Information Frictions in Education and Inequality *

Ana Figueiredo †

November 16, 2018

Abstract

Why does the place where children grow up shape their opportunities in life? I 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 key elements are uncertainty about the skill premium and learning through signals of wages earned by nearby college graduates. In this environment, more exposure to highly educated neighbors brings more information about the skill premium. However, this only translates into more education 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 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.

JEL Classification: D80, E24, I24, J62

Keywords: Intergenerational mobility, education, information frictions, skill premium, uncertainty.

---

*I thank my supervisor Jan Eeckhout for his guidance throughout this project. I am also grateful to Joseph Altonji, Isaac Baley, Cristina Bélles, Andrea Caggese, Davide Debortoli, Jordi Galí, María Lombardi, Christopher Rauh, Cezar Santos, Edouard Schaal, and Laura Veldkamp for their helpful feedback as well as participants at CREi-UPF Macro Breakfast, BGSE PhD Jamboree 2016, XXI Workshop on Dynamic Macroeconomics, PET 16 Rio and Lubramacro 2017 for their suggestions. I acknowledge financial support from Fundação para a Ciência e Tecnologia, PhD Grant SFRH/BD/94516/2013.

†Erasmus School of Economics, figueiredo@ese.eur.nl

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

1 Introduction

College enrollment in the US 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 paper, 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)* there is large uncertainty about the skill premium in the US (Bleemer and Zafar, 2016)[1]; *(ii)* individual perceptions about earnings affect 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 paper 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 effects of better schools,

[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.

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$^{\text{th}}$ 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.

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 in this paper is that 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. 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 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 increases 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.

Related Literature This paper contributes 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 paper 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 paper 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 paper 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 paper 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 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 paper 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 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 paper 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 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.

Outline The paper proceeds as follows. Section 2 presents novel evidence regarding the relationship between neighborhood's characteristics and educational outcomes, and Section 3 explains my empirical findings through a model with local information frictions. In Section 4, I assess the quantitative importance of the proposed mechanism. Section 5 concludes.

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.

[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.

2.1 Data and Descritptive Statistics

For the empirical analysis, I construct a school-district panel with annual frequency that runs from 2008 to 2014. I combine 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 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 attend 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 A.1. 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 A.2

[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/`.

[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.

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 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.

2.2 Empirical Methodology

To formally examine the relationship between the share of college graduates living in a school-district and college enrollment, I estimate the following equation:

$$Enrollment_{ijt} = \beta_0 + \beta_1 College_{ijt} + \beta_2 College_{ijt} \times Y_{ijt} + \beta_3 Y_{ijt} + \delta X_{ijt} + \gamma_j + \gamma_t + \rho_j t + \varepsilon_{ijt} \quad (1)$$

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,

$$\frac{\partial Enrollment_{ijt}}{\partial College_{ijt}} = \beta_1 + \beta_2 \times Y_{ijt} \quad (2)$$

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. Figure A.1 illustrates the effect of college graduates on college enrollment under different signs of the coefficients of interest.

Identification To identify β_1 and β_2, I exploit variation in the share of residents with a college degree and their median earnings of across school-districts within a city over time. To illustrate that there are indeed important differences across school-districts within a city in the magnitude of the main variables of interest: 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 (2) 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 (2) 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[8]; and *(iii)* for unobserved city level characteristics that might be trending over time such as gentrification dynamics or deterioration in housing quality. Given this, the key assumption for causal

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

[8] As a robustness check, I also run a specification that controls for time shocks affecting all school-districts within a city.

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 2.5.

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 (2) that does not include the interaction term, $College_{ijt} \times Y_{ijt}$. Panel A in Table 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, and to address possible concerns over heterogeneous trends, column 6 includes city-specific linear trends. I find that the sign, the magnitude and significance of coefficients remains nearly unchanged.[9] 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).[10]

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 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 heterogeneities along some dimensions such as earnings. I investigate the presence of heterogeneities replicating all specifications in panel A, Table 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 presents OLS estimates of

[9] 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.

[10] To check the coefficients on school quality measures and average ACT score, see Table A.3.

Table 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, College Degree	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.

Equation (2).

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 A.2 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 3), 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.

2.4 Robustness Checks

Next, I provide evidence that my results are robust to *(i)* alternative specifications, *(ii)* different samples, and *(iii)* alternative measures of earnings of college graduates

Quadratic Specification Equation (2) assumes that the correlation betwenn 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 (2) where I consider the effect of college graduates on enrollment to be quadratic in earnings: $\frac{\partial Enrollment_{ijt}}{\partial College_{ijt}} = \beta_1 + \beta_2 \times Y_{ijt}^2$. Column 1 in Table A.5 reports the estimated coefficients, and panel (b) in Figure A.2 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 A.5 and panel (c) in Figure A.2 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 (2) 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 (2) including instead a vector of 2009 school-district level controls, X_{ij2009}. Column 3 in Table A.5 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, 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 12$^{\text{th}}$ students were non-resident students. Nevertheless, to check the robustness of the results to

[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 A.5, column 4.

[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).

the inclusion/exclusion of students that do not live in the school-district they attend, I re-estimate Equation (2) focusing only on school-districts which have a low share of non-resident students attending the 12$^{\text{th}}$ grade (I fix the share threshold at 10%). Column 4 in Table A.5 shows that my findings hold when we exclude from the analysis school-districts with a higher share of non-resident students attending 12$^{\text{th}}$ 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. I re-estimate Equation (2) 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 A.5): the sign of the estimated coefficients is the same as the one in column 6 in Table 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 A.5 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 using median annual earnings of individuals with a post-graduate degree 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 A.5 show that using these two alternatives proxies leaves the magnitude, sign and significance of the coefficients of interest relatively unchanged.

2.5 Discussion

The previous section 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 graudating from in that district. 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. As previously mentioned, this result is net of the effect of school resources and local amenities. In particular, I disentangle the correlation caused by neighbors' characteristics from neighborhood amenities by taking into account (i) cohort and socioeconomic characteristics of the school-district, and (ii) school quality in the OLS estimation. I caution that although specification (2) may provide an improvement to the baseline OLS specification without controls, concerns about omitted variable bias remain. Therefore, following the approach of Oster (2016), who developed a method to examine how robust OLS estimates are to omitted variable bias by studying coefficient movements and movements in R^2 values when including additional controls, I assess the possible degree of omitted variable bias.

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

Omitted Variable Bias Under the assumption that selection on the observables is proportional to selection on the unobservables by a factor δ[14], Oster (2016)'s bias-adjusted coefficient is

$$\beta_i^* = \hat{\beta}_i - \delta(\tilde{\beta}_i - \hat{\beta}_i)\frac{R^{max} - \hat{R}}{\hat{R} - \tilde{R}}, i = 1, 2 \tag{3}$$

where $\hat{\beta}_i$ and $\hat{R}$ are the estimated coefficients and R^2 of column 6 in Panel B of Table 1 and $\tilde{\beta}_i$ and $\tilde{R}$ come from OLS estimation of Equation (2) 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 $[\hat{\beta}_i, \beta_i^*]$.[15] 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 A.3 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.

Mechanisms Why do individuals with a college degree 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, 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 included controls 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 (2) 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

[14] 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.

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

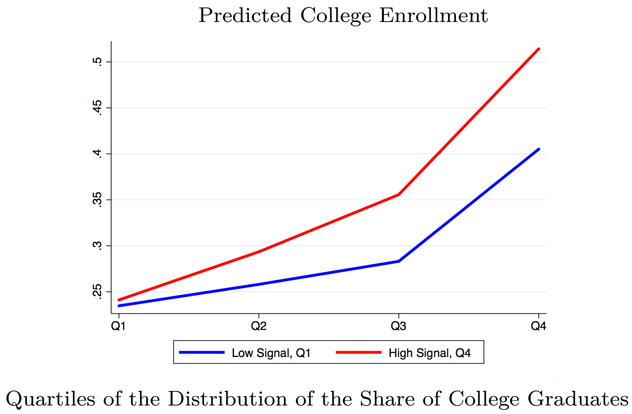
12^{th} grade as a I control, which I previously mentioned, is a good proxy for the ability of the students.

School-resources One may also think that the empirical findings are driven by differences in the resources of public schools across school districts. 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 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: 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).

3 Education Choice with Information Frictions and Local Learning

Section 2 documents robust evidence showing 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 illustrates the role of 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 findings in Section 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 3.6 discusses their implications, and shows that they do not affect the model's key prediction.

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), $k \in \{H, L\}$.[16] Each parent provides, inelastically, one unit of labor in 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, $A_j = \{A_{jH}, A_{jL}\}$[17], and school quality q_j. Both A_j and q_j attributes of each location are taken by individuals as exogenously given.[18] 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 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.[19] The utility for an individual i of type $k \in \{H, L\}$ living in neighborhood j is given by

$$U(c_{i,j}^k, \Phi_{i,j}^k) = \frac{-\exp(-\gamma(c_{i,j}^k))}{\Phi_{i,j}^k} \tag{4}$$

[16] 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 Murphy (1992) and Goldin and Katz (2008).

[17] 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.

[18] Even though this may strike as a strong assumption, in Section 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.

[19] In section 3.6, I show that the model's main prediction is qualitatively robust to this assumption.

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 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.

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 distribution. 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.[20] Parents i with skill level $k \in \{H, L\}$ solves the following program:

$$\operatorname*{Max}_{j} \qquad U(c^k_{i,j}, \Phi^k_{i,j}) = \frac{-\exp(-\gamma \cdot c^k_{i,j})}{\Phi^k_{i,j}} \qquad \text{subject to} \qquad c^k_{i,j} + r_j = w^k_i \tag{5}$$

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

$$\Phi^k_{i,j} = q_j \cdot A_{j,k} \cdot \varepsilon_{i,j} \tag{6}$$

Individual's i idiosyncratic taste for neighborhood j is denoted by $\varepsilon_{i,j}$. I model this heterogeneity following McFadden (1973).[21] For each parent i, I consider that the idiosyncratic taste for neighbor-

[20] 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.

[21] 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.

hood j is drawn from a Fréchet distribution (also called the Type II extreme value distribution):

$$Pr(\varepsilon_{i,j} \leq x) = e^{-x^{-\theta}}, \text{for x} > 0, \text{iid}, \theta > 0 \tag{7}$$

where the parameter θ reflects the amount of variation in the distribution and is treated as common across all parents.[22] In the location choice context, θ governs preference heterogeneity for locations across parents.[23] 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

$$U(w_i^k, r_j, q_j, A_{j,k}, \varepsilon_{i,j}) = \frac{-\exp(-\gamma(w_i^k - r_j))}{q_j \cdot A_{j,k}} \varepsilon_{i,j} \tag{8}$$

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 $\rho_{i,j}^k$:

$$\rho_{i,j}^k = \frac{(q_j A_{j,k})^\theta (\exp(-\gamma(w_i^k - r_j)))^{-\theta}}{\sum_{j' \in \mathcal{J}} (q_j A_{j',k})^\theta (\exp(-\gamma(w_i^k - r_{j'})))^{-\theta}} \tag{9}$$

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

$$M_{k,j} = \sum_{i=1}^{M_k} \rho_{i,j}^k = \rho_j^k M_k \tag{10}$$

where M^k is the exogenous measure of k-type parents living in the city.[24] 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 equal the supply in that location and so:

$$H_j = \rho_j^H M_H + \rho_j^L M_L, \forall j \in \mathcal{J} \tag{11}$$

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 B.1, when amenities are equal across neighborhoods the spatial equilibrium is non-sorted. In this environment, amenities do not react to the characteristics of the

[22] 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.

[23] 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.

[24] See Appendix B for details.

population that chooses to live on it. In Section 3.6, I discuss the implications of relaxing this assumption.

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

$$w_{s,j}^H = w^H + \epsilon_{s,j}^H, \tag{12}$$

where $\epsilon_{s,j}^H$ 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,

$$\epsilon_{s,j}^H \sim \mathcal{N}(0, M_j \sigma_{\epsilon^H}^2), \tag{13}$$

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

$$w_j^H \equiv \frac{1}{M_{jH}} \sum_{i=1}^{M_{jH}} w_{s,j}^H = w^H + \epsilon_j^H, \tag{14}$$

with

$$\epsilon_j^H \equiv \frac{1}{M_{jH}} \sum_{s=1}^{M_{jH}} \epsilon_{s,j}^H \sim \mathcal{N}(0, m_{jH}^{-1} \cdot \sigma_{\epsilon^H}^2), \tag{15}$$

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_j^H. Important for the model's key prediction, the signal's precision, $m_{jH} \cdot \sigma_\epsilon^{-2}$, increases with the share of high-skill parents in the neighborhood.

Learning Initial beliefs are assumed to be identical across all children, $\tilde{w}_i^H \sim \mathcal{N}(\tilde{\mu}, \tilde{\sigma}^2)$.[25] To update these beliefs, they use information gathered by the public signal w_j^H. Children are *passive learners* and cannot take any action to change the quality of this signal: after receiving information from high-skill parents, each 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

$$\hat{\mu}_j = \underbrace{\frac{\sigma_j^2}{\tilde{\sigma}^2 + \sigma_j^2}}_{\text{weight on prior}} \tilde{\mu} + \underbrace{\frac{\tilde{\sigma}^2}{\tilde{\sigma}^2 + \sigma_j^2}}_{\text{weight on signal}} w_j^H, \tag{16}$$

$$\hat{\sigma}_j^2 = \left(\tilde{\sigma}^{-2} + \sigma_j^{-2}\right)^{-1} \tag{17}$$

where σ_j^2, the signal's variance, is equal to $m_{jH}^{-1}\sigma_{\epsilon^H}^2$. 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 (16) and (17), I establish the following:

Lemma 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 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, $m_{j,H}$, 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 16).

[25] 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 3.6 discusses the implications of relaxing this assumption for the model's main prediction

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:

$$V(w^L, \hat{\mu_j}, \hat{\sigma_j}^2, a_i) = \max\{V_j^L(w^L), V_j^H(\hat{\mu_j}, \hat{\sigma}_j^2) - c(a_i, q_j, m_{jH})\}, \tag{18}$$

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

$$V_j^H = \sum_{j' \in \mathcal{J}} \mathbb{E}_{w_i^H}[U(c_{i,j}^H, \Phi_{j'}^H)|\mathcal{I}_j]\rho_{j'}^H \tag{19}$$

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),

$$V^L = \sum_{j' \in \mathcal{J}} U(c_{i,j}^L, \Phi_{j'}^L)\rho_{j'}^L \tag{20}$$

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

$$V_j^H(\hat{\mu}_j, \hat{\sigma}_j^2) - V^L(w^L) = c(a_j^*). \tag{21}$$

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

$$s_{i,j} = 1 - \mathcal{G}(a_j^*). \tag{22}$$

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

$$s_j = \frac{\sum_{i=1}^{M_j} s_{i,j}}{M_j} = 1 - \mathcal{G}(a_j^*) \tag{23}$$

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 $\hat{\boldsymbol{\mu}} = \{\hat{\mu}_1, ..., \hat{\mu}_J\}$ and uncertainty $\hat{\boldsymbol{\sigma}}^{\mathbf{2}}$=$\{\hat{\sigma}_1^2, ..., \hat{\sigma}_J^2\}$, 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 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 (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.

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[26]

$$\triangle V_j = J\left(\rho_{j'}^H \frac{-\exp(-\gamma(\hat{\mu}_j - \gamma(\hat{\sigma}_j^2/2))}{\sum_{j' \in \mathcal{J}} \frac{\Phi_{j'}^H}{\exp(\gamma r_{j'})}} - \frac{-\exp(-\gamma w^L)}{\sum_{j' \in \mathcal{J}} \frac{\Phi_{j'}^L}{\exp(\gamma r_{j'})}}\right). \tag{24}$$

The key variable that drives the optimal investment threshold a_j^* and, as a consequence, the optimal investment decision is beliefs about w_i^H. Combining Equations (21) and (22), I begin by establishing two intuitive properties of the optimal investment decision. All proofs are provided in the Appendix B.3.

Lemma 3. *The ability threshold a_j^* is strictly decreasing in $\hat{\mu}_j$ and strictly increasing in $\hat{\sigma}_j^2$. Hence, the probability of investing in education s_j is strictly increasing in $\hat{\mu}_j$ and strictly decreasing in $\hat{\sigma}_j^2$.*

First, a higher expected value of the high-skill wage $\hat{\mu}_j$ increases the probability that a child will invest in education, holding all else equal. Increasing the expected value of the high-skill wage $\hat{\mu}_j$ increases the perceived skill premium (Equation (B.12)) decreasing a_j^* and, therefore, the fraction of children from neighborhood j that invest in education. Second, greater uncertainty about the high-skill wage ($\hat{\sigma}_j^2$) translates into a lower perception of the skill premium (Equation (B.12)) 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.

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 (17)), and is established in Lemma 1. Panel A in Figure 2 illustrates this effect. Because children are risk averse, lower uncertainty associated with human capital investment translates into a widening

[26] 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 B.2.

of the mass of children that invest in education (a_j^* decreases, thus s_j^* increases), as established in Lemma 3.

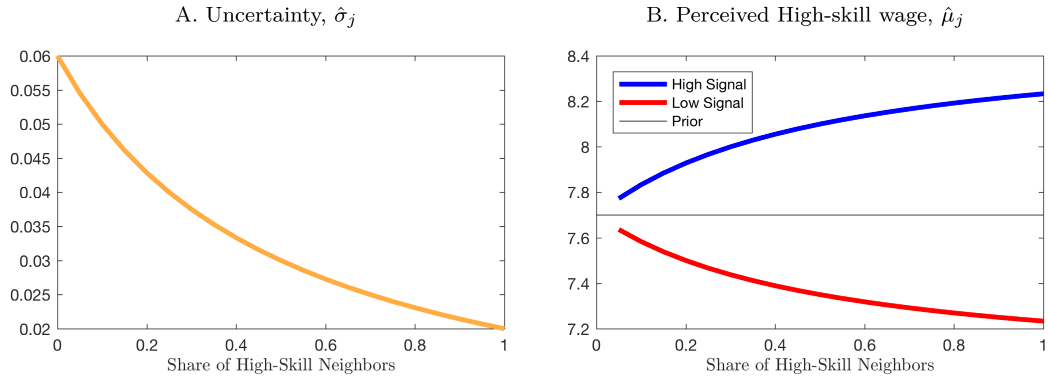
The effect of high-skill neighbors on $\hat{\mu}_j$ is, however, ambiguous. As a reaction to a more precise signal, when estimating w_i^H in a Bayesian fashion (Equation (16)), children place a higher weight on the labor market information disclosed by their neighbors (the public signal w_j^H), as formalized in Lemma 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 $m_{j,H}$. 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 (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 3, a lower/higher $\hat{\mu}_j$ translates into a lower/higher fraction of children investing in education. Panel B in Figure 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 $m_{j,H}$) 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 formalizes this result.

Proposition 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 $w_j^H > w^*$.*

Figure 2: Posterior Mean and Uncertainty: Different Neighborhoods Skill-Mix



The left panel plots the posterior sigma $\hat{\sigma}^2_j$ for different skill compositions of neighborhoods. The right panel plots the posterior mean $\hat{\mu}_j$ for different skill compositions of neighborhoods and different signals. To compute $\hat{\mu}_j$ and $\hat{\sigma}^2_j$, I use the following parameter values $\tilde{\mu} = 7.6$, $\tilde{\sigma}^2 = 0.06$, $\sigma^2_{\epsilon^H} = 0.03$, 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 4.

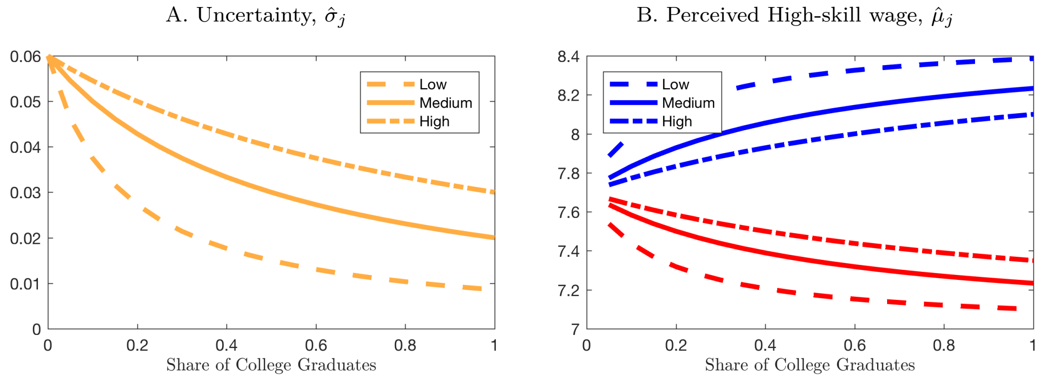
Wage dispersion $(\sigma^2_{\epsilon^H})$ Information transmission from high-skill neighbors as a channel through which children learn about the high-skill wage w_i^H depends on its dispersion $\sigma^2_{\epsilon^H}$. For a given spatial equilibrium, the higher is $\sigma^2_{\epsilon^H}$, the lower is 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 as shown in Equation (15). Figure 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.

School quality (q_j) The quality of the school at each location plays a standard role: 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, holding all else equal, the perceived skill premium is higher, hence the probability of investing in education increases.

To sum up, 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

Figure 3: Posterior Mean and Uncertainty: The Role of Wage Dispersion



Notes: The left panel plots the posterior sigma $\hat{\sigma}_j^2$ for different skill compositions of neighborboods and three different levels os wage dispersion $(\sigma^2_{\epsilon^H})$ - low, medium and high. The right panel plots the posterior mean $\hat{\mu}_j$ for different skill compositions of neighborboods and three different levels os 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 $\tilde{\mu} = 9$, $\tilde{\sigma}^2 = 0.06$, $\tilde{\mu} = 7.6$, $\sigma^2_{\epsilon^H} = 0.03$, 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 4.

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. But, more importantly, 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. This result is consistent with the empirical evidence presented in the previous section.

3.6 Discussion of the Model's Assumptions

I make several assumptions that make my model simpler 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.[27] One could re-define $\Phi_{k,j}$ in 8 as $\Phi_{k,j} = q_j A_{1,j}^{\beta_k} A_{2,j}^{1-\beta_k}$, with $A_{1,j} = m_{j,H}$ begin 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

[27] 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).

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. 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 B.4 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 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 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 a neighborhood j, there will 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. 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

$$s_j = \frac{\sum_{i=1}^{M_j^H} s^H_{i,j} + \sum_{i=1}^{M_j^L} s^L_{i,j}}{M_j} \tag{25}$$

s_j increases with the fraction of high-skill neighbors $m_{j,H}$ if $\frac{\partial s^H_{i,j}}{\partial m_{j,H}} + \frac{\partial s^L_{i,j}}{\partial m_{j,H}} > 0$. Whether $\frac{\partial s^H_{i,j}}{\partial m_{j,H}}$ and $\frac{\partial s^L_{i,j}}{\partial m_{j,H}}$ are greater or lower than zero will depend, respectively, on the magnitude of the signal w^H_j relative to the threshold $w^{*,H}_j$ and $w^{*,L}_j$, as stated in Proposition 1. In Section 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 parental educational 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 $h_i = a_i^{\varphi} \cdot h^{\eta} \cdot q_j^{\kappa} \cdot m_{H,j}^{\rho}$, 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 described in Proposition 1. Appendix B.4 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 eduation is negative – albeit higher than the one in Proposition 1 due to the fact that now individuals do not dislike uncertainty. Hence, under risk neutrality, the magnitude of the signal does 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, consistent with the empirical evidence presented in Section 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 will depend 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.

4 The Importance of Local Learning

Even though imperfect information paired with local learning can reconcile the documented pattern in Section 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 paper to explore 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.

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.[28]

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.[29] 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.[30]

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 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.

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.

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.

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

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 The 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.[31] Assuming that the current allocation of individuals across school-districts is in equilibrium, for any two neighborhoods j and j', the following holds

$$\frac{M_j^k}{M_{j'}^k} = \frac{\Phi_j^k}{\exp(\gamma r_j)} \frac{\exp(\gamma r_{j'})}{\Phi_{j'}^k} \tag{26}$$

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 26.

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

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

[31] 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.

where $\mathbf{W}$ 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 A.6 summarizes all parameters.

4.4 Model Fit

This section discusses the calibrated economy. Table 2 compares the empirical targets for the calibrated parameters and the corresponding moments produced by the model. The left panel in figure 4 depicts the predicted and observed values of the college enrollment rate for the 95 school-districts within Detroit, and the right panel 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 Table 3, 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 paper, local learning, and then I assess which channel matters the most for differences in enrollment across neighborhoods.

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 $\hat{\sigma}_j^2 = \tilde{\sigma}^2$). The left and the right panel in Figure 5 plot, respectively, the perceived skill premium and college

Table 2: Model Fit: 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

Notes: The table 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. Observations are at school-district level. The sample is composed by 95 school-districts within Detroit in the year 2013.

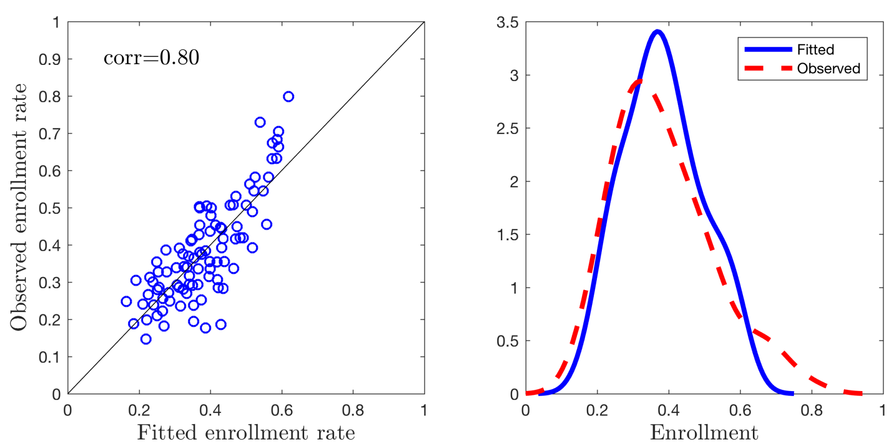
Table 3: Out-of-sample Years

	2009	2010	2011	2012	2014
Correlation	0.50	0.74	0.83	0.80	0.80

Notes: The table reports the correlation between fitted and observed enrollment rate across school-districts in out-of-sample years. Observations are at school-district level. In each year the sample is composed by all school-districts within Detroit.

Figure 4: Model vs. Data

A. Enrollment Rate B. Enrollment Distribution



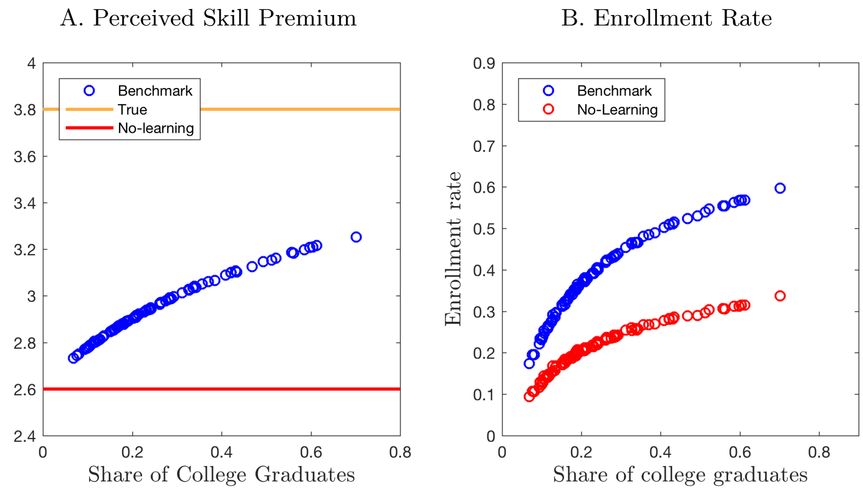
Notes: The left panel plots fitted and observed values for the college enrollment rate across school-districts. Fitted values are on the horizontal axis; observed values are on the vertical axis. Correlation between fitted and observed values is equal to 0.8. Observations are at school-district level. The right panel 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 3% and its uncertainty decreases by 28%, on average. As a consequence, the perceived skill premium rises 6.7%, on average (right panel in Figure 5). This has a significant effect on enrollment as shown in the right panel of Figure 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 22 percentage points lower, on average. This means that instead of having 42% of high-school graduates enrolling in college within 6 months of graduation in Detroit, only 18% would.

Figure 5: Benchmark vs. No-Learning Counterfactual



Notes: The left panel plots the perceived skill premium across school-districts in the benchmark economy (blue) versus the no-learning counterfactual (red). The right panel 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.

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.

4.6 Decomposition: which channel matters the most?

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 spillovers equal to the average value across school-districts.[32] Table 4 reports the 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 4: 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 ($m_{j,H} = \bar{m}$) 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.

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.

[32] 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 paper 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.

4.7.1 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.[33] 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 5 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 5 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 5 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 reallocation policy on the probability of enrolling in college for movers ranges between 0.20 to 0.28.

4.7.2 Disclosure Policy

To understand the potential of an information campaign, I perform a counterfactual analysis where all children are given the an extra signal that informs them about the distribution of the high-skill wage: $\omega = w^H + \epsilon^H$, with $\epsilon_i^H \sim \mathcal{N}(0, \sigma^2_{\epsilon^H})$. Figure 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 perceived 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

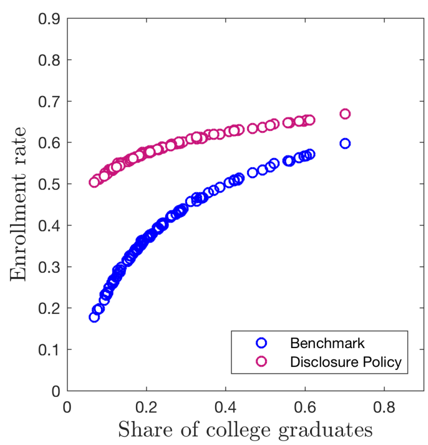
[33] This can be rationalized by the existence of a Government that has land in all locations where it can build public houses. Also note that if amenities were endogenous and depended, for instance, on the share of high-skill neighbors as in Diamond (2016), the effects of this policy would be the same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

educated parents reduces in 62%.

This policy counterfactual exercise is comparable to the recent information experiment by 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).

All together, my results suggest that information campaigns that inform individuals about the wage distribution are an effective in closing the gap between children from different backgrounds. Given the low cost of these campaigns, as pointed out by Bleemer and Zafar (2016), the case to implement them is clear. More so, when the effectiveness of other policies such as subsidies or loans may depend on whether individuals know the true value of education.

Figure 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 $\hat{\sigma}_j^2 = \sigma_{\epsilon^H}$), pink).The sample is composed by the 95 school-districts within Detroit in 2013.

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

Table 5: **Reallocation Policy**

	High-skill neighbors		Enrollment rate		
	1^{st}qtl	4^{th}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 25^{th} percentile of the college graduates distribution to location in the 75^{th} 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

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 A.4 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 A.5 and table A.8 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.

5 Conclusion

Why does the place where children grow up shape their opportunities in life? I have proposed a novel explanation featuring imperfect information and local information transmission: individuals are uncertain about the skill premium and learn about it by observing noisy signals of wage realizations of their neighbors. Spatial location matters because it shapes children 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 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, it is the most important channel in explaining inequality in college enrollment across school-districts. 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.

References

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.

Altonji, J. and R. Mansfield, "Estimating Group Effects Using Averages of Observables to Control for Sorting On Unobservables: School and Neighborhood Effects," 2016.

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.

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

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.

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

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

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

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.

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.

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., 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.

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.

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.

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.

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

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.

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

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.

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.

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

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

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.

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.

A Additional Tables and Figures

Table A.1: 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 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 A.2: 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 A.3: 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-districs 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 A.4: 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, College Degree	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, College Degree	-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 A.5: 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 9 uses the same set of controls as in the baseline estimation (column 6 of table A.4), 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. The latter 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, College Degree (square)	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, College Degree		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 Degree								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 A.6: 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 A.7: 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 25^{th} percentile of the college graduates distribution to location in the 75^{th} 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

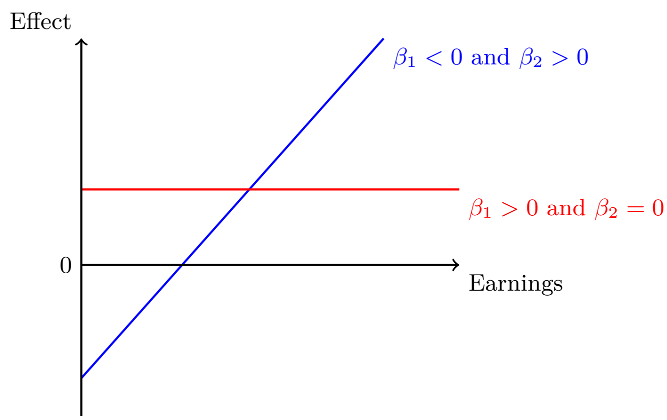
	5% of childrens within 1^{st}qtl					25% of childrens within 1^{st}qtl					50% of childrens within 1^{st}qtl				
	High-skill neighbors		Enrollment rate			High-skill neighbors		Enrollment rate			High-skill neighbors		Enrollment rate		
	1^{st}qtl	4^{th}qtl	Movers	Stayers	Receivers	1^{st}qtl	4^{th}qtl	Movers	Stayers	Receivers	1^{st}qtl	4^{th}qtl	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 A.8: Different Priors: Benchmark Economy vs. Counterfactuals

The table reports several statistics under the benchmark economy and four different scenarios: no local learning, equal human capital spillovers ($m_{j,H} = \bar{m}$) 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.

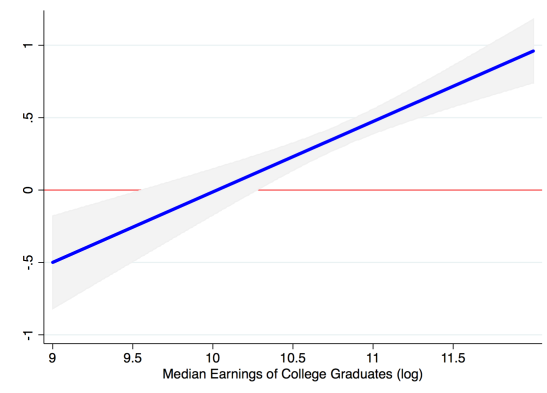
	Benchmark	No-learning	$m_{j,H} = \bar{m}$	$q_j = \bar{q}$	No Frictions
Std. Dev. Enrollment	0.13	0.07	0.10	0.13	0.01

Figure A.1: Interpretation of Coefficients' Signs

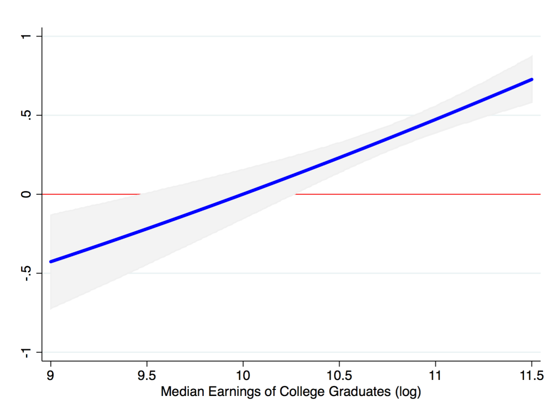


Notes: The graph illustrates the effect of college graduates on college enrollment along the earnings dimension under different signs of the coefficients of interest, β_1 and β_2. The red line displays the effect under the human capital spillovers channel proposed in the literature.

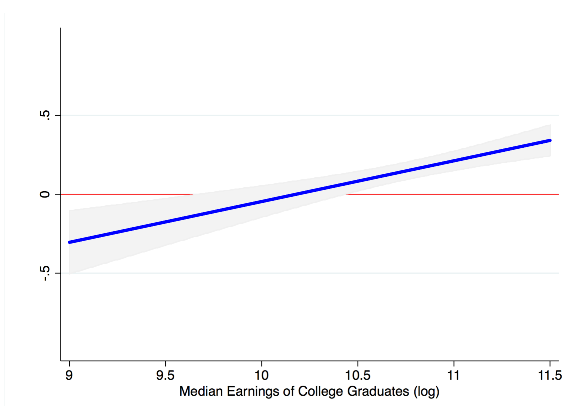
Figure A.2: Correlation between College graduates and Enrollment: Heterogeneity By Earnings



(a) Linear Specification (column 4 in table A.4)



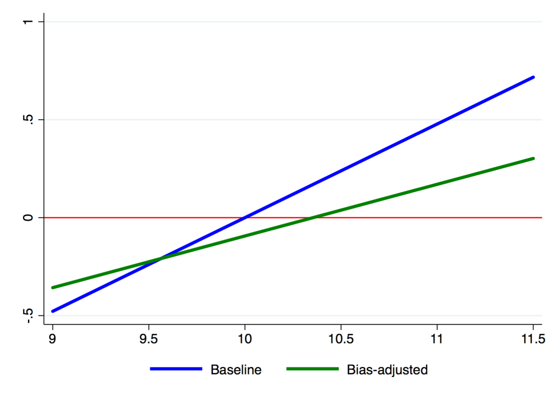
(b) Quadratic Specification (column 1 in table A.5)



(c) Lagged Enrollment (column 2 in table A.5)

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 A.4, while Panels (b) and (c) plot the average marginal effect when I consider a quadratic specification in earnings (column 1 in Table A.5) and control for college enrollment in the previous period (column 2 in Table A.5) . The shaded area represents 95% confidence intervals. The x-axis corresponds to the log median earnings of college graduates in 2010 dollars.

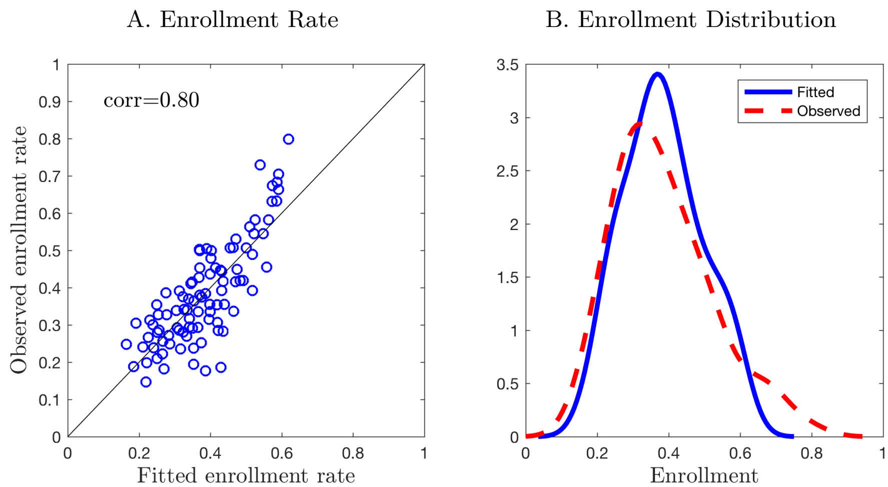
Figure A.3: College graduates and Enrollment: Adjusted-bias Coefficients



$Earnings_{ijt}$

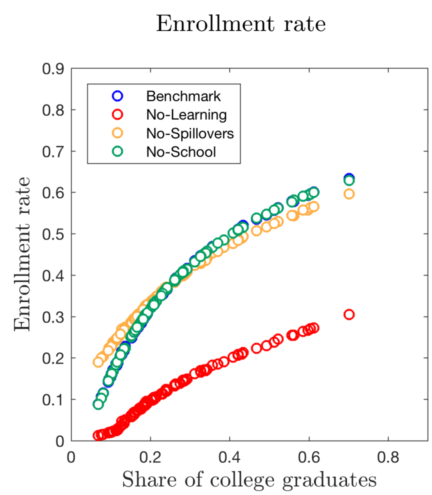
Notes: This figure plots the average marginal effect when I use the estimated coefficients in column 6 in Table A.4 (blue line) and the bias-adjusted coefficients, β_1^* and β_2^* (green line) when the influence of unobservables on the outcome variable is of similar magnitude as the impact of observable 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 A.4 and $\tilde{\beta}$ and $\tilde{R}$ are the estimated coefficients and R^2 of OLS estimation of Equation (2) 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 A.4: Different Priors: Model vs. Data



The left panel 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. The right panel 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 A.5: 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 ($m_{j,H} = \bar{m}$, yellow) and equal school resources ($q_j = \bar{q}$, green). The sample is composed by the 95 school-districts within Detroit in 2013.

B Theoretical Appendix

Location decisions I report additional details for the characterization of parents locations decisions, as described by Equation (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$(\theta, 1)$.[34] Hence, the indirect utility function described by Equation (8) is also Weibull distributed:

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

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.[35] 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

$$Pr[k \in argmin\ X_i] = \frac{\upsilon_k^{-\theta}}{\sum_i \upsilon_i^{-\theta}}, \forall k \in \mathcal{I} \tag{B.2}$$

Combining Equations (B.1) and (B.2), 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:

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

which simplifies to

$$\rho_j^k = \frac{\Phi_{k,j}\exp(\gamma(-r_j))}{\sum_{j'\in\mathcal{J}}\Phi_{k,j'}\exp(\gamma(-r_{j'}))} \tag{B.3}$$

Because $\rho_{i,k,j}$ 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 this, the number of k-skill parents in each neighborhood is

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

B.1 Spatial Equilibrium - An Illustration

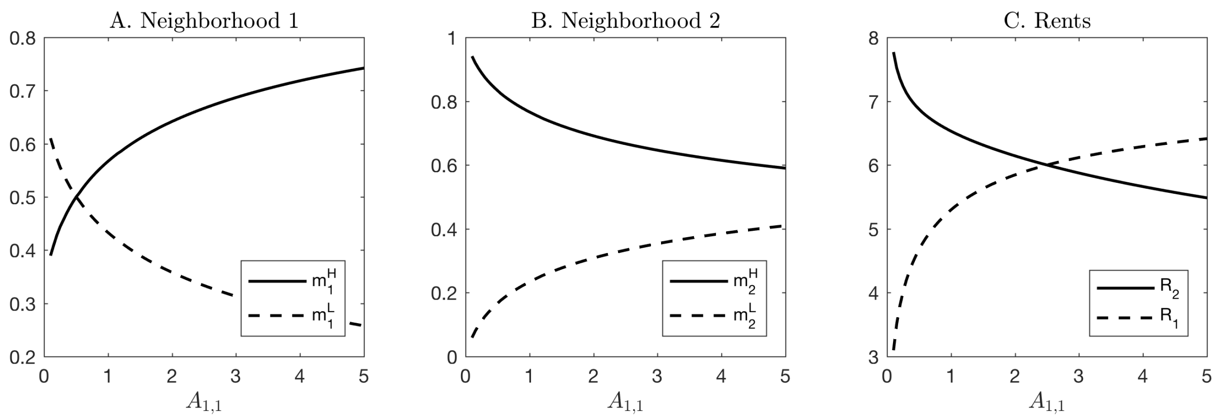
Let's consider the example of a city with two neighborhoods, 1 and 2, each with the same capacity, $H_1 = H_2$. I set $A_{2H} = A_{1L} = A_{2L} = 2.5$, and look to the spatial equilibrium for different values of $A_{1,1}$. Panels A, B and C in Figure B.1 show, respectively, the equilibrium skill-mix 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

[34] 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, $\lambda = 1$.

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

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.

Figure B.1: 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. $H_1 = H_2 = 75$, $M^H = 100$, $M^L = 50$, $\beta_H = 1$, $\beta_L = 0$, $A_{1,2} = A_{1,2} = A_{2,2}$.

B.2 Value Functions

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

$$V_j^H = \sum_{j' \in \mathcal{J}} \underbrace{\Gamma\left(1 + \frac{1}{\theta}\right) \mathbb{E}_{w_i^H}[U(c_{i,j}^H, \Phi_{H,j'})|\mathcal{I}_j]}_{\text{Expected utility of living in location } j' \text{ if } k=H} \quad \rho_{j'}^H \tag{B.4}$$

where $\Gamma\Big(1 + \frac{1}{\theta}\Big)$ 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. $\rho_{j'}^H$ is the probability of living in neighborhood j' conditional on being a high-skill worker.[36] I assume $\theta = 1$

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

for simplicity, hence $\Gamma\Big(1+\frac{1}{\theta}\Big)=1$. Equation (B.4) can be rewritten as

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

$$= \sum_{j'\in\mathcal{J}} \Big[\frac{-\exp(-\gamma(\hat{\mu}_j - \gamma(\hat{\sigma}_j^2/2) - r_{j'}))}{\Phi_{j'}^H}\Big]\rho_{j'}^H =$$

which simplifies to

$$V_j^H = -\exp(-\gamma(\hat{\mu}_j - \gamma(\hat{\sigma}_j^2/2))\left(\frac{J}{\sum_{j'\in\mathcal{J}}\frac{\Phi_{j'}^H}{\exp(\gamma r_{j'})}}\right) \tag{B.5}$$

Equation (B.5) 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

$$V_j^L = \sum_{j'\in\mathcal{J}} \underbrace{\Gamma\left(1+\frac{1}{\theta}\right)U(c_{i,j}^L, \Phi_{j'}^L)}_{\text{Expected utility of living in location } j' \text{ if } k=L} \rho_{j'}^L = \tag{B.6}$$

where $\Gamma\Big(1+\frac{1}{\theta}\Big)$ 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 $\Gamma\Big(1+\frac{1}{\theta}\Big)=1$. Equation (B.6) can be rewritten as

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

which simplifies to

$$V^L = -\exp(-\gamma w^L)\left(\frac{J}{\sum_{j'\in\mathcal{J}}\frac{\Phi_{L,j'}}{\exp(\gamma r_{j'})}}\right) \tag{B.7}$$

Equation (B.7) is equal for all children in the city, regardless of where they live. Hence I suppress j.

B.3 Proofs

Proof of Lemma 3 Given V_j^H (Equation (B.5)) and V^L (Equation (B.7)), the perceived skill premium for a child born in neighborhood j, $\triangle V_j \equiv V_j^H - V^L$, is given by

$$\triangle V_j = J\left(\frac{-\exp(-\gamma(\hat{\mu}_j - \gamma(\hat{\sigma}_j^2/2))}{\sum_{j'\in\mathcal{J}} \frac{\Phi_{j'}^H}{\exp(\gamma r_{j'})}} - \frac{-\exp(-\gamma w^L)}{\sum_{j'\in\mathcal{J}} \frac{\Phi_{j'}^L}{\exp(\gamma r_{j'})}}\right) \tag{B.8}$$

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 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

$$\frac{\partial \varpi}{\partial m_{jH}} = \frac{\partial \varpi}{\partial \triangle V_j}\frac{\partial \triangle V_j}{\partial m_{jH}} + \frac{\partial \varpi}{\partial c(a_i^*)}\frac{\partial c(a_i^*)}{\partial m_{jH}} = = \underbrace{\Upsilon \cdot \frac{\sigma_{\epsilon^H}^2}{m_{jH}}\frac{\tilde{\sigma}^2}{(\tilde{\sigma}^2+\sigma_j^2)^2} \cdot \left[w_j^H - \tilde{\mu_j} + \frac{\gamma}{2}\tilde{\sigma}^2\right]}_{A>0 \text{ or } A<0} + \underbrace{\frac{\partial \varpi}{\partial c(a_i^*)}\frac{\partial c(a_i^*)}{\partial m_{jH}}}_{B>0} \tag{B.9}$$

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.

B.4 Implications of other specifications

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

$$U(w_i^k, r_j, \Phi_j^k, \varepsilon_{i,j}) = w_i^k - r_j + \Phi_j^k + \varepsilon_{i,j} \tag{B.10}$$

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.[37] The distributional assumption on the idiosyncratic taste, ε, allows me to derive a close-form expression for $\rho_{i,j}^k$, as before:

$$\rho_{i,j}^k = \frac{\exp(w_i^k - r_j + \Phi_j^k)}{\sum_{j' \in \mathcal{J}} \exp(w_i^k - r_{j'} + \Phi_{j'}^k)} \tag{B.11}$$

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 (B.10) and (B.11), 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 (B.12):

$$\triangle V_j = \underbrace{\sum_{j' \in \mathcal{J}} \left[\hat{\mu}_j - r_{j'} + \Phi_{j'}^H\right]\rho_{j'}^H}_{V_j^H} - \underbrace{\sum_{j' \in \mathcal{J}} \left[w^L - r_{j'} + \Phi_{j'}^L\right]\rho_{j'}^L}_{V_j^L} \tag{B.12}$$

It can be shown that:

[37] 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 $f(x) = \exp(-\exp(-(x - \mu_\varepsilon)/\sigma_\varepsilon))$.

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.

Uncertainty about Low-Skill Wage If $\sigma^2_{\epsilon L} > 0$, Equation (B.13) can be re-written as

$$\triangle V_j = J\left(\frac{-\exp(-\gamma(\hat{\mu}_j^H - \gamma(\hat{\sigma}^2_{H,j}/2))}{\sum_{j'\in\mathcal{J}} \frac{\Phi^H_{j'}}{\exp(\gamma r_{j'})}} - \frac{-\exp(-\gamma(\hat{\mu}_j^L - \gamma(\hat{\sigma}^2_{L,j}/2))}{\sum_{j'\in\mathcal{J}} \frac{\Phi^L_{j'}}{\exp(\gamma r_{j'})}}\right) \tag{B.13}$$

where are $\hat{\mu}_j^H$ and $\hat{\sigma}^2_{H,j}$ the posterior mean and variance of the beliefs about w_i^H; $\hat{\mu}_j^L$ and $\hat{\sigma}^2_{L,j}$ are the posterior mean and variance of the beliefs about w_i^L for a child born in neighborhood j. Following the same steps as in the proof of lemma 3 above, it can be shown that:

1. $\frac{\partial s_{i,j}}{\partial \hat{\mu}_j^H} > 0$ and $\frac{\partial s_{i,j}}{\partial \hat{\mu}_j^L} < 0$

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

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

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

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

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

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

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 hold: $\frac{\partial s_{i,j}}{\partial m_{j,H}} > 0$ 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.

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.