences in which the parents' utility function depends on the expected value of the child's income. This would entail solving a fixed-point problem when determining the location problem of parents as their utility would depend on the equilibrium skill-mix of the location.

[19] Following McFadden (1973), a long line of models with location decisions using preference heterogeneity has emerged, such as (Bayer et al., 2007), Kennan and Walker (2011), Ferreyra (2009), Busso et al. (2013), Ahlfeldt et al. (2015), Monte et al. (2015), Diamond (2016), among others.

treated as common across all parents.[20] In the location choice context, θ governs preference heterogeneity for locations across parents.[21] The idiosyncratic taste shock implies that when faced with the same rental prices and neighborhood amenities equal parents, with the same skill and wage, may choose to live in different locations.

The indirect utility function of parent i of type $k \in \{H, L\}$ living in neighborhood j can then be represented as

$$U(w_i^k, r_j, q_j, A_{j,k}, \varepsilon_{i,j}) = \frac{-\exp(-\gamma(w_i^k - r_j))}{q_j \cdot A_{j,k}} \varepsilon_{i,j} \tag{1.8}$$

Let $\rho_{i,j}^k$ be the probability that, after observing the vector of $\varepsilon_{i,j}$ (one for each location), parent i with skill level k chooses to live in location j. The distributional assumption on the idiosyncratic taste allows me to derive a close-form expression for $\rho_{i,j}^k$:

$$\rho_{i,j}^k = \frac{(q_j A_{j,k})^\theta (\exp(-\gamma(w_i^k - r_j)))^{-\theta}}{\sum_{j' \in \mathcal{J}} (q_j A_{j',k})^\theta (\exp(-\gamma(w_i^k - r_{j'})))^{-\theta}} \tag{1.9}$$

Other things equal, a type-k parent is more likely to live in a neighborhood the more attractive are j-specific amenities and the lower are rental prices (r_j). Since migration is only allowed in the beginning of the period, $\rho_{i,j}^k$ translate directly into the neighborhood size distribution. The equilibrium number of k-skill parents in neighborhood j, $M_{k,j}$, is given by

$$M_{k,j} = \sum_{i=1}^{M_k} \rho_{i,j}^k = \rho_j^k M_k \tag{1.10}$$

where M^k is the exogenous measure of k-type parents living in the city.[22] Given this, the total population living in neighborhood j is $M_j = M_j^H + M_j^L$. In order for the housing market to clear, the demand for houses in neighborhood j must

[20]The general cumulative distribution function for the Fréchet distribution is $\Pr(X \leq x) = \exp(-(\frac{x-\mu}{\beta})^{-\theta})$ if $x > \mu$, where μ is the location parameter and β is the scale parameter. I am implicitly setting β=1 and μ=0.

[21]The larger is θ, the smaller is taste dispersion: if θ tends to infinite, the variance of idiosyncratic shocks is zero. In that case, only amenities determines neighborhood choice.

[22]See Appendix 1.6.2 for details.

equal the supply in that location and so:

$$H_j = \rho_j^H M_H + \rho_j^L M_L, \forall j \in \mathcal{J} \tag{1.11}$$

The distribution of amenities across neighborhoods determines the skill-mix of each neighborhood and, therefore, whether low-skill households live more or less isolated from the high-skill ones. As shown in the example in Appendix 1.6.2, when amenities are equal across neighborhoods the spatial equilibrium is non-sorted. In this environment, amenities do not react to the characteristics of the population that chooses to live on it. In Section 1.3.6, I discuss the implications of relaxing this assumption.

1.3.3 Children Investment Decision

Children are born to a household of type k, $k \in \{H, L\}$, living in neighborhood j. Besides family background, children differ in their innate ability a, which is known. The distribution of ability is assumed to be the same across neighborhoods and households types and is given by the distribution function $\mathcal{G}(a)$, with support $[\underline{a}, \overline{a}]$. Innate ability together with human capital spillovers from the location skill-mix, as in Bénabou (1993), Bowles et al. (2014) and Kim and Loury (2013), and school resources (Bénabou, 1996a,b; Durlauf, 1996) determine the cost of skill acquisition. The cost function c is continuous and strictly decreasing in innate ability, human capital spillovers and school resources.

Given the cost $c(a_i, q_j, m_{jH})$, all children have to decide whether to invest or not in education. Not investing implies the child to work as a low-skill worker, while investing, implies the payment of the investment cost and working as a high-skill worker. The key and novel feature in this model is that, at the investment stage, children are uncertain about the return to human capital investment, namely, they do not know the true value of the wage they will receive as a high-skill worker, w_i^H. Therefore, children make their investment choice based on their perceptions about the skill premium.

Information Set Spatial location determines the composition of the signals in the children's information set. Children acquire information about w_i^H through social learning. In particular, they learn about it by observing noisy signals of the wage realizations of high-skill parents living in the same neighborhood as them. Each signal from a high-skill neighbor s living in neighborhood j is given by

$$w^H_{s,j} = w^H + \epsilon^H_{s,j}, \tag{1.12}$$

where $\epsilon^H_{s,j}$ denotes the signal noise. Following Fajgelbaum et al. (2016), I assume that the information gathered by each high-skill neighbor in neighborhood j is proportional to its size,

$$\epsilon^H_{s,j} \sim \mathcal{N}(0, M_j \sigma^2_{\epsilon^H}), \tag{1.13}$$

this means that the largest is the neighborhood, the noisier are the signals. Because of the normality assumption, a sufficient statistic for the information provided by high-skill parents living in neighborhood j is the public signal

$$w^H_j \equiv \frac{1}{M_{jH}} \sum_{i=1}^{M_{jH}} w^H_{s,j} = w^H + \epsilon^H_j, \tag{1.14}$$

with

$$\epsilon^H_j \equiv \frac{1}{M_{jH}} \sum_{s=1}^{M_{jH}} \epsilon^H_{s,j} \sim \mathcal{N}(0, m^{-1}_{jH} \cdot \sigma^2_{\epsilon^H}), \tag{1.15}$$

where m_{jH} is the fraction of high-skill parents living neighborhood j. The signal is neighborhood-specific: all children born in j observe the same high-skill parents, hence a common public signal, w^H_j. Important for the model's key prediction, the signal's precision, $m_{jH} \cdot \sigma^{-2}_\epsilon$, increases with the share of high-skill parents in the neighborhood.

Learning Initial beliefs are assumed to be identical across all children, $\tilde{w}^H_i \sim \mathcal{N}(\tilde{\mu}, \tilde{\sigma}^2)$.[23] To update these beliefs, they use information gathered by the public signal w^H_j. Children are *passive learners* and cannot take any action to change the quality of this signal: after receiving information from high-skill parents, each

[23]This assumption can be relaxed by allowing the prior to be heterogeneous along the parent's skill-type, with the prior mean and/or variance of a child born to a low-skill parent, being different than the ones of children in high-skill households. Section 1.3.6 discusses the implications of relaxing this assumption for the model's main prediction

child just updates her prior beliefs using Bayes' rule. The normality assumption about the prior and the signal implies that the posterior belief about w_i^H is also normally distributed with mean $\hat{\mu}_j$ and variance $\hat{\sigma}_j^2$ given by

$$\hat{\mu}_j = \underbrace{\frac{\sigma_j^2}{\tilde{\sigma}^2 + \sigma_j^2}}_{\text{weight on prior}} \tilde{\mu} + \underbrace{\frac{\tilde{\sigma}^2}{\tilde{\sigma}^2 + \sigma_j^2}}_{\text{weight on signal}} w_j^H, \tag{1.16}$$

$$\hat{\sigma}_j^2 = \left(\tilde{\sigma}^{-2} + \sigma_j^{-2}\right)^{-1} \tag{1.17}$$

where σ_j^2, the signal's variance, is equal to $\sigma_{\epsilon^H}^2 / m_{jH}$. The Bayesian estimator of the high-skill wage is an uncertainty-weighted average of the initial belief and the new information given by the public signal w_j^H. Uncertainty about the high-skill wage, defined as the variance of the children beliefs about w_i^H, does not depend on the realization of the public signal but on the fraction of high-skill neighbors m_{jH}, the prior's variance $\tilde{\sigma}^2$, and wage dispersion $\tilde{\sigma}_{\epsilon^H}^2$. From Equations (1.16) and (1.17), I establish the following:

Lemma 1.1 *Uncertainty about w_i^H $\hat{\sigma}_j^2$ decreases in the fraction of high-skill neighbors in the neighborhood m_{jH} but increases with prior uncertainty $\tilde{\sigma}^2$ and wage dispersion $\tilde{\sigma}_{\epsilon^H}^2$.*

Lemma 1.2 *When making their estimates about w_i^H, children living in neighborhoods with a higher fraction of high-skill neighbors m_{jH}, put relatively more weight on the public signal w_j^H.*

Note that because children share a common prior and information is neighborhood-specific, beliefs about w_i^H are common across children living in the same neighborhood. Nevertheless, they may differ across neighborhoods depending on the allocation of high-skill parents across locations. The fraction of high-skill parents in neighborhood j, m_{jH}, plays two roles. On the one hand, it determines uncertainty associated with the returns to educational investment. On the other hand, it determines the weight children put on the public signal: as the fraction of high-skill parents increases, the weight on the prior decreases relative to the weight on the public signal. This implies that those children who have more labor market information, meaning that live in a neighborhood with a higher fraction of high-skill parents, update their beliefs in response to signals to a greater extent than those that have less information: $\Delta\mu_j \equiv \hat{\mu}_j - \tilde{\mu} = \frac{\tilde{\sigma}_j^2}{\tilde{\sigma}^2+\sigma_j^2}(w_j^H - \tilde{\mu})$ increases with m_{jH} (this follows from Equation (1.16)).

Educational choice Given the cost of skill acquisition, $c(a_i, m_{jH}, q_j)$, and beliefs about w_i^H, a child chooses either to invest or not in education. Let $\mathcal{I}_j$ be the information set of any child born to a family living in neighborhood j, $\mathcal{I}_j = \{w_j^H\}$. The optimal policy of a child i born in neighborhood j with innate ability a_i is to invest in education if and only the cost of doing so is lower than the perceived skill premium, conditional on the information set:

$$V(w^L, \hat{\mu_j}, \hat{\sigma_j}^2, a_i) = \max\{V_j^L(w^L), V_j^H(\hat{\mu_j}, \hat{\sigma}_j^2) - c(a_i, q_j, m_{jH})\}, \quad (1.18)$$

where $V_j^H(\hat{\mu}, \hat{\sigma}^2)$ is the perceived value of investing in education for a child born in neighborhood j,

$$V_j^H = \sum_{j' \in \mathcal{J}} \mathbb{E}_{w_i^H}[U(c_{i,j}^H, \Phi_{j'}^H)|\mathcal{I}_j]\rho_{j'}^H \quad (1.19)$$

with $\mathbb{E}_{w_i^H}[U(c_{i,j}^H, \Phi_{j'}^H)|\mathcal{I}_j]$ being the expected utility of being high-skilled and living in location j', and $V_j^L(w^L)$ is the expected value of being a low-skill worker for a child living in neighborhood j (because V_j^L is equal across neighborhoods I will drop the subscript j from now on),

$$V^L = \sum_{j' \in \mathcal{J}} U(c_{i,j}^L, \Phi_{j'}^L)\rho_{j'}^L \quad (1.20)$$

where $\Phi_j^H = q_j \cdot A_{jH}$, $\Phi_j^L = q_j \cdot A_{jL}$ $\mathbb{E}$ is the expectations operator and the expectation is taken over the high-skill wage. $\rho_{j'}^H$ and $\rho_{j'}^L$ are the probability of living in neighborhood j' conditional on being a high-skill worker and the probability of living in neighborhood j' conditional on being an low-skill worker, respectively. I assume children are myopic in the sense that they not consider that their education decision will determine populations and rental prices, so when computing the expected skill premium, they consider that they will pay the same rent as their parents. Note that in this setting I completely abstract from credit constraints. I do it so not because I do not think they might be important for the decision to invest in education, but so I can start with the simplest model possible that allows me to isolate the the implications of the local information transmission channel in human capital formation. This choice is also supported by empirical evidence found by Carneiro and Heckman (2002), who found that credit constraints do not play a significant role in post-secondary education.

Since the skill acquisition cost is decreasing in ability, the child's optimal investment decision takes the form of a cut-off rule $a_j^*(w^L, \tilde{\mu}_j, \tilde{\sigma}^2, m_{jH}, q_j)$ such that a child only invests in education if $a_i \geq a_j^*$. This threshold is defined by the following indifference condition

$$V_j^H(\hat{\mu}_j, \hat{\sigma}_j^2) - V^L(w^L) = c(a_j^*). \tag{1.21}$$

Given this threshold, for a child i born to a household living in neighborhood j, the probability of investing in education is then given by

$$s_{i,j} = 1 - \mathcal{G}(a_j^*). \tag{1.22}$$

Note that $s_{i,j}$ does not depend on the parents' type but only on the optimal threshold, a_j^*, which is equal across all children living in neighborhood j. Hence, the decision to invest in human capital is not linked to the parents' educational attainment directly, but it is rather linked to the skill-mix of the neighborhood: all children living in same neighborhood, with an ability level higher than a_j^* invest in education, independently of their parents' type. This result lies on the fact that the driver for the investment decision is the child's information endowment, which is common across children living within the boundaries of a neighborhood. Given this, the fraction of children investing in education in neighborhood j is

$$s_j = \frac{\sum_{i=1}^{M_j} s_{i,j}}{M_j} = 1 - \mathcal{G}(a_j^*) \tag{1.23}$$

1.3.4 Equilibrium

Given M_H, M_L, the distribution of high-skill and low-skill wages, the distribution of ability a_i, a vector of school quality $\mathbf{q} = \{q_1, ...q_J\}$, and the vector of neighborhood amenities $\mathbf{A} = \{A_1, ...A_J\}$, the equilibrium is defined by an allocation of M_H and M_L over J neighborhoods with an associated vector of housing rental prices $\mathbf{r} = \{r_1...r_J\}$, a vector of cutoff rules $\mathbf{a}^*$=$\{a_1^*, ..., a_J^*\}$, a vector of high-skill wage estimates $\hat{\mu} = \{\hat{\mu}_1, ..., \hat{\mu}_J\}$ and uncertainty $\hat{\sigma}^2$=$\{\hat{\sigma}_1^2, ..., \hat{\sigma}_J^2\}$, value functions $V_j(w^L, \hat{\mu}_j, \hat{\sigma}_j^2, a_i)$, $V_j^L(w^L)$, $V_j^H(\hat{\mu}, \hat{\sigma}^2)$ in each neighborhood j, and a vector with the fraction of children investing in education in each location $\mathbf{s} = \{s_1, ..., s_J\}$ such that:

1. Parents choose a location j within the city boundaries to maximize utility in Equation (1.4) subject to the budget constraint,

2. For each neighborhood j, the value function $V_j(w^L, \hat{\mu}_j, \hat{\sigma}_j^2, a_i)$ solves Equation (1.18), yielding the cutoff rule in a_j^*,

3. Housing market clears in each neighborhood.

Because there are no agglomeration forces (amenities are exogenous), the dispersion forces of the model—inelastic supply of land and taste heterogeneity—ensure the existence of a unique set of rents that clears the housing market, as shown in Bayer et al. (2004). In this environment, the distribution of amenities and school quality across neighborhoods determines the skill-mix of each neighborhood and whether low-skill households live more or less isolated from the high-skill ones. In turn, the spatial allocation of families determines children's inference about the skill premium and, therefore, the optimal decision regarding the investment in education.

1.3.5 Comparative Statics

Taking the expectations over the unknown wage, w_i^H, the perceived skill premium for a child born in neighborhood j, $\triangle V_j \equiv V_j^H(\hat{\mu}_j, \hat{\sigma}_j^2) - V_j^L(w^L)$, conditional on the information set, is[24]

$$\triangle V_j = J\left(\rho^H_{j'}\frac{-\exp(-\gamma\big(\hat{\mu}_j - \gamma(\hat{\sigma}^2_j/2)\big))}{\sum_{j'\in\mathcal{J}}\frac{\Phi^H_{j'}}{\exp(\gamma r_{j'})}} - \frac{-\exp(-\gamma w^L)}{\sum_{j'\in\mathcal{J}}\frac{\Phi^L_{j'}}{\exp(\gamma r_{j'})}}\right). \tag{1.24}$$

The key variable that drives the optimal investment threshold a^*_j and, as a consequence, the optimal investment decision is beliefs about w^H_i. Combining Equations (1.21) and (1.22), I begin by establishing two intuitive properties of the optimal investment decision. All proofs are provided in the Appendix 1.6.2.

Lemma 1.3 *The ability threshold a^*_j is strictly decreasing in $\hat{\mu}_j$ and strictly increasing in $\hat{\sigma}^2_j$. Hence, the probability of investing in education s_j is strictly increasing in $\hat{\mu}_j$ and strictly decreasing in $\hat{\sigma}^2_j$.*

First, a higher expected value of the high-skill wage $\hat{\mu}_j$ increases the probability that a child invests in education, holding all else equal. Increasing the expected value of the high-skill wage $(\hat{\mu}_j)$ increases the perceived skill premium (Equation (1.B13)), decreasing a^*_j and, therefore, increasing the fraction of children from neighborhood j that invest in education. Second, greater uncertainty about the high-skill wage $(\hat{\sigma}^2_j)$ translates into a lower perception of the skill premium (Equation (1.B13)) and, as thus into a lower probability of investing in education, holding all else equal. More uncertainty makes educational investment more risky. Because I consider individuals to be risk-averse, as uncertainty increases, the ability threshold increases and the share of children investing in education decreases. Higher levels of risk aversion amplify this effect.

[24]Note that in contrast to the literature (Bowles et al., 2014; Kim and Loury, 2013), the benefit of human capital investment is not merely the expected wage gap. Instead, in this framework, the benefit of investing in education takes into account differences in amenities and rental prices payed across different skill groups. This is important because high-skill workers tend to live in places with higher housing costs which may offset some of the consumption benefits from higher wages, but they also tend to enjoy better amenities which may compensate for higher housing costs possibly increasing their well-being. The importance of these differences is highlighted by Diamond (2016), who finds that from 1980 to 2000 changes in cities' rents and amenities increased welfare inequality between college and high-school graduates by more than the increase suggested by the wage gap alone. For details, see Appendix 1.6.2.

High-skill Neighbors (m_{jH}) Spatial location matters for the decision to invest in education because high-skill neighbors determine the cost of skill acquisition but also because, in this environment, they shape children's perception about the skill premium through their estimate of the high-skill wage $\hat{\mu}_j$ and its uncertainty $\hat{\sigma}_j^2$. Regarding uncertainty, the role played by high-skill neighbors is straightforward. A higher share of highly educated neighbors living in a given location j means that children born to that location observe more a precise signal (σ_j^2 is lower), holding all else equal. As a consequence, they are less uncertain about the high-skill wage. This result follows from the filtering problem (Equation (1.17)), and is established in Lemma 1.1. Panel A in Figure 1.2 illustrates this effect. Because children are risk averse, lower uncertainty associated with human capital investment translates into a widening of the mass of children that invest in education (a_j^* decreases, thus s_j^* increases), as established in Lemma 1.3.

The effect of high-skill neighbors on $\hat{\mu}_j$ is, however, ambiguous. As a reaction to a more precise signal, when estimating w_i^H using Baye's rule (Equation (1.16)), children place a higher weight on the labor market information disclosed by their neighbors (the public signal w_j^H), as formalized in Lemma 1.2. This implies that those children with more information, i.e. those that live in a neighborhood with a larger fraction of high-skill neighbors, update their beliefs in response to new information to a greater extent than those that have less information: $\Delta\mu_j \equiv \hat{\mu}_j - \tilde{\mu}$ increases with m_{jH}. However, having more information does not necessarily translate into a higher perception about the skill premium: $\Delta\mu_j$ may be positive or negative depending on the size of the signal relative to the prior (Equation (1.16)). If the signal is sufficiently low, living in a neighborhood with a high fraction of high-skill neighbors translates into a lower $\hat{\mu}_j$ than the one in locations with a low share of high-skill neighbors. On the other hand, if the signal is sufficiently high, children from neighborhoods with a larger share of high-skill neighbors will have a higher $\hat{\mu}_j$. As shown in Lemma 1.3, a lower/higher $\hat{\mu}_j$ translates into a lower/higher fraction of children investing in education. Panel B in Figure 1.2 plots the posterior mean$\hat{\mu}_j$ when the public signal w_j^H is sufficiently high and low. First, the magnitude of the posterior mean change increases in the share of high-skill neighbors. However, while at high values of w_j^H, the estimate is higher in the neighborhood with a higher human capital level, the opposite is true when the signal's magnitude is small. In this case, being exposed to high-skill neighbors translates into a lower perception of the skill premium.

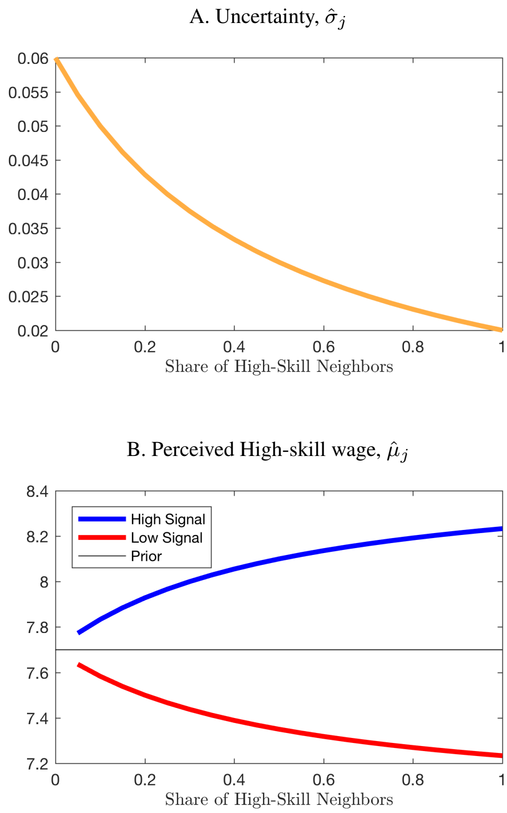
The total effect of high-skill neighbors on the share of children investing in education depends then on the size of the signal. If the signal is sufficiently high such that $\Delta\mu_j > 0$, a higher of high-skill neighbors increases the perceived skill premium increases (the estimate of w_i^H is higher and its uncertainty is lower). Thus, the share of children that decide to invest in education increases. In contrast, if the signal is sufficiently low and $\Delta\mu_j < 0$, there are two opposing forces on the perceived skill premium. High-skill neighbors decrease uncertainty and the cost of skill-acquisition, but they also decrease children's estimate about the high-skill wage. Whether the share of children investing in education increases or decreases depends on which effect dominates. This, in turn, depends on the size of the signal relative to a threshold w^*.

Overall, more information about w_j^H (living in a location with a high m_{jH}) increases the share of children investing in education if and only if $w_j^H > w^*$. Under this condition, perceived skill premium is increasing in the share of high-skill neighbors. Otherwise, if $w_j^H < w^*$, the expected value of w_j^H decreases in the fraction of highly educated parents in the neighborhood, and this effect dominates the fact that uncertainty is lower, reducing the probability of investing in education even though the exposure to high-skill neighbors is higher. Proposition 1.1 formalizes this result.

Proposition 1.1 *Given a spatial allocation of M_H and M_L over J neighborhoods in the city, locations with a higher fraction of high-skill parents, m_{jH}, have a higher fraction of children investing in education s_j if and only if $w_j^H > w^*$.*

Wage dispersion ($\sigma^2_{\epsilon^H}$) Information transmission from high-skill neighbors as a channel through which children learn about the high-skill wage w_i^H depends on its dispersion $\sigma^2_{\epsilon^H}$. For a given spatial equilibrium, the higher is $\sigma^2_{\epsilon^H}$, the lower is the change in the estimate of w_i^H and its uncertainty upon arrival of new information. Therefore, the lower is the potential to learn from high-skill neighbors. This follows from the fact that as $\sigma^2_{\epsilon^H}$ increases, the public signal becomes less precise

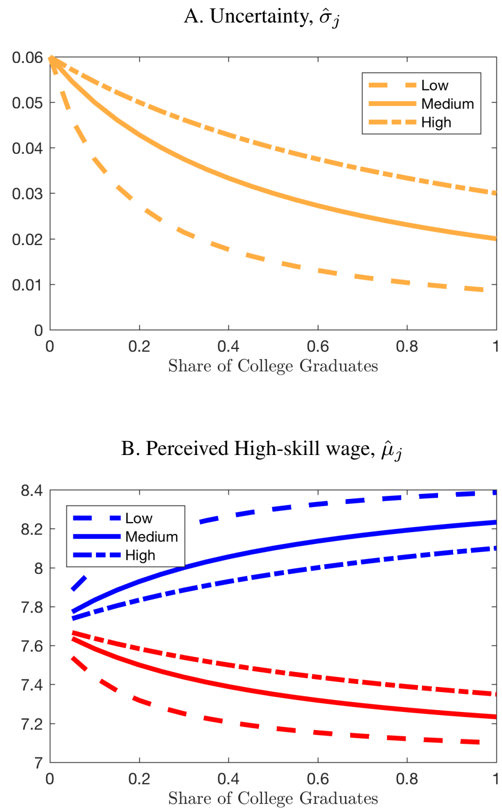
Figure 1.2: Posterior Mean and Uncertainty: Different Neighborhoods Skill-Mix



Notes: Panel A plots the posterior sigma $\hat{\sigma}_j^2$ for different skill compositions of neighborhoods. Panel B the posterior mean $\hat{\mu}_j$ for different skill compositions of neighborhoods and different signals. To compute $\hat{\mu}_j$ and $\hat{\sigma}_j^2$, I use the following parameter values $\tilde{\mu} = 7.6$, $\tilde{\sigma}^2 = 0.06$, $\sigma^2_{\epsilon^H} = 0.03$, the high signal is equal to 8.5 and the low signal is equal to 7. Expect for the signals, these values correspond to the values used in the calibration in Section 1.4.

as shown in Equation (1.15). Figure 1.3 illustrates this effect by plotting the posterior uncertainty (Panel A) and mean (Panel B) across different values of $\sigma^2_{\epsilon^H}$. For the same skill-mix level, the magnitude's change of both $\hat{\mu}_j$ and $\hat{\sigma^2}_j$ is lower for higher values of $\sigma^2_{\epsilon^H}$. The overall effect of wage dispersion on the share of children investing in education depends also on the size of the signal. If the signal is sufficiently high, such that $\Delta\mu_j > 0$, the share of children that decide to invest in education is decreasing in wage dispersion because the perceived skill premium decreases (the estimate of w_i^H is lower and its uncertainty is higher). In contrast, if the signal is sufficiently low, such that $\Delta\mu_j < 0$, there are two opposing effects. Wage dispersion increases uncertainty, but it also increases the estimate. So, the total effect depends on which effect dominates.

Figure 1.3: Posterior Mean and Uncertainty: The Role of Wage Dispersion



Notes: Panel A plots the posterior sigma $\hat{\sigma}_j^2$ for different skill compositions of neighborhoods and three different levels os wage dispersion ($\sigma^2_{\epsilon^H}$) - low, medium and high. Panel B plots the posterior mean $\hat{\mu}_j$ for different skill compositions of neighborhoods and three different levels of wage dispersion ($\sigma^2_{\epsilon^H}$) - low, medium and high. The red and yellow lines correspond to a scenario where the signal is low and high, respectively. To compute $\hat{\mu}_j$ and $\hat{\sigma}_j^2$, I use the following parameter values $\tilde{\mu} = 9$, $\tilde{\sigma}^2 = 0.06$, $\tilde{\mu} = 7.6$, $\sigma^2_{\epsilon^H} = 0.03$, the high signal is equal to 8.5 and the low signal is equal to 7. Expect for the signals, these values correspond to the values used in the calibration in Section 1.4.

School quality (q_j) Holding all else equal, higher values of school quality translate into a lower cost of investing in human capital is lower, hence the probability of investing in education increases.

Low-skill wage (w^L) The wage of low-skill workers also plays a standard role: for lower values of the low-skill wage, the perceived skill premium is higher, hence the probability of investing in education increases, holding all else equal.

All in all, the skill-mix of neighborhoods and the education decision of children are connected through an information channel. The configuration of the city, namely, the distribution of high-skill parents across neighborhoods shapes the public signal w_j^H children observe. Local information diffusion creates inequalities between neighborhoods as their skill-mix generates different perceptions about the skill premium. Under information frictions and social learning, the effect of local interactions in the education decision is not only about being more exposed to high-skill neighbors, as suggested by previous literature, it is also about the labor market information they disclose. More exposure implies more information, but more information does not necessarily increase the probability of investing in education, this will depend on the information that children observe, namely, the magnitude of the public signal about the high-skill wage. This result is consistent with the empirical evidence presented in the previous section.

1.3.6 Discussion of the Model's Assumptions

I make several assumptions that make the model more tractable without affecting its main qualitative result. In this section, I discuss the implications of each assumption for the model's results.

Exogenous amenities I consider that neighborhood's amenities are taken to be exogenous. However, places that attract a higher share of skilled workers may endogenously become more desirable places to live in (see, for instance Diamond, 2016). In line with this, one could consider that neighborhood amenities have two distinct parts: *(i)* an exogenous component that is invariant to the skill-mix of the neighborhood such as the geographic characteristics, and *(ii)* an endogenous component that depends on the share of high-skill workers in the neighborhood.[25] One could re-define $\Phi_{k,j}$ in Equation (1.8) as $\Phi_{k,j} = q_j A_{1,j}^{\beta_k} A_{2,j}^{1-\beta_k}$, with $A_{1,j} = m_{jH}$ being a location attribute that I allow to endogenously respond to the types of families living in the neighborhood, namely the share of high-skill parents. Allowing for endogenous amenities affects the spatial allocation of households across neighborhoods within the city without affecting the role of high-skill neighbors in the decision to invest in education described in Proposition 1.1. Note that the introduction of these agglomeration forces generates the potential for multiple equilibria in the model, if these agglomeration forces are sufficiently strong relative to the exogenous differences in characteristics across locations. However, within each equilibrium the main prediction of the model holds.

Uncertainty about low-skill wage If $\sigma^2_{\epsilon^L} > 0$, children are both uncertain about the high-skill wage and the low-skill wage. In this case, a higher fraction of high-skill neighbors yields more information about the high-skill wage, but less information about the low-skill wage. This amplifies differences in the perceived skill premium across neighborhoods and, therefore, in the share of children investing in education. Appendix 1.6.2 shows that under this scenario there is also a signal threshold w^* below which a higher fraction of high-skill neighbors decreases the fraction of children investing in education. However, in this setting, the magnitude of this threshold also depends on the magnitude of the signal children receive about the low-skill wage.

Common prior In Section 1.3.3, I assume that children share a common prior about w_i^H and update this prior using information at the neighborhood level. This implies that the probability of investing in education, defined in Equation (1.22), is independent of the parent's type. This assumption can be relaxed by allowing the prior to be heterogeneous along the parent's type, with the prior mean and/or variance of a child born to a low-skill parent, being different than the ones of children in high-skill households. If children priors depend on their parent's type, for

[25] A growing literature has considered how amenities change in response the composition of an location residents: Bayer et al. (2007), Card et al. (2008), Guerrieri et al. (2013) and Diamond (2016).

a neighborhood j, there are two ability thresholds that determine the probability of investing in education for children born to high and low-skill families, s_j^H and s_j^L, respectively. The thresholds $a^*_{H,j}$ and $a^*_{L,j}$ are defined by the indifference condition, $V^H_{k,j}(\hat{\mu}^k_j, \hat{\sigma}^2_{k,j}) - V^L(w_u) = c(a^*_{k,j})$, where $V^H_{k,j}(\hat{\mu}^k_j, \hat{\sigma}^2_{k,j})$ is the perceived value of being a high-skill worker for a child born to a k-type household living in neighborhood j. Given this, the probability of investing in education for a child i born to a household of type k living in neighborhood j is $s^k_{i,j} = 1 - \mathcal{G}(a^*_{k,j})$, and the fraction of children investing in education in neighborhood j is

$$s_j = \frac{\sum_{i=1}^{M_j^H} s^H_{i,j} + \sum_{i=1}^{M_j^L} s^L_{i,j}}{M_j} \tag{1.25}$$

s_j increases in the fraction of high-skill neighbors m_{jH} if $\frac{\partial s^H_{i,j}}{\partial m_{jH}} + \frac{\partial s^L_{i,j}}{\partial m_{jH}} > 0$. Whether $\frac{\partial s^H_{i,j}}{\partial m_{jH}}$ and $\frac{\partial s^L_{i,j}}{\partial m_{jH}}$ are greater or lower than zero depends, respectively, on the magnitude of the signal w^H_j relative to the threshold $w_j^{*,H}$ and $w_j^{*,L}$, as stated in Proposition 1.1. In Section 1.4, I relax this assumption and show that my quantitative results remain similar once I allow for different priors.

Correlated human capital across generations The importance of the parents' human capital as an input in the formation of the human capital of the child has been extensively explored theoretically as well as empirically. One can introduce such a feature by allowing the level of human capital of the child to depend on the level of human capital of its parent $h_i = a_i^{\varphi} \cdot h^{\eta} \cdot q_j^{\kappa} \cdot m^{\rho}_{H,j}$, with h being the parent human capital level. Under this specification, the parent affects the child directly: for a given level of innate ability, children born to parents with higher levels of human capital will have a higher level of human capital as well. Importantly, the main prediction of the model is robust to this specification and the threshold level w^* above which the relationship between the share of high-skill neighbors and children investing in education remains unchanged.

Risk-aversion Assuming individuals have CARA utility function over consumption with risk-aversion parameter γ is not crucial for the model's prediction regarding the role of high-skill neighbors in education decisions, as described in

Proposition 1.1. Appendix 1.6.2 shows that if instead individuals are risk neutral with a linear utility function in consumption and amenities, there is also a threshold w^* below which the relationship between the share of high-skill neighbors and children investing in education is negative – albeit higher than the one in Proposition 1.1 due to the fact that now individuals do not dislike uncertainty. Hence, under risk neutrality, the magnitude of the signal has to be higher in order to trigger a positive relationship between the probability of investing in education and the share of high-skill neighbors. This is due to the fact that under risk neutrality a_j^* depends only on the posterior mean $\hat{\mu}_j$ but not on the posterior variance $\hat{\sigma}_j^2$.

To sum up, I have shown that if one takes into account information frictions and local learning, the relationship between the share of high-skill neighbors on the fraction of children investing in education may be negative depending on the earnings of high-skill neighbors, consistent with the empirical evidence presented in Section 1.2. In contrast with the existing literature, in this model, more exposure to high-skill neighbors brings more information, but additional does not necessarily translate into more investment in education. This depends on the labor market information disclosed by highly educated neighbors. In the next section, I estimate the model and assess the quantitative importance of information frictions and local learning as a channel through which neighborhoods affect the decision to enroll in college.

1.4 The Importance of Local Learning

Even though imperfect information paired with local learning can reconcile the documented pattern in Section 1.2, it remains an open question whether this novel mechanism is quantitatively important. To tackle this issue, I calibrate the model to match 2013 data regarding the wage distribution by educational attainment, the distribution of individuals and college enrollment rates across school districts in the city of Detroit. I choose Detroit because it is the largest city in Michigan, with 95 school districts in 2013.

Armed with the calibrated economy, I ask three different questions. First, I ask by how much would college enrollment change if children did not observe any information from high-skill neighbors. Second, which neighborhood channel is more important in explaining differences in college enrollment across school districts? Third, can a disclosure policy that corrects children's perceptions about the skill premium equalize opportunities? It should be noted that a more realistic analysis would nest the learning mechanism within a richer framework. As this is the first study about the contribution of the local information constraints to the accumulation of human capital and, therefore, in persistent inequality, assessing its quantitative potential in a simple model that allows both the theory and the calibration to be fairly transparent is an important first step to subsequently developing more complicated quantitative models.

1.4.1 Definition of Variables in the Model

City A city in the model corresponds to a metropolitan statistical area (MSA) that is a region consisting of a group of counties that have a high degree of economic and social integration with the core county as measured through commuting.[26]

Neighborhoods I define a neighborhood in the model to be a school district. The most commonly definition of a neighborhood is a census tract, a "small, relatively permanent statistical subdivisions of a county", which have generally a size between 1200 and 8000 people.[27] school districts tend to be relatively larger. I pick school districts over census tracts, because school districts are the smallest unit of analysis for which I observe both college enrollment by high-school graduates and socioeconomic characteristiscs of the location such as the % of college graduates, median family income, among others. I use the Geographic Correspondence Engine with Census 2010 from the Missouri Census Data Center to link school districts to MSA's.[28]

High and Low-skill Neighbors I use education to proxy for skills as in Acemoglu and Autor (2011) and Diamond (2016), and define "high-skill" neighbors

[26]More information here https://www.census.gov/geo/reference/gtc/gtc_cbsa.html.
[27]More information here https://www.census.gov/geo/reference/gtc/gtc_ct.html.
[28]The linking file can be download here http://mcdc.missouri.edu/websas/geocorr14.html.

as those individuals living in the school district who have at least a 4-year bachelor's degree while "low-skill" neighbors are those who have less years of education than that.

1.4.2 Functional Forms

The parameterization of the model is as follows: The utility function is CARA with risk aversion parameter γ. Innate ability is assumed to be uniformly distributed between $\overline{a}$ and $\underline{a}$. The cost functions is given by $C(a_i) = \overline{c} - \phi(a_i^{\varphi} \cdot q_j^{\kappa} \cdot m_{j_H}^{\rho})$, where a_i is innate ability, q_j is expenditures per student in school district j and m_{j_H} corresponds to the share of high-skill neighbors living in the school district.

1.4.3 Calibration Strategy

Calibration is proceed in two steps. In the first step, I set parameters that either have a direct counterpart in the data or that have been used in previous literature. In the second step, I use the simulated method of moments to estimate the remaining parameters, which are the ones that characterize the cost function.

I set $\overline{a}$ equal to one, $\underline{a}$ to zero, and θ=1. The number of neighborhoods J equals the number of school districts in Detroit in 2013, 95. As in Babcock et al. (1993), I set the risk aversion parameter of the CARA utility function γ to 0.5.

Wages and prior distributions Wage distributions in the model match the empirical distributions of labor income of full-time workers with different skills from the American Community Survey 2008-2013. Full-time workers are defined to be individuals aged between 25 and 55 years working at least 35 hours per week, 48 weeks per year. For the low-skill wage distribution, I normalize the variance to 0 and calibrate the mean w^L to match the mean of the log monthly-wage distribution of low-skill full time workers. For the mean of the high-skill wage distribution w^H, I match the mean of the distribution of the log monthly-wage distribution of high-skill full-time workers. Because I normalize the variance of the low-skill wage to 0, I set the variance of the high-skill wage distribution equal to the difference between the variance of the labor income distribution of high-skill full-time workers and the variance of the labor income distribution of low-skill full-time workers. I set the mean and variance prior ($\tilde{\mu}$ and $\tilde{\sigma}^2$) such that the average of the perceived skill-premium after observing the signals from the neighbors matches the one in Bleemer and Zafar (2016): 1.63.

Amenities I recover the distribution of A_{jH} and A_{jL} across the school districts from the data. From NCES-EDGE, I observe for each school district: total population M_j, the number of high-skill and low-skill individuals, M_{jH} and M_{jL}, expenditures per student q_j and rents r_j. Following Diamond (2016), as a measure of rents, I use the median gross rent at each school district, which includes both the housing rent and the cost of utilities.[29] Assuming that the current allocation of individuals across school districts is in equilibrium, for any two neighborhoods j and j', the following holds

$$\frac{M_j^k}{M_{j'}^k} = \frac{\Phi_j^k}{\exp(\gamma r_j)} \frac{\exp(\gamma r_{j'})}{\Phi_{j'}^k} \tag{1.26}$$

where M_j^k is the number of type k-individuals that live in j. For high-skilled individuals, $\Phi_j^H = q_j A_{jH}$, for low-skill individuals, $\Phi_j^L = q_j A_{jH}$. I set both Φ_j^L and Φ_j^H equal to one for Detroit City school district, and then back out the level of A_{jH} and A_{jL} for the other school districts using Equation (1.26).

Cost function The parameters without observable counterparts are the cost function parameters, $\bar{c}$, φ, ϕ, κ and ρ. I estimate them using the simulated of method of moments, which picks the parameter vector θ=($\bar{c}$, φ, ϕ, κ, ρ) that minimizes the weighted sum of square deviations between data moments and their

[29]Ideally, I would like to have school district specific rent indices controlling for differences in the quality of housing across school districts following the hedonic-regression approach by Eeckhout et al. (2014). However, because I cannot link individuals in the ACS to the school districts where they live, this is not possible, thus I use the reported median gross rent in NCES-EDGE.

model-generated counterpart:

$$\hat{\theta} = \text{arg min } (y(\theta) - y^*)'\mathbf{W}(y(\theta) - y^*)$$

where **W** is the identity matrix, implying that each moment is equally weighted, y^* is a $t \times 1$ vector of moments observed in the data and $y(\theta)^*$ is a $t \times 1$ of those moments from the model evaluated at a given parameter vector θ. In the estimation, I match data moments of the distribution of college enrollment in 2013 (mean, standard deviation and p75-p50 ratio), the correlation between college enrollment and college graduates and the correlation between college enrollment and expenditures per student. An advantage of estimating the model is the understanding of what features of the data identify each parameter. The mean of college enrollment across districts identifies $\bar{c}$. The school district variation in enrollment identifies φ, while p75-p50 ratio identifies ϕ. Finally, the correlation of enrollment with the share of college graduates and expenditures per student identify ρ and κ, respectively. Table 1.A6 summarizes all parameters.

1.4.4 Model Fit

This section discusses the calibrated economy. Panel A in Table 1.2 compares the empirical targets for the calibrated parameters and the corresponding moments produced by the model. Panel A in Figure 1.4 depicts the predicted and observed values of the college enrollment rate for the 95 school districts within Detroit, and Panel B plots the enrollment distribution in the model and the one observed in the data. The calibrated model reproduces reasonably well the five targeted moments, and the distribution of enrollment is highly correlated with the one observed in the data. The model can also be used to derive enrollment rates in out-of-sample years. I assume that the prior distribution is constant across years, and construct wage and amenities distributions for each year. As shown in Panel B of Table 1.2, the correlation between fitted and observed values of college enrollment for out-of-sample years is high. These results show that the model successfully captures patterns in the data.

In the calibrated economy, the average high-skill family lives in a school district where 24.4% of its population is high-skill, while the average low-skill family lives in a location where the proportion of high-skill is 8 percentage points lower. Differences in the skill-composition of locations as well as differences in school resources translate into differences in the subsequence education decisions of children. The probability of becoming a high-skill worker for a child born to a low-skill household is 8 percentage points lower than the probability of becoming a high-skill worker for a child from a high-skill family. Next, I quantify the role of the novel mechanism proposed in this chapter, local learning, and then I assess which channel matters the most for differences in enrollment across neighborhoods.

Table 1.2: Model Fit

A. Targeted Moments

	Mean	Std. Dev.	p75/p50	Corr. w/ $m_{j,H}$	Corr. w/ q_j
Data	0.38	0.13	1.28	0.84	0.12
Model	0.38	0.10	1.22	0.95	0.15

B. Out-of-sample Years

	2009	2010	2011	2012	2014
Correlation	0.50	0.74	0.83	0.80	0.80

Notes: The table in Panel A reports targeted moments in the estimation. *Corr w/* m_j^H corresponds to the correlation between the share of college graduates and college enrollment. *Corr w/* q_j corresponds to the correlation between expenditures per student and college enrollment. The table in Panel B reports the correlation between fitted and observed enrollment rate across school districts in out-of-sample years. Observations are at school district level. The sample is composed by 95 school districts within Detroit in the year 2013.

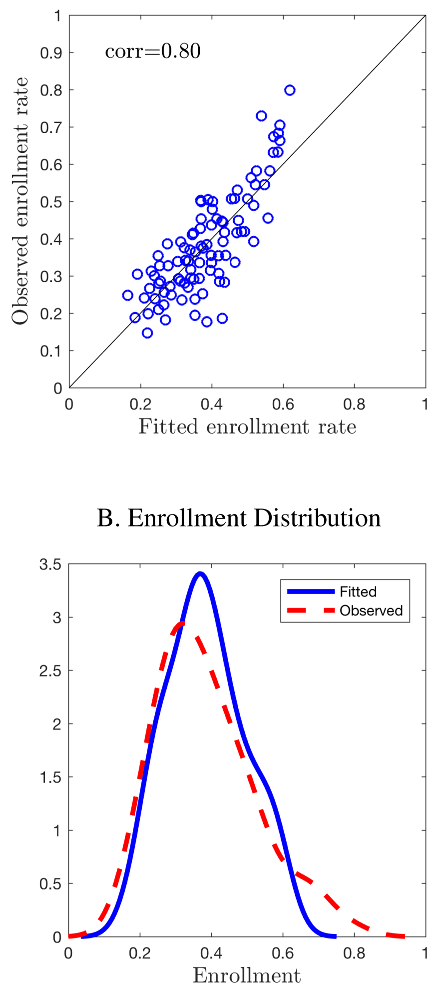
1.4.5 Quantifying Local Learning

Armed with the calibrated economy, I ask the following question: *by how much would college enrollment rate change in the absence of the public signal from high-skill neighbors?* To answer this question, I simulate what would happen if individuals did not update their initial beliefs ($\hat{\mu}_j = \tilde{\mu}$ and $\hat{\sigma}_j^2 = \tilde{\sigma}^2$). Panel A and B in Figure 1.5 plot, respectively, the perceived skill premium and college

Figure 1.4: Model vs. Data

A. Enrollment Rate

B. Enrollment Distribution



Notes: Panel A plots fitted and observed values for the college enrollment rate across school districts. Fitted values are on the horizontal axis; observed values are on the vertical axis. Correlation between fitted and observed values is equal to 0.8. Observations are at school district level. Panel B plots the enrollment distribution simulated in the model and observed in the data. The sample is composed by the 95 school districts within Detroit in 2013.

enrollment in the baseline model (with the learning mechanism, and thus matching the data) versus the no-learning counterfactual across school districts.

I find that high-skill neighbors play an important role in correcting initial beliefs. Before observing any information, children hold beliefs about the high-skill wage that are downward biased ($\tilde{\mu} < w^H$) and more uncertain ($\tilde{\sigma}^2 > \sigma_\epsilon^2$). By observing the public signal w_j^H, children's estimate of the high-skill wage increases by 2.3% and its uncertainty decreases by 21%, on average. As a consequence, the perceived skill premium rises 6.7%, on average (Panel B in Figure 1.5). This has a significant effect on enrollment as shown in the right panel of Figure 1.5. I find that if individuals did not observe any public signal from high-skill neighbors, the college enrollment rate across school districts at the would be 17 percentage points lower, on average. This means that instead of having 38% of high-school graduates enrolling in college within 6 months of graduation in Detroit, only 21% would.

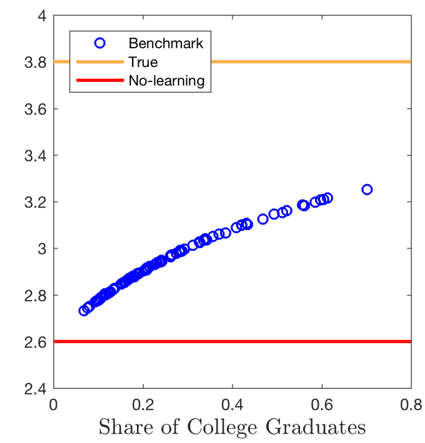
Skill persistence A high-skill family lives in a neighborhood with a share of college graduates that is 8 percentage points higher, on average, than the average neighborhood where low-skill families live. This difference translates into differences in the average perceived skill premium, which in turn translate into different probabilities of investing in education. Namely, a child that is born to a high-skill family has a probability of becoming a high-skill worker that is 22% higher when compared to a child that is born to a low-skill family. By shutting down local learning, I find that differences in perceptions are responsible for 60% of the gap between children from high-skill families and those from low-skill families.

1.4.6 Decomposition: Which channel is more important?

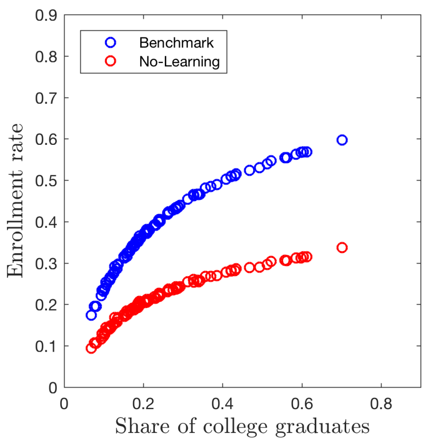
In the calibrated economy, differences in enrollment across neighborhoods arise from three different channels: *(i)* information externalities: school districts that have a higher share of college graduates generate more information about schooling returns; *(ii)* human capital spillovers in the cost function, and *(iii)* expenditures per student. Given this, a natural extension of the main counterfactual exercise is to ask which channel is more important in explaining differences in college enrollment across locations. One way to conduct this decomposition is to start from a counterfactual with information externalities only, and then activate one of the other two channels at a time, by setting school resources and human capital

Figure 1.5: Benchmark vs. No-Learning Counterfactual

A. Perceived Skill Premium



B. Enrollment Rate



Notes: Panel A plots the perceived skill premium across school districts in the benchmark economy (blue) versus the no-learning counterfactual (red). Panel B plots college enrollment rate across school districts in the full model (blue) versus the no-learning counterfactual (red). The sample is composed by the 95 school districts within Detroit in 2013.

spillovers equal to the average value across school districts.[30] Table 1.3 reports the

[30]Alternatively, one could shut down each channels at the time by setting ρ and κ equal to zero. This procedure, however, produces a level effect. Because the focus in this chapter is to understand what is driving inequalities across locations, I eliminate differences produced by each channel by setting their to the mean, and keep ρ and κ unchanged. The implications of each channel for dispersion of college enrollment are similar under these two approaches.

dispersion of enrollment across school districts and the enrollment gap between a child born to a high-skill family and a child born to a low-skill family in the full benchmark as well as when I turn off each channel at a time. This table suggests that local learning is, by far, the most important channel in explaining enrollment inequality across school districts. This channel accounts for 57% of the dispersion in college enrollment across school districts, and it explains 53% of the difference between the probability of being high-skill for a child born to a high-skill family and a child born to a low-skill family.

Table 1.3: Benchmark Economy vs. Counterfactuals

	Data	Benchmark	$m_{jH} = \bar{m}$ $q_j = \bar{q}$	$m_{jH} \neq \bar{m}$ $q_j = \bar{q}$	$m_{jH} \neq \bar{m}$ $q_j \neq \bar{q}$
Std. Dev. Enrollment	0.13	0.10	0.057	0.10	0.10
Enrollment Gap	0.08	0.072	0.043	0.073	0.072

Notes: The table reports the standard deviation of the distribution of college enrollment across school districts and the enrollment gap, defined as the difference between the average college enrollment rate of children with high- and low-skill parents, under the benchmark economy and four different scenarios: no local learning, equal human capital spillovers ($m_{jH} = \bar{m}$) and equal school resources ($q_j = \bar{q}$). Observations are at school district level. The sample is composed by 95 school districts within Detroit in 2013.

1.4.7 Policy Counterfactuals

Imperfect information paired with local information transmission explains more than half of the differences in college enrollment across locations, and more than half of the enrollment gap between children from different backgrounds. This result points in favor for policies that either correct individuals' perceptions about the skill premium, like the information interventions studied by Hoxby and Turner (2015), Bleemer and Zafar (2016) and Hastings et al. (2017), or that change the location where children grow up as a way to improve outcomes for children of parents with a low level of education. In this section, I examine the effects of implementing such policies by simulating a disclosure policy that informs the students about the high-skill wage distribution, and a reallocation program that moves a fraction of children from low-skill parents into a better location.

Relocation Policy I simulate the implementation of a policy that moves "disadvantageous" children (and their parents) into an "advantageous" location. To do this exercise, I assume that the policy is implemented after parents choose where to locate, and that there are extra housing units in the "advantageous" location to accommodate the moves.[31] The simulated policy targets children of low-skill parents living in a location within the first quartile of college graduates distribution, and moves 25% of these children to locations within the last quartile of the college graduates distribution. Such a policy changes the skill-mix of both locations, therefore it will affect (i) targeted children who are moved, (ii) children who live in the receiving location, and (iii) children who remain in the disadvantageous location.

Panel A in Table 1.4 shows the effects of this policy for children who stayed (stayers) in the "disadvantageous" neighborhood, those that moved (movers) and those living in the "advantageous" location (receivers). Two results stand out. First, the policy has a small effect stayers and receivers. For the former, the probability of becoming a high-skill worker drops 5 percentage points, and for the latter it increases 5 percentage points. Second, for the movers the probability of becoming a high-skill worker increases from 0.25 to 0.49. Panel B in Table 1.4 reports the decomposition of the overall effect for the movers with respect to each of the components that characterize a location in the model: information externalities, school quality and spillovers. I find that 70% of the change in the probability of becoming a high-skill worker is due to the information channel of neighborhoods.

The effect of the relocation policy hinges on the change in the locations' skill-mix, therefore it is important to assess its dependency on the size of the population that moves from one location to the other. Table 1.4 reports the policy counterfactual if the policy moves 5%, 25% or 50% of children living locations within the first quartile of the college graduates distribution. I find that the effect of the

[31] This can be rationalized by the existence of a Government that has land in all locations where it can build public houses. Also note that if amenities were endogenous and depended, for instance, on the share of high-skill neighbors as in Diamond (2016), the effects of this policy would be the same if I assume that parents cannot move after the policy implementation and that they do not anticipate it when choosing where to locate.

reallocation policy on the probability of enrolling in college for movers ranges between 0.20 to 0.28.

Table 1.4: Reallocation Policy

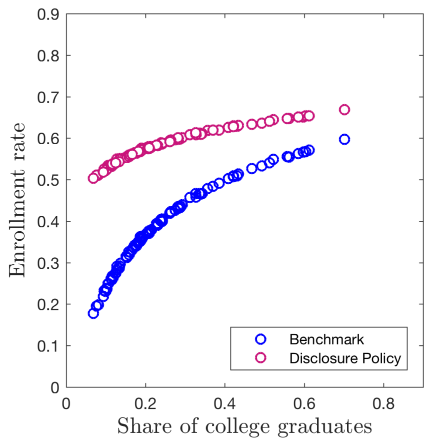
	High-skill neighbors		Enrollment rate		
	1st qtl	4th qtl	Movers	Stayers	Receivers
Panel A: Total Effect					
Benchmark	0.11	0.47	0.25	0.25	0.54
Policy Counterfactual	0.15	0.38	0.49	0.30	0.49
Panel B: Decomposition					
Local learning			0.42		
School quality			0.42		
Spillovers			0.49		

Notes: The table reports the effects for movers, stayers and receivers when a policy that moves 25% of the children living in the 25th percentile of the college graduates distribution to location in the 75th percentile of the college graduates distribution is implemented. *High-skill neighbors* corresponds to the share of high-skill neighbors in both the baseline and the counterfactual

Disclosure Policy To understand the potential of an information campaign, I perform a counterfactual analysis where all children are informed about the $true$ distribution of the high-skill wage: $\omega = w^H + \epsilon^H$, with $\epsilon_i^H \sim \mathcal{N}(0, \sigma^2_{\epsilon^H})$. Figure 1.6 plots college enrollment across school districts under this policy and the benchmark economy, and shows that giving information to children about wage' distribution increases college enrollment substantially in all school districts: 58% of high school graduates would enroll in college, which compares to 38% in the benchmark economy. This result relies on the fact that with the extra signal the perception high-skill wage is increases in 2.8% with respect to the benchmark economy. More important, my results show that by implementing a policy that correct beliefs, while leaving the other sources of inequalities across neighborhoods at work, one can reduce significantly inequalities across locations and between children from different backgrounds: in particular, the enrollment gap between children with low educated parents and those with highly educated parents reduces in 62%.

This policy counterfactual exercise is comparable to the recent information experiment run in Bleemer and Zafar (2016), where a representative sample of US households was informed about the average skill-premium and look to the effect of this intervention in the intention to enroll their children in college. First, they find that non-college graduates update their beliefs to a greater extent than college graduates — as predicted by my model. Second, they find that this intervention increased the intention to enroll their children in college in 5 percentage points. If we believe that intention to enroll in college maps one to one with enrollment rate, then the model estimates are substantially larger. This could be explained by the fact that I do not take into account credit constraints. Given this, my model provides an upper bound estimate of the effect of a policy intervention like the one in Bleemer and Zafar (2016).

Figure 1.6: Disclosure Policy



Notes: The panel plots college enrollment across school districts in a scenario where individuals know the right distribution of high-skill wage ($\hat{\mu}_j = w^H$ and $\hat{\sigma}^2_j = \sigma_{\epsilon^H}$), pink). The sample is composed by the 95 school districts within Detroit in 2013.

1.4.8 Robustness: Different Priors

The results from the previous counterfactual analysis rely on the assumption that children share a common prior regardless of their parents skills. However, it is likely case that growing up with high-skill parents gives children a different perception about the value of education. To assess the implications of this assumption, I relax it by allowing the prior to be different for each type of parents. In particular, I consider that the prior mean of children born to a low-skill parent is lower than the one of children in high-skill households: $\tilde{\mu}^L < \tilde{\mu}^H$, while prior uncertainty remains equal. As before, I discipline these parameters using the distribution of perceived skill premium by educational attainment from the survey conducted by Bleemer and Zafar (2016). This extension of the model improves its fit to the data, namely in explaining dispersion of college enrollment across school districts: it explains 93% of the standard deviation of college enrollment, which compares to 70% in the benchmark economy. Figure 1.A3 shows the fit of the extended model. I simulate the model under all three different scenarios considered previously: (i) no local learning ($\hat{\mu}^L = \tilde{\mu}^L$ and $\hat{\mu}^H = \tilde{\mu}^H$), (ii) no differences in school resources ($q_j = \bar{q}$), and (iii) no differences in human capital spillovers ($m_{j_H} = \bar{m}$). Figure 1.A4 and Table 1.A8 display the results, and show that my findings are robust to different priors depending on parents skills. First, local learning increases the enrollment rate in 23 percentage points, which compares to 22 percentage points in the benchmark model. Second, local learning is, as before, the most important channel in explain differences in college enrollment across school districts: it accounts for 43% of the dispersion in college enrollment across school districts. This magnitude is, however, smaller than the one found previously.

1.5 Conclusion

Why does the place where children grow up shape their opportunities in life? I have proposed a novel explanation featuring imperfect information about the skill premium and local information transmission, i.e. learning about the skill-premium by observing noisy signals of wage realizations of their neighbors. In this environment, spatial location matters because it shapes children's perception about the skill premium. To the best of my knowledge, this mechanism is new in the literature.

I find that imperfect information paired with local learning is able to reconcile novel empirical evidence showing that when earnings of college graduates are sufficiently low, a higher share of college graduates living in a school district is associated with lower college enrollment of students graduating from a high-school in that district. Moreover, I show that it is the most important channel in explaining inequality in college enrollment across school districts ad the enrollment gap. A disclosure policy that is able to correct initial beliefs about the skill premium, while keeping differences in human capital spillovers and school resources across location, has a significant effect in leveling the playing field across children from different backgrounds. These results have important policy implications. In particular, they point in favor of broader information interventions, specially among individuals from lower socio-economic backgrounds, as a tool to address opportunity inequality.

Going forward, it would be interesting to explore the role of local learning interacted with "The Great Divergence" in explaining the geography of upward mobility in the US documented in Chetty et al. (2016). Diamond (2016) shows that, from 1980 to 2000, more productive cities for high skill workers attracted a larger share of these workers, which caused increases in local productivity, boosting all worker's wages, and improved the local amenities. How does this divergence across cities reflect into differences in education decisions, and thus upward mobility? Local learning predicts children in cities more productive for high skill workers to have higher perceptions about the skill premium, hence more likely to enroll in college, potentially feeding the "The Great Divergence" phenomenon. I plan to study this issue in future research.

1.6 Appendix

1.6.1 Data Appendix

This section contains additional tables and figures referred to in the main text.

Table 1.A1: Summary Statistics

The table reports summary statistics for the main variables used in the empirical analysis. Observations are at the school district level and cover the period from 2008 to 2014. *Enrollment in a 4-year College* measures the share of high-school graduates in all public schools that enroll in a 4-year college within 6 months after graduation. *College graduates* is the share of population over 25 years old with 4 or more years of college. *Black* and *white* residents are measured as the share of total population in the school district that are black and white, and the *unemployment rate* is the share of the civilian labor force that is unemployed. *ACT score* is the score in the American College Testing averaged over all high-school graduates in all public schools. *Females* measures the share of high school graduates in all public schools that are females. *Earnings by Educational Attainment* correspond to median annual earnings per education level at the school district level and are expressed in 2010 dollars. Expenditures and revenues per pupil are also expressed in 2010 dollars. Source: CEPI, NCES-EDGE and NCES-CCD.

	Observations	Mean	Std. Dev	Min.	Max.
College Enrollment					
Enrollment in a 4-year College	1847	0.33	0.14	0.05	0.80
Earnings by Educational Attainment					
High School Degree	1851	26462.86	4039.23	12365.45	42366.31
College Degree	1851	46730.47	9205.04	11230.26	85625.00
Post-Graduate Degree	1848	60924.20	12024.40	15378.39	107063.21
Socioeconomic Variables					
College Graduates	1851	0.23	0.13	0.04	0.79
Median Family Income	1851	62700.72	17822.69	19409.57	147755.92
Black Residents	1839	0.08	0.15	0.00	0.93
White Residents	1839	0.86	0.16	0.04	1.00
Unemployment Rate	1851	0.11	0.04	0.03	0.36
Total Population	1839	28834.78	53748.40	2145.00	916133.00
Cohort Variables					
ACT Score	1851	19.10	2.01	12.23	25.93
Females	1845	0.51	0.05	0.33	0.75
School Quality Variables					
Expenditure per student	1840	10820.13	2312.69	7624.06	30499.21
Local revenue per student	1840	3432.45	2034.61	761.49	23402.94
Teachers to student ratio	1842	0.05	0.01	0.03	0.09

Table 1.A2: Correlations between Main Variables

The table reports the correlation pattern between the main variables used in the empirical analysis. The correlations are computed using the 1851 district-year observations in the sample over the period from 2008 to 2014. Source: CEPI, NCES-EDGE and NCES-CCD.

	Enrollment in a 4-year College	College Graduates	Median Earnings, College Grad.	Median Family Income	ACT Score	Expenditure per student	Local Revenue per student
Enrollment in a 4-year College	1						
College Graduates	0.739	1					
Median Earnings, College Grad.	0.427	0.445	1				
Median Family Income	0.701	0.830	0.683	1			
ACT Score	0.749	0.714	0.509	0.789	1		
Expenditure per student	0.047	0.181	-0.028	0.061	-0.127	1	
Local Revenue per student	0.230	0.424	0.148	0.307	0.182	0.587	1

Table 1.A3: College Enrollment and College Graduates

The table reports coefficients from an OLS regression with robust standard errors clustered at the school district level reported in parentheses. The dependent variable is the share of high-school graduates that enroll in a 4-year college within 6 months of graduation, with mean equal to 0.33. Column 2 to 6 control for characteristics of the graduating class (the share of females among the high-school graduates and the average ACT score). Socioeconomic controls include the share of black and white residents, the unemployment rate, the median family income, school district size. The sample includes all school districts within MSA's in Michigan over the period 2008 and 2014. ***, ** and * represent statistical significance at 1%, 5% and 10% levels, respectively. Source: CEPI, NCES-EDGE and NCES-CCD.

Dependent Variable: Share of High-School Graduates that Enroll in a 4-year College

	(1)	(2)	(3)	(4)	(5)	(6)
College Graduates	0.777***	0.437***	0.387***	0.375***	0.367***	0.366***
	(0.029)	(0.035)	(0.050)	(0.053)	(0.054)	(0.053)
ACT Score		0.031***	0.038***	0.038***	0.038***	0.039***
		(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Median Family Income			0.045	0.046	0.047	0.049*
			(0.028)	(0.029)	(0.029)	(0.029)
Expenditure per student				0.004	0.003	0.004
				(0.004)	(0.005)	(0.005)
Local Revenue per student				0.002	0.004	0.003
				(0.005)	(0.005)	(0.005)
Teachers to student ratio				-0.001	-0.000	-0.000
				(0.006)	(0.006)	(0.006)
Observations	1841	1839	1827	1818	1818	1818
Adjusted R^2	0.703	0.786	0.798	0.798	0.803	0.801
Socieconomic controls	N	N	Y	Y	Y	Y
Year FE	Y	Y	Y	Y	N	Y
City FE	Y	Y	Y	Y	N	Y
City-year FE	N	N	N	N	Y	N
City trend	N	N	N	N	N	Y

Table 1.A4: College Enrollment and College Graduates: Heterogeneity by Earnings

The table reports coefficients from an OLS regression with robust standard errors clustered at the school district level reported in parentheses. The dependent variable is the share of high-school graduates that enroll in a 4-year college within 6 months of graduation, with mean equal to 0.33. Column 2 to 6 control for characteristics of the graduating class (the share of females among the high-school graduates and the average ACT score). Socioeconomic controls include the share of black and white residents, the unemployment rate, the median family income, school district size and median annual earnings of high-school graduates. The sample includes all school districts within MSA's in Michigan over the period 2008 and 2014. ***, ** and * represent statistical significance at 1%, 5% and 10% levels, respectively. Source: CEPI, NCES-EDGE and NCES-CCD.

Dependent Variable: Share of High-School Graduates that Enroll in a 4-year College

	(1)	(2)	(3)	(4)	(5)	(6)
College Graduates	-5.989***	-5.508***	-4.763***	-4.783***	-4.771***	-4.708***
	(1.468)	(1.252)	(1.122)	(1.110)	(1.111)	(1.097)
College Graduates × Earnings	0.619***	0.550***	0.478***	0.479***	0.477***	0.471***
	(0.135)	(0.115)	(0.104)	(0.103)	(0.103)	(0.102)
Earnings	-0.008	-0.078***	-0.062**	-0.060**	-0.062**	-0.057**
	(0.030)	(0.026)	(0.025)	(0.025)	(0.026)	(0.026)
Earnings, High-school Degree	0.061**	-0.032	-0.018	-0.022	-0.016	-0.018
	(0.024)	(0.021)	(0.022)	(0.022)	(0.023)	(0.022)
ACT Score		0.030***	0.036***	0.036***	0.036***	0.036***
		(0.002)	(0.003)	(0.003)	(0.003)	(0.003)
Median Family Income			0.020	0.022	0.021	0.021
			(0.028)	(0.028)	(0.029)	(0.028)
Expenditure per student				0.003	0.002	0.003
				(0.004)	(0.004)	(0.004)
Local Revenue per student				0.004	0.005	0.004
				(0.005)	(0.005)	(0.005)
Teachers to student ratio				-0.003	-0.002	-0.003
				(0.006)	(0.006)	(0.006)
Observations	1841	1839	1827	1818	1818	1818
Adjusted R^2	0.737	0.795	0.804	0.805	0.810	0.807
Socieconomic controls	N	N	Y	Y	Y	Y
Year FE	Y	Y	Y	Y	N	Y
City FE	Y	Y	Y	Y	N	Y
City-year FE	N	N	N	N	Y	N
City trend	N	N	N	N	N	Y

Table 1.A5: College Enrollment and College Graduates: Robustness Checks

The table reports coefficients from an OLS regression with robust standard errors clustered at the school district level reported in parentheses. The dependent variable is the share of high-school graduates that enroll in a 4-year college within 6 months of graduation, with mean equal to 0.33. Each column replicates column 6 in table using either a different proxy for high-skill neighbors earnings or a different sample. Column 1 reports results if I assume the marginal effect of $College_{ijt}$ is quadratic in Y_{ijt}: $\frac{\partial Enrollment_{ijt}}{\partial College_{ijt}} = \beta_1 + \beta_2 \times Y_{ijt}^2$. Column 2 includes college enrollment in $t-1$ as a control. Column 3 uses the same set of controls as in the baseline estimation (column 6 of table 1.A4), but measured in 2009. Column 4 restricts the sample to school districts with less than 10% of non-resident students. Column 5 and 6 report estimation results using only the post-Great Recession years (2010-2014). The former uses all the sample of urban school districts, while the latter only uses urban school districts with less than 10% of non-resident students. Column 7 includes school districts in urban and rural areas. In this specification, I also include a dummy variable that equals one if the school district belongs to an urban area. Columns 8 and 9 use, respectively, the median annual earnings of individuals with a post-graduate degree and the average between this variable and median annual earnings of individuals with a college degree. All columns include year and city fixed effects, a city-specific trend and a vector socioeconomic controls, which include the share of black and white residents, unemployment rate, median family income, school district size and median annual earnings of high-school graduates. The sample includes all school districts within MSA's in Michigan over the period 2008 and 2014. ***, ** and * represent statistical significance at 1%, 5% and 10% levels, respectively. Source: CEPI, NCES-EDGE and NCES-CCD.

Dependent Variable: Share of High-School Graduates that Enroll in a 4-year College

	Quadratic Specification	Lagged Enrollment	2009 Controls	Only Resident Students	Only 2010-2014		Urban + Rural SD's	≠ Earnings Measures	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
College Graduates	-2.170***	-2.639***	-5.055***	-4.450***	-4.828***	-4.450***	-3.799***	-4.404***	-4.978***
	(0.551)	(0.690)	(1.231)	(1.432)	(1.191)	(1.432)	(1.031)	(1.221)	(1.330)
College Graduates × Earnings^2	0.022***								
	(0.005)								
Earnings, College Degree	-0.056**	-0.048***	-0.079***	-0.079**	-0.075***	-0.079**	-0.059**		
	(0.025)	(0.018)	(0.029)	(0.036)	(0.028)	(0.036)	(0.024)		
College Graduates × Earnings		0.260***	0.504***	0.452***	0.482***	0.452***	0.381***		
		(0.064)	(0.114)	(0.134)	(0.111)	(0.134)	(0.095)		
College Graduates × Earnings (post-college)								0.431***	
								(0.110)	
College Graduates × Earnings (average)									0.490***
									(0.121)
Observations	1818	1539	1430	876	1424	876	3023	1815	1815
Adjusted R^2	0.807	0.838	0.781	0.792	0.782	0.792	0.689	0.811	0.811

Table 1.A6: Parameters

Description	Parameter	Value	Source/Target
Panel A: Exogenously chosen			
Number of neighborhoods	J	95	Number of school-districts within Detroit in 2013 (CEPI)
Risk-aversion (CARA)	γ	0.5	(Babcock et al., 1993)
Low-skill wage's mean	w^L	7.9	Low-skill workers earnings distribution (ACS 2008-2013)
High-skill wage's mean	w^H	8.8	High-skill workers earnings distribution (ACS 2008-2013)
High-skill wage's variance	σ_ϵ^H	0.03	High-skill workers earnings distribution (ACS 2008-2013)
Prior mean	$\tilde{\mu}^2$	8.2	(Bleemer and Zafar, 2016)
Prior variance	$\tilde{\sigma}^2$	0.06	(Bleemer and Zafar, 2016)
Panel B: Estimated			
Cost function parameter	$\bar{c}$	7.91	Mean of enrollment
Cost function parameter	φ	0.72	Std. deviation of enrollment
Cost function parameter	ϕ	1.46	p75/p50
Cost function parameter	ρ	0.09	Corr. btw. enrollment and college graduates
Cost function parameter	κ	0.21	Corr. btw. enrollment and expenditures per student

Table 1.A7: Reallocation Policy: Different number of movers

The table reports the effects for movers, stayers and receivers when a policy that moves 5%, 25% and 50% of the children living in the 25th percentile of the college graduates distribution to location in the 75th percentile of the college graduates distribution is implemented. *High-skill neighbors* corresponds to the share of high-skill neighbors in both the baseline and the counterfactual

	5% of childrens within 1stqtl					25% of childrens within 1stqtl					50% of childrens within 1stqtl				
	High-skill neighbors		Enrollment rate			High-skill neighbors		Enrollment rate			High-skill neighbors		Enrollment rate		
	1stqtl	4thqtl	Movers	Stayers	Receivers	1stqtl	4thqtl	Movers	Stayers	Receivers	1stqtl	4thqtl	Movers	Stayers	Receivers
Panel A: Total Effect															
Benchmark	0.11	0.47	0.25	0.25	0.54	0.11	0.47	0.25	0.25	0.54	0.11	0.47	0.25	0.25	0.54
Policy Counterfactual	0.12	0.45	0.53	0.25	0.53	0.15	0.38	0.49	0.30	0.49	0.22	0.47	0.45	0.38	0.46
Panel B: Decomposition															
Local learning			0.46					0.42					0.39		
School quality			0.46					0.42					0.39		
Spillovers			0.53					0.49					0.46		

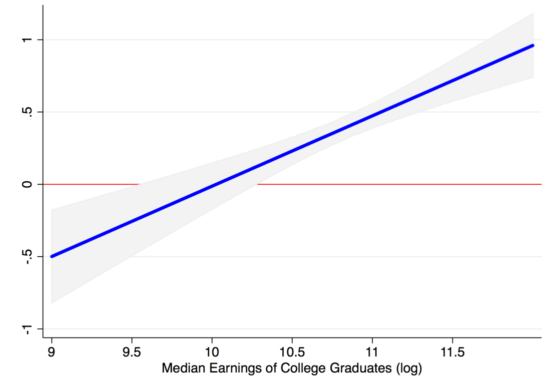
Table 1.A8: Different Priors: Benchmark Economy vs. Counterfactuals

The table reports the standard deviation of the college enrollment distribution across school districts under the benchmark economy and four different scenarios: no local learning, equal human capital spillovers ($m_{jH} = \bar{m}$) and equal school resources ($q_j = \bar{q}$), and no information frictions ($\hat{\mu}_j = w^H$ and $\hat{\sigma}_j^2 = \sigma_{\epsilon^H}$). Observations are at school district level. The sample is composed by 95 school districts within Detroit in 2013.

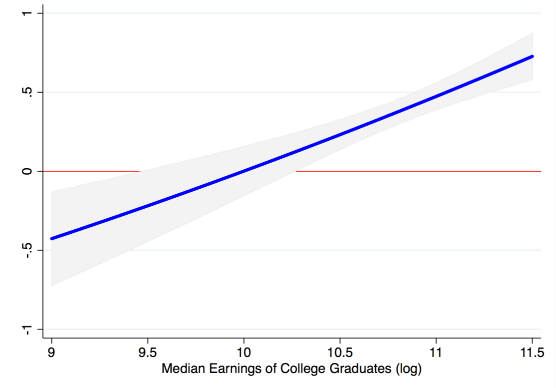
	Benchmark	No-learning	$m_{j,H} = \bar{m}$	$q_j = \bar{q}$	No Frictions
Std. Dev. Enrollment	0.13	0.07	0.10	0.13	0.01

Figure 1.A1: Correlation between College graduates and Enrollment: Heterogeneity By Earnings

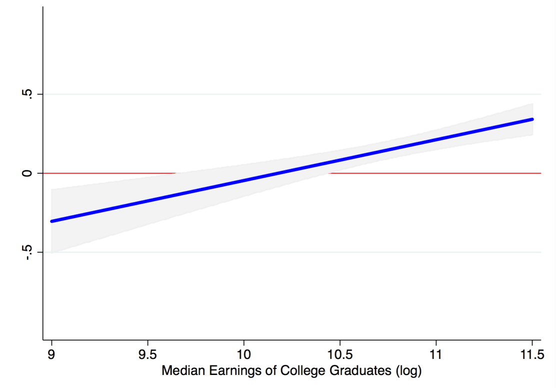
A. Linear Specification (column 4 in Table 1.A4)



B. Quadratic Specification (column 1 in Table 1.A5)

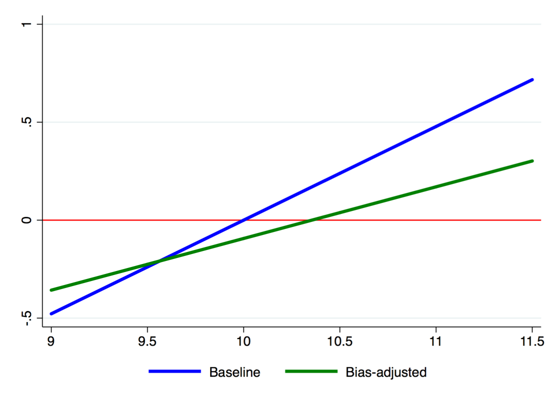


C. Lagged Enrollment (column 2 in Table 1.A5)



Notes: All panels plot the average marginal effect of an increase in the share of college graduates by one unit on the college enrollment rate for different levels of median earnings of college graduates. Panel A plots the average marginal effect from the specification in column 7 in Table 1.A4, while Panels B and C plot the average marginal effect when I consider a quadratic specification in earnings (column 1 in Table 1.A5) and control for college enrollment in the previous period (column 2 in Table 1.A5). The shaded area represents 95% confidence intervals. The x-axis corresponds to the log median earnings of college graduates in 2010 dollars.

Figure 1.A2: College graduates and Enrollment: Adjusted-bias Coefficients

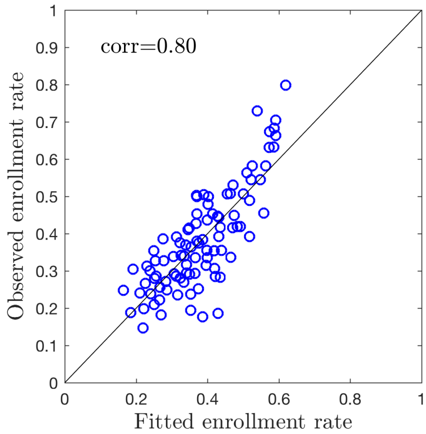


$Earnings_{ijt}$

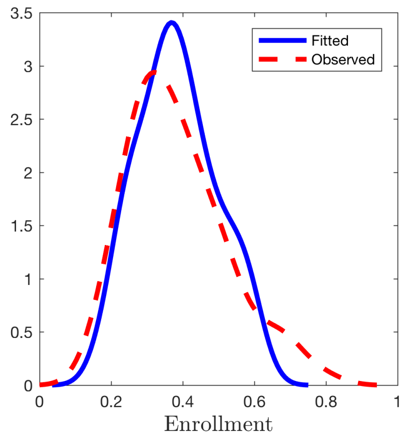
Notes: This figure plots the average marginal effect when I use the estimated coefficients in column 6 in Table 1.A4 (blue line) and the bias-adjusted coefficients, β_1^* and β_2^* (green line) when the influence of unobservables on the outcome variable is of similar magnitude as the impact of observable variables, $\delta = 1$. $\beta_i^* = \hat{\beta}_i - \delta(\tilde{\beta}_i - \hat{\beta}_i)\frac{1-\hat{R}}{\hat{R}-\tilde{R}}$, where $\hat{\beta}$ are the estimated coefficients and R^2 of column 6 in Table 1.A4 and $\tilde{\beta}$ and $\tilde{R}$ are the estimated coefficients and R^2 of OLS estimation of Equation (1.1) with no controls (i.e. not including city and year fixed effects, a city-specific trend and the controls vector X_{ijt}). The x-axis corresponds to the log median earnings of college graduates in 2010 dollars.

Figure 1.A3: Different Priors: Model vs. Data

A. Enrollment Rate

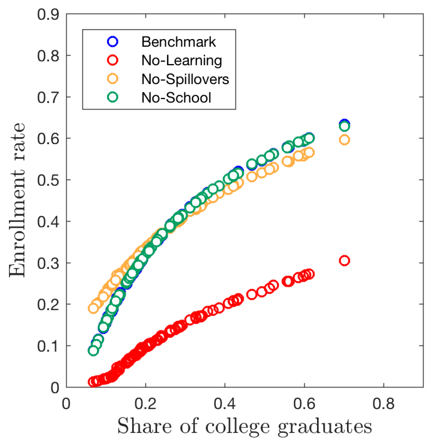


B. Enrollment Distribution



Panel A plots fitted and observed values for the college enrollment rate across school districts. Fitted values are on the horizontal axis; observed values are on the vertical axis. Correlation between fitted and observed values is equal to 0.8. Observations are at school district level. Panel B plots the enrollment distribution simulated in the model and observed in the data. The sample is composed by the 95 school districts within Detroit in 2013.

Figure 1.A4: Different Priors: Benchmark Economy vs. Counterfactuals



The panel plots college enrollment across school districts under the benchmark economy (blue) and three different scenarios: no local learning (red), equal human capital spillovers ($m_{jH} = \bar{m}$, yellow) and equal school resources ($q_j = \bar{q}$, green). The sample is composed by the 95 school districts within Detroit in 2013.

1.6.2 Theoretical Appendix

Location decisions

I report additional details for the characterization of parents locations decisions, as described by Equation (1.9). Given the Fréchet distribution for the idiosyncratic taste, $\varepsilon_{i,j} \sim$ Fréchet$(\theta, 1)$, it follows that $\varepsilon_{i,j}^{-1} \sim$ Weibull$(\theta, 1)$.[32] Hence, the indirect utility function described by Equation (1.8) is also Weibull distributed:

$$\upsilon_{i,k,j}\varepsilon_{i,j} \sim \text{Weibull}(\theta, \upsilon_{i,k,j}) \tag{1.B1}$$

where $\upsilon_{i,k,j} = \frac{-\exp(-\gamma(w_{i,k,j}-R_j))}{\Phi_{k,j}}$, with $\Phi_{k,j} = q_j \cdot A_{jk}$, is a constant.[33] Let $X_1, ..., X_n$ be statistically independent, with each $X_i \sim$ Weibull(θ, υ_i), for $\theta, \upsilon_1, ...\upsilon_n > 0$. Then

$$Pr[k \in argmin\ X_i] = \frac{\upsilon_k^{-\theta}}{\sum_i \upsilon_i^{-\theta}}, \forall k \in \mathcal{I} \tag{1.B2}$$

Combining Equations (1.B1) and (1.B2), and setting $\theta = 1$, the probability that a parent i with skill level k chooses to live in location j out of all possible locations, $\rho_{i,k,n}$, is:

$$\rho_{i,j}^k = Pr[U_{i,k,j} \geq U_{i,k,n'}; \forall j' \in \mathcal{J}], = \frac{\Phi_{k,j}\exp(\gamma(w_i^k - r_j))}{\sum_{j'\in\mathcal{J}} \Phi_{k,j'}\exp(\gamma(w_i^k - r_{j'}))} \tag{1.B3}$$

which simplifies to

$$\rho_j^k = \frac{\Phi_{k,j}\exp(\gamma(-r_j))}{\sum_{j'\in\mathcal{J}} \Phi_{k,j'}\exp(\gamma(-r_{j'}))} \tag{1.B4}$$

Because $\rho_{i,j}^k$ does not depend on the wage, which is the same no matter where the family lives in the city, it is equal across individuals in the same skill group. Given

[32]The cumulative distribution function of the Weibull distribution with parameters θ and λ is Pr$(X \leq x) = 1 - \exp(-(\frac{x}{\lambda})^\theta)$ with $x \geq 0$. The mean is $\lambda\Gamma(1 + 1/\theta)$ and the variance is $\lambda^2[\Gamma(1 + 2/\theta) - \Gamma^2(1 + 1/\theta)]$. Since β, the scale parameter of the Fréchet distribution, is equal to 1, $\lambda = 1$.

[33]If $Y = tX$, where $X \sim$ Weibull$(\theta, 1)$, then Y is Weillbul(θ, t).

this, the number of k-skill parents in each neighborhood is

$$M_{k,j} = \sum_{i=1}^{M_k} \rho_{i,j}^k = \sum_{i=1}^{M^k} \rho_j^k = \rho_{k,j} \cdot M_k \tag{1.B5}$$

Spatial Equilibrium - An Illustration

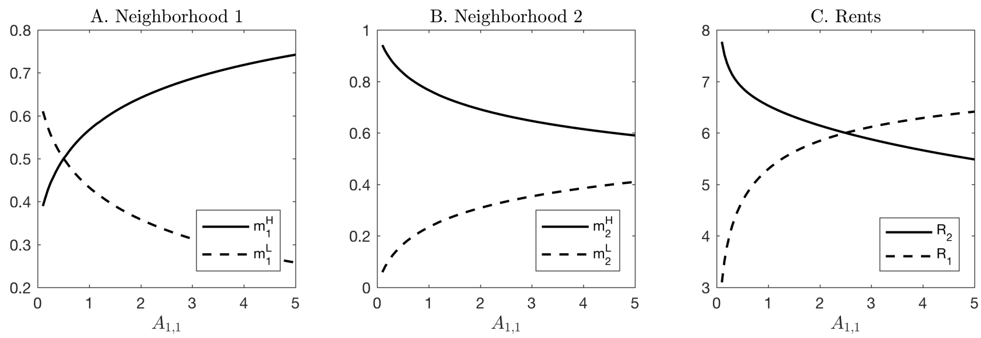
Let's consider the example of a city with two neighborhoods, 1 and 2, each with the same capacity, $H_1 = H_2$. I set $A_{2H} = A_{1L} = A_{2L} = 2.5$, and look to the spatial equilibrium for different values of $A_{1,1}$. Panels A, B and C in Figure 1.B1 show, respectively, the equilibrium skill-mix in neighborhood 1 and 2 and equilibrium rents in both locations, the endogenous variables, as a function of A_{1H}. At low values of A_{1H}, the probability of choosing to live in neighborhood 2, conditional on being a high-skill parent, is high relative to the probability of choosing to live in neighborhood 1. On the other hand, the probability of choosing to live in neighborhood 2, conditional on being a low-skill parent, is very low due to the high rents in this location. This makes neighborhood 2 mainly composed of high-skill households. At high values of $A_{1,1}$, neighborhood 1 becomes more attractive to high-skill families, increasing housing prices in neighborhood 1. Higher rents in neighborhood 1, in turn, make this neighborhood less attractive, and low-skill households transfer to neighborhood 2. Note that when amenities are equal across neighborhoods, rents and the skill-mix of each location is also equal. In this situation, the spatial equilibrium is non-sorted.

Value Functions

For a child born in neighborhood j, the perceived value of being a high-skill worker, V_j^H, is given by

$$V_j^H = \sum_{j' \in \mathcal{J}} \underbrace{\Gamma\left(1 + \frac{1}{\theta}\right) \mathbb{E}_{w_i^H}[U(c_{i,j}^H, \Phi_{H,j'})|\mathcal{I}_j]\, \rho_{j'}^H}_{\text{Expected utility of living in location } j' \text{ if } k=H} \tag{1.B6}$$

Figure 1.B1: Spatial Equilibrium - An Example



Panel A: Equilibrium skill-mix in neighborhood 1 for different levels of A_1 in neighborhood 1, share of high-skill households (solid line) and share of low-skill households (dotted line). Panel B: Equilibrium skill-mix in neighborhood 2 for different levels of A_1 in neighborhood 1, share of skilled families (solid line) and share of unskilled families (dotted line). Panel C: Equilibrium rents in neighborhood 1 (dotted line) and neighborhood 2 (solid line) for different levels of A_1 in neighborhood 1. $H_1 = H_2 = 75$, $M^H = 100$, $M^L = 50$, $\beta_H = 1$, $\beta_L = 0$, $A_{1,2} = A_{1,2} = A_{2,2}$.

where $\Gamma\left(1 + \frac{1}{\theta}\right)$ is the expected value of the idiosyncratic component of utility and $\Gamma(.)$ the gamma function. $\mathbb{E}$ is the expectations operator and the expectation is taken over the high-skill wage. $\rho^H_{j'}$ is the probability of living in neighborhood j' conditional on being a high-skill worker.[34] I assume $\theta = 1$ for simplicity, hence $\Gamma\left(1 + \frac{1}{\theta}\right) = 1$. Equation (1.B6) can be rewritten as

$$\sum_{j' \in \mathcal{J}} \mathbb{E}_{w_i^H} \left[\frac{-\exp(-\gamma(w_i^H - r_{j'}))}{\Phi_{s,n'}} \bigg| \mathcal{I}_j \right] \rho^H_{j'} =$$

$$= \sum_{j' \in \mathcal{J}} \left[\frac{-\exp(-\gamma(\hat{\mu}_j - \gamma(\hat{\sigma}_j^2/2) - r_{j'}))}{\Phi^H_{j'}} \right] \rho^H_{j'} =$$

which simplifies to

$$V_j^H = -\exp(-\gamma(\hat{\mu}_j - \gamma(\hat{\sigma}_j^2/2)) \left(\frac{J}{\sum_{j' \in \mathcal{J}} \frac{\Phi^H_{j'}}{\exp(\gamma r_{j'})}} \right) \tag{1.B7}$$

[34] Since the idiosyncratic taste and the skilled wage are two independent random variables, it follows that $\mathbb{E}[w_i^H \cdot \varepsilon_{i,j}] = \mathbb{E}[w_i^H] \cdot \mathbb{E}[\varepsilon_{i,j}]$

Equation (1.B7) is equal for all children born in neighborhood j, but different across children from neighborhoods as long as the share of skilled individuals differs.

For a child born in neighborhood j, the expected value of becoming an unskilled worker, V_j^L, is given by

$$V_j^L = \sum_{j' \in \mathcal{J}} \underbrace{\Gamma\left(1 + \frac{1}{\theta}\right) U(c_{i,j}^L, \Phi_{j'}^L)}_{\text{Expected utility of living in location } j' \text{ if } k = L} \rho_{j'}^L = \tag{1.B8}$$

where $\Gamma\left(1 + \frac{1}{\theta}\right)$ is the expected value of the idiosyncratic component of utility and $\Gamma(.)$ the gamma function. $\rho_{j'}^L$ is the probability of living in neighborhood j' conditional on being a high-skill worker. I assume $\theta = 1$ for simplicity, hence $\Gamma\left(1 + \frac{1}{\theta}\right) = 1$. Equation (1.B8) can be rewritten as

$$\sum_{j' \in \mathcal{J}} \left[\frac{\exp(\gamma(w^L - r_{j'}))}{\Phi_{j'}^L}\right] \rho_{j'}^L$$

which simplifies to

$$V^L = -\exp(-\gamma w^L) \left(\frac{J}{\sum_{j' \in \mathcal{J}} \frac{\Phi_{L,j'}}{\exp(\gamma r_{j'})}}\right) \tag{1.B9}$$

Equation (1.B9) is equal for all children in the city, regardless of where they live. Hence I suppress j.

Proofs

Proof of Lemma 1.3 Given V_j^H (Equation (1.B7)) and V^L (Equation (1.B9)), the perceived skill premium for a child born in neighborhood j, $\triangle V_j \equiv V_j^H - V^L$, is given by

$$\triangle V_j = J\left(\frac{-\exp(-\gamma\big(\hat{\mu}_j - \gamma(\hat{\sigma}_j^2/2))}{\sum_{j'\in\mathcal{J}}\frac{\Phi^H_{j'}}{\exp(\gamma r_{j'})}} - \frac{-\exp(-\gamma w^L)}{\sum_{j'\in\mathcal{J}}\frac{\Phi^L_{j'}}{\exp(\gamma r_{j'})}}\right) \tag{1.B10}$$

where j indexes the neighborhood where the child lives, and J is the number of neighborhoods in the city. The optimal investment decision takes the form of a cut-off rule. The ability cut-off, a_J^*, is defined by the indifference condition $\triangle V_n = c(a_n^*)$. Defining $\varpi_j \equiv \triangle V_j - c(a_j^*)$, I establish the following:

1. $\frac{\partial s_{i,j}}{\partial \hat{\mu}_j} > 0$. The effect of $\hat{\mu}_j$ on the probability of becoming a high-skill worker, $s_{i,j}$ is given by

$$\frac{\partial s_{i,j}}{\partial \hat{\mu}_j} = \frac{\frac{\partial s_{i,j}}{\partial a_j^*}}{\frac{\partial a_j^*}{\partial \hat{\mu}_j}}$$

By the implicit function theorem, $\frac{\partial a_j^*}{\partial \hat{\mu}_j} = -\frac{\frac{\partial \varpi}{\partial a_j^*}}{\frac{\partial \varpi}{\partial \hat{\mu}_j}} < 0$, because $\frac{\partial \varpi}{\partial a_j^*} > 0$ and $\frac{\partial \varpi}{\partial \hat{\mu}_j} > 0$. Since $\frac{\partial s_{i,j}}{\partial a_j^*} < 0$ and $\frac{\partial a_j^*}{\partial \hat{\mu}_j} < 0$, one can conclude that $\frac{\partial s_{i,j}}{\partial \hat{\mu}_j} > 0$.

2. $\frac{\partial s_{i,j}}{\partial \hat{\sigma}_j^2} < 0$. The effect of $\hat{\sigma}_j^2$ on the probability of becoming a high-skill worker $s_{i,j}$ is given by

$$\frac{\partial s_{i,j}}{\partial \hat{\sigma}_j^2} = \frac{\frac{\partial s_{i,j}}{\partial a_j^*}}{\frac{\partial a_j^*}{\partial \hat{\sigma}_j^2}}$$

By the implicit function theorem, $\frac{\partial a_j^*}{\partial \hat{\sigma}_j^2} = -\frac{\frac{\partial \varpi}{\partial a_j^*}}{\frac{\partial \varpi}{\partial \hat{\sigma}_j^2}} > 0$, because $\frac{\partial \varpi}{\partial a_j^*} > 0$ and $\frac{\partial \varpi}{\partial \hat{\sigma}_j^2} < 0$. Since $\frac{\partial s_{i,j}}{\partial a_j^*} < 0$ and $\frac{\partial a_j^*}{\partial \hat{\sigma}_j^2} > 0$, one can conclude that $\frac{\partial s_{i,j}}{\partial \hat{\sigma}_j^2} < 0$.

Proof of Proposition 1.1 The effect of m_{jH} on the probability of investing in education $s_{i,j}$ is given by

$$\frac{\partial s_{i,j}}{\partial m_{jH}} = \frac{\frac{\partial s_{i,j}}{\partial a_j^*}}{\frac{\partial a_j^*}{\partial m_{jH}}}$$

By the implicit function theorem, $\frac{\partial a_j^*}{\partial m_{jH}} = -\frac{\frac{\partial \varpi}{\partial a_j^*}}{\frac{\partial \varpi}{\partial m_{jH}}}$. The numerator is higher than zero, the denominator is given by

$$\begin{aligned}\frac{\partial \varpi}{\partial m_{jH}} &= \frac{\partial \varpi}{\partial \triangle V_j}\frac{\partial \triangle V_j}{\partial m_{jH}} + \frac{\partial \varpi}{\partial c(a_i^*)}\frac{\partial c(a_i^*)}{\partial m_{jH}} = \\ &= \underbrace{\Upsilon \cdot \frac{\sigma^2_{\epsilon^H}}{m_{jH}}\frac{\tilde{\sigma}^2}{(\tilde{\sigma}^2+\sigma_j^2)^2}\cdot \left[w_j^H - \tilde{\mu_j} + \frac{\gamma}{2}\tilde{\sigma}^2\right]}_{A>0 \text{ or } A<0} \\ &+ \underbrace{\frac{\partial \varpi}{\partial c(a_i^*)}\frac{\partial c(a_i^*)}{\partial m_{jH}}}_{B>0}\end{aligned}$$

where $\Upsilon = J \cdot \gamma \frac{\exp(-\gamma(\hat{\mu}_j - \gamma(\hat{\sigma}_j^2/2)))}{\sum_{j' \in \mathcal{J}} \frac{\Phi_{H,j'}}{\exp(\gamma r_{j'})}}$.

If $w_j^H > \tilde{\mu}_j - \frac{\gamma}{2}\tilde{\sigma}^2$, then $A > 0$ and $\frac{\partial s_{i,j}}{\partial m_{jH}} > 0$. If $w_j^H < \tilde{\mu}_j - \frac{\gamma}{2}\tilde{\sigma}^2$, then $A < 0$. If w_j^H is sufficiently low such that $|A| > B$, the positive effect through the cost function does not compensate the negative effect through the information channel, $\frac{\partial s_{i,j}}{\partial m_{jH}}$¡0. The signal threshold below which $\frac{\partial s_{i,j}}{\partial m_{jH}} < 0$ is lower than the one in the case with no human capital spillovers in the cost function.

Implications of Different Specifications for the Model

Risk neutrality Consider that individuals have a linear indirect utility function given by

$$U(w_i^k, r_j, \Phi_j^k, \varepsilon_{i,j}) = w_i^k - r_j + \Phi_j^k + \varepsilon_{i,j} \tag{1.B11}$$

where $\Phi_j^k = q_j A_{j,k}$ and the utility shock $\varepsilon_{i,j}$ follows the extreme value type 1 distribution with parameters μ_ε and σ_ε.[35] The distributional assumption on the idiosyncratic taste, ε, allows me to derive a close-form expression for $\rho_{i,j}^k$, as before:

$$\rho_{i,j}^k = \frac{\exp(w_i^k - r_j + \Phi_j^k)}{\sum_{j' \in \mathcal{J}} \exp(w_i^k - r_{j'} + \Phi_{j'}^k)} \tag{1.B12}$$

Other things equal, as before, a type-j parent is more likely to live in a neighborhood the more attractive are j-specific amenities and the lower are rental prices (r_j). Since migration is only allowed in the beginning of the period, $\rho_{i,j}^k$ translate directly into the neighborhood size distribution. The equilibrium number of j-skill parents in neighborhood j, M_j^k, is given by

$$M_{j,k} = \textstyle\sum_{i=1}^{M_k} \rho_{i,j}^k = \rho_j^k M_k$$

Using Equations (1.B11) and (1.B12), I can compute the perceived expected value of being a high-skill worker, V_j^H and the expected value of being a low-skill worker, V_j^L functions, and the perceived skill premium for a child born in neighborhood j, $\triangle V_j$, Equation (1.B13):

$$\triangle V_j = \underbrace{\sum_{j' \in \mathcal{J}} \left[\hat{\mu}_j - r_{j'} + \Phi_{j'}^H\right] \rho_{j'}^H}_{V_j^H} - \underbrace{\sum_{j' \in \mathcal{J}} \left[w^L - r_{j'} + \Phi_{j'}^L\right] \rho_{j'}^L}_{V_j^L} \tag{1.B13}$$

It can be shown that:

1. $\frac{\partial s_{i,j}}{\partial \hat{\mu}_j} > 0$,
2. $\frac{\partial s_{i,j}}{\partial \hat{\sigma}_j^2} = 0$, this follows from the fact that children are risk neutral, and,
3. $\frac{\partial s_{i,j}}{\partial m_{jH}} > 0$ if $w_j^H > \tilde{\mu}$,

as before.

[35] The extreme value type 1 distribution is commonly used in the discrete-choice literature. The density of the extreme value type 1 distribution with parameters parameters μ_ε and σ_ε is $f(x) = \exp(-\exp(-(x - \mu_\varepsilon)/\sigma_\varepsilon))$.

Uncertainty about Low-Skill Wage If $\sigma^2_{\epsilon^L} > 0$, Equation (1.B14) can be re-written as

$$\triangle V_j = J\left(\frac{-\exp(-\gamma\big(\hat{\mu}^H_j - \gamma(\hat{\sigma}^2_{H,j}/2)\big)}{\sum_{j'\in\mathcal{J}}\frac{\Phi^H_{j'}}{\exp(\gamma r_{j'})}} - \frac{-\exp(-\gamma\big(\hat{\mu}^L_j - \gamma(\hat{\sigma}^2_{L,j}/2)\big)}{\sum_{j'\in\mathcal{J}}\frac{\Phi^L_{j'}}{\exp(\gamma r_{j'})}}\right) \tag{1.B14}$$

where are $\hat{\mu}^H_j$ and $\hat{\sigma}^2_{H,j}$ the posterior mean and variance of the beliefs about w^H_i; $\hat{\mu}^L_j$ and $\hat{\sigma}^2_{L,j}$ are the posterior mean and variance of the beliefs about w^L_i for a child born in neighborhood j. Following the same steps as in the proof of lemma 1.3 above, it can be shown that:

1. $\frac{\partial s_{i,j}}{\partial \hat{\mu}^H_j} > 0$ and $\frac{\partial s_{i,j}}{\partial \hat{\mu}^L_j} < 0$

2. $\frac{\partial s_{i,j}}{\partial \hat{\sigma}^2_{H,j}} < 0$ and $\frac{\partial s_{i,j}}{\partial \hat{\sigma}^2_{L,j}} > 0$

Naturally, the higher is the expected value of the low-skill wage, the lower is the probability to invest in education, since the perceived skill-premium is lower, holding all else constant. On the other hand, because individuals are risk-averse, higher uncertainty about the low-skill wage, increases the perceived skill-premium, hence the probability of investing in education.

As before, the effect of m_{jH} on the probability of investing in education $s_{i,j}$ is given by

$$\frac{\partial s_{i,j}}{\partial m_{jH}} = \frac{\frac{\partial s_{i,j}}{\partial a^*_j}}{\frac{\partial a^*_j}{\partial m_{jH}}}$$

By the implicit function theorem, $\frac{\partial a^*_j}{\partial m_{jH}} = -\frac{\frac{\partial \varpi}{\partial a^*_j}}{\frac{\partial \varpi}{\partial m_{jH}}}$. The numerator is higher than zero, the denominator is given by

$$\frac{\partial \varpi}{\partial m_{jH}} = \underbrace{\frac{\partial \varpi}{\partial \triangle V_j}}_{>0} \underbrace{\left[\frac{\partial V_j^H}{\partial m_{jH}} + \frac{\partial V_j^L}{\partial m_{jH}}\right]}_{?} + \underbrace{\frac{\partial \varpi}{\partial c(a_i^*)} \frac{\partial c(a_i^*)}{\partial m_{jH}}}_{>0} =$$

$$= J \cdot \Big[\underbrace{\Upsilon_H \cdot \frac{\sigma_{\epsilon H}^2}{m_{jH}} \frac{\tilde{\sigma}_H^2}{(\tilde{\sigma}_H^2 + \sigma_{H,j}^2)^2}}_{>0} \cdot \underbrace{\left[w_j^H - \tilde{\mu}_j^H + \frac{\gamma}{2}\tilde{\sigma}_H^2\right]}_{A}$$

$$= -\underbrace{\Upsilon_L \cdot \frac{\sigma_{\epsilon L}^2}{m_{jH}} \frac{\tilde{\sigma}_L^2}{(\tilde{\sigma}_L^2 + \sigma_{L,j}^2)^2}}_{>0} \cdot \underbrace{\left[-w_j^L + \tilde{\mu}_j^L - \frac{\gamma}{2}\tilde{\sigma}_L^2\right]}_{B} \Big] + \underbrace{\frac{\partial \varpi}{\partial c(a_i^*)} \frac{\partial c(a_i^*)}{\partial m_{jH}}}_{>0}$$

where $\Upsilon_H = \gamma \frac{\exp(-\gamma(\hat{\mu}_j^H - \gamma(\hat{\sigma}_{H,j}^2/2)))}{\sum_{j' \in \mathcal{J}} \frac{\Phi_{H,j'}}{\exp(\gamma r_{j'})}}$, and $\Upsilon_L = \gamma \frac{\exp(-\gamma(\hat{\mu}_j^L - \gamma(\hat{\sigma}_{L,j}^2/2)))}{\sum_{j' \in \mathcal{J}} \frac{\Phi_{L,j'}}{\exp(\gamma r_{j'})}}$.

1. If $w_j^L = \tilde{\mu}_j^L - \frac{\gamma}{2}\tilde{\sigma}_L^2$ such that $B = 0$, the results in proposition 1.1 hold: $\frac{\partial s_{i,j}}{\partial m_{jH}} > 0$ if $w_i^H > \tilde{\mu}_j^H + \frac{\gamma}{2}\tilde{\sigma}_H^2$.

2. If $w_j^L < \tilde{\mu}_j^L - \frac{\gamma}{2}\tilde{\sigma}_L^2$ such that $B > 0$, then the threshold below which $\frac{\partial s_{i,j}}{\partial m_{jH}} < 0$ is higher than the one in in Proposition 1.1.

3. If $w_j^L > \tilde{\mu}_j^L - \frac{\gamma}{2}\tilde{\sigma}_L^2$ such that $B < 0$, then the threshold below which $\frac{\partial s_{i,j}}{\partial m_{jH}} < 0$ is lower than the one in in Proposition 1.1.

Chapter 2

Mismatch Cycles

Joint with Isaac Baley (Universitat Pompeu Fabra and Barcelona GSE) and Robert Ulbricht (Toulouse School of Economics)

2.1 Introduction

The function of a labor market is to allocate each worker to the right job. However, information and search frictions might prevent firms and workers from finding their best match. In this chapter, we ask: How do business cycles affect the allocation of workers to jobs? Do workers end up more mismatched when jobs are scarce, or is it the opposite? Using a measure of mismatch recently developed by Guvenen et al. (2018), we present new empirical evidence that mismatch is procyclical: in recessions, workers skills are more aligned with job requirements; whereas in expansions mismatch increases. This pattern, however, masks important heterogeneities along the flows of job destruction and job creation. In particular, our results suggest that during recessions highly mismatched jobs are destroyed but also created.

To explain our empirical findings, we build a model of learning about unobserved skill mismatch. The novel feature is that recessions are characterized by lower aggregate productivity but also a higher fraction of worker-firm pairs with high mismatch uncertainty. Consistent with the empirical findings, we show that negative productivity shocks destroy (perceived) high mismatched worker-job pairs, but at the same time large information frictions create undetected worker-firm matches with high levels of skill mismatch. We explore the role of occupational switching as a source of the countercyclical uncertainty and document suggestive evidence pointing towards this channel.

Cleansing or Sullying? Economic theory provides two opposing predictions for the cyclical behavior of worker-occupation mismatch. On the one hand, the matching model with endogenous separations in Mortensen and Pissarides (1994) suggests that mismatch is procyclical. In downturns, reservation match quality increases; low quality matches are destroyed while only high quality matches are formed, decreasing average mismatch. This is known as the *cleansing effect* of recessions. On the other hand, the matching model in Barlevy (2002), which allows for on-the-job search, points in favor of countercyclical mismatch because in recessions workers in ongoing job relationships reallocate to better matches (climb the ladder) more slowly, and get stuck in worst matches. This is referred to as the *sullying effect* of recessions. In addition to this, as in recessions firms post fewer vacancies a different type of sullying effect may arise. Moscarini (2001) suggests that unemployed job seekers accept less desirable jobs due to higher competition among them, which increases mismatch.

To assess which of these two effects dominates, we study the cyclical behavior of mismatch using a worker-level panel from the 1979 National Longitudinal Study of Youth (NLSY79), that runs from 1979 to 2012, combined with occupational-level data from O*NET and data on aggregate unemployment. We adopt the mismatch index developed in Guvenen et al. (2018) as a direct measure of skill mismatch. This measure is defined as the difference between a worker's abilities in different skills and how intensive these skills are required by a job. As such, it can be interpreted as the lack of match quality: the larger is this difference, the lower is the quality of a match.[1] In order to estimate the effect of business cycle conditions on mismatch, our identification strategy takes advantage of

[1]Guvenen et al. (2018) use the index of mismatch to study the impact of match quality on wages and patterns of occupational switching. They find that mismatch decreases wages, but increases the probability of switching occupations. Thus, they argue that this mismatch index can be interpreted as a signal of the lack of job match quality.

within-individual variation in the unemployment rate across months the individual is employed, using the monthly unemployment rate at the national level as a proxy for the macroeconomic shocks. Our results show a robust negative association between mismatch and the aggregate unemployment rate, i.e. mismatch is procyclical, consistent with the *cleansing* hypothesis. In particular, during a typical recession, mismatch seems to decrease from its high to its low point.

What job flow drives this pattern, job destruction or job creation? Are recessions times in which better matches are created or times when the worst matches are destroyed? Or both? One advantage of using Guvenen et al. (2018)'s mismatch index is that it allows us to isolate the effect of business cycle conditions on the mismatch of ongoing job relationships from their effect on the mismatch of newly formed relationships at a given point in time. In doing so, we uncover important differences. While for workers in ongoing job relationships mismatch is negatively associated with unemployment in line with the *cleansing effect* of recessions; for new hires from unemployment an increase in unemployment is associated with an increase in mismatch, consistent with the *sullying effect*. Thus, recessions destroy but also create highly mismatched jobs, with the former being the dominant effect. Using quantile regressions, we further show that, for both job stayers and new hires from unemployment, the relationship between economic conditions and mismatch differs significantly across the mismatch distribution. For instance, for job stayers that are perfectly matched, the correlation between the unemployment rate and mismatch is not statistically different from zero; whereas for job stayers that are poorly matched this correlation is 2.4 times larger than the average. Digging deeper, we decompose Guvenen et al. (2018)'s mismatch index into a measure of positive and negative mismatch, which capture, respectively, the extent to which workers' abilities are higher (over-qualification) or lower (under-qualification) than the skill requirements, and examine whether one of the two is driving our results. Surprisingly, we find that while the decrease in mismatch for job stayers is only driven by negative mismatch — workers are less under-qualified in bad times — the increase in mismatch for new hires from unemployment arises from an increase in both over- and under-qualification.

Overall, our results show strong evidence that even tough both the cleansing and the sullying hypothesis are present in recessions, the former dominates, i.e.

average mismatch is procyclical, and that this effect is stronger for poorly matched workers. These findings hold true across industries, occupations, different measures of economic activity and a variety of alternative specifications.

Revisiting job tenure cyclicality Because match quality is not observed, it has been traditionally been proxied by employment duration and wages. This literature has found that matches starting in recessions are shorter and have lower wages, concluding that job match quality is lower in recessions (Bowlus, 1995; Kahn, 2010; Oreopoulos et al., 2012). Armed with a direct measure of mismatch, we revisit the cyclical behavior of job tenure. To do so, we estimate the hazard rate of separation as a function of the current unemployment rate, the unemployment rate at the start of the job spell, as in Bowlus (1995), as well as our mismatch measure. Two results stand out. First, mismatch is positively associated with the hazard rate of separation, and this effect increases with the current level of unemployment. Second, conditional on mismatch, a match that starts in a recession has a shorter duration. This result is surprising as according to the definition of match quality as an experience good, first suggested by Jovanovic (1979), once we control for mismatch, there should be no relationship between business cycle conditions at the start of the job and job tenure.

A Model of Mismatch Cycles What forces can reconcile *(i)* the opposite behavior of mismatch for new hires from unemployment and job stayers, and *(ii)* the variation in job duration for different initial economic conditions? Because these findings are difficult to reconcile with current labor market theories, we provide a theoretical framework that gives a natural role to cyclical information frictions.

Building on Jovanovic (1979) and Moscarini (2005), we develop a model of learning about worker-firm mismatch augmented with fixed adjustment costs (it is costly to break and create relationships) and aggregate shocks. Each match between a worker and a firm is characterized by a skill mismatch level that is unobserved. This is defined by the difference workers' ability in a given skill, which is not know, and how intensive this skill is required by the job. The worker-firm pair uses Baye's law to learn about mismatch from a stream of signal realizations. In this setup, there are two drivers of the reallocation process. Workers who become sufficiently pessimistic about the quality of the match separate to unemployment and continue searching as unemployed. Workers can also leave in response to a new aggregate productivity shock. In particular, the worker selection policy takes the form of an inaction region that varies with uncertainty, as in Baley and Blanco (2018), and aggregate productivity. Regarding the former, as uncertainty decreases over time, the inaction region decreases as well, and only lower levels of perceived mismatch are tolerated. We show that newer, hence more uncertain, relationships are more likely to separate. This is because with higher uncertainty the volatility of the estimates is larger and overcomes the option effect embedded in a wider inaction region. Hence, the probability of separating from a job declines with worker's tenure as in Jovanovic (1979).

A key feature of the model is the interaction between aggregate productivity shocks and uncertainty. More specifically, in recessions lower aggregate productivity is accompanied a higher uncertainty. While there are several potential mechanisms that would generate countercyclical mismatch uncertainty, we consider uncertainty arising from occupational switching. In the model, an employed worker only learns about her abilities in the skills required by an occupation. Therefore, when a match starts in an new occupation than the one held previously, worker and firms are initially more uncertain about the mismatch. With large uncertainty, it is plausible that workers and firms create relationships that are perceived as highly mismatched but with high uncertainty surrounding their estimates. The mechanism that yields countercyclical uncertainty is that in recessions a higher fraction of unemployed job seekers switch occupations. We provide evidence supporting countercyclical uncertainty arising from occupational switching. First, we show that the likelihood of occupational switching for new hires is countercyclical. Second, keeping everything else constant, the separation hazard is higher for new hires from unemployment that switched occupation upon reemployment. Third, mismatch dispersion increases in recession. Lastly, in recessions the correlation between skill requirements and workers' abilities decreases for new matches, and this is true for all skills.

The model accounts for the documented empirical facts. First, following a negative productivity shock, mismatch becomes less tolerable; this destroys worker-firm matches with high levels of perceived mismatch. At the same, there are more matches with high levels of mismatch uncertainty. This countercyclical information friction reconciles the fact that in recessions, jobs with high mismatch are destroyed, while matches with high mismatch are created. Second, newer (more uncertain) relationships are more likely to separate. Because, on average, in matches that start in a recession, firms and workers have higher uncertainty about the mismatch level, the model reconciles the fact that matches that start in a recession are shorter.

2.1.1 Related Literature

Over the business cycle, economies face a large amount of reallocation: firms enter and exit, plants are built and destroyed, and workers change jobs and occupations. How do recessions affect resource allocation? This question has long attracted the attention of economists. According to the conventional Schumpeterian view (Schumpeter, 1939), recessions give rise to a more efficient allocation of resources by driving out less productive producers. However, several theories have contested this view. For instance, Barlevy (2002) argues that in recessions is harder for workers to reallocate from low to high productivity jobs and, thus, the misallocation of resources may increase; while Ouyang (2009) suggests that recessions may lower average productivity by increasing the exit of infant and potentially productive firms before they realize their potential.

This chapter contributes to this debate by showing that during recessions highly mismatched jobs are destroyed but also created, with the former being the dominant effect. In doing so, it draws on the empirical literature that examines the cyclical behavior of match quality and the scarring effects of recessions. Overall, this literature has concluded that job match quality is procyclical. Using a sample from the NLSY79 that runs from 1979 to 1988, the pioneering work by Bowlus (1995) proxies match quality with job duration and finds that job relationships that start in downturns are shorter, concluding that mismatch is countercyclical. Kahn (2010) and Oreopoulos et al. (2012) instead offer causal evidence that initial labor market conditions significantly affect wages. In particular, they show that recessions are associated with lower wages in the short- and long-term. Related to their work, Liua et al. (2016) provide evidence that lower wages from entering the labor market in a recession arise due to a lower likelihood of finding a job in an industry that is well-matched to the workers' field of study. A common feature of this literature is that its findings hinge on a sample of newly hired workers at a giving point in time. This means that it focuses on the flow of job creation and, therefore, ignores an alternative channel through which economic fluctuations may affect the allocation of workers into jobs: job destruction. We contribute to this literature by showing that both margins are important. In line with the existing evidence, we find that new hires from unemployment are more mismatched in bad times, yet we also find that for workers in ongoing job relationship mismatch decreases. This result suggests that in downturns highly mismatched jobs are created but also destroyed. Interestingly, we show that the latter effect is stronger.

An advantage of using Guvenen et al. (2018)'s skill mismatch index, when compared to wages or job duration, is that by characterizing separately workers and jobs in terms of their abilities and requirements along a set of skills, it allows us to assess the behavior of the mismatch distribution. Given this, we move this literature a step forward by showing first that recessions affect mainly highly mismatched workers, and second that new hires from unemployment are both more over- and under-qualified increase in recessions, whereas for job stayers the decrease in mismatch is only driven by under-qualification. So, matches at most risk of separation during an economic contraction are the ones between "low-type" workers and "high-type" jobs. This finding is consistent with recent work by Lise and Robin (2017). Using an empirical application of a tractable model of equilibrium search with two-sided ex ante heterogeneity and aggregate shocks, they show that when the economy enters a recession, low-type workers are fired, particularly those matched with high-type firms.

Our work also relates to previous contributions searching for identification and quantification of sorting in labor markets using realized wages. Conclusions in this literature are mixed. On the one hand, Eeckhout and Kircher (2011) argue that using wage data alone we cannot identify the sign of the sorting, but we can identify its strength. On the other hand, Hagedorn et al. (2017) show that assortative matching is indeed recoverable with wages and, by applying their framework to matched employer-employee data from Germany, they find signs of strong positive assortative matching. Closely related to this chapter, Crane et al. (2018) rank workers and firms using wages and provide evidence that higher unemployment is associated with *(i)* a shifting in the employment distribution towards high productivity workers, and *(ii)* a stronger positive correlation between worker ranks and firm ranks. Our results are consistent with their findings. Nonetheless, our approach and findings are complementary to theirs. Unlike these previous studies, we build on recent contributions that exploit the unique properties of NLSY79 and O*NET (Guvenen et al., 2018; Lise and Postel-Vinay, 2016), and rank workers and occupations using data on workers' abilities and the skill requirements of occupations instead of wages. Moreover, by relying on these datasets, we are able to use a definition of mismatch that is multidimensional: while in this literature workers and firms are heterogeneous in one-dimension, we characterize them in 4 different skill dimensions. The importance of this feature has been recently highlighted by Lindenlaub and Postel-Vinay (2016), who show in simulation exercises that approximating workers' and jobs' true multidimensional characteristics by a one-dimensional index in empirical applications may translate into quantitatively biased results about the extent of mismatch. Overall, we add to this literature by showing, through a different identification strategy, that there is a positive correlation between workers and occupations and, more importantly, that this correlation varies cyclically, being stronger in recessions for job stayers but weaker for new hires from unemployment.

Finally, this chapter contributes to a body of literature that has built on Jovanovic (1979)'s idea that firms and workers learn about the quality of the match as it is experienced. An example is Pries (2004) who shows that when interacted with exogenous shocks (separation probability), this mechanism can explain high job finding rate combined with a persistent high unemployment rate. More recently,Borovickova (2016) interacts firm productivity shocks with unobservable match quality to explain the fact that long-tenure workers in growing firms have a higher separation rate. A common feature to this literature is that the learning experience about mismatch is the same over the cycle. In our model, learning experiences vary in the cycle as during recessions average prior uncertainty rises due to a larger fraction of matches with high uncertainty driven by more occupational switching.

While there is a vast literature on uncertainty shocks, this literature is still scant with respect to the role of uncertainty on labor market flows. Few examples are (Lin, 2014), Leduc and Liu (2016), Pries (2016) and Schaal (2017). Pries (2016) models uncertainty as a noisy component in a firm's initial signal about job productivity when they are considering creating a new job. In this setting, a more noisy environment decreases job creation because firms face a higher risk of making a mistake in deciding to create a job that will turn out to be unprofitable. However, he assumes that job's productivity is fully revealed after the firm makes the start-up investment and it does not capture the decreasing hazard of separation as our model does. Schaal (2017) shows that time-varying idiosyncratic risk is important to explain the magnitude of fluctuations in aggregate unemployment for past US recessions. None of these papers explore the implications of changes in the uncertainty distribution for the learning process of match quality. We fill this gap in the literature.

The remainder of this chapter proceeds as follows. The next section presents the empirical strategy and the estimation results. Section 2.3 builds a model of learning about skill mismatch that is able to account for our empirical findings, and Section 2.4 concludes.

2.2 Empirical Facts: Mismatch and Job Duration over the Cycle

In this section we document a set of new empirical facts about the cyclical behavior of worker-occupation mismatch and job tenure. Using a worker-level panel from the NLSY79 combined with occupation-level data from O*NET and aggregate data on unemployment, we find that *(1)* mismatch is procyclical for workers in ongoing job relationships, but countercyclical for new hires from unemployment; *(2)* for job stayers, the decrease in mismatch is driven by negative mismatch, while for new hires from unemployment both positive and negative mismatch increase in downturns; *(3)* conditional on mismatch, a match that starts in a recession has a shorter duration; and *(4)* mismatch decreases job duration, and this effect is increasing in the level of current unemployment.

2.2.1 Measuring Skill Mismatch

To measure the extent to which a worker's skills are aligned with the skills required by a job, we adopt the skill mismatch index recently developed by Guvenen et al. (2018). This measure is defined by the difference between a worker's abilities in different skills and the intensity these skills are required by a job. We interpret this measure as the *lack* of match quality: the larger is this difference, the lower is the quality of a match.[2]

Skill Mismatch Index Consider that jobs and workers are characterized by J skill dimensions, $j = \{1, .., J\}$. Let $a_{i,j}$ be the measured score of worker i's ability in skill dimension j, and $r_{c_t,j}$ be the measured score of the job requirement in skill dimension j by the occupation held at time t, c_t. At a given point in time, the mismatch between individual i and his occupation c_t is measured as the weighted sum of the absolute value of the difference between the worker's skills and the skill requirements in each dimension:

$$m_{i,c_t} \equiv \sum_{j=1}^{J} \omega_j \, |a_{i,j} - r_{c_t,j}|, \qquad \sum_{j=1}^{J} \omega_j = 1, \tag{2.1}$$

where ω_j is is the weight of skill j. In the empirical analysis, $a_{i,j}$ will correspond to the the percentile rank of worker i's ability in skill dimension j and $r_{c_t,j}$ to the percentile rank of the job requirement in skill dimension j. Therefore, m_{i,c_t} ranges between 0 and 100, with 0 indicating a perfect match between a worker's abilities and the job skill requirements, and 100 the highest level of mismatch.[3]

Asymmetric Mismatch A worker is said to be mismatched at time t whenever m_{i,c_t} is higher than zero. This may happen either because in one of the J dimensions, the worker is under-qualified, i.e. has a level of ability that falls short of the

[2]Because mismatch is negatively correlated with wages, Guvenen et al. (2018) argue that their mismatch index can be interpreted as a signal of the lack of job match quality. In the next chapter, I replicate this result.

[3]In the empirical analysis, we set equal weights for all dimensions. Guvenen et al. (2018) instead use the factor loadings from the first principal component. The results are robust to using different weighting strategies.

job's skill requirement ($a_{i,j} < r_{c_t,j}$), or because he is over-qualified, meaning that his abilities are higher than the job's skill requirements ($a_{i,j} > r_{c_t,j}$). Given this, m_{i,c_t} can be decomposed into a measure of over-qualification, positive mismatch m^+_{i,c_t}:

$$m^+_{i,c_t} \equiv \sum_{j=1}^{J} \omega_j \max\{a_{i,j} - r_{c_t,j}, 0\}, \tag{2.2}$$

and a measure of under-qualification, negative mismatch m^-_{i,c_t}:

$$m^-_{i,c_t} \equiv \sum_{j=1}^{J} \omega_j \left|\min\{a_{i,j} - r_{c_t,j}, 0\}\right| \tag{2.3}$$

Relevant Skills To measure mismatch, we need to define the set of relevant skills. Empirical evidence suggests that both cognitive and non-cognitive abilities have important implications for labor market outcomes, namely for wages and occupational choice (see, for example, Heckman, Stixrud and Urzua, 2006; Lindqvist and Vestman, 2011). Therefore, for our empirical exercise, we compute the skill mismatch index in Equation (2.1) and its decomposition, Equations (2.2) and (2.3) using workers' and jobs' scores in four different skill dimensions (J=4). In particular, we capture cognitive skills through individual's abilities in verbal, math and technical skills, and non-cognitive skills through a social dimension. A similar definition of skills has been adopted by several recent papers studying the education and labor market effects of different worker abilities (Boehm, 2015; Guvenen et al., 2018; Prada and Urzúa, 2016; Lise and Postel-Vinay, 2016; Speer, 2017). Note that to capture cognitive skills, Guvenen et al. (2018) use only the verbal and math dimensions. These are the two components of the Armed Forces Qualifications Test (AFQT), a score that has been extensively used as proxy of cognitive ability in the literature. We add the technical component because it has been shown to be an important determinant of labor market outcomes (Prada and Urzúa, 2016). Our results are robust to using the 3-skill mismatch index.

2.2.2 Data

Our empirical exercise relies on a worker-level panel from the NLSY79 combined with occupational-level data from O*NET and aggregate data on business cycle conditions.

NLSY79 The NLSY79 is a nationally representative longitudinal survey whose respondents were between the ages of 14-22 when they were first interviewed in 1979, and have been followed through the present. We focus on a sub-sample of males and females from the NLSY79 cross-sectional sample, which includes 2,991 individuals and runs from 1979 and 2012. The complete description of our sample selection criteria is in Appendix 2.5.1. Using the Work History Data File, we construct a monthly frequency panel reporting information on individuals' labor market history, including wages, occupation and industry for each employment spell. Although individuals may hold more than one job, we focus on the main job. As standard in the literature, we define a main job as the one at which an individual spends the most hours working within a given month. With this panel at hand, we are able to identify job-to-job transitions and transitions from non-employment to employment.

Besides reporting the labor market history of each respondent, the NLSY79 also has information on the Armed Services Vocational Aptitude Battery (ASVAB) test scores. Following Guvenen et al. (2018), these scores are used to measure individual ability in each skill dimension ($a_{i,j}$). The procedure to convert the different components of the test into ability scores for the four skills is described in Appendix 2.5.1. $a_{i,j}$ corresponds to the percentile of the distribution of ability in skill dimension j across individuals in our sample. Panel A in Table 2.A5 reports the correlation of workers' abilities across skill dimensions. The observed pattern suggests that workers with high abilities in one skill dimension tend to have high ability in the other three, in line with Guvenen et al. (2018) and Lise and Postel-Vinay (2016). Appendix 2.5.1 contains further details on the construction of the panel and the methodology used to measure abilities in each skill dimension.

O*NET An occupation is defined by Dorn (2009)'s three-digit occupational classification system. This has the advantage of being consistent over time. Examples of occupations in our sample are nurses, teachers, lawyers or bartenders. For each occupation, we compute the skill requirements in each dimension ($r_{c_t,j}$) using O*NET. This is a database that describes occupations in terms of skill and knowledge requirements.[4] The importance score of 32 out of the 277 descriptors provided by O*NET are transformed into skill requirements employing Guvenen et al. (2018)'s methodology, as detailed in Appendix 2.5.1.[5] $r_{c_t,j}$ corresponds to the percentile of the skill requirement distribution across all occupations in O*NET. To check whether the constructed variables characterize occupations reasonably, Table 2.A4 presents the percentile rank scores for selected occupations. For instance, economists require the use of the math skill more intensively, whereas lawyers require a higher the use of the verbal skill and elevator installers require mostly technical skills. These scores are consistent with the ones presented by Speer (2017) and Lise and Postel-Vinay (2016). The occupation skill requirements are merged with the worker-level panel using Dorn (2009)'s three-digit occupational codes. Panel B in Table 2.A5 presents the correlation between occupation skill requirements and workers' ability in each skill dimension. We observe that while workers tend to select themselves into jobs that fit their skill bundles best, sorting is far from perfect.

[4]The O*NET database is the successor to the Dictionary of Occupational Titles, which classified the types of tasks necessary to work in a particular occupation. The O*NET expands upon this, by providing quantitative information on several descriptors that are organized into 9 broad categories: skills, abilities, work activities, work content, experience/education level required, work values, job interests, knowledge and work styles. The scores for each descriptor are built using questionnaires that ask workers to rate their own occupation in terms of a subset of the O*NET descriptors, and a survey of occupation analysts who are asked to rate others descriptors. More information is available at http://www.onetcenter.org.

[5]We follow Guvenen et al. (2018) and use 26 descriptors that were considered by the Defense Manpower Data Center to be most related to the ASVAB component tests and another 6 descriptors related to the social skills. The descriptors used are the following: oral comprehension, written comprehension, deductive reasoning, inductive reasoning, information ordering, mathematical reasoning, number facility, reading comprehension, mathematics skill, science, technology design, equipment selection installation, operation and control, equipment maintenance, troubleshooting, repairing, computers and electronics, engineering and technology, building and construction, mechanical, mathematics knowledge, physics, chemistry, biology, english language, social perceptiveness, coordination persuasion, negotiation instructing, service orientation.

Business Cycle Conditions The primary indicator of the state of the economy is the unemployment rate, measured by the monthly, national, civilian unemployment rate for ages 16+ obtained from the Bureau of Labor Statistics (BLS). This is a widely accepted proxy of macroeconomic shocks. We also investigate the robustness of the results to alternatives measures of economic conditions such as the difference of the unemployment rate from its Hodrick-Prescott filter, the composite Help-Wanted Index developed by Barnichon (2010), a proxy for the number of job openings at a given point in time, and the Industrial Production Index.

We summarize the main variables used in the empirical analysis in Table 2.A2.

2.2.3 Empirical Framework

To formally examine the dynamics of mismatch over the business cycle, we focus on the set of existing matches at time t, and estimate the following regression:

$$m_{i,c_t} = \beta_0 \ + \ \beta_1 U_t \ + \ \gamma' x_{i,t} \ + \ \delta_i \ + \ \delta_m \ + \ \delta_y \ + \ \varepsilon_{i,t} \tag{2.4}$$

where m_{i,c_t} is the mismatch level of worker i in the occupation held at time t, $x_{i,t}$ is a set of individual controls, which includes the region of residence, occupation, industry, and a quadratic polynomial in age; U_t is the demeaned aggregate unemployment rate in month m and year y; δ_i, δ_m, and δ_y are individual, monthly and yearly fixed effects, respectively; and $\varepsilon_{i,t}$ is the error term, which includes all unobserved determinants of mismatch for worker i at time t. Standard errors are clustered at the individual level to allow for serial correlation.[6] We run separate regressions for total (m_{i,c_t}), positive (m^+_{i,c_t}) and negative (m^-_{i,c_t}) mismatch, and also for mismatch in each skill dimension separately, to examine whether sensitivity to the business cycle is heterogeneous across skills.

Under the standard exogeneity restrictions, the effect of macroeconomic conditions in month m and year y on the level of mismatch of existing worker-job

[6]By clustering the standard errors at the individual level, observations may be correlated within each individual, but must be independent across individuals. However, common shocks such as the business cycles may induce correlation between individuals at a moment in time. The results are robust to the inclusion of standard errors clustered at the month-year level and double clustered at the month-year and individual level instead.

matches is identified by β_1. If $\beta_1 > 0$, mismatch increases during downturns, i.e. mismatch is countercyclical, consistent with the *sullying effect* of recessions. Otherwise, if $\beta_1 < 0$, mismatch is procyclical, in line with the *cleansing effect* of recessions.

Identification OLS estimation of Equation (2.4) hinges on a sample of individuals that are employed in month m and year y. Thus, we face an endogeneity problem related to the decision of working. If the distribution of the workers' skills changes systematically with the business cycle for other reasons not related to macroeconomic shocks, that could generate a positive (negative) association between economic conditions and mismatch, which would be mistakenly interpreted as mismatch being procyclical (countercyclical). For instance, Solon et al. (1994) showed that the pool of employed shifts towards high-ability workers in expansions. To tackle the problem of selection into employment, our empirical strategy exploits within-individual variation in business cycle conditions across months the individual is reported to be employed. Under the assumption that the selection process across individuals is constant over time, the inclusion of individual fixed effects restores the orthogonality condition violated by the operation of the selection process. OLS estimates of β_1 further control for unobserved factors that may influence mismatch and are associated with the region of residence, industry and occupation (at the one-digit level[7]) as well as for time varying shocks affecting all individuals and individuals' age. The later is particularly important as mismatch decreases non-linearly with age, as shown in Figure 2.A1.

2.2.4 Mismatch Dynamics over the Cycle

Panel A in Table 2.1 presents OLS estimates of Equation (2.4). Column 1 shows a negative relationship between economic conditions and mismatch, i.e mismatch is procyclical. This is consistent with the *cleansing effect* of recessions as suggested

[7]NLSY79 reports industry codes in the Census 1970 Industry Classification Code before 1994, in the Census 1980 Industry Classification Code for the year 1994, and in the Census 2000 Industry Classification Code after 2000. We use Guvenen et al. (2018)'s crosswalk to convert them into the Census 1970 one-digit industry code, and use these one-digit level codes to create industry dummies.

by the Mortensen and Pissarides (1994)'s matching model with endogenous separations: in downturns bad matches are destroyed and only high quality matches are created, thus mismatch decreases. Our coefficient of interest (β_1) is robust to the inclusion of *(i)* time and month fixed effects (column 2), so as to control for year and monthly shocks affecting all individuals, *(ii)* region fixed effects (column 3), and *(iii)* occupation and industry controls (column 4), this alleviates concerns over the change in the composition of occupations and industry over the cycle. The latter is our preferred specification. To illustrate the magnitude of the point estimates, an increase in the unemployment rate from the 50^{th} to the 90^{th} percentile is associated with a 1.65% decrease in mismatch, on average.

Digging deeper, columns 5 and 6 of Panel A show that mismatch between workers and jobs diminishes because workers are less under-qualified for the job: for positive mismatch (column 2), our measure of over-qualification, the estimated coefficient on the unemployment rate is statistically insignificant, while for negative mismatch (column 3), that captures under-qualification, the coefficient is negative and statistically significant. These findings suggest that following a negative productivity shock, the matches which are at most risk of separation are those between high-type jobs and low-type workers, and workers in ongoing job-results are less under-qualified.

Panel A in Figure 2.1 compares the evolution of the aggregate unemployment rate to a measure of mismatch overlaid on shaded areas indicating NBER recession periods. The mismatch series represents the average partial residuals:

$$M_t = \frac{1}{N}\sum_{i=1}^{N} \hat{\varepsilon}_{i,t} + \hat{\beta}_1 U_t \tag{2.5}$$

with $\hat{\varepsilon}_{i,t}$ and $\hat{\beta}_1$ being the residuals and the coefficient estimated from Equation (2.4). M_t captures the relationship between unemployment and mismatch, keeping all other variables in the model constant. Two important facts can be drawn from that picture. First, as shown in Table 2.1, mismatch is procyclical. Second, mismatch decreases as recessions unfold. In particular, a typical recession encompasses periods in which mismatch is decreasing from its high point to its low point, while unemployment increases from its low point to its high point.

Table 2.A6 show estimation results when we use as dependent variable total, positive and negative mismatch in each skill dimension separately. Interestingly, we find that the decrease in total and negative mismatch following an increase in unemployment is driven by the decrease in mismatch along the math and technical skills. In contrast, for the verbal and social skills, the estimated coefficient on the unemployment rate is statistically insignificant and substantially smaller. A potential explanation for the difference in the cyclical behavior of the different skills may be that math and technical skills are easier to observe by employers, while verbal and social skills capture more subtle traits.

Heterogeneity by Previous Employment Status The set of existing matches at time t is composed of worker-job pairs that were created before t and those that were created at time t. By estimating Equation (2.4), we are implicitly assuming that macroeconomic shocks affect mismatch of workers in on-going job relationships and new hires equally. However, recessions affect the allocation of workers to jobs both through *job destruction* and *job creation*, and the effect along these two margins is potentially different. On the one hand, the *cleansing* hypothesis suggests that recessions are times of productivity-enhancing reallocation. Bad matches are destroyed and only good matches are created, meaning that mismatch decreases both for workers in on-going relationships and new hires. On the other hand, the *sullying* hypothesis tell us that recessions distort the reallocation of resources, increasing mismatch both for job stayers and new hires: Barlevy (2002) argues that workers climb the job ladder more slowly, hence get stuck in poor matches, while Moscarini (2001) predicts that unemployed job seekers accept less desirable jobs due to higher competition. Consistent with the latter, the empirical literature that uses indirect measures of mismatch, such as earnings (Kahn, 2010; Oreopoulos et al., 2012) or job duration (Bowlus, 1995; Baydur and Mukoyama, 2015), and focus only on the flow of new matches suggests that match quality is procyclical.

To better understand what is behind the evidence of mismatch procyclicality shown in Panel A of Table 2.1, it is then important to isolate the mismatch behavior of new hires from that of workers in on-going job relationships. To do so, we estimate a version of Equation (2.4) that allows for a separate new hire effect for workers coming from non-employment and workers making direct job-to-job transitions:

$$m_{i,c_t} = \beta_0 \ + \ \beta_1 U_t \ + \ \beta_2 EE'_{i,t} \ + \ \beta_3 UE_{i,t} \ + \ \beta_4 (U_t \cdot EE'_{i,t}) \\ + \ \beta_5 (U_t \cdot UE_{i,t}) \ + \ \gamma' x_{i,t} \ + \ \delta_i \ + \ \delta_y \ + \ \delta_m \ + \ \varepsilon_{i,t}, \qquad (2.6)$$

where $UE_{i,t}$ equals one for workers with an intervening spell of non-employment at t, meaning that the worker was not working at time $t-1$[8] but reported to be employed at time t; and $EE'_{i,t}$ is a dummy for whether the worker i is making a job-to-job transition at t, which we define to be a situation where the worker was employed at time $t-1$ and t, but with a different employer. Over the period from 1974 to 2012, individuals in our sample start a new job 13.77 times, on average. Around 44% of these changes are job-to-job transitions and the remaining correspond to transitions from non-employment to employment. Appendix 2.5.1 provides further details on how we identify each transition.[9]

Our coefficients of interest are β_1, which measures mismatch cyclicality for job stayers; β_4 and β_5 which correspond, respectively, to the differential in mismatch cyclicality between job stayers and job-to-job transition and between job stayers and new hires from unemployment. In light of the results in Panel A of Table 2.1, which suggest that $\beta_1 < 0$, mismatch of new hires coming from employment and coming from unemployment is countercyclical if $\beta_1 + \beta_4 > 0$ and $\beta_1 + \beta_5 > 0$, respectively.

Panel B in Table 2.1 provides points estimates of the coefficients of interest. Three results stand out. First, mismatch is, on average, larger for new hires from unemployment (across all specifications estimated in columns 1 to 4 we observe that $\beta_3 > \beta_2 > 0$). Second, an increase in unemployment is associated with a decrease in mismatch for job-stayers, but with an increase in mismatch for new hires from unemployment: β_1 remains negative, and the sum $\beta_1 + \beta_5$ is positive and statistically different from zero, as reported in Panel C. Interestingly, our results

[8]We define a worker to be in non-employment if she reported to be not working, unemployed or out of the labor force.

[9]Transitions from non-employment to employment include *recalls*, workers that return to their previous employer after a jobless spell. For robustness, we also consider different measures of what constitutes a new hire from non-employment.

show that for job-to-job transitions mismatch is acyclical, i.e. the sum $\beta_1 + \beta_4$ is not statistically different from zero (Panel C, Table 2.1). Third, columns 5 and 6 show that while the decrease in mismatch for job stayers is only driven by negative mismatch, i.e. as unemployment increases, workers in going-job relationships are less under-qualified, the increase in mismatch for new hires from unemployment is supported by an increase in both over- and under-qualification.

Panel B in Figure 2.1 plots aggregate unemployment rate against mismatch of job stayers and new hires, along with shaded areas that correspond to NBER recessions. This plot strongly suggests that NBER recession periods coincide with a decrease in mismatch for workers in ongoing job relationships, as proposed by the *cleansing* hypothesis, and an increase in mismatch for new hires, in line with the *sullying* effect of recessions. The latter reaches its peak at the end of the downturn. Interestingly, the recession that followed the 2007-2008 collapse of the financial markets seems to be different from the previous ones, with mismatch for job stayers and new hires decreasing during the downturn.

Table 2.A7 reports the results for mismatch in each skill separately. We find that for workers in ongoing job relationships the decrease in mismatch in downturns is mainly due to a decrease in mismatch in the math and technical skill dimensions, while for new hires from unemployment mismatch increases across all skills. One potential interpretation for this difference is that for workers in ongoing job relationships, firms know the worker's abilities in the skills that are easily observable, such as math, and therefore fire those workers that are more mismatch in those dimensions. In contrast, for new hires from unemployment, nor the firms or the workers have less information about each others characteristics, and therefore, mismatch is higher across all skills.

Overall, our findings suggest that even though, on the aggregate, the cleansing effect dominates the sullying effect, both are present in the data. In particular, an increase in the unemployment rate from the 50$^{\text{th}}$ to the 90$^{\text{th}}$ percentile, is associated with a 1.86% decrease in mismatch for workers in ongoing job relationships, and a 2.56% increase in mismatch for new hires from unemployment, on average.[10] To sum up, during economic contractions highly mismatched jobs are

[10] I emphasize that this exercise is only meaningful under the assumption that the estimates reflect a causal relationship.

Table 2.1: Mismatch and the Business Cycle

	m_{i,c_t}				m^{+}_{i,c_t}	m^{-}_{i,c_t}
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: No Heterogeneity						
(β_1) Unemployment$_t$	-0.191***	-0.138***	-0.139***	-0.141***	-0.041	-0.099***
	(0.044)	(0.051)	(0.051)	(0.050)	(0.037)	(0.035)
Observations	510788	510788	510788	510788	510788	510788
Adjusted R^2	0.488	0.490	0.490	0.500	0.771	0.763
Panel B: Heterogeneity						
(β_1) Unemployment$_t$	-0.206***	-0.158***	-0.158***	-0.159***	-0.050	-0.109***
	(0.044)	(0.051)	(0.051)	(0.050)	(0.038)	(0.035)
(β_2) EE'$_{i,t}$,	0.244*	0.236*	0.238*	0.245*	0.159*	0.086
	(0.135)	(0.135)	(0.135)	(0.133)	(0.096)	(0.096)
(β_3) UE$_{i,t}$,	0.669***	0.457***	0.459***	0.412***	0.474***	-0.062
	(0.142)	(0.139)	(0.139)	(0.135)	(0.096)	(0.092)
(β_4) EE'$_{i,t}$, $\times$ Unemployment$_t$	0.169*	0.147	0.146	0.147	0.072	0.074
	(0.094)	(0.094)	(0.094)	(0.093)	(0.068)	(0.064)
(β_5) UE$_{i,t}$, $\times$ Unemployment$_t$	0.417***	0.411***	0.408***	0.383***	0.168***	0.216***
	(0.083)	(0.086)	(0.086)	(0.084)	(0.061)	(0.056)
Panel C: Mismatch Cylicality						
($\beta_2 + \beta_4$) EE'$_{i,t}$	-0.036	-0.011	-0.012	-0.012	0.022	-0.035
	(0.100)	(0.101)	(0.101)	(0.099)	(0.075)	(0.070)
($\beta_3 + \beta_5$) UE'$_{i,t}$	0.211**	0.253***	0.249***	0.224**	0.118*	0.106*
	(0.088)	(0.092)	(0.092)	(0.091)	(0.067)	(0.060)
Observations	510788	510788	510788	510788	510788	510788
Adjusted R^2	0.488	0.490	0.490	0.500	0.771	0.763

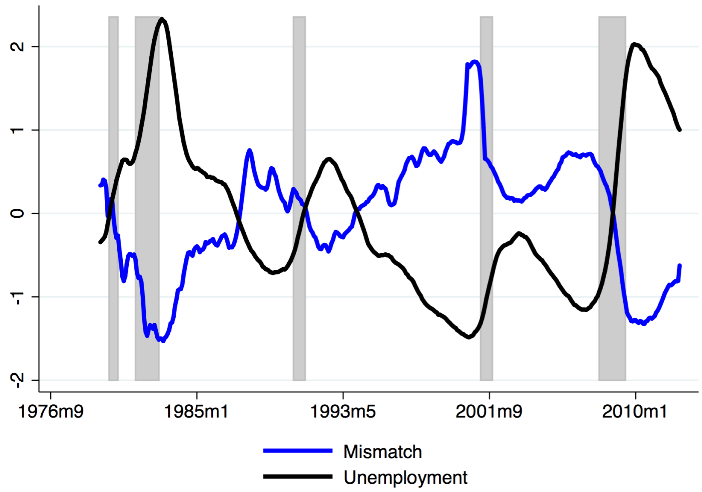
Notes: The table reports coefficients from an OLS regression with robust standard errors clustered at the individual level reported in parentheses. Panel A and B report, respectively, estimation results of Equation (2.4) and (2.6). Panel C reports statistics for $\beta_1 + \beta_4$ and $\beta_1 + \beta_5$ in Equation (2.6). All columns include a quadratic polynomial in age and individuals fixed effects. The sample includes all worker-job matches between 1979 and 2012. ***, ** and * represent statistical significance at 1%, 5% and 10% levels, respectively.

destroyed but also created.

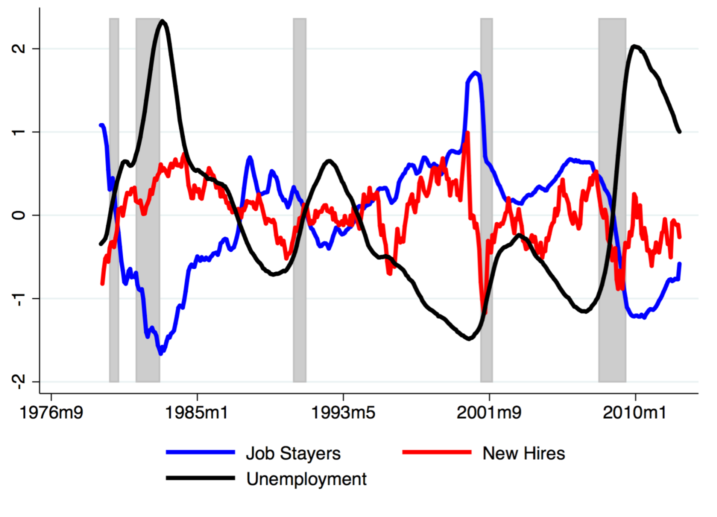
Heterogeneity along the Mismatch Distribution Our results so far show that the relationship between aggregate unemployment and *average* mismatch is negative for job stayers but positive for new hires from unemployment. However, the magnitude of this correlation may be different at the different locations of the mismatch distribution. Naturally, one could expect worker-job relationships char-

Figure 2.1: Aggregate Mismatch and Unemployment

A. Mismatch



B. Heterogeneity by Employment Status



Notes: Data are shown in standard deviations. *Unemployment* is the aggregate civilian unemployment rate at a monthly frequency. In Panel A, *Mismatch* displays the average partial residuals, $M_t = \frac{1}{N}\sum_{i=1}^{N}\hat{\varepsilon}_{i,c_t} + \hat{\beta}_1 U_t$, estimated from Equation (2.4). In Panel B the blue line corresponds to mismatch of job stayers and the red line to mismatch for new hires. Shaded areas correspond to NBER recessions. Source: NLSY79, BLS, NBER and authors' calculations.

acterized by high levels of mismatch to be more at risk of separation in bad times when compared to perfectly matched workers. To understand how the relationship between mismatch and economic conditions changes along the distribution of mismatch, we measure the correlation between the unemployment rate and mismatch at different quantiles of the mismatch distribution conditional on the explanatory variables.

$$Q_\theta(m_{i,c_t}|U_t, EE'_{i,t}, UE_{i,t}, x_{i,t}) = \beta_0^\theta + \beta_1^\theta U_t + \beta_2^\theta EE'_{i,t} + \beta_3^\theta UE_{i,t} + \beta_4^\theta(U_t \cdot EE'_{i,t}) + \beta_5^\theta(U_t \cdot UE_{i,t}) + \delta^{\theta\prime} x_{i,t} \tag{2.7}$$

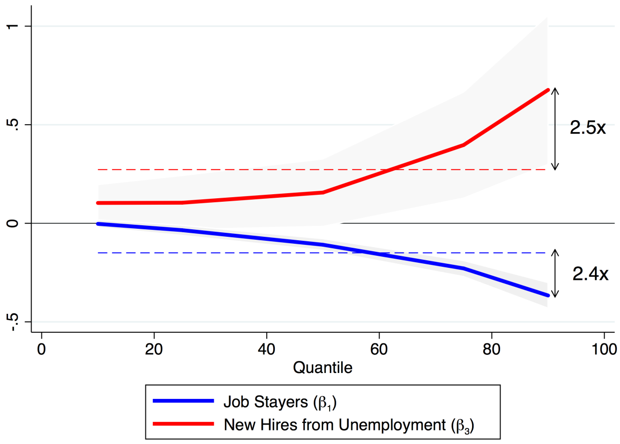
where $\theta \in (0,1)$ and $x_{i,t}$ is a quadratic polynomial in age.[11] Figure 2.2 plots the OLS estimates of the conditional mean effect (dashed lines), and the quantile regression estimates for θ ranging from the 10^{th} to the 95^{th} quantile (solid lines). The shaded gray areas depict a 95% pointwise confidence band for the quantile regression estimates. We can observe that for job stayers (these correspond to the blue lines in the graph) an increase in unemployment is associated with lower mismatch, nonetheless, the quantile estimations show that this correlation is much weaker in the lower quantiles of the mismatch distribution and considerably larger in the upper tail of this distribution. In particular, for workers at the 95^{th} quantile, the correlation between unemployment and mismatch is negative and 2.4 times larger than the OLS estimate; whereas for the lowest mismatched workers, those located at the 10^{th} quantile, this correlation is not statistically different from zero. The same happens for new hires from unemployment. Given this, we can conclude that business cycle conditions affect mainly the upper tail of the mismatch distribution, i.e. those workers that are poorly matched.

2.2.5 Robustness Checks

This section provides evidence showing our results are robust across different specifications. First, Figures 2.A2 and 2.A3 show that our results hold across different industries and occupations at the one digit level. This suggests that the doc-

[11] Because the OLS estimate of β_1 has a small variation when we exclusion individual, industry and occupation fixed effects, for the purpose of the quantile regression analysis we do not include them.

Figure 2.2: Quantile and OLS Estimates



Notes: The dashed lines display point estimates based on OLS estimation of Equation (2.6), the solid lines correspond to the quantile regression estimates based on Equation (2.7) for θ ranging from the 10^{th} to the 95^{th} quantile. Shaded areas correspond to the latter 95% confidence intervals.

umented facts are not driven by any industry or occupation in particular. Next, we explore the sensitivity of our results to *(i)* different definitions of new hires from non-employment to address concerns regarding short jobless spells, *(ii)* changes in the composition of occupations over the cycle, *(iii)* expanding the set of controls, *(iv)* using alternative measures of economic conditions, *(v)* a reduced sample to include only males and the years before the Great Recession, and finally *(vi)* using different methods to compute the mismatch index.

Redefining new hires from non-employment In the baseline case presented in Table 2.1, we used the broadest definition of new hires from non-employment: independent of how long the unemployment spell was, all workers who did not reported a job at time $t-1$ (i.e reported to be not working, unemployed or out of the labor force) and are working at time t were considered to be new hires from unemployment. This definition includes $recalls$ — workers that return to their previous employer after a jobless spell — as new hires from non-employment.[12] The concern with using a broad definition that does not take into account unem-

[12]Using data from the Survey of Income and Program Participation, **?** document that in the US over 40% of the employed workers who separate into unemployment return, after the jobless spell, to their last employer.

ployment duration is that new hires from unemployment with short jobless spells may be in fact be job-changers taking a short break between jobs. To address this issue, we recode workers with jobless spells equal or smaller than 1, 2 and 3 months as workers making job-to-job transitions instead of transitions from non-employment to employment. Under these new definitions, $recalls$, i.e. those workers that return to their previous employer within 1, 2 and 3 months, are redefined as job stayers. We can observe in Table 2.A8 that our results are robust to these different definitions.

Job skill requirements over the cycle We observe that in recessions mismatch increases for new hires. This pattern could be driven by worst labor market conditions (Moscarini, 2001), but also by a systematic change in the skill supply and demand distribution over the business cycle. If there is a substantial difference between the skills that are needed and those that are supplied by unemployed job seekers, then the labor market can only clear under considerable mismatch, even in the absence of search frictions (Lindenlaub, 2017). By exploiting within-individual variation, we keep the skill supply distribution fixed. In addition, by controlling for one-digit level occupations, we mitigate concerns over the change in the skills demanded over the cycle. In addition to this, I now estimate a version of Equations (2.4) and (2.6) with occupation-year fixed effects, so as to control for shocks affecting an occupation in a given year. Table 2.A9 shows our results remain unchanged. Thus, we can conclude that the increase in mismatch is not driven by an increase in the difference between the skills demanded and the skills supplied.

Additional controls We investigate whether the results are robust to introducing additional controls. In particular, we introduce in Equations (2.4) and (2.6) *(i)* lagged unemployment rate, U_{t-1}, *(ii)* wages, measured using the log real hourly earnings, and *(iii)* a quadratic polynomial of labor market experience instead of age. Table 2.A10 shows that expanding the set of controls with the unemployment rate in the previous month as well as with wages and labor market experience has a small effect on the estimates.

Alternative measures of the state of the economy We replicate the estimation of Equations (2.4) and (2.6) using three different measures of business cycle conditions at the national level. First, we use the composite Help-Wanted Index developed by Barnichon (2010). This measure of vacancy posting captures the behavior of total — print and online — help-wanted advertising, thus is an important alternative indicator of labor market conditions. Second, we use the Industrial Production Index published by the Federal Reserve Board. This economic indicator measures real output for manufacturing, mining, and electric, and gas utilities at a monthly frequency. Finally, we replace demeaned unemployment by its deviations from the Hodrick-Prescott (HP) filtered unemployment rate. As shown in Panel A of Table 2.A11, an increase in vacancy postings, i.e. an improvement in economic conditions, is associated with (i) an increase in mismatch between worker's abilities and job skill requirements (column 1), and (ii) an increase in under-qualifications (columns 5). We also find important heterogeneity along previous employment status: for job stayers, mismatch increases when the number of vacancies increases; but for new hires from unemployment, mismatch decreases; and for new hires from employment, mismatch is acyclical (column 2). We observe the same pattern when using the Industrial Production Index, as shown in Panel B of Table 2.A11. Finally, Panel C shows that our estimates remain barely unchanged when we use deviations of unemployment from the HP filtered unemployment rate. Thus, we conclude that the main results are robust to using alternative state variables.

Great Recession The sample used in the empirical analysis covers the period between 1979 and 2012, which includes the period of the Great Recession. According to the National Bureau of Economic Research, the Great Recession began in December 2007 and ended in June 2009. To assess the extent to which our results are driven by this particular period, we re-estimate Equation (2.4) and (2.6) restricting the sample to the years before this period, 1979-2006. We can observe in Table 2.A12 that the sign and significant of the estimated coefficients are the same as the ones reported in Table 2.1 and that their magnitude is relatively unchanged. Hence, the documented pattern hold if we exclude the Great Recession.

Only males The OLS estimates shown in Table 2.1 hinge on a sub-sample of males and females from the cross-sectional sample of the NLSY79. However, most of the studies on match quality and wage cyclicality restrict the sample to males. First, because it is a more homogeneous group. Second, because of the sharp transitional dynamics of female participation in the past decades. To check that our results are not driven by the sample of females, Table 2.A13 provides OLS estimates of Equations (2.4) and (2.6) using a sample restricted to males, and shows that our key findings remain unchanged.

Alternative mismatch measures In the empirical analysis, we measured mismatch as an unweighted average of the mismatch along 4 skill dimensions (math, verbal, social and technical) and that uses factor analysis to identify the set of underlying factors used to compute the skill scores, as in Guvenen et al. (2018). Table 2.A14 shows that our findings are robust to four different versions the mismatch index. First, Panel A uses a mismatch index taking into account only 3 skill dimensions (math, verbal and social) as in Guvenen et al. (2018). Second, estimates in Panel B rely on a mismatch index that is a weighted average of the mismatch along 3 skills, in which I use the same weights as in Guvenen et al. (2018): (verbal, math,social) = (0.43, 0.43, 0.12). Third, we use a mismatch index that follows Speer (2017)'s methodology in the computation of the worker's abilities and the job skill requirements.[13]. Finally, we also measure mismatch in terms of mean squared deviation between worker's abilities and job skill requirements: $m_{i,c_t} \equiv \left(\sum_{j=1}^{J} (a_{i,j} - r_{c_t,j})^2 / J \right)^{0.5}$.

All in all, we provide robust evidence that during recessions mismatch decreases for job stayers consistent with the cleansing hypothesis, but it increases for new

[13] Speer (2017)'s methodology differs in two dimensions. First, instead of collapsing the several categories into 3 skill dimensions using PCA, he computes the score of each skill as the mean score across the different components of the test. Given this, the math score is the mean of the mathematics knowledge and arithmetic reasoning tests, verbal the mean of word knowledge and paragraph comprehension and the technical score is the mean of general science and electronics information. Second, while Guvenen et al. (2018) run PCA on the whole set of O*NET descriptors, and do not rely on an priori judgment of which descriptors are measures of skills, Speer (2017) chooses a subset of O*NET descriptors for each skill, and takes the mean as the score for the job skill requirement in each skill dimension.

hires from unemployment in line with the sullying effect. So, business cycles affect the allocation of workers to jobs both through job destruction and job creation: when the economy contracts, highly mismatched jobs are destroyed but also created. Furthermore, while job stayers become less under-qualified, new hires from unemployment become both more under- and over-qualified during downturns.

2.2.6 Revisiting Job Tenure Cyclicality

Because match quality is hard to observe, it has been traditionally been proxied by employment duration. This approach has its roots in Jovanovic (1979)'s interpretation of a match as an experience good: in a context of imperfect information, the productivity of the match only becomes known as the match is experienced, hence a match that lasts longer signals better quality. In light of this, the pioneering work by Bowlus (1995) finds that jobs that start in recessions have a shorter duration, which shes interpret as evidence that match quality is procyclical. In line with these findings, we find that, for new hires from unemployment, and increase in aggregate unemployment is associated with higher mismatch. Having an alternative proxy for match quality, a natural question then is whether the relationship between job duration and business cycle conditions at the start of the job still holds when we control for the mismatch index.

Empirical Framework To address this issue, we estimate a discrete proportional hazard model using the complementary log model,

$$h_i(\tau) = \alpha_0 + \alpha_1 U_0 + \alpha_2 U_\tau + \alpha_3 m_{i,c_0} + \delta x_{i,\tau} + h_0(\tau) + \varepsilon_{i,\tau}, \quad (2.8)$$

where $h_i(\tau)$ be the probability that individual i's job ends at date τ given that it lasted until τ, $h_0(\tau)$ is the baseline hazard, which is parameterized as $\ln(\tau)$, U_0 is the unemployment rate at the start of the job relationship, U_τ is the current unemployment rate, m_{i,c_0} is the mismatch level on the current job, and $x_{i,\tau}$ is a set of controls that includes the following variables: a quadratic polynomial in age, current wage, education, race, gender, one-digit level occupation and one-digit level industry.

Results Table 2.2 gives the parameter estimates. As a preliminary step, we replicate Bowlus (1995)'s proportional hazard model specification. Column 1 shows that, by using the same sample but relying on a larger period (Bowlus (1995) sample runs from 1979 to 1988), we recover coefficient estimates that are similar to hers: the initial unemployment rate is positively associated with the likelihood of separation, while the current unemployment has a non-linear relationship with the likelihood of separation.

Columns 2 to 5 present the results from estimating the proportional hazard model specified in Equation (2.8). Column 3 shows that the observed pattern remains unchanged once we control for mismatch. This means that when comparing two jobs with a similar level of mismatch between the abilities of the worker and the job skill requirements, the one that started during times with higher unemployment is more likely to end, however the one that is currently active in times of higher unemployment it is less likely to end. Further, as expected, the estimated coefficients show that mismatch increases the likelihood of separation. This result implies that the better is the alignment between workers' skills and job requirements, the less likely is the match to end. Following Jovanovic (1979), this result suggests that better matches are the ones where skill mismatch is smaller.

Heterogeneity by U_t In Section 2.2.4, we presented robust evidence that in recessions mismatch decreases for workers in ongoing job relationships, meaning highly mismatched job are destroyed. Given this, the relationship between mismatch and the hazard rate of separation should be increasing in the level of current unemployment. Column 4 reports the estimates a version of Equation 2.8 that includes an interaction of m_{i,c_0} with U_τ. Figure 3.A2 plots the marginal effect of mismatch for different levels of the current unemployment rate. Our results shows that a higher mismatch level is associated with an increase in the likelihood of separation, with this relationship being stronger as the level of current unemployment increases. The observed pattern is in line with the *cleansing effect* of recessions: as the unemployment rate increases, mismatch decreases through the destruction of worker-occupation matches that are more mismatched. Under this specification, the role of initial unemployment remains unchanged, i.e. jobs started in downturns are more likely to end, keeping all else constant. This pattern

Table 2.2: Job Tenure Cyclicality

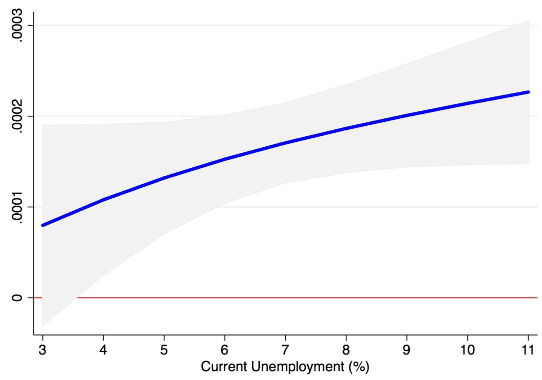
Dependent Variable: Hazard rate of Separation					
	(1)	(2)	(3)	(4)	(5)
Unemployment$_{\tau_0}$	0.0218***	0.0236***	0.0234***	0.0228***	0.0222***
	(0.0062)	(0.0062)	(0.0063)	(0.0063)	(0.0063)
Unemployment$_{\tau}$	-0.5820***	-0.0666***	-0.0659***	-0.0951***	-0.0958***
	(0.0271)	(0.0058)	(0.0059)	(0.0099)	(0.0098)
Unemployment$_{\tau}$ squared	0.0356***				
	(0.0018)				
Mismatch$_{i,\,c_0}$			0.0039***	-0.0030	-0.0030*
			(0.0005)	(0.0018)	(0.0018)
Unemployment$_{\tau}$ $\times$ Mismatch$_{i,\,c_0}$				0.0010***	0.0010***
				(0.0003)	(0.0003)
UE_{i,τ_0}					0.1139***
					(0.0143)
Observations	595395	596372	592792	592792	592792

Notes: The table reports coefficients from the proportional hazard model specified in Equation (2.8), except column 1 that replicates Bowlus (1995)'s specification, with robust standard errors clustered at the individual level reported in parentheses. The dependent variable is the probability that the worker-job pair ends in τ given that it lasted until τ. *Unemployment*$_{\tau_0}$ corresponds to the aggregate unemployment rate when the match was created, and *Unemployment*$_{\tau}$ measures current aggregate unemployment rate. UE is a dummy variable that equals one if the worker was hired out of unemployment. All columns include the following controls: quadratic polynomial in age, current wage, education, race, gender, one-digit industry, one-digit occupation. The sample includes all worker-job matches between 1980 and 2012. * ***, ** and * represent statistical significance at 1%, 5% and 10% levels, respectively.

could however be capturing compositional bias arising from the difference in the types of new hires (from employment or from unemployment). For instance, it is known that in times of low unemployment, there are more job-to-job transitions, and these matches potentially last longer.

Column 5 in Table 2.2 further controls for whether the worker-job match was originated from a hire from unemployment or a worker making a job-to-job transition. While previous results remain unchanged, interestingly the estimated coefficient on the dummy variable is negative and statistically significant meaning that, keeping all else constant including mismatch, new hires from unemployment have lower job duration as compared to matches that arise from job-to-job transitions: a worker that was hired out of unemployment is 11% more likely to separate.

Figure 2.3: Marginal Effect of m_{i,c_0} on the Hazard Rate of Separation



Notes: The graph plots the marginal effect of m_{i,c_t} on the hazard of separation for different levels of current unemployment U_τ computed using estimated parameters reported in column 4 of Table 2.2 The shaded area corresponds to 95% confidence interval.

2.2.7 Mechanisms

The previous sections presented a set of new empirical facts about the cyclical behavior of worker-occupation mismatch and its implications for job tenure:

Fact 1 Mismatch is procyclical for workers in ongoing job relationships, acyclical for workers making job-to-job transitions, and countercyclical for new hires from unemployment.

Fact 2 For job stayers, the decrease in mismatch during recessions is driven by negative mismatch, while for new hires from unemployment both positive and negative mismatch support the increase in mismatch as the unemployment rate increases.

Fact 3 Conditional on mismatch, a match that starts in a recession is more likely to end.

Fact 4 Mismatch increases probability of separation into unemployment, and this effect is increasing in the level of current unemployment.

Overall, our findings are difficult to explain simultaneously by current theories. According to the matching model with endogenous separations in Mortensen and Pissarides (1994), recessions *cleanse* the labor market of low quality matches. The idea behind this framework is that following a negative productivity shock reservation match quality increases; low quality matches are destroyed while only high quality matches are formed, decreasing average mismatch. Given this, while Mortensen and Pissarides (1994)'s framework can explain the negative association between mismatch of workers in ongoing job relationships and aggregate unemployment, it cannot account for the increase in mismatch for new hires from unemployment, as it predicts lower mismatch also for newly created matches in downturns.

Where is the sullying force coming from? In Moscarini (2001)'s version of the Roy model with search and matching frictions, as unemployment rises, workers pay less attention on locating the specific occupations which are best suited to their skills and focus more on remaining employed. As a consequence, there are fewer unemployed workers who direct their search to the sector that provides them the highest value in the market, and more unemployed workers who accept any job that comes along, i.e in downturns unemployed workers sort more randomly. This increases the potential for workers to be more under-qualified, increasing mismatch in recessions for new hires. At the same time, more recently, Lise and Robin (2017) provide evidence that in recessions there is an increase in positive mismatch for new hires from unemployment driven by an increase in hiring from medium and high type firms of medium and high type unemployed workers. Nonetheless, we find that for new hires from unemployment the increase in mismatch during recessions is driven by an in increase in both over- and under-qualification (**Fact 2**).

Another surprising result is **Fact 3**. According to the notion of match quality as an experience good, first suggested by Jovanovic (1979), once we control for mismatch, there should be no relationship between business cycle conditions at the start of the job and job tenure.

What mechanism can reconcile *(i)* the opposite behavior of mismatch for new hires from unemployment and job stayers, and *(ii)* the variation in job duration for different initial economic conditions, conditional on mismatch? In the next section, we provide a theoretical framework that yields a natural role to cyclical information frictions. As in Jovanovic (1979), workers and firms learn about mismatch while the match is active. The key feature is that in recessions negative productivity shocks destroy (perceived) highly mismatched worker-job pairs, but at the same time an increase in uncertainty creates undetected highly mismatched worker-job pairs. As a source of countercyclical uncertainty, we explore the role of occupational switching. Upon switching occupation, workers start learning about skills not previously used, hence uncertainty at the start of a match is higher. A larger fraction of larger fraction of unemployed job seekers switching occupations during recessions creates an increase in mismatch due to the fact that it is hard to estimate whether the worker is going to be a good match or not. This novel mechanism is supported by four empirical regularities. First, occupational switching for new hires is countercyclical. Second, keeping all else constant, the separation hazard is higher for new hires from unemployment that switched occupation upon reemployment. Third, mismatch dispersion increases in recession. Lastly, in recessions the correlation between skill requirements and workers' abilities decreases for new matches, and this is true for all skills.

Using the sample of the previous sections, we first estimate a linear probability model for the event that a worker hired at time t is observed to be working in a different occupation from the one in the previous job. Occupational mobility is defined as a movement across three-digit occupations codes from the time consistent occupation system developed by Dorn (2009).[14] The set of controls include a quadratic polynomial in age, one-digit level occupation and industry, individual, year and month fixed effects. Robust standard errors are provided, clustered at the individual level to allow for serial correlation. Results in Table 2.3 show statistically significant evidence for countercyclical occupation switching conditional on being a new hire. This result is in line with Kambourov and Manovskii (2008) and Huckfeldt (2016). The former uses the Panel Study of Income Dynamics (PSID) over the period from 1968 to 1997, and, using all types of work flows, finds that occupational mobility is countercyclical for young and old workers. The latter

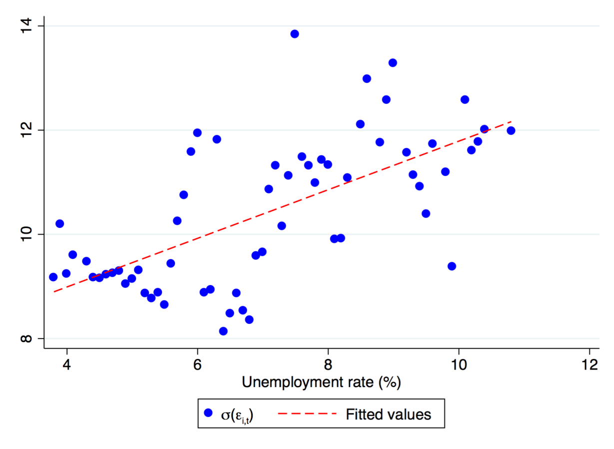
[14]The NLSY79 reports the three-digit Census occupation code. Because this classification system changed over time, we converted all the occupational codes across the years into the *occ1990dd* occupation system developed by Dorn (2009) that is time-consistent.

uses the Displaced Worker Supplement from the Current Population Survey (CPS) and finds evidence for countercyclical movements of displaced workers across occupations. In contrast, Moscarini and Thomsson (2007) and Carrillo-Tudela and Visschers (2014) use, respectively, the CPS and the SIPP, and provide suggestive evidence that occupational mobility is larger in expansions. Note, however, that Carrillo-Tudela and Visschers (2014) do not consider issues of statistical inference. Further, Moscarini and Thomsson (2007) only consider switches among workers who were employed two months in a row, while Carrillo-Tudela and Visschers (2014)'s analysis considers only unemployed job seekers upon reemployment and focus on mobility across broad occupational categories (one-digit level).

Next, we show that new hires from unemployment that switch occupations have lower job durations. We consider the separation hazard model in Equation (2.8), including *(i)* a dummy for occupational switcher, and *(ii)* an indicator for transitions from unemployment to employment and occupational switcher. Column 3 in Table 2.3 shows that, conditional on mismatch, the separation hazard is around 9% higher for new hires from unemployment with a different occupation than the held in the previous job. This can be reconciled with a model of learning about skill mismatch between a worker and a job: when workers switch occupation, both the firm and the worker are more uncertain about how the worker's skills fit the requirements in the new occupation, in turn higher uncertainty translates into a higher probability of breaking the match.

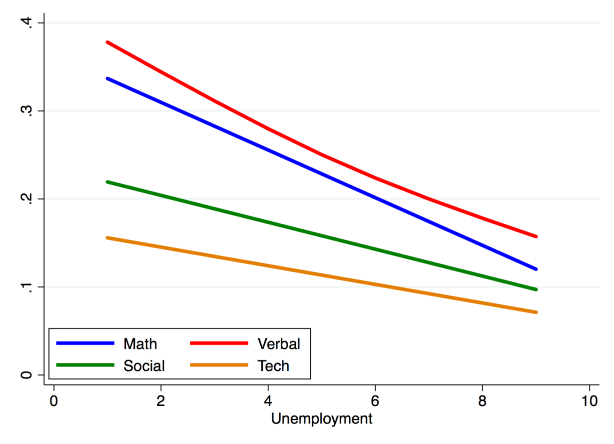
Finally, Figure 2.4 plots the standard deviation of residuals estimated from a version of Equation (2.4) that does not include the regressor corresponding to the state of the economy. This plots strongly suggest that mismatch is more dispersed when the unemployment rate is high. Figure 2.5 pictures the correlation between worker's abilities and job skill requirements for new hires against the aggregate unemployment rate. This plot shows that as the unemployment rate increases, worker's tend to sort into jobs where their skill bundle is less fitted.

Figure 2.4: Mismatch Dispersion and Unemployment



Notes: This graph plots the aggregate unemployment rate against the dispersion in estimated residuals from a version Equation (2.4) without controlling for the state of the economy. Source: NLSY79, BLS and author's calculations.

Figure 2.5: Correlation between abilties and skill requirements and Unemployment



Notes: This graph plots the correlation between worker's abilities and job skill requirements for new hires along four skill dimensions against aggregate unemployment rate. Source: NLSY79, BLS and author's calculations.

Table 2.3: Occupational Switching, Hazard Rate and the Business Cycle

Occupational Switch			Hazard Rate	
	(1)	(2)		(3)
Unemployment$_t$	0.024**	0.028**	Unemployment$_{\tau_0}$	0.0193***
	(0.011)	(0.012)		(0.0064)
Prev. Mismatch	0.002***	0.002***	Unemployment$_\tau$	-0.0945***
	(0.000)	(0.000)		(0.0102)
UE$_{i,t}\times$ Unemployment$_t$		-0.005	UE$_{i,\tau_0}$	0.05404**
		(0.006)		(0.0252)
			UE$_{i,\tau}$ $\times$ Switcher$_{i,\tau_0}$	0.0989***
				(0.0306)
Observations	19229	19229		574862
Ajusted R^2	0.161	0.161		-

Notes: The table reports coefficients with robust standard errors clustered at the individual level reported in parentheses. In columns 1 and 2, the dependent variable is the probability of occupational switching. Both columns include a quadratic polynomial in age, and the following fixed effects: individual, month, year, region, one-digit industry and one-digit occupation. *Prev. Mismatch* corresponds to mismatch in the previous job. UE_{i,τ_0} is a dummy for whether individual i is a new hire from unemployment. In column 3, the dependent variable is the probability that the worker-job pair ends given that it lasted until τ. It includes the following controls: quadratic polynomial in age, current wage, education, race, gender, one-digit industry, one-digit occupation. $Switcher_{i,\tau_0}$ is a dummy for whether individual i switched occupation at the start of the match in τ_0. The sample includes all new hires between 1980 and 2012. ***, ** and * represent statistical significance at 1%, 5% and 10% levels, respectively.

2.3 A Model of Mismatch Cycles

In this section we develop an inaction model with learning that builds on Jovanovic (1979) and Baley and Blanco (2018). The key new element is that in recessions, lower aggregate productivity is also accompanied by higher uncertainty; in turn, higher average uncertainty arises by a higher fraction of worker that switches occupations and thus needs to learn about a new skill. Within this framework, learning experiences about the match will depend on the business cycle conditions, on average. When a match starts in an unfamiliar occupation (more likely in a recession), worker and firms are initially more uncertainty about their mismatch. As the pair observes signals about mismatch, their uncertainty decreases and they produce more precise estimates.

2.3.1 Environment

Time is continuous and the future is discounted by all agents at a rate ρ. There is a continuum of risk-neutral workers and potential firms. There is a large number of occupations, indexed by $k \in \{1, 2, ..., K\}$. Workers are ex-ante heterogeneous in their abilities in each skill dimensions. An occupation is defined by the single skill it uses for production, implying an equal number of occupations and skills. A job in career k is a single firm-worker pair. Firms within an occupation differ on how intensively they use skill k. The productivity of a worker-job pair is determined by the skills of the worker and the skill intensity of the job.

Workers Each worker $i \in [0, 1]$ is endowed with different abilities in each skill dimension. Denote **a** as the time-invariant vector that characterizes individual i's abilities at different skills. **a**$= [a_1, ..., a_K]$ is drawn from a Log-normal distribution with mean $\bar{a}_{K\text{x}1}$ and variance $S_a = I_{K\text{x}K} \cdot \sigma_a^2$. These skills are not directly observable to either the worker or the employer, but the distribution of the initial skills is public information.

Production Within a occupation k, there is unit mass of firms. A firm j within occupation k is characterized by a match-specific requirement r_j^k which is drawn (and publicly observed) from some distribution $G(r)$. Each firm consists of only one job that can be either vacant or filled. Production is subject to an aggregate shock z that uniformly affects output at all occupations, and is reduced by mismatch. Mismatch between firm j in occupation k and worker i ability in skill dimension k is defined as

$$m_{ijk} = r_j^k - a_i^k. \tag{2.9}$$

The per-period output of a worker-firm pair is

$$dy_{ij,t} \quad = \quad (z_t - \phi m_{ijk}^2)dt.$$

The firm-worker only observes the aggregate productivity shock z_t, but the value of the match-specific productivity m_{ijk} is unobserved because a_i^k is not know. The worker-firm pair infers the value of m_{ijk} by solving a filtering problem, described in Section 2.3.2.

Labor market and Matching While unemployed, workers receive flow unemployment benefits (or value of home production) equal to bdt and search for jobs randomly across the K occupations. A key assumption is that firms in different occupations do not congest each other in the matching process. Each occupation labor market has the DMP structure. Firms within an occupation meet all unemployed workers through a standard matching function $M(u, v)$ which converts unemployed workers u and vacancies v into matches. Denote the job finding probability with $p(\theta)$ and the vacancy filling rate with $q(\theta) = \theta p(\theta)$ where $\theta = v/u$ is labor market tightness. We assume that all labor markets have the same matching technology. Firms pay a flow cost cdt to keep a vacancy open. Free entry of firms at each occupation determines endogenously the number of entrants. Search is costless and unemployed workers search in all labor markets at the same time. This means that one can think of unemployed workers as sending an application for a job in each occupation. This setting is similar to Carrillo-Tudela and Visschers (2014) where workers randomly search across different occupations. For simplicity, we consider that when a firm hires a worker out of unemployment, it offers a wage w to make the worker indifferent between taking the offer and staying unemployed, i.e. we set the bargaining power of the worker to zero. Given this wage-protocol, the value of unemployment does not depend on the vacancy distribution and the match surplus only depends on time through z_t, as show in Appendix 2.5.2[15]

A worker-firm match can end with an exogenous (and constant) probability δdt, but can also end if the worker and the firm decide to do so, as we describe below. As in Borovickova (2016), in case of an endogenous separation, the firm pays a cost κ to end the relationship (severance pay, or other type of costs). We model this cost as a fee that the firm pays to the government every time a match breaks. This means that only created matches are potentially subject to the payment of κ. Hence, if a firm contacts a worker but decide not to form a match,

[15]Even though this may strike as a strong assumption, this wage-setting for firms hiring workers out of unemployment holds in a model with on-the-job search and wages determined by sequential auctions as in Postel-Vinay and Robin (2002).

firms do not incur in a cost. Once the match is broken, workers go back to unemployment and firms decide whether to post a vacancies or not. Importantly, if the firm finds a new worker, it redraws its requirement level r from its unconditional distribution; in this way, the separation value is fully pinned down by free entry and there is no need to keep track of (π_t, Γ_t) for the purpose of computing value of filled positions.

Reallocation Upon separation, unemployed workers switch occupation with probability π_t and search for jobs in a different occupation than the one they previously held. Because abilities are not correlated across occupations, working at an occupation k does not reveal the worker's aptitude at other occupations that utilize different skills. Also, for simplicity we assume no recall of abilities once the worker has left his occupation as in Carrillo-Tudela and Visschers (2014). Thus, whenever an unemployed worker is reallocated to another occupation, upon matching the firm and the worker set their beliefs about the worker's abilities equal to the unconditional prior $(\bar{a}, S_a)$. With probability $1 - \pi_t$, a worker retains his current career path, and when matched to a firm in an identical occupation as his previous employer, the new match is initialized at their current beliefs about the worker's abilities $(\hat{a}, \Sigma)$.

Aggregate state The aggregate state in this economy consists of a triple $\Omega_t \equiv (z_t, \pi_t, \Gamma_t)$, where Γ is the distribution over $(a, \hat{a}, \Sigma)$ for filled vacancies, and over $(a, \hat{a}, \Sigma)$ among unemployed workers. Aggregate productivity z_t and the switching probability π_t may take two different values: $z_t \in \{z_L, z_H\}$ and $\pi_t \in \{\pi_L, \pi_H\}$. In line with the empirical evidence presented in Section 2.2.7, we assume that π_t and z_t are negatively correlated so that we have two aggregate states (z_L, π_H) and (z_H, π_L). The switches between (z_L, π_H) and (z_H, π_L) are Poisson, with arrival rate λ_L and λ_H per unit of time, respectively. Finally, we normalize $z_1 \leq z_2$ and assume $\pi_1 \geq \pi_2$.

2.3.2 Information and Learning

Signals Following Jovanovic (1979), Moscarini (2005) and Borovickova (2016), among others, firms and workers do not observe match-specific mismatch m_{ijk} (or a_i^k) directly. However, in contrast with the existing literature, where firms and workers observe output flow, we assume that agents do not observe flow output $dy_{ij,t}$ directly but receive continuous noisy observations about mismatch, denoted by s_t, which evolve according to

$$ds_t = m\, dt + \sigma_s dW_t \tag{2.10}$$

where the signal noise W_t follows a Wiener process. Nonetheless, this setup is isomorphic (up to a constant shift in z of $-\psi\sigma_s^2$) to a case in which flow output is observed by agents but contains transitory shocks to mismatch which generate noise around its true value.

Information Set The information set at time t is given by the σ-algebra generated by the history of signals s and aggregate productivities z:

$$\mathcal{I}_t = \sigma\{s_r, z_r; r < t\} \tag{2.11}$$

Filtering Firms and workers estimate mismatch in a Bayesian way by optimally weighing new information from signals in Equation (2.10) against old information from prior estimates. This is a passive learning technology in the sense that firms process the information that is available to them, but they cannot take any action to change the quality of the signals. Let $\mu_t \equiv \mathbb{E}[m|\mathcal{I}_t]$ be the best estimate (in a mean-squared error sense) of the mismatch and let $\Sigma_t \equiv \mathbb{E}[(m_t - \mu_t)^2|\mathcal{I}_t]$ be its variance. We call Σ_t *mismatch uncertainty*. Proposition 2.1 establishes the laws of motion for mismatch estimates and uncertainty.

Proposition 2.1 *Let the signal evolve according to* (2.10)*, and consider the information set in* (2.11)*. Then the posterior distribution of mismatch is Gaussian*

$m_t|\mathcal{I}_t \sim \mathcal{N}(\mu_t, \Sigma_t)$ *where the first two moments evolve as follows:*

$$d\mu_t = \frac{\Sigma_t}{\sigma_s} d\hat{W}_t \tag{2.12}$$

$$d\Sigma_t = -\left(\frac{\Sigma_t}{\sigma_s}\right)^2 dt \tag{2.13}$$

The filtered process (news) is a Brownian motion under the information set of the pair $\mathcal{I}_t$:

$$d\hat{W}_t = \frac{1}{\sigma_s}\left(m_t - \mu_t\right) dt + dW_t \tag{2.14}$$

According to the Proposition 2.1, the mismatch estimate μ_t follows a Brownian motion. Mismatch uncertainty Σ_t decreases with tenure as mismatch is revealed through the observation of the signal s_t. Due to Bayesian updating, when uncertainty is high, estimates optimally put more weight on signals instead of the prior. Learning is faster, but it also brings more white noise into the estimation. As such, estimates become more volatile with high uncertainty.

2.3.3 Equilibrium Separation Policy

With the filtering problem at hand, we now derive the optimal match break up decision. The wage-setting rule implies that the separation decision depends only on the match surplus and not the wage itself. We can therefore focus on the value function for the match surplus.

Proposition 2.2 (Separation problem) *Let $J(\mu, \Sigma, z)$ be the surplus of a match. Also let κ be the cost of ending a relationship. Then the optimal separation time τ solves the following problem*

$$J(\bar{a}, S_a, z_0) = \max_{\tau} \mathbb{E}\left[\int_0^{\tau} e^{-(\rho+\delta)t}\left(z_t - \psi(\mu_t^2 + \Sigma_t) - b\right) dt - e^{-(\rho+\delta)\tau}\kappa\right]$$

subject to the filtering Equations in (2.12) and (2.13) and aggregate productivity z_t.

Inaction Region The solution to the separation problem is characterized by an inaction region $\mathcal{R}$ such that the optimal time to break the match is given by the first time that the state falls outside such a region. The inaction region is three-dimensional because the worker-firm pair has three states: μ_t, Σ_t and z_t. Let $\bar{\mu}(\Sigma, z)$ denote the inaction region's border as a function of uncertainty and the aggregate productivity. The inaction region is described by the set $\mathcal{R}_t = \{(\mu, \Sigma, z) : |\mu| < \overline{\mu}(\Sigma, z)\}$. Its symmetry around zero is inherited from the specification of the stochastic process and the quadratic payoffs. A match between a worker and a firm breaks if $(\mu_t, \Sigma, z) \notin \mathcal{R}_t$, otherwise it continues. Note that, as in Baley and Blanco (2018), the inaction region refers to mismatch estimates and not the true mismatch. As a result, there are relationships that are destroyed because they are perceived to be highly mismatched, when the true mismatch is actually low. This feature is key to account for evidence showing that, conditional on mismatch, younger relationships and relationships starting in bad times are more likely to break.

This inaction problem is non-standard because it is three-dimensional. Following Baley and Blanco (2018), in order to provide sufficient conditions of optimality, we impose the Hamilton-Jacobi-Bellman equation, the value matching condition, and the standard smooth pasting condition for the three states (μ, Σ, z). Proposition 2.3 formalizes this.

Proposition 2.3 ***(HJB Equation, Value Matching and Smooth Pasting)*** *τ is the optimal separation time if:*

1. *Inside the inaction region, it satisfies the solves the Hamilton-Jacobi-Bellman (HJB) for $(\mu_t, \Sigma, z) \in \mathcal{R}$*

$$\begin{aligned}(\rho+\delta)J(\mu,\Sigma,z^i) &= z^i - \psi(\mu^2+\Sigma) - b + \left(\frac{\Sigma}{\sigma_s}\right)^2\left(\frac{J_{\mu\mu}}{2} - J_\Sigma\right) - \delta\kappa \\ &+ \lambda_j\left[J(\mu,\Sigma,z^j) - J(\mu,\Sigma,z^i)\right] \qquad (2.15)\end{aligned}$$

   where $\mu = r^k - \hat{a}$ is the mismatch estimate, b is the flow of unemployment benefits, δ is the intensity of exogenous separation, and $z - \psi(\mu^2 + \Sigma)$ is the expected output flow from the match with $z^i \in \{z_L, z_H\}$.

2. *At the border of the inaction region, it satisfies value matching for $\mu = \pm\bar{\mu}$,*

which sets the value of separating equal to the value of not separating:

$$J(\mu, \Sigma, z^i) = -\kappa \tag{2.16}$$

3. *At the border of the inaction region, it satisfies the smooth pasting conditions for $\mu = \pm\bar{\mu}$*

$$J_\mu(\mu, \Sigma, z^i) = J_\Sigma(\mu, \Sigma, z^i) = J_z(\mu, \Sigma, z^i) = 0 \tag{2.17}$$

Note that the value of preserving the match (Equation 2.15) equals the expected net flow return $z - \psi(\mu^2 + \Sigma) - b$, the gain from learning reflected in the second-derivative terms and the net benefit from receiving a new productivity shock. Interestingly, in a setting with *(i)* on-the-job search and *(ii)* wages determined by the sequential auction model of Postel-Vinay and Robin (2002), as in Borovickova (2016), the value of match would be equal to the one in Equation (2.15). This is because when wage contracts are renegotiated sequentially by mutual agreement the benefit from searching while on the job does not appear in the value function: outside offers do not affect the surplus of the match, but only how it is shared between the firm and the worker. Thus, the following results regarding the optimal separation policy to unemployment hold in a more general setting incorporating search on-the-job with wages determined through the sequential auction framework by Postel-Vinay and Robin (2002).

Comparative Statics: Inaction Region Following Baley and Blanco (2018), it can be shown that:

1. *Uncertainty widens the inaction region*, holding all else constant. This feature captures the well known *option value effect* in Dixit (1991): due to high uncertainty, firms and workers are more tolerant to higher mismatch levels and delay any adjustment. As a result of uncertainty dynamics, the option value is time varying and the inaction region is time dependent. In particular, for a given worker-firm pair, mismatch uncertainty decreases over time due to the filtering problem described in Section 2.3.2, so the inaction region shrinks with match tenure.

2. *Separation hazard decreases with tenure* as in Jovanovic (1979). The elasticity of the inaction region with respect to mismatch uncertainty Σ_t is positive (option value effect) and lower than 1. This implies that the "volatility effect" — the fact that mismatch estimates in younger matches are more volatile following the Bayesian updating — dominates the "option value effect", thus newer (more uncertain) job relationships are more likely to separate.

3. *The cost of separation κ increases the width of the inaction region* The role of the firing cost is straightforward: when κ increases it is more costly to fire a worker, thus keeping everything else constant, higher levels of perceived mismatch becomes more tolerable.

4. *The sensitivity of output to mismatch ψ decreases the inaction region.* If ψ is higher, the decrease in output following an increase in mismatch is larger, hence highly mismatched jobs are less tolerated.

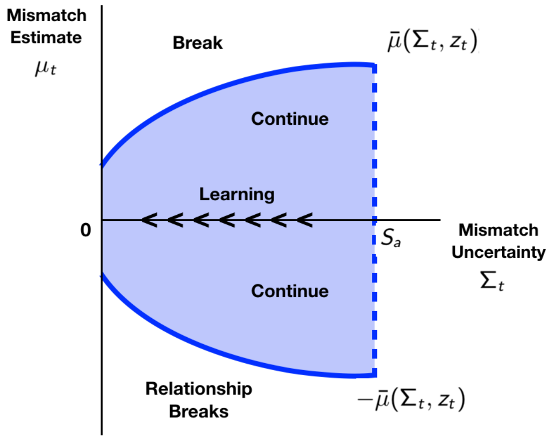
5. *For any level of uncertainty Σ, the inaction region increases in aggregate productivity z.* When aggregate productivity z increases, the mismatch estimate μ that makes the worker and the firm indifferent between breaking the match or continuing is higher, hence the inaction region widens for all levels of uncertainty Σ.

In this setup, a firm and worker prefer to break up when the match prospects are bad. This depends on two different forces. On the one hand, workers and firms who become sufficiently pessimistic about their match quality, i.e. perceive mismatch between workers abilities and job skill requirements to be high, decide to break up: workers go back to unemployment and continue searching as unemployed firms post vacancies. On the other hand, workers and firms can also decide to separate in response to negative productivity shocks. Thus, the inaction region $\mathcal{R}$ is both time and state dependent.

2.3.4 Distribution of Beliefs

Let $\Gamma(r^k, \hat{a}, \Sigma)$ be the measure over skill-requirements and worker-specific beliefs $(\hat{a}, \Sigma)$ among active worker-firm matches in an occupation k. For $(r^k, \hat{a}, \Sigma)$ such

Figure 2.6: The Inaction Region for Mismatch Estimates



Notes: This graph pictures the inaction region for a given aggregate productivity level z_t, separation cost κ and ψ. A firm and a worker decide to break a match whenever the mismatch estimate μ_t falls outside the inaction region (the blue area) otherwise it continues. The inaction region is wider for larger mismatch uncertainty. Due to Bayesian learning, mismatch uncertainty decreases with match tenure, as such the option value decreases, implying that the inaction region shrinks over time.

that a worker-firm pair does not exist, the measure $\Gamma(r^k, \hat{a}, \Sigma)$ is zero. For the others, the Kolmogorov forward equation describes the dynamics of this measure:

$$\begin{aligned}\frac{\partial \Gamma(r^k, \hat{a}, \Sigma)}{\partial t} &= \frac{\partial \Gamma^{act}(r^k, \hat{a}, \Sigma)}{\partial t} \\ &+ \frac{\partial \Gamma^{in}(r^k, \hat{a}, \Sigma)}{\partial t} - \frac{\partial \Gamma^{out}(r^k, \hat{a}, \Sigma)}{\partial t} \end{aligned} \tag{2.18}$$

where the first term balances all flows that are due to learning, given by

$$\frac{\partial \Gamma^{act}(r^k, \hat{a}, \Sigma)}{\partial t} = \left(\frac{\Sigma}{\sigma_s}\right)^2 \Gamma_\Sigma + \frac{2\Sigma}{\sigma_s} + \frac{1}{2}\left(\frac{\Sigma}{\sigma_s}\right)^2 \Gamma_{\mu^2}, \tag{2.19}$$

the second term corresponds to the inflow of workers from unemployment conditional to belief $(r^k, \hat{a}, \Sigma)$ and is equal to

$$\frac{\partial \Gamma^{in}(r^k, \hat{a}, \Sigma)}{\partial t} = p(\theta_t) \cdot \mathbf{1}\{J(r^k - \hat{a}, \Sigma, z) > 0\} \cdot \Lambda(\hat{a}, \Sigma) \cdot G(r^k), \tag{2.20}$$

where $\Lambda(\hat{a}, \Sigma)$ is measure over unemployed workers; and finally the third term corresponds to outflows at $(r^k, \hat{a}, \Sigma)$ due exogenous separations and caused new productivity shocks and is given by

$$\frac{\partial \Gamma^{out}(r^k, \hat{a}, \Sigma)}{\partial t} = \Gamma_t(r^k, \hat{a}, \Sigma) \cdot [\delta + (1-\delta)\mathbf{1}\{J(r^k - \hat{a}, \Sigma, z) \leq -\kappa\}]. \quad (2.21)$$

2.3.5 Vacancy Posting

To hire new workers, firms have to post vacancies. Their decision to post vacancies depends on the expected value from randomly meeting a worker. Any vacancy that does not deliver a contact with a worker is lost and generates no continuation value. Also any contact with a worker that does not end up in a match is lost and has zero value. The expected value of a vacancy, $\bar{J}$ is then given by

$$\bar{J} = \int \Big[(1-\pi_t) \underbrace{\int J^*(\hat{\mu}, \Sigma, z) d\Lambda(\hat{a}, \Sigma)}_{\text{no switch}} + \pi_t \underbrace{J^*(r^k - \bar{a}, S_a, z)}_{\text{switch}} \Big] dG^k(r)$$

where $J^* = \max\{J, 0\}$, with J being the surplus of a match. With probability π_t the firm contacts an unemployed job seeker that switched occupations, the match is formed if the match surplus is positive and upon matching the firm and the worker set their beliefs about the worker's abilities equal to the unconditional prior with mean $\bar{a}$ and variance S_a. In turn, if a contact is with an unemployed worker that did not switch occupation happens with probability $1 - \pi_t$, in which case the match begins with priors at their current beliefs about the worker's abilities with mean $\hat{a}$ and variance Σ. As long as the expected value from meeting a worker is positive, the firm posts vacancies up to the point where the marginal cost of posting a vacancy equals the expected value from filling the vacancy. Thus, in equilibrium the number of advertised job opportunities is determined by equating the marginal cost to the expected value of a job opening,

$$c = q(\theta_t)\bar{J}. \quad (2.22)$$

2.3.6 Equilibrium

Given the distribution of workers abilities, the exogenous stochastic process for idiosyncratic noise W_t and for each aggregate state Ω_t, an equilibrium consists of mismatch estimates μ and uncertainty Σ, a match surplus $J(r^k, \hat{a}, \Sigma)$, a hiring cost c, and labor market tightness θ across all occupations such that equations (2.12)- (2.22), are satisfied.

2.3.7 Implications

We now turn to explore the main implications of our theory. In this set up, recessions are characterized by *lower* aggregate productivity and *higher* occupational switching, in line with the documented evidence in Section 2.2.7. These two forces have distinct effects in the economy.

On the one hand, following a negative aggregate productivity shock the inaction region shrinks, as shown in the left panel in Figure 2.7. The intuition is the following. From Equation (2.15), it follows that, conditional on μ and Σ, higher productivity increases the surplus from the match. Thus, when aggregate productivity is lower, the worker and the firm are less willing to tolerate a match with high levels of *perceived* mismatch, for any level of uncertainty. The perceived mismatch at the threshold that makes the worker and the firm indifferent between breaking up the match or continuing decreases and, therefore, relationships with a higher level of $perceived$ mismatch are destroyed. This is the *"cleansing effect"* of recessions. On the other hand, because the probability of switching occupation increases there is a larger fraction of unemployed workers switching occupation. For these workers, mismatch uncertainty in the new occupation is equal to the unconditional prior S_a, hence mismatch uncertainty is larger. This creates a larger number of worker-firm pairs with *undetected* high mismatch. As such, the uncertainty distribution changes in recessions towards higher uncertain matches, as shown in the right panel in Figure 2.7. We call this the *"sullying effect"*. This countercyclical information friction combined with aggregate productivity shocks reconciles the fact that in recessions, jobs with high mismatch are destroyed, while matches with high mismatch are created.

One important feature to note is that our model with *perfect information* nests the standard search and matching model of the labor market in the tradition of Diamond, Mortensen and Pissarides. Recall that in Mortensen and Pissarides (1994)'s matching model with endogenous separations during recessions reservation match quality increases, low quality jobs are destroyed and only high quality jobs are created, hence mismatch decreases. Under our framework, when mismatch is perfectly observed, $\Sigma = 0$, the sullying effect is shut down, because even if there is more occupational switching in recessions firms and workers know perfectly the mismatch level. Given this, only the cleansing effect of recessions exists: following a negative productivity shock reservation mismatch decreases; high mismatched worker-firm pairs are destroyed while only high quality matches are formed, decreasing average mismatch.

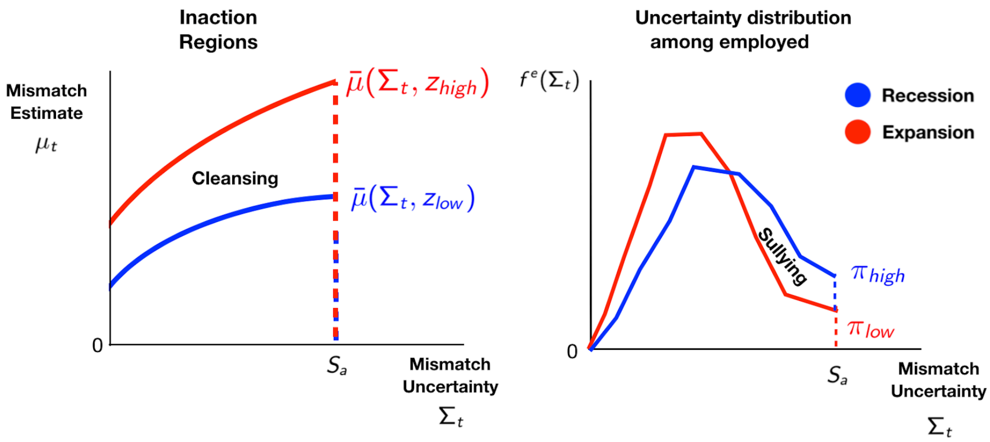
Another result that follows from our set up is that conditional on mismatch, worker-firm relationships starting in recessions are more likely to separate, on average. Due to higher occupational switching, workers and firms in matches that start in a recession are, on average, more uncertain about the mismatch level. As such, mismatch estimates $\hat{\mu}$ are more volatile. As previously mention, the elasticity of the inaction region with respect to mismatch uncertainty Σ_t is positive (option value effect) and lower than 1 (Baley and Blanco, 2018). Given this, the "volatility effect" — the fact that mismatch estimates in younger matches are more volatile due to the filtering problem described in Section 2.3.2 — dominates the "option value effect" — the fact that the inaction region is wider for higher levels of uncertainty. As such, more uncertain job relationships are more likely to separate, meaning that on average matches that start in a recession are more likely to end.

All in all, the proposed mechanism is able to account for the *(i)* the opposite behavior of mismatch for new hires from unemployment and job stayers, and *(ii)* the variation in job duration for different initial economic conditions.

2.4 Conclusion

This chapter has explored the role of business cycles in the allocation of workers to jobs. Using a measure of skill mismatch developed by Guvenen et al. (2018) combined with a worker-level panel from the NLSY79 over the period 1979-2012, we

Figure 2.7: Inaction Region and Uncertainty Distribution over the Cycle



Notes: The left graph pictures the inaction region when the aggregate productivity level is high (red) or low (blue). The downward shift of the inaction region boundary corresponds to matches that are destroyed in a recession because they are no longer profitable. The right graph changes in the distribution of uncertainty across employed workers in an expansion (red) and in a recession (blue). The shift of the distribution to the right is driven by an increase the fraction of unemployed workers that switch occupation when the aggregate productivity level is low: upon switching to a new occupation, these workers re-start learning as such their beliefs equal to the unconditional prior $(\bar{a}, S_a)$.

document robust evidence that during economic contractions mismatch decreases, meaning that in recessions workers' skills are more aligned with the jobs' skill requirements. Surprisingly, we show that this pattern masks important differences along the flow of job creation and the flow of job destruction. For job stayers mismatch decreases, meaning that workers that lose their jobs are more mismatched than average, consistent with the cleansing effect of recessions, but new hires from unemployment are more mismatched in line with the sullying effect of recessions. Therefore, both the cleansing and the sullying effects of bad times are present in the labor market, but the first dominates. We further show that following a negative productivity shock, job stayers are less under-qualified, while new hires from unemployment are both more under- and over-qualified. Digging deeper, we have also examined the implications of mismatch fluctuations over the cycle for the cyclical behavior of job duration. We showed that conditional on mismatch, job duration is lower for matches that start in times of high unemployment. This is a surprising result as according to the notion of match as an experience good, first suggested by Jovanovic (1979), once we control for mismatch, there should be no relationship between business cycle conditions at the start of the job and job tenure.

We developed a model of learning about unobserved skill mismatch that constitutes a first step to understand the potential role to cyclical information frictions in reconciling our empirical findings. The key novelty of the model is that recessions are characterized by lower aggregate productivity but also a larger fraction of matches with high uncertainty about mismatch. Consistent with the documented patterns, we showed that following a negative productivity shock, (perceived) high mismatched worker-job pairs are destroyed, but at the same time large information frictions create undetected worker-firm matches with high levels of skill mismatch. In this chapter, we have explored the role of occupational switching as a source of the countercyclical uncertainty and document suggestive evidence pointing towards this channel. In a current project, we endogenize the decision to switch occupations by consider a directed search framework in which unemployed workers choose to search in the occupation that gives then the highest value.

2.5 Appendix

2.5.1 Data Appendix

Sample Selection and Construction of Variables

Sample selection The NLSY79 is a nationally longitudinal survey of 12,696 individuals who were between 14 and 22 years when they were first interviewed in 1979. This dataset consists of three sub-samples: *(i)* a cross-sectional sample; *(ii)* a supplemental sample designed to oversample civilian Hispanic, black, and economically disadvantaged non-black/non-Hispanic youths; and *(iii)* a military sample designed to represent the youths enlisted in the active military forces as of September 30, 1978. We focus on a sub-sample of males and females from the cross-sectional sample, because many members of supplemental and military samples were dropped from the NLSY79. The cross-sectional sample has 6,111 respondents and was designed to represent the non-institutionalized civilian segment of people living in the United States in 1979 with ages 14-22 as of December 31, 1978. Following standard procedures in the literature, we further drop individuals who were more than two years in the military force, individuals who displayed weak labor market attachment, i.e. individuals spent more than 10 years out of the labor force, individuals that were already working in 1979, and those that do not have information on the Armed Services Vocational Aptitude Battery (ASVAB) test scores. Our final sample is composed of 2,991 individuals. Descriptive statistics for the sample are reported in Panel A of Table 2.A1.

Worker's employment history The NLSY79 interviewed individuals on an annual basis in the years from 1979 to 1993, and on a biannual basis for the period 1994-2012. Information on labor force status is recorded at a weekly frequency throughout the sample period, even in the later period where interviews were at biannual frequency. To construct a monthly panel for our main analysis, we use the NLSY79's Work History Data file. This file is a week-by-week record of the working history for each respondent, which contains information about weekly labor status and hours worked. While an individual may hold more than one job, we focus on the primary job at a given month, which is defined as the one for which an individual worked the most hours in a given month. For each primary job, we retain information on the hourly wage, occupation and industry codes. Before merging occupation and industry information with the employment panel, we clean occupational and industry titles following Guvenen et al. (2018)'s approach: to each job, we assign the occupation and industry code that is most often observed during the employment spell. In the NLSY79, occupation titles are described by the three-digit Census occupation code. Because this classification system changed over time[16], before cleaning we converted all the occupational codes across the years into the *occ1990dd* occupation system developed by (Dorn, 2009), which has the advantage of being time-consistent.[17] Wages correspond to the hourly wage, which include tips, overtime and bonuses, and are measured in 2000 dollars (we use the consumer price index from the BLS to deflate wages).

Worker's employment transitions We identify a *job-to-job* transition when the primary job for an individual at month t is different from the one reported in the previous month, and a *non-employment to employment* transition if the worker was unemployed in month $t-1$ (i.e. reported to be not working, unemployed or out of the labor force) and employed in month t, meaning that she reported a job.[18] We consider that new hires from non-employment include also $recalls$, workers that return to their previous employer after a jobless spell. Additionally, we define a worker making an occupational switch when the occupation at month t is different from the one in the last reported job. Panel B in Table 2.A1 reports descriptive statistics about employer and occupational mobility. We observe that, from 1979 and 2012, individuals change employer, on average, 13.77 times (including job-to-job and non-employment to employment transitions), out of which 7 they also change occupation. Annual occupational mobility in our sample is 21.49% compared with 15.79% reported in Guvenen et al. (2018) and 18.48% re-

[16]Until 2000, NLSY79 reports occupation codes in the Census 1970 three-digit occupation code. After this year, occupation codes are reported in the Census 2000 three-digit occupation code.

[17]The crosswalk files between the Census classification codes and the *occ1990dd* occupation aggregates created by (Autor and Dorn, 2013) can be found at http://www.ddorn.net/data.

[18]The NLSY79 provides a mapping that links jobs across consecutive interviews, which allows us to build employment spells for each job reported by the respondent.

ported in Kambourov and Manovskii (2008) who use the Panel Study of Income Dynamics (PSID) for the period 1968-1997.

Worker's abilities The NLSY79 has information on the ASVAB test scores, which was taken by individuals between ages 14 and 24. The ASVAB is a general test that measures knowledge and skills in 10 different components.[19] We focus on a subset of six components (arithmetic reasoning, mathematics knowledge, paragraph comprehension, word knowledge, mechanical comprehension, general science and electronics information) which are linked to the 3 skill counterparts considered in the empirical analysis: math, verbal and technical. To measure individuals' skills in each dimension, $a_{i,j}$, we follow Guvenen et al. (2018)'s approach: the ASVAB categories are reduced into the 3 skill dimensions using Principal Component Analysis (PCA). For the social dimension, we proceed in the same fashion using the individual scores in two different tests provided by the NLSY79: the Rotter Locus of Control Scale and the Rosenberg Self-Esteem Scale. The Rotter Locus of Control Scale measures the degree of control individuals feel they possess over their life, and the Rosenberg Self-Esteem Scale aims at reflecting the degree of approval or disapproval towards oneself. These measures have been commonly used in previous studies as measures of non-cognitive skills (Speer, 2017; Lise and Postel-Vinay, 2016; Guvenen et al., 2018). For more details, see Heckman et al. (2006).

Note that to adjust for differences in the test-taking age, before proceeding with PCA, we normalize the mean and the variance of each ASVAB test score according to their age-specific values. Also, once we have the raw scores in each skill dimension, we convert them into percentile rank scores, $a_{i,j}$ in Equation (2.1).[20]

[19]The components are arithmetic reasoning, mathematics knowledge, paragraph comprehension, word knowledge, general science, numerical operations, coding speed, automotive and shop information, mechanical comprehension, and electronics information.

[20]Because the raw scores that result from PCA do not have any meaning, we transform them into percentile rank scores, as in Guvenen et al. (2018). This allows us to have a clear interpretation of the scores and compare two different scores. The percentile rank is the percentage of scores that fall below a given score. For example, if an individual's raw score in *math* is transformed into a percentile rank score of 50, it means that the individual is better than 50% of the sample in *math*.

Job skill requirements To obtain measures of the skill requirements in each occupation, $r_{c_t,j}$, we use the O*NET database, that collects that on a list of 277 descriptors, with the ratings of importance level and relevance, for 974 different occupations. As in Guvenen et al. (2018), we use 26 O*NET descriptors from the Knowledge, Skills and Abilities categories that were identified by the Defense Manpower Data Center (DMDC) to be related to each ASVAB category; and other six descriptors to describe the social dimension.[21] Following Guvenen et al. (2018)'s methodology, for each occupation, we build a score comparable to each ASVAB category, and then we collapse the seven ASVAB categories analogues into the 3 skill dimensions (verbal, math and technical) by applying PCA. For the social dimension, we also collapse six O*NET descriptors related to social skills into a single dimension by taking the first principal component. Finally, we rescale the scores by converting them into percentile rank scores, $r_{c_t,j}$ in Equation 2.1. Panel B in Table 2.A2 reports summary statistics of the measures of job skill requirements. To check whether the constructed variables characterize occupations reasonably, we report the mean percentile rank score of each main occupation category of the *occ1990dd* occupation system from Dorn (2009) in Table 2.A3. *Managerial occupations* require more verbal and math skills than *Repair occupations*, which have a higher requirement of the technical skill. As expected, within each broad category there is a large variation in job skill requirements as shown in table 2.A4. For instance, economists require the use of the math skill more intensively, whereas lawyers, within the same broad category, require a higher the use of the verbal skill but use the technical skill less intensively.

To compute job skill requirements, we use O*NET 21.1, which was released in November 2016. Because our panel data starts in 1979, one might be concerned that the computed scores do not reflect the change in the requirements of which occupation over time. To mitigate these concerns, we computed the job skill requirements using the first version of O*NET. We find the correlation between

[21]The descriptors used are the following oral comprehension, written comprehension, deductive reasoning, inductive reasoning, information ordering, mathematical reasoning, number facility, reading comprehension, mathematics skill, science, technology design, equipment selection installation, operation and control, equipment maintenance, troubleshooting, repairing, computers and electronics, engineering and technology, building and construction, mechanical, mathematics knowledge, physics, chemistry, biology, english language, social perceptiveness, coordination persuasion, negotiation instructing, service orientation.

these and the scores we use in the main analysis is very strong (0.80).

Merge Once we have the percentile rank scores in each skill dimension on the occupation and worker-side, we merge the panel of worker-level data with the occupation data using using three-digit occupational codes from Dorn (2009)'s classification system. Note that O*NET uses SOC codes from 2000, which are more detailed than the occupational codes in the NLYS79, based on the three-digit Census occupation codes, hence several occupations in NLSY79 have more than one score. Using a crosswalk to identify each SOC code with a Census code, we take an unweighted average over all the SOC codes that map to the same code in the census three-digit level occupation classification, and then we use Dorn (2009)'s crosswalk to convert occupational codes to a time consistent classification system.

Table 2.A1: Descriptive Statistics of the Sample

The table reports descriptive statistics of the sample used in the empirical analysis, which is a sub-sample of 2,991 individuals from the cross-sectional sample of the NLSY79 and runs over the period from 1979 and 2012. *Job mobility* is defined as the fraction of individuals who start a new job in a given month, including *job-to-job* and *non-employment to employment* transitions. *Occupational mobility* is defined as the fraction of individuals who switch occupations in a given month. Source: NLSY79 and author's calculations.

	Mean	Std. Dev	p10	p90
Panel A: Sample characteristics				
Female (% of total)	48.51			
Male (% of total)	51.49			
African-American (% of total)	11.27			
Hispanic (% of total)	7.15			
Age at interview	33.70	10.03	20.00	47.00
Panel B: Work history				
# Job Transitions per individual	13.77	7.65	5.00	24.00
# Job to Job Transitions	6.03	4.89	1.00	12.00
# Non-employment to Job Transitions	7.74	4.63	2.00	14.00
# Occupation Transitions per individual	7.28	4.77	2.00	14.00
Job mobility (per month, % of total)	3.24	2.66	0.85	7.07
Occupation mobility (per month, % of total)	1.78	1.50	0.50	3.91
Job tenure (months)	14	21	1	35
Occupation tenure (months)	39	65	2	120
Unemp. duration (months)	7	11	1	17

Additional Tables and Figures

This section contains additional tables and figures referred to in the main text.

Table 2.A2: Summary Statistics

The table reports summary statistics for the main variables used in the empirical analysis. Panel A presents the statistics for the measure of worker's abilities in the different skills dimensions, $a_{i,j}$. The sample includes respondents in the NLSY79 dataset that satisfy the selection criteria in Appendix 2.5.1. Panel B reports the statistics for the measures of job skill requirements, $r_{c,j}$, at the three-digit occupational code level constructed by Dorn (2009). Panel C presents the statistics for the job mismatch measures. *Mismatch$_t$* is defined as $m_{i,c_t} \equiv \sum_{j=1}^{J} \omega_j \left|a_{i,j} - r_{c_t,j}\right|$; *Positive mismatch$_t$* as $m^{+}_{i,c_t} \equiv \sum_{j=1}^{4} \omega_j \max\{a_{i,j} - r_{c_t,j}, 0\}$; and *Negative mismatch$_t$* as $m^{-}_{i,c_t} \equiv \sum_{j=1}^{4} \omega_j \left|\min\{a_{i,j} - r_{c_t,j}, 0\}\right|$. The sample consists of unique occupations observed in NLSY79 with occupational characteristics in O*NET. Panel D reports summary statistics of the business cycle indicators. *Unemployment Rate$_t$* is the monthly unemployment rate at the national level published by BLS. *Vacancies Index$_t$* is the Composite Help-Wanted index developed by Barnichon (2010) which captures the behavior of total - print and online - help-wanted advertising, a proxy for the number of job openings at a given point in time. *Industrial Production$_t$* is the monthly industrial production index. Source: NLSY79, O*NET, BLS and author's calculations.

	Observations	Mean	Std. Dev	Min.	Max.
Panel A: Worker's abilities					
Pctl. rank of the verbal skill	2991	49.81	28.44	1	100
Pctl. rank of the math skill	2991	50.15	28.76	1	100
Pctl. rank of the mechanical skill	2991	50.36	28.87	1	100
Pctl. rank of the social skill	2991	49.95	28.86	1	100
Panel B: Job Skill Requirements					
Pctl. rank of the verbal skill	324	50.47	28.92	1	100
Pctl. rank of the match skill	324	50.46	28.92	1	100
Pctl. rank of the technical skill	324	50.37	28.96	1	100
Pctl. rank of the social skill	324	50.44	28.94	1	100
Panel C: Job Match Quality					
Mismatch$_t$	581027	27.43	14.33	1.25	91.25
Positive mismatch$_t$	581027	13.74	15.69	0.00	91.25
Negative mismatch$_t$	581027	13.69	14.48	0.00	82.00
Panel D: Business Cycle Indicators					
Unemployment Rate$_t$	408	0.06	0.02	0.04	0.11
Vacancies Index$_t$	408	2.75	0.44	1.70	3.90
Industrial Production$_t$	408	77.17	18.40	48.47	105.33

Table 2.A3: Mean Percentile Scores of Job Skill Requirements for Broad Occupation Classes

The table reports the mean percentile rank scores, $r_{c_t,j}$, along the four skill dimensions considered in the empirical analysis for the main occupation categories of *occ1990dd* occupation system from Dorn (2009). Source: O*NET, ASVAB and author's calculations.

Occupation	**Percentile rank score**			
	Verbal	Social	Math	Technical
Economists	91	65	96	10
Elevator Installers and Repairers	52	45	53	100
Lawyers	100	89	72	6
Waiters	71	29	7	13

Table 2.A4: Percentile Scores of Job Skill Requirements for Selected Occupations

The table reports the percentile rank scores, $r_{c_t,j}$, along the four skill dimensions considered in the empirical analysis for selected three-digit occupations in the O*NET dataset. Source: O*NET, ASVAB and author's calculations.

Occupation	Percentile rank score			
	Verbal	Social	Math	Technical
Agents and Business Managers of Artists, Performers, and Athletes	93	99	64	3
Economists	91	65	96	10
Elevator Installers and Repairers	52	45	53	100
Helpers–Installation, Maintenance, and Repair Workers	30	29	16	92
Lawyers	100	89	72	6
Painting Workers	4	14	9	62
Tour and Travel Guides	51	73	31	18
Waiters	71	29	7	13

Table 2.A5: Correlation between Worker's Abilities and Job Skill Requirements

The table reports the correlation pattern between the percentile rank scores of worker's abilities, $a_{i,j}$, and the percentile scores of job skill requirements, $r_{c_t,j}$, across 4 skill dimensions: verbal, math, technical and social. The values in bold capture the sorting pattern between worker's abilities and job skill requirements. In Panel A, correlations are computed using the sample of individuals in the sample. The correlations in Panel B are computed using the individual-month observations for employed individuals in the sample. Source: NLSY79, O*NET and author's calculations.

	$a_{i,v}$	$a_{i,m}$	$a_{i,t}$	$a_{i,s}$
Panel A: Worker's abilities				
$a_{i,v}$	1	0.785	0.728	0.319
$a_{i,m}$	0.785	1	0.760	0.317
$a_{i,t}$	0.728	0.760	1	0.295
$a_{i,s}$	0.319	0.317	0.295	1
Panel B: Job Skill Requirements				
$r_{c_t,v}$	**0.315**	0.362	0.316	0.189
$r_{c_t,m}$	0.271	**0.338**	0.313	0.168
$r_{c_t,t}$	0.114	0.198	**0.277**	0.0951
$r_{c_t,s}$	0.311	0.299	0.181	**0.179**

Table 2.A6: Mismatch and the Business Cycle: Effects by Skill

The table reports coefficients with robust standard errors clustered at the individual level reported in parentheses. The dependent variable is the level of total (Panel A), positive (Panel B) and negative (Panel C) mismatch in each skill dimension: math (m^m_{i,c_t}), verbal (m^v_{i,c_t}), technical (m^t_{i,c_t}) and social (m^s_{i,c_t}). The mismatch measure in skill j is defined as $m^j_{i,c_t} \equiv |a_{i,j} - r_{c_t,j}|$, where $a_{i,j}$ is the worker i's ability in skill j and $r_{c_t,j}$ the job requirements of skill j. All columns include a quadratic polynomial in age, and the following fixed effects: individual, month, year, region, one-digit industry and one-digit occupation. The sample includes all worker-job matches between 1979 and 2012. ***, ** and * represent statistical significance at 1%, 5% and 10% levels, respectively.

Dependent Variable:	m^m_{i,c_t}	m^v_{i,c_t}	m^t_{i,c_t}	m^s_{i,c_t}
	(1)	(2)	(3)	(4)
Panel A: Total Mismatch				
Unemployment$_t$	-0.202^{***}	-0.097	-0.230^{***}	-0.034
	(0.071)	(0.071)	(0.067)	(0.071)
Observations	510788	510788	510788	510788
Adjusted R^2	0.498	0.501	0.547	0.604
Panel B: Positive Mismatch				
Unemployment$_t$	-0.058	-0.026	-0.085^{*}	0.004
	(0.049)	(0.048)	(0.051)	(0.045)
Observations	510788	510788	510788	510788
Adjusted R^2	0.752	0.780	0.773	0.765
Panel C: Negative Mismatch				
Unemployment$_t$	-0.143^{***}	-0.071	-0.146^{***}	-0.037
	(0.045)	(0.044)	(0.043)	(0.052)
Observations	510788	510788	510788	510788
Adjusted R^2	0.747	0.771	0.740	0.812

Table 2.A7: Mismatch and the Business Cycle: Heterogeneous Effects By Skill

The table reports coefficients with robust standard errors clustered at the individual level reported in parentheses. In column 1, the dependent variable is total mismatch, m_{i,c_t} (Equation 2.1). In columns 2-5, the dependent variable is the level of mismatch in each skill dimension: math (m^m_{i,c_t}), verbal (m^v_{i,c_t}), technical (m^t_{i,c_t}) and social (m^s_{i,c_t}). The mismatch measure in skill j is defined as $m^j_{i,c_t} \equiv |a_{i,j} - r_{c_t,j}|$, where $a_{i,j}$ is the worker i's ability in skill j and $r_{c_t,j}$ the job requirements of skill j. All columns include a quadratic polynomial in age, and the following fixed effects: individual, month, year, region, one-digit industry and one-digit occupation. The sample includes all worker-job matches between 1979 and 2012. ***, ** and * represent statistical significance at 1%, 5% and 10% levels, respectively.

Dependent Variable:	m_{i,c_t}	m^m_{i,c_t}	m^v_{i,c_t}	m^t_{i,c_t}	m^s_{i,c_t}
	(1)	(2)	(3)	(4)	(5)
Unemployment$_t$	−0.159***	−0.222***	−0.123*	−0.242***	−0.048
	(0.050)	(0.071)	(0.072)	(0.068)	(0.071)
EE'$_{i,t}$ × Unemployment$_t$	0.146	0.044	0.165	0.174	0.203
	(0.093)	(0.132)	(0.135)	(0.127)	(0.132)
UE$_{i,t}$ × Unemployment$_t$	0.378***	0.486***	0.566***	0.202*	0.260**
	(0.085)	(0.120)	(0.119)	(0.114)	(0.124)
Observations	510788	510788	510788	510788	510788
Adjusted R^2	0.500	0.498	0.501	0.547	0.604

Table 2.A8: Redefining New hires from Unemployment

The table reports coefficients from an OLS regression with robust standard errors clustered at the individual level reported in parentheses. m_{i,c_t} is total mismatch, m^+_{i,c_t} is positive mismatch and m^-_{i,c_t} negative mismatch. Estimation relies on a sample in which new hires from unemployment with jobless spells lower than 1 (columns 1-3), 2(columns 4-6) and 3 columns 7-9) months where redefined as job-to-job transitions. $EE'_{i,t}$ is a dummy for whether individual i is a new hire from employment and $UE_{i,t}$ is a dummy for whether individual i is a new hire from unemployment. All columns include a quadratic polynomial in age, and the following fixed effects: individual, month, year, region, one-digit industry and one-digit occupation. The sample includes all worker-job matches between 1979 and 2012. ***, ** and * represent statistical significance at 1%, 5% and 10% levels, respectively.

	1 month			2 months			3 months		
	$m_{i,t}$	$m^+_{i,t}$	$m^-_{i,t}$	$m_{i,t}$	$m^+_{i,t}$	$m^-_{i,t}$	$m_{i,t}$	$m^+_{i,t}$	$m^-_{i,t}$
Unemployment$_t$	-0.159***	-0.050	-0.109***	-0.158***	-0.049	-0.109***	-0.158***	-0.049	-0.109***
	(0.050)	(0.038)	(0.035)	(0.050)	(0.038)	(0.035)	(0.050)	(0.038)	(0.035)
$EE'_{i,t}$ × Unemployment$_t$	0.153*	0.078	0.075	0.165**	0.092	0.073	0.188**	0.093	0.096*
	(0.086)	(0.063)	(0.058)	(0.081)	(0.059)	(0.055)	(0.078)	(0.057)	(0.053)
$UE_{i,t}$ × Unemployment$_t$	0.442***	0.195***	0.247***	0.467***	0.188**	0.279***	0.483***	0.209**	0.275***
	(0.097)	(0.070)	(0.064)	(0.109)	(0.078)	(0.072)	(0.121)	(0.087)	(0.079)
Observations	510788	510788	510788	510788	510788	510788	510788	510788	510788
Adjusted R^2	0.500	0.771	0.763	0.500	0.771	0.763	0.500	0.771	0.763

Table 2.A9: Occupation-Year FE

The table reports coefficients from an OLS regression with robust standard errors clustered at the individual level reported in parentheses. In columns 1-2, the dependent variable is mismatch, m_{i,c_t}, in columns 3-4 is positive mismatch, m^{+}_{i,c_t}, and in columns 5-6 is negative mismatch, m^{-}_{i,c_t}. $EE'_{i,t}$ is a dummy for whether individual i is a new hire from employment and $UE_{i,t}$ is a dummy for whether individual i is a new hire from unemployment. All columns include a quadratic polynomial in age, and the following fixed effects: individual, month, region, one-digit industry and one-digit occupation-year. The sample includes all worker-job matches between 1979 and 2006. ***, ** and * represent statistical significance at 1%, 5% and 10% levels, respectively.

	m_{i,c_t}		m^{+}_{i,c_t}		m^{-}_{i,c_t}	
	(1)	(2)	(3)	(4)	(5)	(6)
Unemployment$_t$	−0.127**	−0.145***	−0.035	−0.044	−0.091***	−0.101***
	(0.049)	(0.050)	(0.037)	(0.037)	(0.035)	(0.036)
EE'$_{i,t}$, × Unemployment$_t$		0.147		0.077		0.070
		(0.093)		(0.068)		(0.064)
UE$_{i,t}$, × Unemployment$_t$		0.365***		0.161***		0.203***
		(0.085)		(0.062)		(0.056)
Observations	510788	510788	510788	510788	510788	510788
Adjusted R^2	0.503	0.503	0.773	0.773	0.765	0.765

Table 2.A10: Additional Controls

The table reports coefficients from an OLS regression with robust standard errors clustered at the individual level reported in parentheses. Panel A adds as a control Unemployment in the previous month, Panel B wages, and Panel C substitutes individual's age by labor market experience. In columns 1-2, the dependent variable is mismatch, m_{i,c_t}, in columns 3-4 is positive mismatch, m^+_{i,c_t}, and in columns 5-6 is negative mismatch, m^-_{i,c_t}. $EE'_{i,t}$ is a dummy for whether individual i is a new hire from employment and $UE_{i,t}$ is a dummy for whether individual i is a new hire from unemployment. All columns include a quadratic polynomial in age, and the following fixed effects: individual, month, year, region, one-digit industry and one-digit occupation. The sample includes all worker-job matches between 1979 and 2012. ***, ** and * represent statistical significance at 1%, 5% and 10% levels, respectively.

Dependent Variable:	m_{i,c_t}		m^+_{i,c_t}		m^-_{i,c_t}	
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Lagged Unemployment						
Unemployment_t	-0.175^{***}	-0.194^{***}	-0.050	-0.060	-0.125^{***}	-0.135^{***}
	(0.062)	(0.063)	(0.047)	(0.048)	(0.040)	(0.041)
$\text{EE'}_{i,t}, \times \text{Unemployment}_t$		0.147		0.072		0.074
		(0.093)		(0.068)		(0.064)
$\text{UE}_{i,t}, \times \text{Unemployment}_t$		0.383^{***}		0.168^{***}		0.216^{***}
		(0.084)		(0.061)		(0.056)
Observations	510788	510788	510788	510788	510788	510788
Adjusted R^2	0.500	0.500	0.771	0.771	0.763	0.763
Panel B: Wages						
Unemployment_t	-0.148^{***}	-0.167^{***}	-0.056	-0.064^{*}	-0.093^{***}	-0.103^{***}
	(0.051)	(0.051)	(0.038)	(0.038)	(0.036)	(0.036)
$\text{EE'}_{i,t}, \times \text{Unemployment}_t$		0.152		0.078		0.074
		(0.094)		(0.068)		(0.065)
$\text{UE}_{i,t}, \times \text{Unemployment}_t$		0.387^{***}		0.165^{***}		0.222^{***}
		(0.085)		(0.063)		(0.057)
Observations	497262	497262	497262	497262	497262	497262
Adjusted R^2	0.503	0.504	0.773	0.773	0.764	0.764
Panel C: Experience						
Unemployment_t	-0.132^{***}	-0.167^{***}	-0.032	-0.064^{*}	-0.100^{***}	-0.103^{***}
	(0.049)	(0.051)	(0.037)	(0.038)	(0.035)	(0.036)
$\text{EE'}_{i,t}, \times \text{Unemployment}_t$		0.152		0.078		0.074
		(0.094)		(0.068)		(0.065)
$\text{UE}_{i,t}, \times \text{Unemployment}_t$		0.387^{***}		0.165^{***}		0.222^{***}
		(0.085)		(0.063)		(0.396)
Observations	510788	497262	510788	497262	510788	497262
Adjusted R^2	0.500	0.504	0.771	0.773	0.763	0.764

Table 2.A11: Alternative Measures for the State of the Economy

The table reports coefficients from an OLS regression with robust standard errors clustered at the individual level reported in parentheses. In columns 1-2, the dependent variable is mismatch, m_{i,c_t}, in columns 3-4 is positive mismatch, m^{+}_{i,c_t}, and in columns 5-6 is negative mismatch, m^{-}_{i,c_t}. Panel A uses as the business cycle indicator the Composite Help-Wanted index developed by Barnichon (2010), Panel B uses the industrial production index, and Panel C uses the deviations from the Hodrick-Prescott filtered unemployment rate. $EE'_{i,t}$ is a dummy for whether individual i is a new hire from employment and $UE_{i,t}$ is a dummy for whether individual i is a new hire from unemployment. All columns include a quadratic polynomial in age, and the following fixed effects: individual, month, year, region, one-digit industry and one-digit occupation. The sample includes all worker-job matches between 1979 and 2012. ***, ** and * represent statistical significance at 1%, 5% and 10% levels, respectively.

Dependent Variable:	m_{i,c_t}		m^{+}_{i,c_t}		m^{-}_{i,c_t}	
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Vacancies						
Vacancies$_t$	0.241**	0.276**	0.056	0.075	0.186**	0.201**
	(0.112)	(0.113)	(0.082)	(0.082)	(0.081)	(0.082)
EE'$_{i,t}$ × Vacancies$_t$		−0.297		−0.044		−0.253
		(0.290)		(0.216)		(0.201)
UE$_{i,t}$ × Vacancies$_t$		−0.969***		−0.595***		−0.374**
		(0.284)		(0.210)		(0.180)
Observations	510788	510788	510788	510788	510788	510788
Adjusted R^2	0.500	0.500	0.771	0.771	0.763	0.763
Panel B: Industrial Production						
Industrial Prod$_t$	0.034***	0.034***	0.014*	0.014*	0.019**	0.020**
	(0.012)	(0.012)	(0.009)	(0.009)	(0.009)	(0.009)
EE'$_{i,t}$, × Industrial Prod$_t$		−0.001		−0.001		0.001
		(0.010)		(0.007)		(0.007)
UE$_{i,t}$, × Industrial Prod$_t$		−0.022**		−0.001		−0.020***
		(0.009)		(0.006)		(0.006)
Adjusted R^2	0.500	0.500	0.771	0.771	0.763	0.763
Observations	510788	510788	510788	510788	510788	510788
Panel C: Unemployment (HP-filtered)						
Unemployment$_t$	−0.111**	−0.135**	−0.019	−0.031	−0.092**	−0.104***
	(0.052)	(0.053)	(0.040)	(0.040)	(0.036)	(0.037)
EE'$_{i,t}$ × Unemployment$_t$		0.214*		0.045		0.168*
		(0.127)		(0.095)		(0.086)
UE$_{i,t}$ × Unemployment$_t$		0.481***		0.264***		0.217***
		(0.114)		(0.085)		(0.073)
Adjusted R^2	0.500	0.500	0.771	0.771	0.763	0.763
Observations	510788	510788	510788	510788	510788	510788

Table 2.A12: Great Recession

The table reports coefficients from an OLS regression with robust standard errors clustered at the individual level reported in parentheses. In columns 1-2, the dependent variable is mismatch, m_{i,c_t}, in columns 3-4 is positive mismatch, m^{+}_{i,c_t}, and in columns 5-6 is negative mismatch, m^{-}_{i,c_t}. $EE'_{i,t}$ is a dummy for whether individual i is a new hire from employment and $UE_{i,t}$ is a dummy for whether individual i is a new hire from unemployment. All columns include a quadratic polynomial in age, and the following fixed effects: individual, month, year, region, one-digit industry and one-digit occupation. The sample includes all worker-job matches between 1979 and 2006. ***, ** and * represent statistical significance at 1%, 5% and 10% levels, respectively.

Dependent Variable:	m_{i,c_t}		m^{+}_{i,c_t}		m^{-}_{i,c_t}	
	(1)	(2)	(3)	(4)	(5)	(6)
Unemployment$_t$	−0.147**	−0.169***	−0.030	−0.040	−0.117***	−0.129***
	(0.061)	(0.062)	(0.047)	(0.048)	(0.042)	(0.042)
EE'$_{i,t}$ × Unemployment$_t$		0.121		0.065		0.056
		(0.095)		(0.070)		(0.065)
UE$_{i,t}$ × Unemployment$_t$		0.374***		0.157**		0.217***
		(0.088)		(0.064)		(0.058)
Observations	457246	457246	457246	457246	457246	457246
Adjusted R^2	0.505	0.505	0.778	0.778	0.764	0.764

Table 2.A13: Only Males

The table reports coefficients from an OLS regression with robust standard errors clustered at the individual level reported in parentheses. In columns 1-2, the dependent variable is mismatch, m_{i,c_t}, in columns 3-4 is positive mismatch, m^{+}_{i,c_t}, and in columns 5-6 is negative mismatch, m^{-}_{i,c_t}. $EE'_{i,t}$ is a dummy for whether individual i is a new hire from employment and $UE_{i,t}$ is a dummy for whether individual i is a new hire from unemployment. All columns include a quadratic polynomial in age, and the following fixed effects: individual, month, year, one-digit industry and one-digit occupation. The sample includes all worker-job matches between 1979 and 2012 for males. ***, ** and * represent statistical significance at 1%, 5% and 10% levels, respectively.

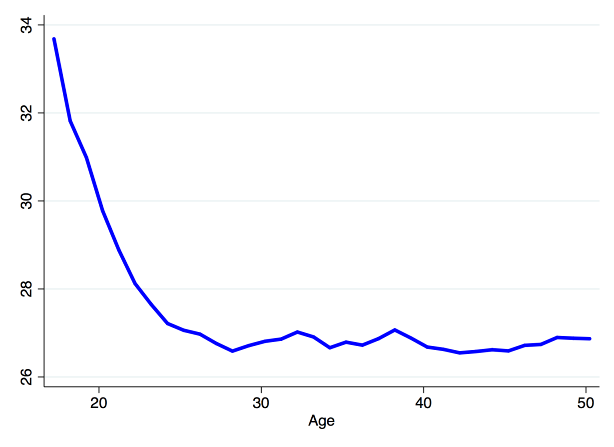
Dependent Variable:	m_{i,c_t}		m^{+}_{i,c_t}		m^{-}_{i,c_t}	
	(1)	(2)	(3)	(4)	(5)	(6)
Unemployment$_t$	−0.221***	−0.243***	−0.088	−0.096*	−0.134**	−0.147***
	(0.075)	(0.076)	(0.055)	(0.056)	(0.052)	(0.053)
EE'$_{i,t}$ × Unemployment$_t$		0.249*		0.131		0.118
		(0.131)		(0.097)		(0.087)
UE$_{i,t}$ × Unemployment$_t$		0.396***		0.120		0.275***
		(0.127)		(0.092)		(0.082)
		(0.203)		(0.146)		(0.131)
Observations	266459	266459	266459	266459	266459	266459
Adjusted R^2	0.487	0.488	0.766	0.766	0.760	0.760

Table 2.A14: Alternative Mismatch Measures

The table reports coefficients from an OLS regression with robust standard errors clustered at the individual level reported in parentheses. In columns 1-2, the dependent variable is mismatch, m_{i,c_t}, in columns 3-4 is positive mismatch, m^+_{i,c_t}, and in columns 5-6 is negative mismatch, m^-_{i,c_t}. Panel A uses a version version of the mismatch index with only 3 skill dimensions (math, verbal and social). In Panel B, each skill has a different weight in the computation of the mismatch index. The weights are the ones used in Guvenen et al. (2018): (verbal, math, social) = (0.43, 0.43, 0.12). Panel C uses a mismatch index computed as in Speer (2017). Panel D uses a mismatch measure in terms of mean squared deviation between worker's abilities and job skill requirements: $m_{i,c_t} \equiv \left(\sum_{j=1}^{J}(a_{i,j} - r_{c_t,j})^2/J\right)^{0.5}$. $EE'_{i,t}$ is a dummy for whether individual i is a new hire from employment and $UE_{i,t}$ is a dummy for whether individual i is a new hire from unemployment. All columns include a quadratic polynomial in age, and the following fixed effects: individual, month, year, one-digit industry and one-digit occupation. The sample includes all worker-job matches between 1979 and 2012. ***, ** and * represent statistical significance at 1%, 5% and 10% levels, respectively.

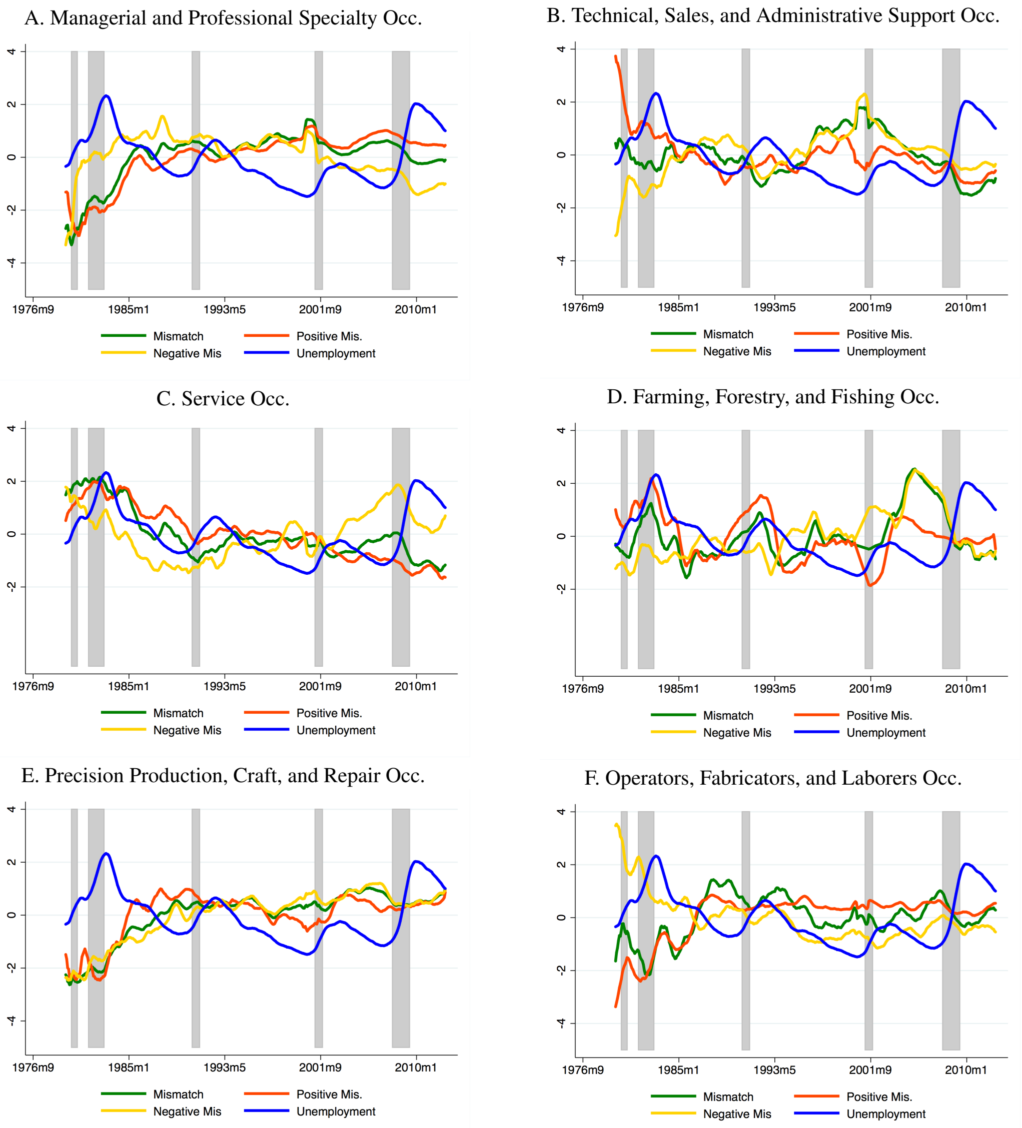
Dependent Variable:	m_{i,c_t}		m^+_{i,c_t}		m^-_{i,c_t}	
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Guvenen et al. (2015), unweighted						
Unemployment$_t$	−0.110**	−0.130**	−0.027	−0.036	−0.083**	−0.094***
	(0.050)	(0.051)	(0.036)	(0.036)	(0.036)	(0.036)
EE'$_{i,t}$ × Unemployment$_t$		0.136		0.063		0.073
		(0.093)		(0.065)		(0.065)
UE$_{i,t}$ × Unemployment$_t$		0.433***		0.201***		0.232***
		(0.086)		(0.060)		(0.058)
Adjusted R^2	0.510	0.511	0.770	0.770	0.779	0.779
Panel B: Guvenen et al. (2015), weighted						
Unemployment$_t$	−0.132**	−0.154***	−0.036	−0.047	−0.097**	−0.047
	(0.057)	(0.058)	(0.040)	(0.041)	(0.038)	(0.041)
EE'$_{i,t}$ × Unemployment$_t$		0.114		0.053		0.053
		(0.107)		(0.073)		(0.073)
UE$_{i,t}$ × Unemployment$_t$		0.483***		0.236***		0.236***
		(0.097)		(0.066)		(0.066)
Adjusted R^2	0.485	0.485	0.768	0.768	0.762	0.768
Panel C: Speer (2017)						
Unemployment$_t$	−0.158***	−0.175***	−0.072*	−0.084**	−0.085***	−0.094***
	(0.043)	(0.043)	(0.038)	(0.038)	(0.027)	(0.028)
EE'$_{i,t}$ × Unemployment$_t$		0.089		0.086		0.041
		(0.083)		(0.070)		(0.052)
UE$_{i,t}$ × Unemployment$_t$		0.375***		0.236***		0.187***
		(0.075)		(0.064)		(0.045)
Adjusted R^2	0.517	0.517	0.781	0.781	0.848	0.848
Panel D: Alternative measure						
Unemployment$_t$	−0.137***	−0.157***	−0.048	−0.057	−0.090**	−0.099**
	(0.050)	(0.051)	(0.040)	(0.040)	(0.039)	(0.040)
EE'$_{i,t}$ × Unemployment$_t$		0.177*		0.071		0.066
		(0.092)		(0.074)		(0.069)
UE$_{i,t}$ × Unemployment$_t$		0.394***		0.181***		0.200***
		(0.086)		(0.065)		(0.062)
Adjusted R^2	0.519	0.519	0.785	0.785	0.772	0.772
Observations	510788	510788	510788	510788	510788	510788

Figure 2.A1: Mismatch and Age



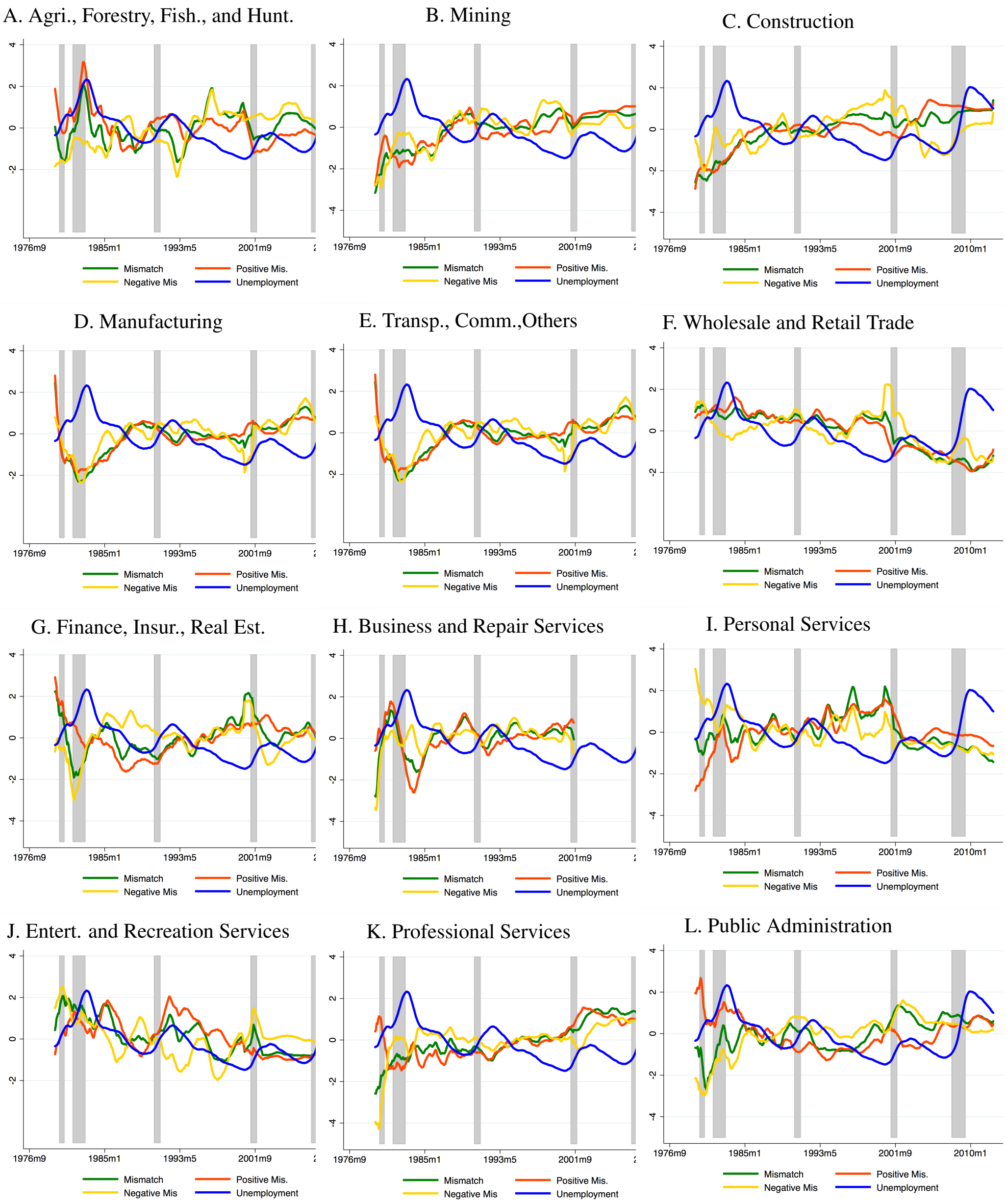
Notes: The graph plots average mismatch across employed individuals in the sample for each age.

Figure 2.A2: Mismatch by Occupation and Unemployment



Notes: Data are shown in standard deviations. *Unemployment* is the monthly unemployment rate at the national level. *Mismatch* is average mismatch M_t in Equation (2.5). *Positive Mis.* and *Negative Mis.* are, respectively, average positive and negative mismatch as defined in Section 2.2.1 and constructed the same way as M_t. Shaded areas correspond to NBER recessions.

Figure 2.A3: Mismatch by Industry and Unemployment



Notes: Data are shown in standard deviations. *Unemployment* is the monthly unemployment rate at the national level. *Mismatch* is average mismatch M_t in Equation (2.5). *Positive Mis.* and *Negative Mis.* are, respectively, average positive and negative mismatch as defined in Section 2.2.1 and constructed the same way as M_t. Shaded areas correspond to NBER recessions.

2.5.2 Theoretical Appendix

For notation simplicity, we index value functions with time t, to express their dependence on the aggregate state, i.e. $V_t(\cdot) \equiv V(\cdot, \Omega_t)$.

Value of Unemployment Let $U_t(\hat{a}, \Sigma)$ and $W_t(\hat{a}, \Sigma)$ be the values of unemployment and employment for a given worker, respectively. Upon separation, unemployed workers change occupation with probability π_t, meet a vacancy in career k' with probability $p(\theta_t)$, and receive $W(r^{k'} - \bar{a}, S_a, z_t)$; with probability $1 - \pi_t$ they remain in the same occupation and receive $W((r^{k'} - \hat{a}, \Sigma, z_t)$.

$$\begin{aligned} \rho U_t(\hat{a}, \Sigma) &= b + \pi_t \int p(\theta_t) \int \mathbb{E}_t \Big[\underbrace{W_t(r^{k'} - \bar{a}, S_a) - U_t(\hat{a}, \Sigma)}_{\text{switch}} \Big] \, dG^{k'}(r) \, dk' \\ &+ (1 - \pi_t) p(\theta_t) \int \mathbb{E}_t \Big[\underbrace{W_t(r^k - \hat{a}, \Sigma) - U_t(\hat{a}, \Sigma)}_{\text{no switch}} \Big] \, dG^k(r) \quad \text{(2.A1)} \end{aligned}$$

We assume that the unemployed have zero bargaining power. As such, whether they switch or not occupation, firms hire unemployed workers at a wage w which sets the value of the match for the worker to the value of unemployment:

$$W_t(r^k - \bar{a}, S_a) = W_t(r^k - \hat{a}, \Sigma) = U_t(\hat{a}, \Sigma_i) \quad \text{(2.A2)}$$

. This implies that that the value of unemployment does not depend on the vacancy distribution.

Value and Surplus of a Match Let $P(\mu, \Sigma, z^i)$ the value of a match (which corresponds to the joint value for the worker and the firm) with mismatch estimate μ and uncertainty Σ, and aggregate productivity z. Inside the continuation region (if the match continues) P satisfies the following HJB equation:

$$\begin{aligned} \rho P(\mu, \Sigma, z^i) &= z^i - \psi(\mu^2 + \Sigma) \left(\frac{\Sigma}{\sigma_s} \right)^2 \left(\frac{P_{\mu\mu}}{2} - P_\Sigma \right) \\ &- \delta(P(\mu, \Sigma, z^i) - U_t(\hat{a}, \Sigma) + \kappa) \quad \text{(2.A3)} \\ &+ \lambda_j \left[P(\mu, \Sigma, z^j) - P(\mu, \Sigma, z^i) \right] \quad \text{(2.A4)} \end{aligned}$$

where $\mu = r^k - \hat{a}$, $z^i - \psi(\mu^2 + \Sigma)$ is the expected output flow from the match with $z^i \in \{z_L, z_H\}$, and δ is the intensity of exogenous separation. The first term corresponds to the output flow of the match plus the value of learning about the mismatch level. The term in 2.A3 captures the capital loss following exogenous separation at rate: a match is destroyed at an exogenous rate δ, in which case the firm obtains the value of a vacancy which is zero due to free entry and pays the firing cost κ and the unemployed worker receives the value of unemployment, U_t. The last term reflects the net benefit from a new aggregate productivity shock. Finally, let the joint surplus of a match be defined as $J_t(\mu, \Sigma, z^i) \equiv P_t(\mu, \Sigma, z^i) - U_t - V_t$. Given that V_t equals zero by free entry, by subtracting the value of unemployment to the value of the match we obtain the following

$$\begin{aligned}(\rho + \delta) J(\mu, \Sigma, z^i) &= z^i - \psi(\mu^2 + \Sigma) - b + \left(\frac{\Sigma}{\sigma_s}\right)^2 \left(\frac{J_{\mu\mu}}{2} - J_\Sigma\right) - \delta\kappa \\ &+ \lambda_j \left[J(\mu, \Sigma, z^j) - J(\mu, \Sigma, z^i)\right] \qquad (2.A5)\end{aligned}$$

where b is the flow of unemployment benefits or home production. Given Equation (2.A5), we can observe that as in Lise and Robin (2017), the match surplus at time t depends on time only through the current aggregate productivity shock z and does not depend on the distributions of vacancies, unemployed workers.

Chapter 3

Wage Cyclicality and Mismatch

3.1 Introduction

Macroeconomists have long been interested in the dynamics of the price of labor over the business cycle. Aggregate wage data suggests relatively little movement in real wages as compared to output and unemployment; whereas an extensive empirical literature relying on panel data shows that new hires wages are more cyclical than the ones of workers in ongoing job relationships. This literature, however, has not yet been able to assess whether these movements in wages over the cycle capture wage cyclicality or instead confounding variation in the wages new hires due to workers moving to better jobs during expansions, i.e. the procyclicality of match quality. This chapter addresses this issue by relying on the skill mismatch measure developed by Guvenen et al. (2018) to account for cyclical dynamics of match quality. I provide evidence that excess wage cyclicality of workers making job-to-job transitions goes beyond skill mismatch cyclicality, and that skill mismatch amplifies wage cyclicality. These results bring important insights to the ongoing debate about what wage setting protocol is consistent with the observed behavior of wages and the role of wage rigidity in search and matching models.

Search and matching models of the labor market in the tradition of Diamond, Mortensen and Pissarides have become a leading model of unemployment in macroeconomics. Under this framework, wages are not competitively determined due to matching frictions, thus most of the literature considers that they are determined by a particular solution to a bargaining problem between workers and firms: the Nash bargaining solution. The outcome is the "Nash wage equation", according to which wages correspond to a linear combination of the match productivity and the worker's returns from search and other non-market activities As such, wages increase when productivity is high. This feature, however, reduces incentives to create jobs. Given this, search and matching models fail to generate the empirical volatility of the vacancy-unemployment ratio under common parameter value, as documented by Shimer (2005).

Motivated by aggregated time-series evidence showing that real wages exhibit relatively little variation as compared to output, Shimer (2005) and Hall (2005) show that by incorporating wage rigidity instead of period-by-period Nash bargaining over wages, they are able to account for unemployment fluctuations. However, evidence based on aggregate data ignores the variation in the composition of the workforce over the cycle, and assumes that the dynamics of real wages and over the cycle is the same for all individuals or groups of individuals. In recent years, the presence of compositional effects have been shown to be important.

Starting with Bils (1985), an extensive literature based on individual-level panel data points toward a procyclical behavior of real wages, and finds that wages of newly hired workers are more cyclical than wages of workers in ongoing employment relationships (Solon et al., 1994; Solon, 1994; Barlevy, 2001; Shin and Solon, 2007; Carneiro et al., 2012, for example). Following this micro-level evidence, Pissarides (2009) calls into question efforts to incorporate wage rigidity into macroeconomic models arguing that what matters for job creation are the wages of new matches. As such, "...explanations of the unemployment volatility puzzle have to preserve the cyclical volatility of wages." (Pissarides, 2009, 1339). More recently, Gertler et al. (2016) contest this view. Using the SIPP, they find evidence that wages of workers making job-to-job transitions are more cyclical than those in ongoing job relationships, but no evidence of excess wage cyclicality among new hires from unemployment. Following evidence showing jobs starting in recessions are shorter and that graduates that enter the labor market during a recession receive lower wages (Bowlus, 1995; Kahn, 2010; Oreopoulos et al., 2012), Gertler et al. (2016) argue that wage cyclicality of new hires captures instead variation in new hire wages that is due to workers moving to better jobs during expansions, i.e. the procyclicality of match quality across new hires from employment. Therefore, supporting the introduction of wage rigidity into search and matching models. However, using a measure of skill mismatch between worker's skills and job skill requirements to proxy match quality, a recent work by Baley et al. (2018) show that skill mismatch is countercyclical for new hires from unemployment, i.e. match quality decreases in recessions for workers coming from unemployment, but acyclical for workers making job-to-job transitions.

This chapter contributes to the ongoing debate about wage dynamics over the business cycle. While I rely on a dataset and empirical framework that are common in the literature, the key novelty is that I account for cyclical movements in match quality.[1] To address this issue, I measure the sensitivity of wages to changes in aggregate labor market conditions using a worker-level panel from the 1979 National Longitudinal Study of Youth. Using this dataset allows me to separately estimate wage cyclicality of new hires from unemployment and those making job-to-job transitions, but more importantly it allows me to measure match quality through the lens of the skill mismatch measure developed by Guvenen et al. (2018). This measure is defined as the difference between a worker's abilities in different skills and how intensive these skills are required by a job, thus it can be interpreted as the *lack* of match quality: the larger is this difference, the lower is the quality of a match. In order to estimate the semi-elasticity of wages with respect to aggregate economic conditions, my identification strategy takes advantage of within-individual variation in the unemployment rate across months the individual reported to be employed, using the monthly unemployment rate at the national level to describe the state of the economy.

I first show that with the worker-level panel at hand I can replicate the results of the existing literature. On the one hand, when I do not separate between job switchers and new hires from unemployment, I find that new hire wages are more cyclical than the wages of continuing workers. In particular, I recover a wage semi-elasticity of -3.1%, which compares to -3.0% in Bils (1985), Barlevy (2001)

[1] Carneiro et al. (2012) control for job (permanent) unobserved heterogeneity using job fixed effects. The empirical exercise in this chapter differs from them in two ways. First, in the strategy to control for job match quality. Second, in the context in which this issue is studied: they use a matched employer-employee data from Portugal, while I use survey data from the U.S.

and Kudlyak (2014). On the other hand, consistent with Gertler et al. (2016), when I distinguish between job-to-job transitions and new hires coming from unemployment, I find that for the latter the excess wage cyclicality disappeears; whereas for the former wages are more cyclical than those of job stayers. While Gertler et al. (2016) argue that movements in wages of job switchers capture match quality dynamics over the cycle, their dataset does not allow them to disentangle wage cyclicality from fluctuations in match quality. To account for the dynamics of match quality over the cycle, I then introduce skill mismatch in the baseline regression. Two results can be drawn from this analysis. First, excess wage cyclicality for workers making job-to-job transitions goes beyond the dynamics of mismatch over the cycle. Second, mismatch amplifies wage cyclicality. For workers in ongoing job relationships and job switchers, I find that when the unemployment rate increases by one percentage point, workers in the 95$^{\text{th}}$ percentile of mismatch face a decrease in wages that is 2.1 times larger than the one faced by perfectly matched workers. More interestingly, I show that the extent to which wages of new hires from unemployment exhibit excessive cyclicality depends on mismatch. For perfectly matched workers, wages of new hires from unemployment are no more cyclical than those of job stayers consistent with Gertler et al. (2016); in contrast, for the 95$^{\text{th}}$ percentile of mismatch the wage semi-elasticity for new hires from unemployment to unemployment is 21% higher when compared to workers in ongoing job relationships. Thus, while recent research in macroeconomics has incorporated some form of wage rigidity to improve the ability of search and matching models to account for unemployment fluctuations (Hall, 2005; Gertler and Trigari, 2009; Blanchard and Galí, 2010), my findings seem to point against this explanation.

The remainder of the chapter proceeds as follows. The next section describes the data used in the empirical exercise. Section 3.3 outlines the empirical framework and Section 3.4 presents the main results as well as their sensitivity to alternative specifications. Section 3.5 concludes.

3.2 Data

To explore how wages evolve over the business cycle, I exploit a worker-level panel from the 1979 National Longitudinal Study of Youth (NLSY79) combined with aggregate data on economic conditions. While other studies exploring wage cyclicality use the Current Population Survey (CPS) (Haefke et al., 2013), which has a much larger number of observations and is representative of the U.S. economy — recall that the NLSY79 constitutes a representative sample of a cohort that were between 14 and 22 years old when they were first interviewed in 1979 — or the Panel Study of Income Dynamics (PSID) (Barlevy, 2001), I opt for the NLSY79 for three main reasons. First, by tracking individuals over the years, I can isolate the individual-specific fixed effects such as unobserved ability. This is not possible using the CPS because respondents are not followed over time. Second, if a worker held more than one job at a given point in time, the NLSY79 kept a separate record for each job, as opposed to PSID data that report the average wage in such cases. Third, the NLSY79 allows me to compute a measure of match quality, as described below. This is key to disentangle wage cyclicality from confounding variation in match quality over the cycle, the main purpose of our empirical exercise.

Wages and Employment I use data on wages and employment from the NLSY79. This survey tracks information on the labor market history of a representative sample of individuals that were between 14 to 21 years old at the time of the first year interview from 1979 until today. As in Baley et al. (2018), I restrict my focus to a sub-sample of males and females from the cross-sectional sample of the NLSY79 that covers the period 1979-2012. This sample consists of 2,991 individuals. Appendix 2.5.1 in the previous chapter documents the steps taken to derive the sample and reports descriptive statistics about it.

From the Work History Data file, a detailed week-by-week work history data from the NLSY79, I build a monthly panel at the worker level as in Baley et al. (2018). I opt for a monthly frequency data because it mitigates concerns about missing transitions when there is a high turnover. I can then observe individuals' labor market status, and whenever they are employed I have information about wages measured by the hour rate of pay. These include tips, overtime and bonuses, are corrected for outliers[2], and deflated using the annual consumer price index of the year the observation refers to from the Bureau of Labor Statistics (BLS). In addition to this, I observe *(i)* three-digit level occupation and industry, *(ii)* the extent of skill mismatch between worker$i$'s abilities in skill j (a_i, j) and job requirements in the occupation[3] held at time t ($r_{c_t,j}$) in 4 skill dimensions ($J = \{$verbal, math, technical, social$\}$) computed as in Guvenen et al. (2018) and Baley et al. (2018)[4],

$$m_{i,c_t} \equiv \sum_{j=1}^{J} \frac{1}{J} |a_{i,j} - r_{c_t,j}|, \tag{3.1}$$

and *(iii)* job mobility. This allows me to identify workers making job-to-job transitions and newly hired workers out of unemployment. In particular, I identify an individual to be a *job switcher* if she was employed at time t and $t-1$, but with a different employer; and to be a *new hire from unemployment* if she was unemployed in month $t-1$ (i.e. reported to be not working, unemployed or out of the labor force) but employed in month t. To identify these transitions, I focus on the main job an individual works at in a given month. As common in the literature using the NLSY79, the *main job* corresponds to the job at which an individual spends the most hours working within a given month. In an average month, 5% of employed individuals found their job within the current month, out of which 2.1% are classified as job switchers and the remaining 2.9% as new hires from non-employment.

State of the Economy To describe economic conditions, I use the monthly unemployment rate. This is measured using the civilian unemployment rate at the national level from the BLS. During the period covered by our sample, 1979-2012, this indicator varied from 3.8% to 10.8%.

[2]As in Guvenen et al. (2018), I trim wages at the top and bottom 0.1% of observations

[3]As in the previous chapter, an occupation is defined by Dorn (2009)'s three-digit occupational classification system.

[4]Appendix 2.5.1 provides a detailed description of how the skill mismatch index described in Equation (3.1) is computed.

Table 3.A1 reports descriptive statistics of the main variables I use in the empirical analysis.

3.3 Empirical Framework

To assess the extent to which wages move with the business cycle, I use the semi-elasticity of wages with respect to the aggregate unemployment rate. This is a commonly used measure of cyclicality in the literature studying wage dynamics over the cycle (Pissarides, 2009). To estimate wage semi-elasticity, I adopt Gertler et al. (2016)'s framework. This allows me to isolate the wage behavior of new hires from unemployment from workers switching jobs and those in ongoing job relationships. Specifically, the estimating regression takes the following form:

$$\begin{aligned} w_{i,t} = \lambda_0 \; + \; \lambda_1 U_t \; + \; \lambda_2 EE'_{i,t} \; + \; \lambda_3 UE_{i,t} \; + \; \lambda_4 (U_t \times EE'_{i,t}) + \; \lambda_5 (U_t \times UE_{i,t}) \\ + \; \lambda_6 m_{i,c_t} \; + \; trend \; + \; \gamma' x_{i,t} \; + \; \delta_i \; + \; \delta_m \; + \; \varepsilon_{i,t} \end{aligned} \tag{3.2}$$

where $w_{i,t}$ is the log real hourly earnings of individual i at time t. As in Gertler et al. (2016) and Baley et al. (2018), $UE_{i,t}$ and $EE'_{i,t}$ correspond to dummies for whether the worker i is a new hire from unemployment and a job switcher at time t, respectively. The coefficients of interest are λ_1, which measures the semi-elasticity of real wages with respect to the unemployment rate for job stayers; λ_4 that corresponds to the differential in the semi-elasticity of wages to the unemployment rate between job stayers and job-to-job transition; and λ_5 which captures excess wage cyclicality for new hires from unemployment. I also run a regression in which I pool the two types of new hires: instead of including $UE_{i,t}$ and $EE'_{i,t}$, I include a dummy variable $NH_{i,t} = UE_{i,t} + EE'_{i,t}$ which equals one if the individual i started a new job at time t. This is the standard framework used in the literature. The key novelty in specification (3.2) is that I am able to take into account cyclical variations in the quality of the match, i.e. job (permanent) unobserved heterogeneity, by using skill mismatch (m_{i,c_t}) as a control variable. Because this measure is defined by the extent to which workers abilities differ from the skills required by the job, it can be interpreted as the *lack* of match quality (Guvenen et al., 2018; Baley et al., 2018). Equation (3.2) also includes a set of time-varying individual characteristics $x_{i,t}$ including age, job tenure (a quadratic term), education level, region of residence, one-digit level industry and occupation; a time trend; and individual, monthly and yearly fixed effects. $\varepsilon_{i,t}$ corresponds to the error term, which includes all unobserved determinants of $w_{i,t}$. I cluster the standard errors at the individual level to allow for serial correlation.

Identification Taking advantage of the panel nature of the data, our identification strategy exploits within-individual variation in the unemployment rate across months when the individual is employed, using the aggregate unemployment rate of the month as a proxy for the state of the economy. To illustrate that there is indeed substantial variation in unemployment over the considered sample period, Figure 3.A1 plots unemployment rate over the years covered by our sample. The fixed individual effect allows me to take into account for systematic differences in the types of individuals who move over the business cycle. This is important as Solon et al. (1994) provides evidence that the employment pool shifts towards high-ability workers during expansions. In this context, the underlying assumption for unbiased estimation of the coefficients of interest, λ_1, λ_4 and λ_5, is that the selection process is constant over time. Following Bils (1985), the common approach in the literature to control for individual unobserved (permanent) heterogeneity has been to use first differences. However, taking first differences of individual wages restricts the exercise to workers that were employed both in the current and in the previous month (those making job-to-job transitions and workers in ongoing job relationships), as new hires that were unemployment or out of the labor force in the previous month did not receive a wage. Given this, I choose to define the wage equation in levels and use individual fixed effects to drop individual so that I can estimate separately wage cyclicality for both job switchers and new hires from unemployment.

To account for occupation and industry changes over the cycle, I use dummies that describe one-digit level industries and occupations. OLS estimation of Equation (3.2) also controls for individual characteristics such as age and job tenure (quadratic polynomial), region of residence, time shocks (monthly and yearly) common to all individuals, and unobserved factors that might be trending over time and that could be correlated with both unemployment and earnings. Note, however, that while Equation (3.2) controls for worker observed and unobserved (permanent) heterogeneity and job (permanent) unobserved heterogeneity, by including skill mismatch, time-varying controls and individual fixed effects, it does not control for firm heterogeneity over the cycle: wages can be lower in recessions because these are times when low-wage paying firms hire more. However, Carneiro et al. (2012), who are able to control for worker, firm and worker heterogeneity, show that by adding the control of firm permanent unobserved heterogeneity has a small effect on the estimates of the semi-elasticity of wages with respect to the unemployment rate for stayers, concluding that their results do not seem to support the hypothesis that workers move from low-paying to high-paying firms during expansions, and vice-versa during recessions. Given this, concerns regarding changes in the composition of firms over the cycle are small. If any, given Carneiro et al. (2012)'s evidence, the estimated coefficients are slightly downward biased. Thus, they constitute a lower bound of wage semi-elasticity with respect to economic conditions.[5]

3.4 Results

This section presents the estimation results. I first show that with the worker-level panel at hand I can obtain the results in the literature. Then, I introduce skill mismatch as a control into the baseline framework in order to take into account the dynamics of match quality over the cycle, and I show that mismatch has important amplification effects. Finally, I explore the sensitivity of the results to different specifications.

3.4.1 Main Findings

Table 3.A2 provides OLS estimates of the real wage semi-elasticity with respect to the aggregate unemployment rate.

[5] Carneiro et al. (2012) find that by including firm fixed effects the real wage sensitivity to the unemployment rate increases from 1.61% to 1.85% (in absolute value).

Replicating results in the literature Column 1 in Table 3.A2 reports OLS estimates of the regression that has been popular in the literature after the seminal paper by Bils (1985) to examine the response of individual level wages to changes in aggregate labor market conditions,

$$w_{i,t} = \lambda_0 + \lambda_1 U_t + \lambda_2 NH_{i,t} + \lambda_3 (U_t \times NH_{i,t}) + trend \\ + \gamma' x_{i,t} + \delta_i + \delta_m + \varepsilon_{i,t}, \tag{3.3}$$

where $NH_{i,t}$ equals to one if individual i started a new job at time t, i.e. it pools together workers making job-to-job transitions and new hires from unemployment. The coefficient λ_1 can be interpreted as the semi-elasticity of wages with respect to unemployment, while $\lambda_1 + \lambda_3$ captures the semi-elasticity of new hires. The key result in the literature is that both λ_1 and λ_3 are negative, suggesting that wages of new hires are more sensitive to aggregate labor market conditions. From the estimated coefficients in column 1, we can observe that for every percentage point rise in unemployment, the wages of workers in ongoing job relationships decrease by about 2.18%, whereas for new hires wages are 3.13% lower. The estimated semi-elasticity is significant at the 1% level for continuing workers and the new hire differential is significant at the 1% level as well.

Studies of wage cyclicality in the U.S. seem to find wages semi-elasticities with respect to the contemporaneous unemployment rate between -1.0% and -2.0% for job stayers, while the consensus in the literature for the cyclicality of the wages of newly hired workers is -3.0% for new hires. For instance, using annual NLSY data from 1966-1980, Bils (1985) reports a wage semi-elasticity of -3.0% for changers; Barlevy (2001), using both PSID and NLSY through 1993, recovers a semi-elasticity of -3.0% for job changers. More recently, Kudlyak (2014) relies on a sample from the NLSY79 over the period 1978-2004, and finds that a one percentage point increase in the unemployment rate is associated with a 3% decrease in wages of newly hired workers and a 1.78% decrease in wages of all workers. For a complete review of this literature, see Pissarides (2009).

The estimated coefficients in column 1 of Table 3.A2 recover the semi-elasticity of wages for new hires, however for job stayers my estimates are slightly at the high end of the literature range. I speculate that the higher estimates might be due to *(i)* the high-frequency of the data: the regressions are based on a monthly panel rather than a yearly panel as most of the literature, *(ii)* the fact that I focus on the main job of each respondent, whereas some literature restrict their focus to workers with only one job, or *(iii)* that in my sample there are young workers — the minimum age observed in the sample is 14. In the Section 3.4.2, I explore the sensitivity of my results to different age intervals.

In a recent paper, Haefke et al. (2013) argue that the key hiring flow to generate unemployment volatility in search and matching models with sticky wages is that of workers coming from unemployment, not that of workers making job-to-job transitions. As such, Gertler et al. (2016) move the literature a step forward by estimating a regression that allows to separate wage cyclicality of new hires from unemployment from movements in wages of workers making job transitions: Gertler et al. (2016)'s specification equals to the one in Equation (3.2) without controlling for skill mismatch. Their key finding is that for new hires coming from unemployment, wages are no more cyclical than those for workers in ongoing job relationships, but that there is significant evidence of procyclical changes in wages of job switchers. Column 2 shows that I can replicate their finding. When I distinguish between new hires coming from unemployment and job-to-job transitions, I find that for the former, the excess wage cyclicality disappears; whereas for the latter wages are more cyclical than those of job stayers: an increase in unemployment by one percentage point is associated with a fall in the real wage of 2.18% for on-going workers and new hires from unemployment, and of 3.75% for job-to-job transitions.

Disentangling wage cyclicality from match quality On the grounds that "it is hard to rationalize a bargaining mechanism whereby wages for new hires from employment are flexible, but wages for new hires from unemployment are not", Gertler et al. (2016) interpret their results as indicative of procyclical match quality for job switchers. This means that wages of job changers are not more cyclical than those of job stayers, instead they just reflect movements in the quality of the match over the cycle. Nonetheless, given the features of their data, they are not able to assess how much of the movements of job switchers' wages are due to variations of match quality. I now address this issue.

One can control for the dynamics of match quality over the cycle in two different ways: *(i)* by including job fixed effects and exploit within-job variation in wages, as in Carneiro et al. (2012), or *(ii)* by including a measure of match quality as a control. In this chapter, I focus on the latter approach using a unique feature of the data at hand that allows me to compute a multidimensional measure of skill mismatch in Equation (3.1). As previously, mentioned this measure corresponds to the unweighted average of the difference between workers abilities in different skills dimensions and how insensitively these skills are required by the job, thus it can be interpreted as the *lack* of match quality.

Column 3 reports OLS estimates of Equation (3.2). First, in line with Guvenen et al. (2018)'s findings, mismatch is negatively associated with wages. To illustrate the point estimates, the 90th percentile worst-matched workers face 7.16% lower wages when compared with perfectly matched workers. Second, column 3 shows that, conditional on mismatch, wages remain more cyclical for new hires from employment, when compared to existing workers; whereas for workers making transitions from non-employment to employment wages are no more cyclical than those for existing workers: the coefficient λ_5 in Equation (3.2) is not statistically significant. Note that by adding mismatch as a control, the magnitude of the coefficient λ_1 decreases while λ_4 and λ_5 slightly increase. This pattern captures the fact that for job stayers mismatch decreases with unemployment and that, in comparison to job stayers, mismatch of new hires is higher when unemployment increases, as shown in Baley et al. (2018).[6] In sum, I conclude that excess wage cyclicality of job switchers is not capturing the dynamics of mismatch. This findings are not surprising as Baley et al. (2018) provide evidence that skill mismatch is acyclical for workers making job-to-job transitions.

Mismatch amplification effect I now examine whether skill mismatch amplifies wage cyclicality. To do so, I estimate two different versions of Equation (3.2). One that also includes the interaction between the unemployment rate and mismatch $(U_t \cdot m_{i,c_t})$, and another that adds the following interaction terms:

[6] Recall that using the same sample, Baley et al. (2018) estimate a version of Equation (3.2) with mismatch as a dependent variable and find that λ_1 is negative while λ_4 and λ_5 are positive, albeit the former is not statistically significant.

$(U_t \cdot EE'_{i,t} \cdot m_{i,c_t})$, and $(U_t \cdot UE_{i,t} \cdot m_{i,c_t})$. Columns 4 and 5 provide the point estimates of such specifications. Figure 3.A2 plots wage semi-elasticity with respect to unemployment for different levels of mismatch and the three types of workers, and Figure 3.A3 reports both wage semi-elasticity estimates and the respective confidence intervals.

Two important results stand out. First, for workers in ongoing job relationships and those making job-to-job transitions, mismatch amplifies wage cyclicality. If the unemployment rates increases by one percentage point, the 95$^{\text{th}}$ worst-matched workers face a decrease in wages that is 2.1 times larger than the one faced by perfectly matched workers. Second, I can observe that, for new hires from unemployment, the extent to which their wages exhibit excessive cyclicality depends on the level of mismatch between worker's abilities and job skill requirements. For perfectly matched workers, wages of new hires from unemployment are no more cyclical than those of job-stayers; in contrast, for the 95$^{\text{th}}$ percentile of mismatch the wage-unemployment semi-elasticity is 19$\%$ percentage points lower for new hires from unemployment. Furthermore, across the different flow types, at the upper part of the mismatch distribution, the OLS estimates of the wage semi-elasticity in column 3 table lie outside the confidence intervals of the estimates reported in column 4, meaning that these coefficient estimates are significantly different from each other, and thus that mismatch has important amplification effects on wage cyclicality.

3.4.2 Robustness Checks

In this section, I show that the observed pattern is robust across different specifications. First, I use a different definition of what constitutes a new hire from non-employment that takes into account the length of the jobless spell. Second, I replicate Table 3.A2 taking into account occupation-industry fixed effects. Third, I restrict my focus to employed individuals older than 20 years old.

New hires from non-employment As in Gertler et al. (2016), the baseline coefficient estimates reported in Table 3.A2 rely on the broadest definition of new hires from non-employment. This means that regardless of how long the unemployment spell was, all workers who did not reported a job at time $t-1$ and are working at time t were considered to be new hires from unemployment. Nonetheless, new hires from unemployment with short non-employment spells may be instead job switchers taking a short break between jobs. To mitigate these concerns, I redefine transitions from non-employment. More specifically, I assume that workers with jobless spells equal or smaller than 1 month are workers making job-to-job transitions instead of transitions from non-employment to employment. Under these new definitions, $recalls$, i.e. those workers that return to their previous employer within 1 month, are recoded as job stayers. Table 3.A3 shows that the baseline results remain unchanged.

Occupation-Industry FE In the baseline estimation, I included dummies for one-digit level occupations and one-digit level industries so as to control for changes in the composition of occupations and industry over the cycle. I now explore the sensitivity of the results presented in the previously to the inclusion of occupation-industry fixed effects to account for the changes in the composition of occupation-industry pairs along the business cycle. This is important as one can find some of the occupations across different industries. We can observe in Table 3.A4 that the observed patterns are robust to this specification: the magnitude, sign and significance of coefficients remain unchanged.

Age at interview The point estimates reported in Table 3.A2 hinge on a subsample of individuals from the cross-sectional sample of the NLSY79 that over the sample period are between 14 and 55 years old at the time of the interview, as shown in Table 3.A1. Following Bils (1985) among others, I now restrict my sample to observations for individuals between older than 20 years old. Table 3.A5 shows that the sign of the coefficients of interest does not change. However, the magnitude of the wage semi-elasticity with respect to unemployment for job stayers decreases, and now lies between the range reported in the existing literature. In particular, if I do not distinguish between new hires and job stayers I recover a semi-elasticity of wages of 1.79% for all workers, which compares to 1.78% found by Kudlyak (2014), who also relies on the NLSY79. As before, we find that excess wage cyclicality for workers making job-to-job transitions does not capture changes in match quality, as measured by skill mismatch, and that mismatch amplifies wage cyclicality.

Overall, I find robust evidence that excess wage cyclicality of workers making job-to-job transitions goes beyond the dynamics of mismatch, while for new hires from unemployment whether their wages are more cyclical or not than those in ongoing job relationships depends how different are workers abilities from the job skill requirements. These results are in line with the findings of Carneiro et al. (2012). Using matching employer-employee data from Portugal, they control for job heterogeneity using job fixed effects and show that still wages procyclical and that wages of new hires are more cyclical than those of job stayers. This evidence points in favor of Pissarides (2009)'s argument that a good explanation for the unemployment volatility puzzle needs to be consistent with flexible wages.

3.5 Conclusion

In this chapter, I have revisited the issue of wage cyclicality. In line with earlier studies, I find that wages of newly hired workers are more cyclical than those of workers in ongoing relationships. Further, separating new hires who change jobs between employers and new hires coming from a jobless spell, we find that wage cyclicality of newly hired workers is driven by the former as in Gertler et al. (2016). While they interpret this evidence as capturing changes in match quality over the cycle, I take advantage of a unique feature of NLSY79 to disentangle movements in wages due to business cycle fluctuations from wages changes driven by match quality dynamics. This is the key novelty in my empirical exercise. To account for unobserved (permanent) heterogeneity at the job level, I use Guvenen et al. (2018)'s skill mismatch measure, which can be interpreted as the inverse of match quality. Controlling for mismatch, I show that wage dynamics of job switchers goes beyond skill mismatch fluctuations. This result is not surprising as Baley et al. (2018) have shown that skill mismatch is acyclical for workers making job-to-job transitions. Additionally, my results show that mismatch amplifies wage cyclicality.

3.6 Appendix

Table 3.A1: Descriptive Statistics

The table reports summary statistics for the main variables used in the empirical analysis. Panel A presents the statistics for individual characteristics: age at interview (years), job tenure in months, wages, and skill mismatch, defined as $m_{i,c_t} \equiv \sum_{j=1}^{J} \frac{1}{J} |a_{i,j} - r_{c_t,j}|$ with $a_{i,j}$ and r_{c_t} corresponding, respectively, to workers' abilities and skill requirements of the occupation held at time t along 4 different dimensions: $J = \{verbal, math, technical, social\}$. The sample constitutes a subsample of 2,991 individuals from the cross-sectional sample of the NLSY79 and runs over the period from 1979 and 2012. Panel B reports summary statistics of the business cycle indicator. *Unemployment Rate$_t$* is the monthly unemployment rate at the national level published by BLS. Source: NLSY79, O*NET, BLS and author's calculations.

	Observations	Mean	Std. Dev	Min.	Max.
Panel A: Individual Characteristics					
Age (years)	560187	31.15	9.21	14.00	55.00
Job Tenure (months)	560187	32.72	36.69	1.00	357.00
Mismatch$_t$	560187	27.38	14.30	1.25	91.25
Log real hourly earnings	560187	7.09	0.69	2.24	10.24
Panel B: Business Cycle Indicators					
Unemployment Rate$_t$	408	0.06	0.02	0.04	0.11

Table 3.A2: Cyclicality of Hourly Wages and Mismatch

The table reports coefficients from an OLS regression with robust standard errors clustered at the individual level reported in parentheses. $NH_{i,t}$ is a dummy variable that equals one if individual i is a new hire in period t, $EE'_{i,t}$ is a dummy for whether individual i is a new hire from employment, and $UE_{i,t}$ is a dummy for whether individual i is a new hire from unemployment. All columns include age, education level, job tenure (quadratic term), time trend, occupation, industry, region, month and individual fixed effects. The sample includes all worker-job matches between 1979 and 2012. ***, ** and * represent statistical significance at 1%, 5% and 10% levels, respectively.

Dependent Variable: Log real hourly earnings

	(1)	(2)	(3)	(4)	(5)
Unemployment_t	−0.0218***	−0.0218***	−0.0219***	−0.0122***	−0.0125***
	(0.002)	(0.002)	(0.002)	(0.004)	(0.004)
$\text{NH}_{i,t}, \times \text{Unemployment}_t$	−0.0096***				
	(0.002)				
$\text{EE'}_{i,t}, \times \text{Unemployment}_t$		−0.0161***	−0.0156***	−0.0151***	−0.0150***
		(0.003)	(0.003)	(0.003)	(0.004)
$\text{UE}_{i,t}, \times \text{Unemployment}_t$		−0.0021	−0.0017	−0.0010	0.0029
		(0.003)	(0.003)	(0.003)	(0.004)
$\text{Mismatch}_{i,t}$			−0.0015***	0.0009	0.0009
			(0.000)	(0.001)	(0.001)
$\text{Mismatch}_{i,t} \times \text{Unemployment}_t$				−0.0004***	−0.0003***
				(0.000)	(0.000)
$\text{EE'}_{i,t}, \times \text{Unemployment}_t \times \text{Mismatch}_{i,t}$					−0.0000
					(0.000)
$\text{UE}_{i,t}, \times \text{Unemployment}_t \times \text{Mismatch}_{i,t}$					−0.0001**
					(0.000)
Observations	500170	500170	497130	497130	497130
Adjusted R^2)	0.560	0.560	0.561	0.561	0.562

Table 3.A3: Cyclicality of Hourly Wages and Mismatch: Redefining new-hires from non-employment

The table reports coefficients from an OLS regression with robust standard errors clustered at the individual level reported in parentheses. $NH_{i,t}$ is a dummy variable that equals one if individual i is a new hire in period t, $EE'_{i,t}$ is a dummy for whether individual i is a new hire from employment, and $UE_{i,t}$ is a dummy for whether individual i is a new hire from unemployment. All columns include age, education level, job tenure (quadratic term), time trend, occupation, industry, region, month and individual fixed effects. The sample includes all worker-job matches between 1979 and 2012. ***, ** and * represent statistical significance at 1%, 5% and 10% levels, respectively.

Dependent Variable: Log real hourly earnings

	(1)	(2)	(3)	(4)	(5)
Unemployment$_t$	−0.0218***	−0.0218***	−0.0219***	−0.0123***	−0.0125***
	(0.002)	(0.002)	(0.002)	(0.004)	(0.004)
NH$_{i,t}$, × Unemployment$_t$	−0.0095***				
	(0.002)				
EE'$_{i,t}$, × Unemployment$_t$		−0.0142***	−0.0137***	−0.0131***	−0.0126***
		(0.003)	(0.003)	(0.003)	(0.003)
UE$_{i,t}$, × Unemployment$_t$		0.0003	0.0005	0.0011	0.0048
		(0.004)	(0.004)	(0.004)	(0.004)
Mismatch$_{i,t}$			−0.0015***	0.0009	0.0009
			(0.000)	(0.001)	(0.001)
Mismatch$_{i,t}$ × Unemployment$_t$				−0.0004***	−0.0003***
				(0.000)	(0.000)
EE'$_{i,t}$, × Unemployment$_t$ × Mismatch$_{i,t}$					−0.0000
					(0.000)
UE$_{i,t}$, × Unemployment$_t$ × Mismatch$_{i,t}$					−0.0001**
					(0.000)
Observations	500170	500170	497130	497130	497130
Adjusted R^2)	0.560	0.560	0.561	0.562	0.562

Table 3.A4: Cyclicality of Hourly Wages and Mismatch: Occupation-Industry FE

The table reports coefficients from an OLS regression with robust standard errors clustered at the individual level reported in parentheses. $NH_{i,t}$ is a dummy variable that equals one if individual i is a new hire in period t, $EE'_{i,t}$ is a dummy for whether individual i is a new hire from employment, and $UE_{i,t}$ is a dummy for whether individual i is a new hire from unemployment. All columns include age, education level, job tenure (quadratic term), time trend occupation-industry, region, month and individual fixed effects. The sample includes all worker-job matches between 1979 and 2012. ***, ** and * represent statistical significance at 1%, 5% and 10% levels, respectively.

Dependent Variable: Log real hourly earnings

	(1)	(2)	(3)	(4)	(5)
Unemployment_t	-0.0215^{***}	-0.0215^{***}	-0.0216^{***}	-0.0113^{***}	-0.0117^{***}
	(0.002)	(0.002)	(0.002)	(0.004)	(0.004)
$\text{NH}_{i,t}$, $\times$ Unemployment_t	-0.0094^{***}				
	(0.002)				
$\text{EE'}_{i,t}$, $\times$ Unemployment_t		-0.0164^{***}	-0.0159^{***}	-0.0153^{***}	-0.0148^{***}
		(0.003)	(0.003)	(0.003)	(0.004)
$\text{UE}_{i,t}$, $\times$ Unemployment_t		-0.0017	-0.0013	-0.0006	0.0042
		(0.003)	(0.003)	(0.003)	(0.004)
$\text{Mismatch}_{i,t}$			-0.0015^{***}	0.0010	0.0010
			(0.000)	(0.001)	(0.001)
$\text{Mismatch}_{i,t}$ $\times$ Unemployment_t				-0.0004^{***}	-0.0004^{***}
				(0.000)	(0.000)
$\text{EE'}_{i,t}$, $\times$ Unemployment_t $\times$ $\text{Mismatch}_{i,t}$					-0.0000
					(0.000)
$\text{UE}_{i,t}$, $\times$ Unemployment_t $\times$ $\text{Mismatch}_{i,t}$					-0.0001^{***}
					(0.000)
Observations	500170	500170	497130	497130	497130
Adjusted R^2)	0.568	0.568	0.569	0.569	0.569

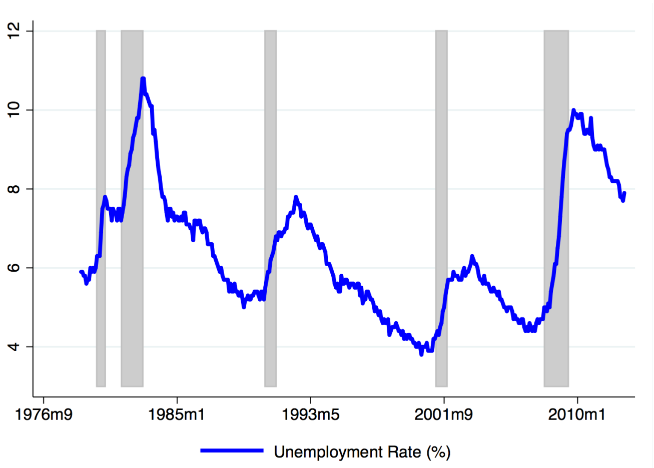
Table 3.A5: Cyclicality of Hourly Wages and Mismatch: Older than 20

The table reports coefficients from an OLS regression with robust standard errors clustered at the individual level reported in parentheses. $NH_{i,t}$ is a dummy variable that equals one if individual i is a new hire in period t, $EE'_{i,t}$ is a dummy for whether individual i is a new hire from employment, and $UE_{i,t}$ is a dummy for whether individual i is a new hire from unemployment. All columns include age, education level, job tenure (quadratic term), time trend occupation-industry, region, month and individual fixed effects. The sample includes all worker-job matches between 1979 and 2012. ***, ** and * represent statistical significance at 1%, 5% and 10% levels, respectively.

Dependent Variable: Log real hourly earnings

	(1)	(2)	(3)	(4)	(5)
Unemployment_t	-0.0175^{***}	-0.0175^{***}	-0.0177^{***}	-0.0078^{**}	-0.0079^{**}
	(0.002)	(0.002)	(0.002)	(0.004)	(0.004)
$\text{NH}_{i,t}$, × Unemployment_t	-0.0100^{***}				
	(0.003)				
$\text{EE'}_{i,t}$, × Unemployment_t		-0.0174^{***}	-0.0167^{***}	-0.0163^{***}	-0.0157^{***}
		(0.004)	(0.004)	(0.004)	(0.004)
$\text{UE}_{i,t}$, × Unemployment_t		−0.0019	−0.0014	−0.0009	0.0012
		(0.004)	(0.004)	(0.004)	(0.005)
$\text{Mismatch}_{i,t}$			-0.0009^{***}	0.0015	0.0015
			(0.000)	(0.001)	(0.001)
$\text{Mismatch}_{i,t}$ × Unemployment_t				-0.0004^{***}	-0.0004^{***}
				(0.000)	(0.000)
$\text{EE'}_{i,t}$, × Unemployment_t × $\text{Mismatch}_{i,t}$					−0.0000
					(0.000)
$\text{UE}_{i,t}$, × Unemployment_t × $\text{Mismatch}_{i,t}$					−0.0001
					(0.000)
Observations	452380	452380	449643	449643	449643
Adjusted R^2)	0.555	0.555	0.556	0.556	0.556

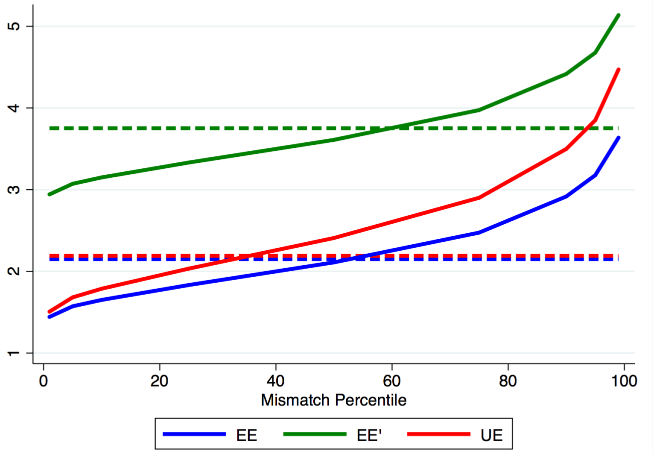
Figure 3.A1: Unemployment Rate



Notes: This graph shows seasonally adjusted civilian unemployment from the BLS. Shaded areas correspond to NBER recessions. Source: BLS and NBER.

Figure 3.A2: Wage Cyclicality

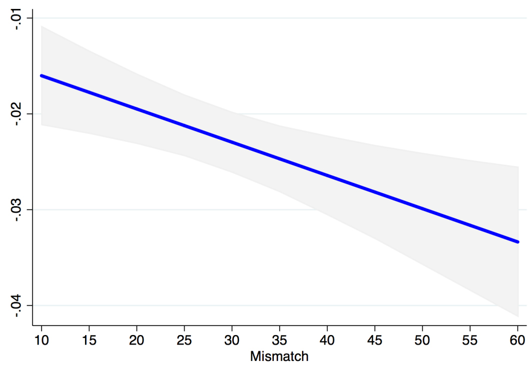
(Coefficient on $U_t \times -100$)



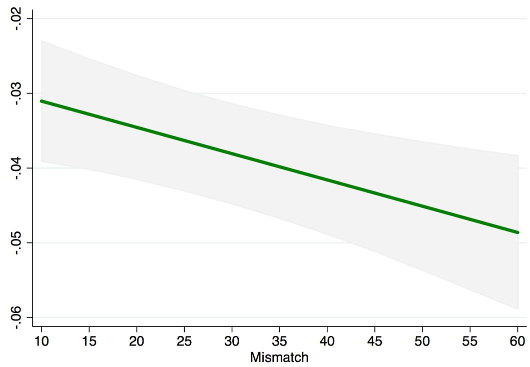
Notes: This graph plots the % change in wages when the unemployment rate increases by one percentage points for different percentiles of mismatch and three types of workers: workers in ongoing job relationships (blue), job-to-job transitions (green) and new hires from unemployment (red). The dashed and solid and lines correspond to the wage-unemployment semi-elasticity calculated using the estimated coefficients reported in column 3 and 5 of Table 3.A2. Figure 3.A3 shows that these estimates are statistically significant.

Figure 3.A3: Wage Cyclicality and Mismatch

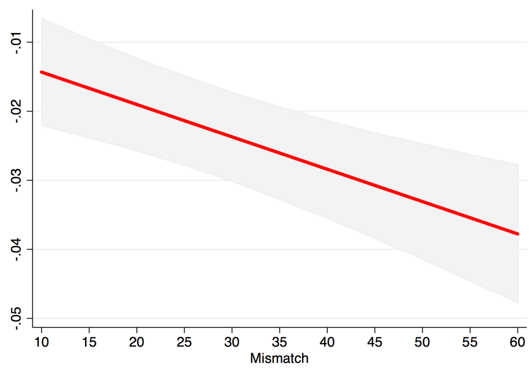
A. Job Stayers



B. Job Changers



C. New Hires from Unemployment



Notes: Each panel plots estimates of wage semi-elasticity with respect to unemployment based on coefficient estimates reported in column 5 of Table 3.A2.

Bibliography

Acemoglu, A. and D. Autor, "Skills, Tasks and Technologies: Implications for Employment and Earnings," *Handbook of Labor Economics*, 2011, *4* (chapter 12), 1043–1171.

Ahlfeldt, G., S. J. Redding, D. Sturm, and N. Wolf, "The Economics of Density: Evidence from the Berlin Wall," *Econometrica*, 2015, *83* (6), 2127–2189.

Akerlof, G., "Social Distance and Social Decisions," *Econometrica*, 1997, *65* (5), 1005–1027.

_ and R. Kranton, "Economics and Identity," *Quarterly Journal of Economics*, 2000, *115* (3), 715–753.

_ and _ , "Identity and Schooling: Some Lessons for the Economics of Education," *Journal of Economic Literature*, 2002, *40* (2), 1167–1201.

Attanasio, O. and K. Kaufmann, "Education Choices and Returns to Schooling: Intrahousehold Decision Making, Gender and Subjective Expectations," *Journal of Development Economics*, 2014, *109*, 203–216.

Autor, D. and D. Dorn, "The Growth of Low Skill Service Jobs and the Polarization of the U.S. Labor Market.," *American Economic Review*, 2013, *103* (5), 1553–1597.

Babcock, B., E. Kwan Choi, and E. Feinerman, "Risk and Probability Premiums for CARA Utility Functions," *Journal of Agricultural and Resource Economics*, 1993.

Baley, I., A. Figueiredo, and R. Ulbritcht, “Mismatch Cycles,” *Working Paper*, 2018.

Baley, Isaac and Julio A Blanco, “Firm Uncertainty Cycles and the Propagation of Nominal Shocks,” *AEJ: Macro*, 2018, (Forthcoming).

Barlevy, G., “Why Are the Wages of Job Changers So Procyclical?,” *Journal of Labor Economics*, 2001, *19* (4), 837–78.

_ , “The Sullying Effect of Recessions,” *Review of Economic Studies*, 2002, *69* (1), 65–96.

Barnichon, R., “Building a composite Help-Wanted Index,” *Economic Letters*, 2010, *109* (3), 175–178.

Baydur, I. and T. Mukoyama, “Employment Duration and Match Quality over the Business Cycle,” *Working paper*, 2015.

Bayer, P., F. Ferreira, and R. McMillan, “A Unified Framework for Measuring Preferences for Schools and Neighborhoods,” *Journal of Political Economy*, 2007, *115* (4), 588–638.

_ , R. McMillan, and K. Rueben, “An Equilibrium Model of Sorting in an Urban Housing Market,” *NBER working paper 10865*, 2004.

Belfield, C., T. Boneva, C. Rauh, and J. Shaw, “Money or Fun? Why Students Want to Pursue Further Education,” *Working paper*, 2016.

Bénabou, R., “Workings of a City: Location, Education and Productions,” *Quarterly Journal of Economics*, 1993, *108* (3), 619–652.

_ , “Equity and Effiency in Human Capital Investment: The Local Connection,” *Review of Economic Studies*, 1996, *63* (2), 237–264.

_ , “Heterogeneity, Stratification and Growth: Macroeconomic Implications of Community Structure and School Finance,” *American Economic Review*, 1996, *86* (3), 584–609.

Bils, Mark J., "Real Wages over the Business Cycle: Evidence from Panel Data," *Journal of Political Economy*, 1985, *93* (4), 666–89.

Blanchard, O. and J. Galí, "Labor Markets and Monetary Policy: A New Keynesian Model with Unemployment," *American Economic Journal: Macroeconomics*, 2010, *2* (2), 1–30.

Bleemer, Z. and B. Zafar, "Intended College Attendance: Evidence from an Experiment on College Returns and Costs," *Working paper*, 2016.

Boehm, M., "The Price of Polarization: Estimating Task Prices under Routine-Biased Technical Change," *Working paper*, 2015.

Borovickova, K., "Job flows, worker flows and labor market policies," 2016.

Bowles, S., G. Loury, and R. Sethi, "Group Inequality," *Journal of the European Economics Association*, 2014, *12* (1), 129–152.

Bowlus, A., "Matching Workers and Jobs: Cyclical Fluctuations in Match Quality," *Journal of Labor Economics*, 1995, *13* (2), 257–273.

Busso, M., J. Gregory, and P. Kline, "Assessing the Incidence and Efficiency of a Prominent Place Based Policy," *American Economic Review*, 2013, *103* (2), 897–947.

Card, D., A. Mas, and J. Rothstein, "Tipping and the Dynamics of Segregation," *The Quarterly Journal of Economics*, 2008, *123* (1), 177–218.

Carneiro, A., P. Guimarães, and P. Portugal., "Real Wages and the Business Cycle: Accounting for Worker, Firm, and Job Title Heterogeneity," *American Economic Journal: Macroeconomics*, 2012, *4* (2), 133–52.

Carneiro, P. and J. Heckman, "The Evidence on Credit Constraints in Post-Secondary Schooling," *The Economic Journal*, 2002, *112* (482), 705–734.

Carrillo-Tudela, C. and L. Visschers, "Unemployment and Endogenous Reallocation over the Business Cycle," *CESifo working paper 4079*, 2014.

Cavalcanti, T. and G. Giannitsarou, "Growth and Human Capital: A Network Approach," *Working paper*, 2013.

Chetty, R. and N. Hendren, "The Impacts of Neighborhoods on Intergenerational Mobility I: Childhood Exposure Effects," *NBER working paper 23001*, 2017.

_ , _ , and L. Katz, "The Effects of Exposure to Better Neighborhoods on Children: New Evidence from the Moving to Opportunity Experiment," *American Economic Review*, 2016, *106* (4), 855–902.

Crane, L., H. Hyatt, and S. Murray, "Cyclical Labor Market Sorting," *Working paper*, 2018.

Delevande, A. and B. Zafar, "University choice: the role of expected earnings, non-pecuniary outcomes, and financial constraints," *FRB of New York Staff Report, 683*, 2014.

Diamond, R., "The Determinants and Welfare Implications of US Workers' Diverging Location Choices by Skill: 1980-2000," *American Economic Review*, 2016, *106* (3), 479–524.

Dixit, A., "Analytical Approximations in Models of Hysteresis," *The Review of Economic Studies*, 1991, *58* (1), 141–151.

Dorn, D., "Essays on Inequality, Spatial Interaction, and the Demand for Skills," *Dissertation University of St. Gallen no. 3613*, 2009.

Duflo, E. and E. Saez, "The Role of Information and Social Interactions in Retirement Plan Decisions: Evidence from a Randomized Experiment," *Quarterly Journal of Economics*, 2003, *118*, 815–842.

Durlauf, S., "A Theory of Persistent Income Inequality," *Journal of Economic Growth*, 1996, *1* (1), 75–93.

_ , "Neighborhood Effects," *Handbook of Regional and Urban Economics*, 2004, pp. 2174–2230.

Eeckhout, J. and P. Kircher, "Identifying Sorting - In Theory," *The Review of Economic Studies*, 2011, *78* (3), 872.

— **, R. Pinheiro, and K. Schmidheiny**, "Spatial Sorting," *Journal of Political Economy*, 2014, *122* (3), 554–620.

Fajgelbaum, P., E. Schaal, and M. Taschereau-Dumouchel, "Uncertainty traps," *Quarterly Journal of Economics*, 2016, *forthcoming.*

Férnandez, R., "Cultural Change as Learning: The Evolution of Female Labor Force Participation over a Century," *American Economic Review*, 2013, *103* (1), 472–500.

— **and R. Rogerson**, "Income Distribution, Communities and the Quality of Public Education," *Quarterly Journal of Economics*, 1996, *111* (1), 135–164.

Ferreyra, M., "An Empirical Framework for Large-Scale Policy Analysis, with an Application to School Finance Reform in Michigan," *American Economic Journal: Economic Policy*, 2009, *1* (1), 147–180.

Fogli, A. and L. Veldkamp, "Nature or Nurture? Learning and the Geography of Female Labor Force Participation," *Econometrica*, 2011, *79* (4), 1103–1138.

Gertler, M. and A. Trigari, "Unemployment Fluctuations with StaggeredNash Wage Bargaining," *Journal of Political Economy*, 2009, *117* (1), 38–86.

— **, G. Huckfeldt, and A. Trigari**, "Unemployment Fluctuations, Match Quality and the Wage Cyclicality of New Hires," *NBER Working Paper 22341*, 2016.

Gibbons, S., O. Silva, and F. Weinhardt, "Everybody needs good neighbours? Evidence from students' outcomes in England," *The Economic Journal*, 2013, *123* (September), 831–874.

Glaeser, Edward, Giacomo Ponzetto, and Yimei Zou, "Urban networks: Connecting markets, people, and ideas," *Papers in Regional Science*, 2016, *95* (1), 17–59.

Goldin, C. and L. Katz, "The Race Between Education and Technology.," *Press for Harvard University Press.*, 2008.

Guerrieri, V., D. Hartley, and E. Hurst, "Endogenous gentrification and housing price dynamics," *Journal of Public Economics*, 2013, *100*, 45–60.

Guvenen, F., B. Kuruscu, S. Tanaka, and D. Wiczer, "Multidimensional Skill Mismatch," *Working Paper 2018-02 - Stony Brook University*, 2018.

Haefke, C., M. Sonntag, and T. van Rens, "Wage rigidity and job creation," *Journal of Monetary Economics*, 2013, *60* (8), 887–899.

Hagedorn, M., T. Hann Law, and I. Manovskii, "Identifying Equilibrium Models of Labor Market Sorting," *Econometrica*, 2017, *85* (1), 29–65.

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

Hastings, J., C. A. Neilson, A. Ramirez, and S. D. Zimmerman, "(Un)informed college and major choice: Evidence from linked survey and administrative data," *Economics of Education Review*, 2016, pp. 1–16.

_ , _ , and S. D. Zimmerman, "The Effects of Earnings Disclosure on College Enrollment Decisions," *NBER working paper 21300*, 2017.

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

Hoxby, C. and C. Avery, "The Missing "One-Offs": The Hidden Supply of High-Achieving, Low-Income Students," *Brookings Papers on Economic Activity*, 2014.

_ and S. Turner, "Expanding College Opportunities for High-Achieving, Low Income Students," *Working paper*, 2014.

_ and _ , "What High-Achieving Low-Income Students Know About College," *American Economic Review*, 2015, *105* (5), 514–517.

Hsieh, C. T., E. Hurst, C. I. Jones, and P. J. Klenow, "The allocation of talent and U.S. economic growth," *Working paper*, 2018.

Huckfeldt, C., "Understanding the scarring effect of recessions.," *Working paper*, 2016.

Jensen, R., "The (perceived) returns to education and the demand for schooling," *The Quarterly Journal of Economics*, 2010, *125* (2), 515–548.

Jovanovic, B., "Job Matching and the Theory of Turnover," *Journal of Political Economy*, 1979, *87* (5), 972–990.

_ , "Misallocation and Growth," *American Economic Review*, April 2014, *104* (4), 1149–71.

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

Kambourov, G. and I. Manovskii, "Rising Occupational and Industry Mobility in the United States: 1968–97," *International Economic Review*, 2008, *49* (1), 41–79.

Katz, L. and K. Murphy, "Changes in Relative Wages, 1963-1987: Supply and Demand Factors.The," *The Quarterly Journal of Economics*, 1992, *107* (1), 35–78.

Kaufmann, K., "Understanding the Income Gradient in College Attendance in Mexico: The Role of Heterogeneity in Expected Returns," *Quantitative Economics*, 2014, *forthcoming.*

Kennan, J. and J. R. Walker, "The Effect of Expected Income on Individual Migration Decisions," *Econometrica*, 2011, *79* (1), 211–251.

Kim, Y. and G. Loury, "Social Externalities, Overlap and the Poverty Trap," *Journal of Economic Inequality*, 2013, *12* (4), 535–554.

Kling, J., J. Liebman, and L. Katz, "Experimental Analysis of Neighborhood Effect," *Econometrica*, 2007, *75* (1), 83–119.

Kudlyak, M., "The cyclicality of the user cost of labor," *Journal of Monetary Economics*, 2014, *68* (C), 53–67.

Leduc, S. and Z. Liu, "Uncertainty shocks are aggregate demand shocks," *Journal of Monetary Economics*, 2016, *82*, 20–35.

Lee, S., Y. Shin, and D. Lee, "The Option Value of Human Capital: Higher Education and Wage Inequality," *Working paper*, 2017.

Lin, T.-T. T., "The Role of Uncertainty in Jobless Recoveries," 2014.

Lindenlaub, I., "Sorting Multidimensional Types: Theory and Application," *The Review of Economic Studies*, 2017, *84* (2), 718–789.

_ **and F. Postel-Vinay**, "Multidimensional Sorting under Random Search," *Working paper*, 2016.

Lindqvist, E. and R. Vestman, "The Labor Market Returns to Cognitive and Noncognitive Ability: Evidence from the Swedish Enlistment," *American Economic Journal: Applied Economics*, 2011, *3* (1), 101–28.

Lise, J. and F. Postel-Vinay, "Multidimensional Skills, Sorting, and Human Capital Accumulation," *Working paper*, 2016.

_ **and J. Robin**, "The Macrodynamics of Sorting between Workers and Firms," *American Economic Review*, 2017, *107* (4), 1104–1135.

Liua, K., K. G. Salvanesb, and E. Sørensen, "Good skills in bad times: Cyclical skill mismatch and the long-term effects of graduating in a recession," *European Economic Review*, 2016, *84*, 3–17.

McFadden, D., "Conditional Logit Analyisis of Qualitative Choice Behavior," *Frontiers in Econometrics*, 1973, *ed. P. Zarembka*, 105–142.

Monte, F., S. J. Redding, and E. Rossi-Hansberg, "Commuting, Migration and Local Employment Elasticities," *NBER Working Paper No. 21706*, 2015.

Mookherjee, D., S. Napel, and D. Ray, "Aspirations, segregation and occupational choice," *Journal of European Economic Association*, 2010, *8* (1), 139–168.

Mortensen, D. and C. Pissarides, "Job Creation and Job Destruction in the theory of unemployment," *Review of Economic Studies*, 1994, *61* (3), 397–415.

Moscarini, G., "Excess Worker Reallocation," *Review of Economic Studies*, 2001, *68*, 593–612.

_ , "Job Matching and the Wage Distribution," *Econometrica*, 2005, *73* (2), 481–516.

_ **and K. Thomsson**, "Occupational and job mobility in the US," *Scandinavian Jornal of Economics*, 2007, *109* (4), 807–836.

Munshi, K., "Social learning in a heterogeneous population: technology diffusion in the Indian Green Revolution," *Journal of Development Economics*, 2004, *73*, 185–213.

_ **and J. Mayaux**, "Social Norms and the Fertility Transition," *Journal of Development Economics*, 2006, *80*, 1–38.

Murphy, K., A. Shleifer, and R. Vishny, "The Allocation of Talent: Implications for Growth*," *The Quarterly Journal of Economics*, 1991, *106* (2), 503–530.

Oreopolous, P., "The long-run consequences of living in a poor neighborhood," *The Quarterly Journal of Economicse*, 2003, *118* (4), 1533–1575.

Oreopoulos, P., T. von Wachter, and A. Heisz, "The short-and long-term career effects of graduating in a recession," *American Economic Journal: Applied Economics*, 2012, *4* (1), 1–29.

Oster, E., "Unobservable Selection and Coefficient Stability:," *Journal of Business Economics and Statistics*, 2016, *Forthcoming.*

Ouyang, M., "The Scarring Effects of Recessions," *Journal of Monetary Economics*, March 2009, *56* (2), 184–199.

Pissarides, C., "The Unemployment Volatility Puzzle: Is Wage Stickinessthe Answer?," *Econometrica*, 2009, *77* (5), 1339–1369.

Postel-Vinay, F. and J. Robin, "Equilibrium Wage Dispersion with Worker and Employer Heterogeneity," *Econometrica*, 2002, *70* (7), 2295–2350.

Prada, M. and S. Urzúa, "One Size does not Fit All: Multiple Dimensions of Ability, College Attendance and Wages," *Journal of Labor Economics*, 2016, *Accepted for publication.*

Pries, M., "Uncertainty-driven labor market fluctuations," *Journal of Economic Dynamics and Control*, 2016, *73*, 181–199.

Pries, Michael J, "Persistence of employment fluctuations: A model of recurring job loss," *The Review of Economic Studies*, 2004, *71* (1), 193–215.

Rauh, C. and T. Boneva, "Socio-Economic Gaps in University Enrollment: The Role of Perceived Pecuniary and Non-Pecuniary Returns," 2017.

Sanbonmatsu, L., J. Kling, G. Duncan, and J. Brooks-Gunn, "Neighborhoods and academic achievement: results from the moving to opportunity experiment," *Journal of Human Resources*, 2008, *41* (4), 649–691.

Schaal, E., "Uncertainty and Unemployment," *Econometrica*, 2017, *Forthcoming.*

Schumpeter, J., *Business Cycles*, McGraw-Hill, 1939.

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

Shin, D. and G. Solon, "New Evidence on Real Wage Cyclicality within Employer-Employee Matches," *Scottish Journal of Political Economy*, 2007, *54*, 648–60.

Solon, D., "Cyclicality of Real Wages among Young Men," *Economics Letters*, 1994, *46*, 137–42.

Solon, G., R. Barsky, and J. Parke, "Measuring the Cyclicality of Real Wages: How Important is Composition Bias," *The Quarterly Journal of Economics*, 1994, *109* (1), 1–25.

Speer, J., "Pre-Market Skills, Occupational Choice, and Career Progression," *Journal of Human Resources*, 2017, *52* (1), 187–246.

Stinebrickner, R. and T. R Stinebrickner, "A Major in Science? Initial Beliefs and Final Outcomes for College Major and Dropout," *The Review of Economic Studies*, 2014, *81* (1), 426–472.

Topa, G. and Y. Zenou, "Neighborhood versus network effects," *Handbook of Regional and Urban Economics*, 2015, *5A*, 561–624.

Wiswall, M. and B. Zafar, "Determinants of college major choice: Identification using an information experiment," *The Review of Economic Studies*, 2015, *82* (2), 891–824.