Uncertainty in education: The role of communities and social learning *

Ana Figueiredo †

February 2017

PRELIMINARY and INCOMPLETE

Abstract

This paper studies the role of information frictions and social learning as a novel channel through which neighborhood affect children educational outcomes. I write a model of educational choice with two novel features: *(i)* individuals are uncertain about the returns to education, and *(ii)* they learn about the them by observing nearby skilled neighbors. In contrast to prior models of human capital formation with local interactions, in a model with information frictions and social learning, it is not only about being exposed skilled neighbors, but also about their wage distribution: skilled neighbors only increase schooling investment if the labor market information they disclose leads to an increase in the perceived education premium. Using school-district data from Michigan over the period 2008-2014, I find evidence supporting the model's prediction. I also calibrate the model to Detroit data in 2013 and through counterfactual experiments show that social learning plays an important role in reducing the underlying uncertainty about college returns, however it also increases inequalities across neighborhoods.

Keywords: educational choice, uncertainty, information frictions, Bayesian updating.

JEL Classification: D83, I24, J24

*I thank Jan Eeckhout for his advice and Isaac Baley for his helpful feedback. I also thank Davide Debertoldi, Steven Durlauf, Maria Lombardi, Andrii Parkhomenko, Christopher Rauh as well as participants at CREI Macrobreakfast, Barcelona GSE PhD Jamboree (2015), SSSI 2015 at the Becker Friedman Institute of University of Chicago, XXI Workshop on Dynamic Macroeconomics and PET 2016 for their comments. I acknowledge financial support from Fundacão para a Ciência e Tecnologia (SFRH/BD/94516/2013).

†PhD candidate, Department of Economics and Business, Universitat Pompeu Fabra, Carrer Ramon Trias Fargas 25-27, Barcelona 08005 Spain. (e-mail: ana.gomesfigueiredo@upf.edu)

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

1 Introduction

Every discussion about residential segregation is intimately tied to race. However, race is not the only way in which individuals are spatially segregated. College graduates and non-college graduates are also likely to live in different neighborhoods. I call this form of residential segregation skill-segregation. Exploring data patterns in US cities, I find that, in comparison to non-college graduates, college graduates live in neighborhoods where the share of college graduates is 10 percentage points higher, on average.[1] The basic concern with residential segregation is that it creates inequalities between neighborhoods, and the neighborhood context is critical for children's educational decisions and future outcomes. These neighborhood effects have been widely documented in the literature: for instance, Chetty et al. [2016] find that children that moved with their families to better neighborhoods when they were young have higher college attendance rates and earnings. In this paper, I focus on the role of information constraints at the neighborhood level as a novel explanation for the positive correlation between children and neighbors' outcomes. This approach builds on recent studies showing that (*i*) individuals are uncertain about schooling returns and that (*ii*) neighborhoods may play have a very important function as a source of information.

When deciding whether to invest in education, parents and children often lack relevant information. On the one hand, there is uncertainty about the conditions that will prevail in the labor market upon graduation, either about wages or about employment rates (Bleemer and Zafat [2015] and Belfield et al. [2016]).[2] On the other hand, there is also uncertainty about

[1] See Appendix A for details on calculations.

[2] Bleemer and Zafat [2015] asked to a group of over a thousand US individuals how much more did they think a college graduate earned as compared to a non-college graduate. A college graduate was thought to earn 63% more than a non-college graduate, on average, 17 percentage points less than reported using US census data

the child's ability and her probability of completing a post-secondary education (Bobba and Frisancho [2014], Arcidiacono et al. [2016] and Kunz and Staub [2016]).[3] More importantly, recent studies show that students from less advantaged backgrounds are those who are most affected by these informational barriers (Hoxby and Avery [2014]).[4] Under this context of uncertainty and lack of information, if the assessment of the education value depends on the distributions of educational levels and incomes observed in a community, neighbors may play a role in educational decisions of children.[5] In locations where low-income households live isolated, children have limited exposure to higher levels of education, hence they have limited information on earnings and unemployment rates for different schooling scenarios, specially higher education. This may result in a poverty cycle in which inaccurate beliefs translate into insufficient investments in education, conditioning labor market outcomes and thus reducing their likelihood of moving up in the income ladder.

I explore this hypothesis by writing a model of residential and educational choice that accounts both for (i) uncertainty at the investment stage and (ii) information constraints imposed by the neighborhood. To the best of my knowledge, the vast theoretical literature that studies the effects of local interactions on human capital formation has not yet taken into account the role of uncertainty and information. In contrast, these models usually assume that parents and children know future returns along the different educational trajectories. This is potentially an important omission as perceptions about the education value and information constraints have been shown to have significant impacts on different educational decisions (e.g. Jensen [2010], Attanasio and Kaufmann [2014], Kaufmann [2014], Bleemer

[3]Kunz and Staub [2016] show that subjective beliefs about completing a post-secondary education are important determinantes of post-secondary educational aspirations, enrollment and completion.

[4]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. More important, the authors find that these low-income and high-achievement students attend schools and live in neighborhoods that lack others from older cohorts who have attended selective colleges.

[5]For instance, Yamuchi [2007] shows that individuals learn about schooling rewards from income realizations of their neighbors. Belfield et al. [2016] show that the socioeconomic background, as proxied by parents' education and income, determines students' beliefs about returns and the consumption value of education.

and Zafat [2015], 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).[6] This paper aims to fill this gap in the literature.

The model has two novel features. First, individuals are uncertain regarding the about the return to human capital investment. Second, learning happens at the local level: individuals learn about the schooling returns by observing nearby skilled individuals. I adopt the concept of *Bayesian uncertainty*: individuals use Baye's rule to update their beliefs about schooling returns, and uncertainty is defined as the variance of the probability distribution that defines these beliefs. During their lives, individuals make two decisions. As young, individuals gather information on schooling returns from their neighbors, update their beliefs and choose whether to invest or not in education. As adults, they work, consume and choose a neighborhood to live within a city. Each neighborhood is characterized by a bundle of skilled and unskilled specific amenities, which are taken by individuals as exogenously given. Preferences for each type of amenity are allowed to vary across skilled and unskilled parents, namely in the relative valuation of type-specific amenities.

The simplicity of the model developed allows me to clearly highlight the mechanism at play.[7] The distribution of skilled and unskilled amenities across neighborhoods determines

[6]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 Zafat [2015] 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.

[7]To highlight the potential role of neighbors through uncertainty and social learning, I abstract from other channels through which neighbors may affect children education outcomes such as the local funding of schools.

the skill-mix of each neighborhood. Different skill compositions will produce different inferences about the value of education. The role of skilled neighbors is twofold. Through the social learning mechanism, as the share of skilled neighbors increases, children become less uncertain about education returns. This is the uncertainty effect. But, they also rely more on the information provided by their neighbors. This does not necessarily translate into a higher perception of the returns to education. There is an information threshold: the public signal has to be sufficiently high so that exposure to skilled individuals leads to an increase in the perceived education premium and induces children to invest in education. If the share of college graduates is high but the observed public signal is sufficiently low, children have very low uncertainty about education returns, but they update downwards their perceived value of education which translates into lower incentives to pursue further education. In contrast to previous models that study the role of neighborhoods in educational investment, under information frictions and social learning, neighborhoods' effects go beyond the exposure to skilled neighbors, in particular their wage distribution also matters. More exposure implies more information, but more information only increases the share of children investing in education if it leads to an increase in the perceived well-being gap.

Empirical evidence is consistent with the proposed theory. Using school-district data from Michigan, I study college enrollment of high-school graduates over the period 2008-2014 and uncover a threshold effect, which to the best of my knowledge is novel in the empirical literature on neighborhood effects. In particular, I find that having a higher share of college graduates living in the school-district only increases college enrollment if their median earnings are sufficiently high, as predicted by the model. These results are robust to controlling for socio-economic characteristics of neighborhood residents, school resources and cohort characteristics. Importantly, my results are not masking the fact that students from richer families have a higher ability. One interesting finding is that the role played by neighborhoods goes beyond the school resource mechanism proposed by Bénabou [1993], Bénabou [1996a], Bénabou [1996b], Férnandez and Rogerson [1996] and Durlauf [1996]: when I control for school quality proxies such as expenditure, local revenue and teacher-pupil ratios, I find that the coefficients for these controls are statistically insignificant, and the coefficient on college graduates, my variable of interest, remains unchanged.

To assess the importance of the proposed mechanism, I calibrate the model using data at the school-district level for Detroit in 2013. The model provides a reasonable fit for the data, which lends credibility to the counterfactual simulations. I find that uncertainty plays an important role in the decision to enroll in college, in particular it decreases college enrollment at the city level by 13%. I also find that, through social learning, neighbors mitigate underlying uncertainty about educational returns by 80%, on average. This translates into an improvement in college enrollment by 14.5% at the city level. The aggregate effect hides important differences across neighborhoods: social learning has a higher impact in neighborhoods with a higher share of college graduates. These heterogeneous effects imply an increase dispersion in college enrollment across neighborhoods by 14.1%. So, even though neighbors, through social learning, play an important role in reducing uncertainty about college returns, which translates into more college enrollment at the city level, they also increase inequality across neighborhoods.

What are the implications of my results? For a policy-maker interested in addressing equality of opportunity, understanding how students make their educational decisions is key to understand the origins of persistence inequality between social groups. Conventional wisdom tell us that living around college graduates improves the educational attainment of children from low-educated families. Because of this, policy-makers have designed housing reallocation programs such as the Moving To Opportunity in the US.[8] In contrast, my results support information interventions, as the one studied by Hoxby and Turner [2015], which make the returns to human capital investment more salient in disadvantaged neighborhoods in order to create incentives to invest in education, especially among individuals from lower socio-economic backgrounds.

Related literature. This paper contributes to the literature that studies local interactions as the mechanism behind poverty traps. Models of local interactions and educational outcomes have focused on three main channels. First, Bénabou [1993], Bénabou [1996a], Bénabou [1996b], Férnandez and Rogerson [1996] and Durlauf [1996] propose models where the local financing of public schools induces endogenous segregation of wealthier families into homogenous communities. In these models, poor children attend schools with lower resources, therefore their prospects of moving upwards are lower when compared to children from better backgrounds. Another strand of the literature focuses on the role of identity: in neighborhoods where few parents are well educated, high education can render an individual the feeling of being alienated from those with whom he wants to share an identity (Arkelof and Kranton [2000]). Finally, Bowles et al. [2014], Cavalcanti and Gisnnitarou [2013] and Kim and Loury [2013] build models where what matters is the social network of the individual. In these models, what creates the link between social influences and educational outcomes is the skill acquisition technology, which depends on the average level of human capital of the individual's social network.

A common feature of these models is the assumption that individuals effectively take into

[8] The Moving To Opportunity (MTO) was a large experiment of the U.S. Department of Housing and Urban Development conducted between 1994 and 1998 in five large U.S. cities. Participant families were randomly assigned to one of the three groups: an experimental voucher group that was offered housing voucher to move to a neighborhood with a poverty rate below 10%, a Section 8 voucher group that was offered a housing voucher with no geographical restrictions, and a control group that did not received a voucher. The vast empirical literature assessing the effects of MTO has found mixed results in several outcomes. For an overview of results from the several MTO studies see Topa and Zenou [2015].

account current and future net returns of education. Departing from the existing literature, I introduce uncertainty and a social learning mechanism into a human capital formation model with local interactions. In doing so, I show that under information frictions and social learning, highly educated neighbors only increase incentives for education investment if the labor market information they disclose leads to an increase of the perceptions about education returns. This contrasts with previous models, which predict a positive relationship between highly educated neighbors and education investment, independently of the wage distribution.

This paper builds on the theoretical literature that studies how information and social learning can affect economic agents' decision making such as Férnandez [2013] (female labor participation), Allen [2014] (international trade) Baley et al. [2016] (international trade) and Fajgelbaum et al. [2016] (firm investment decisions). 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. My model introduces a similar learning environment in a model of human capital investment with local interactions and uncertainty.

Finally, I also provide new insights to the vast empirical literature that exploits the relationship between neighborhoods' composition and educational outcomes (for a review, see 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. I uncover a novel threshold effect: the % of college graduates in the school-district may have a positive or negative effect on college enrollment depending on the earnings distribution. If the median earnings are sufficiently low, the effect is negative. Otherwise, it is positive.

The paper proceeds as follows. Section 2 outlines the model and section 3 presents empirical evidence supporting the mechanism proposed. In section 4, I assess the quantitative importance of information frictions and social learning. Section 5 concludes.

2 The model

Population. There are L families living in a city. Each family is composed by a parent and a child. Parents are of two types, skilled (s) and unskilled (u), $j \in \{s, u\}$, and they are endowed with one unit of labor that they supply inelastically. The city is closed thus the population of skilled and unskilled parents in the city, L_s and L_u respectively, are exogenously given.[9]

City. The city is composed by a set of discrete neighborhoods, indexed by $n = \{1, ..., N\}$. 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. I consider that each neighborhood is characterized by an amenity vector, α_n, which includes two type of amenities, skilled and unskilled, $\alpha_n = \{\alpha_{u,n}, \alpha_{s,n}\}$. The divide between skilled and unskilled amenities aims to capture the idea that the definition of what makes a neighborhood attractive may differ across skill groups: while the road network or culture supply may be a desirable features for individuals with higher income, for low-income families the public transport network or close by low-cost grocery stores may be the features that make a location attractive. Both skilled and unskilled amenities at each location are taken by individuals as exogenously given.[10] Higher amenities denote more desirable locations. All

[9] 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].

[10] Even though this may strike as a strong assumption, in section 2.3 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 results of the paper.

residents within the neighborhood have access to these amenities simply by choosing to live there.

Even though the city's skilled and unskilled populations are exogenous, the quantity of skilled and unskilled families living in a given neighborhood n, $L_{s,n}$ and $L_{u,n}$ respectively, are endogenously determined equilibrium outcomes. The city as sufficient land capacity that everyone can reside on it, but I consider that each location is endowed with an inelastic supply of land H_n. All land is owned by a zero measure of absentee landlords and it will be assumed that the total stock of land available is for residential use only. Houses are identical across the city.

Labor market. There are two occupations in the city. One that requires skills and another that can be performed by unskilled workers. While the wage of an unskilled worker is known, w_u[11], I consider that the wage of the skilled worker is the sum of two components: a common component, ω, as well as an idiosyncratic component ϵ_i,

$$w_{i,s} = \omega + \epsilon_i$$

The innovation is normally distributed and independent across individuals,

$$\epsilon_i \sim \mathcal{N}(0, \sigma_\epsilon^2)$$

and I consider that the skilled wage follows a log-normal distribution

$$w_{i,s} = log(\omega_{i,s}) \sim \mathcal{N}(\omega, \sigma_\epsilon^2)$$

Preferences. All individuals have preferences over consumption, c, and amenities. The consumption good is a tradable numeraire good with price normalized to one. Individuals live two periods. When young, they face a discrete educational choice. As adults, individuals

[11] In section 4, I augment the model with uncertainty about the unskilled wage, considering that it is also the sum of common component and an idiosyncratic term.

work, consume and choose residential location within the city. For simplicity, I consider that individuals only consume when old. I assume that all individuals have constant absolute risk aversion (CARA) utility over consumption with risk-aversion parameter γ:

$$U(c_{i,n}, \alpha_{i,n}) = \frac{-\exp(-\gamma(c_{i,n}))}{\alpha_{i,n}}$$

where $c_{i,n}$ is consumption of parent i living in neighborhood n. Following Diamond [2016], $\alpha_{i,n}$ maps the amenities in neighborhood n to the individuals' utility value for them.

2.1 Parents' residential location choice

Parents are endowed with one unit of labor that they supply inelastically and have heterogeneous preferences for neighborhoods. Before their children decide whether to invest or not in education, parents simultaneously choose a neighborhood n to live in such that they maximize their utility taking as given the choices of others and labor income. I make no assumption regarding altruism, therefore parents are not concerned about their children expected utility or income. Given this, each parent i solves the following program:

$$\underset{n}{\text{Max}} \qquad U(c_{i,n}, \alpha_{i,n}) = \frac{-\exp(-\gamma(c_{i,n}))}{\alpha_{i,n}} \qquad \text{s.t.} \qquad c_{i,n} + R_n = w_{i,j}$$

where $w_{i,j}$ is the wage of an adult i of type j and R_n is the rent payed to live in neighborhood n. I consider that

$$\alpha_{i,n} = ({\alpha_{s,n}}^{\beta_j} \alpha_{u,n}^{1-\beta_j})\varepsilon_{i,n}$$

where β_j measures parents' relative taste for skilled amenities and is skill-specific, $0 \leq \beta_j \leq 1$. This aims to capture the idea that different types of workers tend to prefer different types of amenities as in Glaeser et al. [2016] and Diamond [2016].[12] Each parent has an individual idiosyncratic taste for neighborhoods, which is measured by $\varepsilon_{i,n}$. I model this heterogeneity following McFadden [1973]. In particular, I consider that the idiosyncratic taste is drawn from a Fréchet distribution:

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

The idiosyncratic taste implies that when faced with the same rental prices and neighborhood amenities equal parents, with the same skill and wage, may choose different residential locations. The indirect utility function of parent i living in neighborhood n can now be represented as

$$U(c_{i,n}, a_{i,n}) = \frac{-\exp(-\gamma(w_{i,j} - R_n))}{({\alpha_{s,n}}^{\beta_j} \alpha_{u,n}^{1-\beta_j})\varepsilon_{i,n}} \tag{1}$$

In the beginning of the period, after the vector of ε_i is revealed (one for each location), families choose the location that yields the highest utility, as described by 1. From now on, I assume $\beta_u = 0$ and $\beta_s = 1$. This means that skilled parents only derive utility from skilled-amenities while unskilled parents only derive utility from unskilled-amenities. This assumption can be relaxed. Also, for simplicity, I assume that $\theta = 1$. Let $\rho_{j,n}$ be the probability that, after observing $\varepsilon_{i,n}$, a j-type worker chooses to live in n. Using the distribution of the idiosyncratic taste, one can derive a close-form expression from this the probability that skilled and unskilled parents choose to live in neighborhood n is, respectively

$$\rho_{s,n} = \frac{\alpha_{s,n}}{\exp(-\gamma(w_{i,s} - R_n))} \upsilon_s^{-1} = \frac{\alpha_{s,n}}{\exp(\gamma R_n) \sum_{f=1}^{N} \frac{\alpha_{s,f}}{\exp(\gamma R_f)}} \tag{2}$$

[12] 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 skilled groups. In line with these specifications, Bayer et al. [2007] find evidence of considerable preference heterogeneity across different types of households. Focusing on the heterogeneity in tastes for school quality, they find that a household's willingness-to-pay for better schools increases with education.

$$\rho_{u,n} = \frac{\alpha_{u,n}}{\exp(-\gamma(w_u - R_n))}\upsilon_u^{-1} = \frac{\alpha_{u,n}}{\exp(\gamma R_n)\sum_{f=1}^{N}\frac{\alpha_{u,f}}{\exp(\gamma R_f)}} \tag{3}$$

where $\upsilon_j^{-1} = \sum_{f=1}^{N}\frac{\alpha_{j,f}}{\exp(-\gamma(w_j - R_f))}$. Other things equal, 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_n). The total number of j-type parents in neighborhood n in equilibrium is given by

$$L_{j,n} = \rho_{j,n} L_j$$

where L_j is the exogenous measure of j-type families living in the city. Then the total population in neighborhood n is $L_{u,n} + L_{s,n}$. I assume that absentee landlords own all the land and rent it to residents. Families live in one and only house and houses are identical across neighborhoods. Land market clearing determines the equilibrium price of housed across all neighborhoods (R_n):

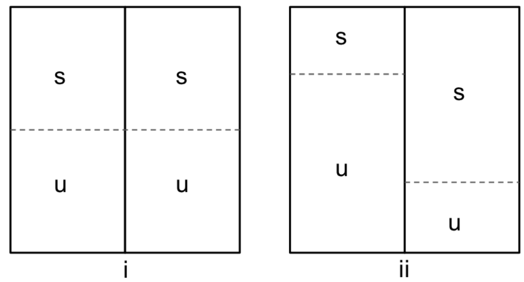
$$H_n = \rho_{s,n} L_s + \rho_{s,n} L_u$$

Definition 1. *Given L_s, L_u, wages (w_u, $w_{i,s}$) and skill specific amenities ($\alpha_{s,n}$, $\alpha_{u,n}$), the spatial equilibrium is an alocation of L_s and L_u over neighborhoods N with an associated vector of housing rental prices $R = \{R_1 ... R_n\}$ such that*

1. *Each parent optimally chooses a neighborhood within the city*

2. *Land market clears in each neighborhood.*

In this context, the distribution of amenities across neighborhoods determines the skill-mix of each neighborhood, hence whether unskilled families live more or less isolated from the skilled ones. Figure 1 illustrates an example of the equilibrium city structure. For the sake of this example, I assume $H_n = H$. City i, for example, is a city where skilled and unskilled amenities are equal across neighborhoods: $\alpha_{s,1} = \alpha_{s,2}$ and $\alpha_{u,1} = \alpha_{u,2}$. This implies



Note: Each square represents a city, within which the solid line divides the city in two neighborhoods, while the dashed line aims at illustrating the skill-mix in each city. In each square, neighborhood 1 is the one to the left of the solid line.

Figure 1: An illustration of the city equilibrium

that the probability of residing in neighborhood 1 or 2 are equal, hence the skill-mix of each neighborhood is the same. The equilibrium in the city is non-sorted. On the other hand, city *ii* is an example of a partially sorted equilibrium: in some neighborhoods, skilled (unskilled) amenities are higher (lower) while in others skilled amenities are lower (higher). In this particular case $\alpha_{s,1} < \alpha_{s,2}$ and $\alpha_{u,1} > \alpha_{u,2}$, which implies that for unskilled individuals the probability of living in 1 (2) is higher (lower) while for skilled individuals is lower (higher). In this case, the skill-mix differs across neighborhoods: one has a higher share of skilled individuals while the other has a lower share.

2.2 Children' educational choice

Children are born to a family of type j, $j \in \{s, u\}$, living in neighborhood n. Besides family background, children differ in their innate ability, a. Each child knows their innate ability. The distribution of ability is assumed to be the same across neighborhoods and family types. Given this, any difference in the education decision arise endogenously in the model, and cannot be traced back to any differences in fundamentals. The distribution of ability is given by the distribution function $\mathcal{G}(a)$, with support $[\underline{a}, \overline{a}]$. Innate ability determines the cost of acquiring education $c(a_i)$. As in Bowles et al. [2014] and Kim and Loury [2013], the function

c is continuous and strictly decreasing in a. The intuition behind this assumption is that the lower one's innate ability, the more mentally stressful the skill acquisition process is or the more materials he must spend on the achievement. In contrast to previous models of human capital formation with local interactions, the skill acquisition cost function abstracts completely from any human capital spillovers through community or parental effects. I do not aim to claim that these effects are not important, I rather do this so I can isolate the role played by information as a driver of educational decisions.

All children have to decide whether to invest or not in education $e \in \{0, 1\}$. Not investing, $e = 0$, implies the child to work as an unskilled worker, while investing, $e = 1$, implies the payment of the skill acquisition cost and working as a skilled worker. As previously said, the key feature of this model is that, at the investment stage, children are uncertain about the future return of the educational investment: in particular, they do not know the true value of the skilled wage they will earn as adults, $w_{i,s}$. Therefore, the children educational decision depends on their beliefs about the unobserved component of the skilled wage, ω. They learn about ω from skilled individuals. In particular, I consider that learning happens at the local level: children learn about ω by observing the wage distribution of skilled parents living in the same neighborhood as them. Since they cannot distinguish ω from the idiosyncratic term, ϵ_i, each signal from a skilled individual, $\omega_{i,s}$, gives noisy information about ω:

$$\omega_{i,s} = \omega + \varepsilon_i \tag{4}$$

Following Fajgelbaum et al. [2016], I consider that the information gathered by each skilled neighbor in neighborhood n its proportional to its size:

$$\varepsilon_i \sim \mathcal{N}(0, L_n \sigma_\epsilon^2)$$

Because of the normality assumption, a sufficient statistic for the information provided by all the skilled parents living in neighborhood n is the public signal

$$\omega_n \equiv \frac{1}{L_{s,n}} \sum_{i=1}^{L_{s,n}} \omega_{i,s} = \omega + \varepsilon_{\omega_n} \tag{5}$$

with

$$\varepsilon_{\omega_n} \equiv \frac{1}{L_{s,n}} \sum_{i=1}^{L_{s,n}} \epsilon_i \sim \mathcal{N}(0, l_{s,n}^{-1}\sigma_\epsilon^2) \tag{6}$$

where $l_{s,n}$ is the share of skilled parents living neighborhood n. Since the public signal depends on the skilled parents in the neighborhood, ω_n is neighborhood specific: all children born in n observe the same skilled parents, hence they share a common public signal, ω_n. More important, the signal's precision increases with the proportion of skilled parents in the neighborhood.

Bayesian updating. After receiving information from nearby skilled parents, each child updates their prior belief using Bayes' rule. I assume that all children, independently of their parent type, share a common prior, $w_{i,s} \sim \mathcal{N}(\tilde{\mu}, \tilde{\sigma}^2)$. Under this assumption, and because all the information is public but neighborhood specific, beliefs are common across children living in the same neighborhood. Even though in this model there is no dispersion in beliefs within a neighborhood, if the proportion of skilled individuals differ across neighborhoods, then there is dispersion in beliefs across neighborhoods. Given, the normality assumption about the prior and the signal, the posterior belief about $w_{i,s}$ is also normally distributed with mean $\hat{\mu}_n$ and variance $\hat{\sigma}_n^2$ given by

$$\hat{\mu}_n = \underbrace{\frac{\sigma_{\omega_n}^2}{\tilde{\sigma}^2 + \sigma_{\omega_n}^2}}_{\text{weight on prior}} \tilde{\mu} + \underbrace{\frac{\tilde{\sigma}^2}{\tilde{\sigma}^2 + \sigma_{\omega_n}^2}}_{\text{weight on signal}} \omega_n \tag{7}$$

$$\hat{\sigma}_n^2 = \left(\tilde{\sigma}^{-2} + \sigma_{\omega_n}^{-2}\right)^{-1} \tag{8}$$

where $\sigma_{\omega_n}^2$, the signal's variance, is equal to $\sigma_\varepsilon^2 l_{s,n}^{-1}$. It can be seen from equation 7 that the Bayesian estimator of the skilled wage is a linear combination of the prior and the public signal. Uncertainty about the future return of education is defined as the variance of the children beliefs about $w_{i,s}$, which depends only on $l_{s,n}$ and the prior's variance. Using equation 8, we can establish the following:

Lemma 1. *For a child born in neighborhood n, uncertainty about $w_{i.s}$, $\hat{\sigma}_n^2$, decreases in the share of skilled parents living in the neighborhood, $l_{s,n}$, and increases in prior uncertainty, $\tilde{\sigma}^2$.*

Lemma 2. *Children rely relatively more on the public signal, ω_n, as the share of skilled parents living in the neighborhood, $l_{s,n}$, increases.*

The share of skilled neighbors, $l_{s,n}$, 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 share of skilled neighbors 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 share of skilled neighbors, update their beliefs in response to signals to a greater extent than those that have less information.

Educational choice. Given the cost of skill acquisition, $c(a_i)$, and beliefs about w_s, a child can choose either to invest in education and become a skilled worker or not to invest in education and be an unskilled worker. Let $\mathcal{I}_n$ be the information set of any child born to a family living in neighborhood n. The child's choice will be a function of his information set, which is only composed by the public signal observed, which is neighborhood specific: $\mathcal{I}_n = \{\omega_n\}$. Note that 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 role played by information frictions in educational decisions. 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. In this setting, the optimal policy of a children i born in neighborhood n with innate ability a_i, $e^*_{i,n}$, is to invest in education if and only the cost of doing so is lower than the expected well being premium, conditional on the information set:

$$e^*_{i,n}(a_i, l_{s,n}) = \begin{cases} 1 & \text{if } V_n^S - V^U \geq c(a_i) \\ 0 & \text{if } V_n^S - V^U < c(a_i) \end{cases} \tag{9}$$

where V_n^S is the expected value of being a skilled worker for a child living in neighborhood n and V^U is the expected value of being an unskilled worker and given are, respectively, by the following

$$V_n^S = \sum_{k=1}^{N} \underbrace{\Gamma\left(1 + \frac{1}{\theta}\right) \mathbb{E}_{w_{i,s}}[U(c_{i,s}, \alpha_{s,k})|\mathcal{I}_n]}_{\text{Expected utility from living in neighborhood } k} \quad \rho_{s,k} \tag{10}$$

$$V^U = \sum_{k=1}^{N} \qquad \underbrace{\Gamma\left(1 + \frac{1}{\theta}\right) U(c_{i,u}, \alpha_{u,k})}_{\text{Expected utility from living in neighborhood } k} \qquad \rho_{u,k} \tag{11}$$

where $\Gamma\left(1 + \frac{1}{\theta}\right)$ is the expected value of the idiosyncratic component of utility and $\Gamma(.)$ the gamma function. $\mathbb{E}$ is the expectations operator and the expectation is taken over the skilled wage. $\rho_{s,n}$ and $\rho_{u,n}$ are the probability of living in neighborhood k conditional on being a skilled worker and the probability living in neighborhood k conditional on being an unskilled worker, respectively. I assume children are myopic in the sense that they not consider that their education decision will determine populations and land prices, so when computing expected well-being gap, they consider that they will pay the same rent as their parents.

Note that in contrast to the literature (Bowles et al. [2014] and Kim and Loury [2013]), the benefit of human capital investment is not simply the expected wage gap. Instead, in this framework, the benefit of investing in education is defined by the expected well-being gap, which takes into account differences in amenities and rental prices payed across different skill groups. This is important because skilled 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. For instance, Diamond [2016] finds that from 1980 to 2000 changes in cities's rents and amenities increased inequality between college and high-school graduates by more than the increase suggested by the wage gap alone.

Since the skill acquisition cost is decreasing in ability, the child's investment decision takes the form of a cut-off rule: there is an ability threshold, a_n^*, which above which children decide to invest in education. This threshold is defined by the following indifference condition

$$V_n^S - V^U = c(a_n^*) \tag{12}$$

Given this threshold, for an individual born to a family of type j living in neighborhood n, the probability of investing in education is then given by equation 13.

$$s_n = 1 - \mathcal{G}(a_n^*) \tag{13}$$

where a_n^* is the ability threshold for all children living in neighborhood n. Importantly, the decision of the younger generation to invest in human capital is not directly linking to their family's background. Rather it is linked to the socioeconomic environment of their neighborhood: all children living in same neighborhood, with an ability level higher than a_n^* invest in education independently of their parents' type. The intuition for this result lies on the fact that the driver for the investment decision is the information endowment at the neighborhood level, namely, the public signal ω_n. All children living in the same neighborhood share the same information set, hence all children with $a > a_n^*$ invest in education while all children with $a < a_n^*$ do not invest in education, independently of their parents' education. In this model, I abstract from any direct source of parent-child correlation, such as cultural or genetic traits, or any economic linkage mechanism across generations, be it bequest or human capital externalities in the skill acquisition cost or technology. Instead, parents affect their children through the place where they decide to live. Parents' residential location determines the neighborhood's skill-mix, hence the child's information endowment, which in turn will determine his investment on education.

Comparative statics. Taking the expectations over the unknown wage, $w_{i,s}$, the expected well-being gap for a child born in neighborhood n, $\triangle V_n$, conditional on the information set, is given by equation 14

$$\triangle V = N\Gamma\Big(1 + \frac{1}{\theta}\Big)\left(\frac{-\exp(-\gamma\big(\hat{\mu}_n - \gamma(\hat{\sigma}_n^2/2))}{\sum_{k=1}^{N}\frac{\alpha_{s,k}}{\exp(\gamma R_k)}} - \frac{-\exp(-\gamma w_u)}{\sum_{k=1}^{N}\frac{\alpha_{u,k}}{\exp(\gamma R_k)}}\right) \tag{14}$$

The key variable that drives the optimal investment threshold, a_n^*, is beliefs. Given equation 14 and equation 13, I establish two main properties of the schooling decision:

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

Naturally, the higher is the expected value of $w_{i,s}$, $\hat{mu}$, the higher is the expected net benefit of the investment, therefore the higher is the probability of investing in education. In turn, more uncertainty about the wage as a skilled worker (higher $\hat{\sigma}_n^2$) translates into a lower expected benefit of the educational investment, hence a higher ability threshold and a lower probability of investing, holding all else equal. The intuition behind this result lies on the fact that more uncertainty about education returns makes education investment more risky. As I consider individuals to be risk-averse, the probability of investment is lower.[13]

What role do the skilled neighbors play in the investment decision? This answer is not straightforward as they have two opposing effects on educational investment. Proposition 1 below establishes the relationship between the children's optimal decision and the share of skilled neighbors living in the neighborhood. A higher share of highly educated neighbors, $l_{s,n}$, means that children observe more a precise signal ($\sigma_{\omega_n}^2$ is lower), thus have lower uncertainty about the returns to education. Because children are risk averse, lower uncertainty associated with human capital investment translates into a widening of the mass of children that invest in education widens (a_n^* decreases). I call this effect the "uncertainty effect". As a reaction to a more precise signal, children place a higher weight on the labor market information disclosed by their neighbors, ω_n (equation 7). This implies that those children with more information (a high $l_{s,n}$) update their beliefs in response to signals to a greater extent than those that have less information. However, having more information does not necessarily imply a higher perception about the education value. This will depend on the the

[13]See Appendix B for details.

size of the public signal relative to an information threshold, ω^*. Having more information (a high $l_{s,n}$) has a positive effect on the share of children investing in education (s_n) if and only if it the perceived returns to education increase. Otherwise, the effect is negative as children are more certain that the returns to education are lower than what they previously estimated. I call this the "threshold effect". Overall, skilled neighbors have a positive effect on the children's optimal decision if the threshold effect is positive or if the uncertainty effect compensates a negative threshold effect.

Proposition 1. *There is an information threshold, ω^*, such that the ability threshold a_n^* is strictly decreasing (increasing) in the share of skilled parents living in n, $l_{s,n}$, if $\omega_n > (<)\omega^*$, where $\omega^* = \tilde{\mu} - \frac{\gamma\tilde{\sigma}^2}{2}$. Under these conditions, the share of children investing in education, s_n, is strictly increasing (decreasing) in $l_{s,n}$.*

The urban equilibrium is then key for the child's investment decision. The configuration of the city, namely, the distribution of skilled parents across neighborhoods shapes the public signal ω_n children receive, hence their perception about the well-being gap between skilled and unskilled workers. Hence, the skill-mix of neighborhoods and the education decision of children are connected through an information channel. Within this setting, the effect of neighborhoods in the education decision is not only about more exposure to skilled neighbors, as suggested by the literature, it is also about the wage distribution of these skilled neighbors. More exposure implies more information but more information does not necessarily increase the probability of investing in education.

2.3 Discussion of assumptions

I make several assumptions that make my model simpler without affecting its qualitative results. First, 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]). Allowing for endogenous amenities would complicate the model, without affecting the main mechanism I have described. When introducing endogenous amenities[14], the skill sorting within the city would be intensified but the role of skilled neighbors through social learning would remain as previously described.

Second, I consider that when deciding where to live within the city, parents do not take into account the neighborhood externalities in the education decision of their children through social learning. However, this would also only affect the sorting pattern within the city (in particular, it would intensify the isolation of skilled families) without changing the main implications of information frictions and social learning for the education decision of children.

Finally, one important feature of the model is the fact that the signal's precision increases in the share of skilled neighbors (lemma 1). This result comes from the assumption that the information gathered by each skilled neighbor is proportional to the size of the neighborhood, i.e., the larger is the neighborhood, the noisier is the signal of each skilled individual. Qualitatively, this assumption does not affect the most important result in the paper (proposition 1). Without this assumption, the signal's precision would increase in the number of skilled neighbors and not in the share of skilled neighbors. Hence, one should re-write proposition 1 in the following way:

Proposition 1 (re-written). *There is an information threshold, ω^*, such that the ability threshold a_n^* is strictly decreasing (increasing) in the* ***number of skilled parents*** *living in n, $L_{s,n}$, if $\omega_n > (<)\omega^*$, where $\omega^* = \mu_0 - \frac{\gamma\sigma_0^2}{2}$. Under these conditions, the probability of investing in education s_n is strictly increasing (decreasing) in $L_{s,n}$.*[15]

[14]For instance, 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 skilled workers in the neighborhood, which would be a good proxy for man-made features of amenities such as restaurants, schools, among others)

[15]Lemma 1 and 2 should also be re-written in this context by replacing the share of skilled neighbors with the number of skilled neighbors, meaning that the signal's uncertainty decreases with the number of skilled neighbors while the weight on the public signal decreases.

This completes the discussion of my model. In the next section, I confront its predictions with the data looking to education decisions of high-school graduates in Michigan, namely the decision of enrolling in college. In section 4, I estimate the model and assess the quantitative importance of information frictions and social learning as channel through which neighborhoods affect the decision to enroll in college.

3 Empirical evidence on the model's prediction

This section presents evidence supporting the role played by information frictions and social learning in linking children educational outcomes to the skill composition of neighborhoods. To explore my model's prediction, I define a neighborhood to be a school-district and I study college enrollment by high-school graduates in Michigan over the period 2008-2014.[16]

3.1 Data

For the empirical analysis, I use school-district level data from Michigan combining three sources of information.

College enrollment data comes from the Michigan Department of Education (CEPI). I measure college enrollment as the share of high-school graduates that enroll in a 4-year college within 6 months in a given school-district. This data also provides with information on student counts, cohort characteristics and average ACT score at the school-district level.

Socio-demographics data comes from the Education Demographic and Geographic Estimates of the National Center for Education Statistics (NCES - EDGE). This data has

[16] The most commonly definition of a neighborhood is a census tract, which have generally a size between 1200 and 8000 people. I define a neighborhood to be school-district because I do not have data on college enrollment by high-school graduates at the census tract level, only high-school dropouts. School-districts tend to be relatively larger, however, in my sample, 41% of school-districts have a population lower than 8000.

rich information on socio-economic conditions of school-districts such as racial composition, family median income and unemployment rate. More important, this data has two variables closely tied to my theoretical model. First, it has data on median annual earnings by education level, hence I observe the median earnings of individuals with a bachelor degree and with more than 4 years of college. This corresponds to the signal (ω_n) that in my model children observe when inferring schooling returns. Second, it has data on the education level of population over 25 years old which allows me to calculate the share of population over 25 with a college degree. This is a good proxy for the signal's precision ($l_{s,n}\sigma_\epsilon^2$).

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 by source (state, federal and local), this data has information on K-12 enrollment and number of teachers.

My dataset is an unbalanced panel of urban school-districts covering the period 2008-2014, which only includes public high-schools. Table 1 summarizes the basic statistics on the different variables I use in my analysis. Before proceeding to the empirical strategy, I take a first look to the raw data. I divided the sample of school-districts into quartiles of the following distributions: (*i*) share of college graduates, (*ii*) the median annual earnings of individuals with bachelor degree and (*iii*) the local revenue per student. Panel A in Figure 2 plots the average of college enrollment in each quartile of the distribution of the share of college graduates. The dashed line represents school-districts in the last quartile of the distribution of local revenue per student, these are school-districts with a high level of school resources, and the solid line corresponds to school-districts in the first quartile of the same distribution, hence includes school-districts with a low resources. One can see that as the share of college graduates (signal's precision) increases, the average of college enrollment increases. However, there seems to be no significant difference in average college

Table 1: Summary statistics

	Observations	Mean	Std. Dev	Min.	Max.
Main variables					
college enrollment (%)	1845	0.33	0.14	0.05	0.80
college graduates (%)	1849	0.23	0.13	0.04	0.79
annual median earnings, bachelor (log)	1849	10.76	0.21	9.41	11.36
annual median earnings, graduates (log)	1846	11.03	0.22	9.65	11.56
Socio-demographic variables					
median family income (log)	1849	11.04	0.28	9.92	11.89
median earnings, high sch graduates (log)	1849	10.20	0.15	9.48	10.64
black residents (%)	1837	0.08	0.15	0.00	0.93
white residents (%)	1837	0.86	0.16	0.04	1.00
unemployment rate (%)	1849	0.11	0.04	0.03	0.36
total population (log)	1837	9.78	0.90	7.67	13.73
Cohort variables					
average ACT score	1849	19.10	2.01	12.23	25.93
females (%)	1843	0.51	0.05	0.33	0.75
School quality variables					
expenditure per student (dollars)	1564	11162.52	2397.67	7945.48	30499.21
local revenue per student (dollars)	1564	3539.53	2056.57	813.51	24536.86
teachers to student ratio (dollars)	1564	0.05	0.01	0.03	0.09

Notes: Observations are at the school-district level and cover all the years from 2008 to 2014. *College enrollment* is measured as the number of high-school graduates that enroll in a 4-year college within 6 months after graduation as a % of high-school graduates in the school-district. *College graduates* is the % of population over 25 years old with 4 or more years of college. Black and white residents are measured as a % of total population in the school-district. The *unemployment rate* is the number of unemployed as a percentage of the civilian labor force. *Females* is the number of high-school graduates that are females as a % of high-school graduates in the school-district. Source: CEPI, NCES-EDGE and NCES-CCD.

enrollment between school-districts with low school resources and school-districts with high-school resources. Panel B in figure 2 plots this difference and the respective 95% confidence interval. One can see that for the first and second quartile the difference is not significantly different from zero. For the third quartile, the difference is significantly different from 0, but is negative which goes against models of local funding of education devolped by Bénabou

[1996a], Bénabou [1996b], Férnandez and Rogerson [1996] and Durlauf [1996]. This evidence suggests that college enrollment is not driven by school resources, but instead seems driven by the presence of college graduates in the school-district. Hence theories that focuses on the role of social exposure to educated individuals as the mechanism through which neighborhoods affect educational outcomes seem to provide a better explanation for the observed pattern.

Panel C also plots average enrollment in each quartile of the share of college graduates' distribution. However, the dashed line now corresponds to school-districts in the first quartile of the median annual earnings of college graduates distribution (these are locations where children observe a high signal about skilled wage), and the solid line represents school-districts in the last quartile of this distribution. Panel D plots the difference between the dashed and the solid line and its 95% confidence intervals. When the share of college graduates is low, the difference in college enrollment between locations where college graduates have higher earnings and locations where median earnings of college graduates are low is not significantly different from 0. From this, we can conclude that the decision to enroll in college does not seem driven by resource constraints: whether the earnings are high or low, when there are few college graduates, college enrollment is low. As the share of college graduates increases, the difference between both groups of school-districts widens substantially. In particular, when the share of college graduates is very high, school-districts where college graduates have high earnings have an average college enrollment that is higher by 10 percentage points when compared to school-districts where college graduates have low earnings. This difference is significantly different from 0. This evidence seems in line with the social learning mechanism proposed in section refsec:section3 As high-school graduates have more labor market information, i.e. a higher share of college graduates, they not only are less uncertain about college benefits but they also place more weight on the public signal they observe. Assuming this signal is higher than the students' prior, on average, then in places where college graduates earn more, the effect of exposure to individuals with a college degree on college enrollments is higher, as predicted by the model.

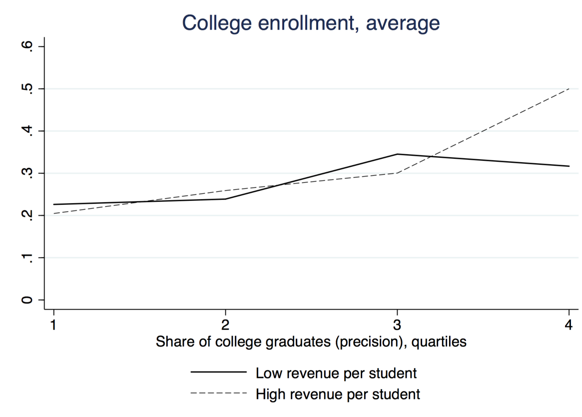
Panel E shows a similar story by plotting the average college enrollment for each quartile of the median earnings distribution. Now, the dashed line corresponds to school-districts with a high share of college graduates and the solid line to school-districts with a low share of college graduates. One can see that in this last group, the average college enrollment is constant across quartiles of earnings. This evidence supports the mechanism proposed here: when students have low information about labor market conditions, i.e. low signal's precision, increased signal size, meaning the income of college graduates, doest have much effect on college enrollment as students put relatively more weight on their prior, not generating an incremental update about college benefit.

3.2 Empirical framework

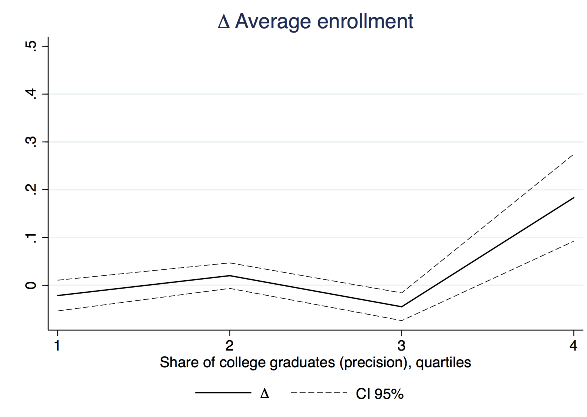
To formely test my model's prediction, I estimate the following

$$E_{ijt} = \alpha + \beta_1 X_{ijt} + \beta_2 X_{ijt} Y_{ijt} + \beta_3 Y_{ijt} + \beta_4 Z_{ijt} + \mu_j + \gamma_t + \varepsilon_{ijt} \tag{15}$$

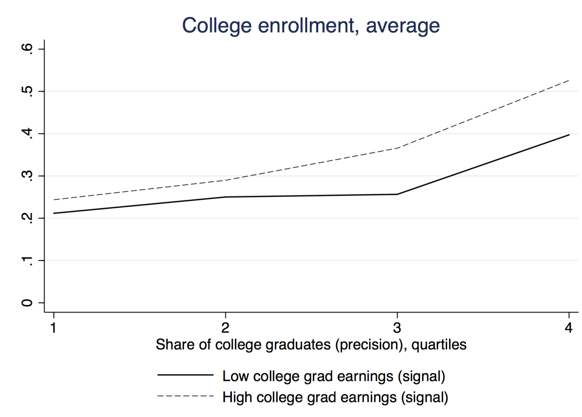
where the dependent variable, E_{ijt}, is the share of high-school graduates that enroll in a 4-year college within 6 months of graduation in school-district i that belongs to city j on year t. X_{ijt} is the share of individuals over 25 years old with a college education (bachelor degree and more) and Y_{ijt} represents median annual earnings of college graduates. Z_{ijt} is a vector of controls. I control for the socio-demographic conditions at the school-district level using the share of black and white residents, the median family income, the unemployment rate, total population and the median earnings of high-school graduates. I also control for the characteristics of the cohort that graduated from high-school that year by including the share of females in the 12th grade and the average ACT score of the graduating class. Because since 2007 the ACT test is mandatory for all high-school students in Michigan, the



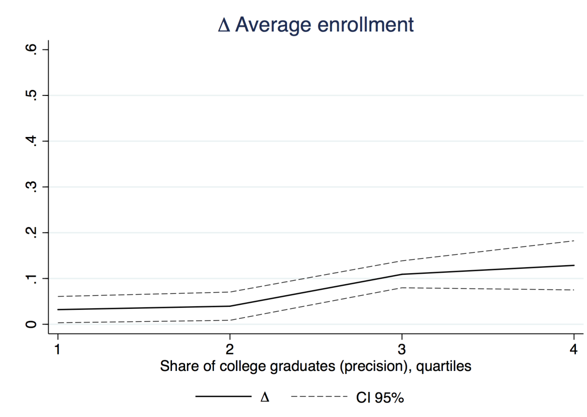
Panel A: Share of college graduates, quartiles



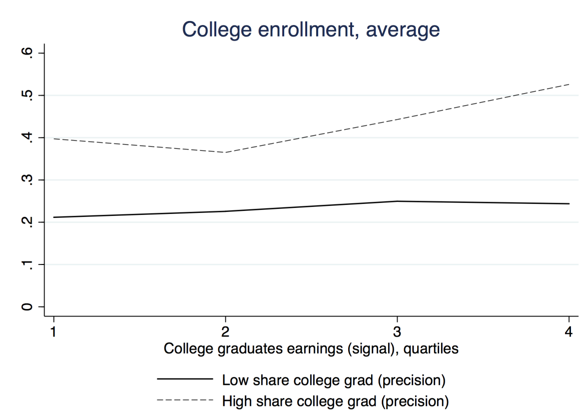
Panel B: Difference btw dashed and solid lines, quartiles



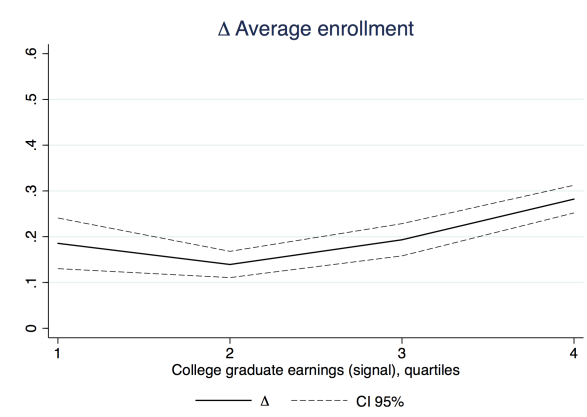
Panel C: Share of college graduates (σ_n^2), quartiles



Panel D: Difference btw dashed and solid lines, quartiles



Panel E: Median earnings, bachelor (ω_n), quartiles



Panel F: Difference btw dashed and solid lines, quartiles

Note: Panels A, C and E plot the average college enrollment in 4-year colleges for each quartile of a given distribution. College enrollment is measured as the % of high-school graduates that enroll in 4-year colleges within 6 months after graduation. In each oh these panels, the dashed line represents school-districts in the last quartile of the distribution refereed in the graph's legend and the solid line represents school-districts in the first quartile of the same distribution. Panels B, D and F plot the difference between the dashed and the solid line of the panel on their left. The dashed lines in these panels represent bootstrap confidence intervals at the 95% level. Source: CEPI, NCES-EDGE and NCES-CCD and author's calculations (2008-2013).

Figure 2: College enrollment in 4-year colleges (average)

average of the ACT score is a good proxy for the ability of high-school graduates. Including this variable allows me to control for the fact that it may be the case that highly educated parents have children with higher ability, hence more likely to enroll in college. Finally, the vector Z_{ijt} also includes school quality proxied by expenditures, revenues and teachers per student. In the US, residential location is closely tied to schools in the neighborhood, for instance Bayer et al. [2007] find that high income are willing to pay higher housing costs for better schools. Hence, the coefficient on X_{ijt} may be capturing the effect of better schools instead of an exposure effect to highly educated people. Including school quality measures allows me to control for this. μ_j and γ_t are MSA and year fixed-effect, respectively, and ε_{ijt} is an error term. I cluster standard errors at the school-district level.

The coefficients of interest are β_1, the coefficient on the share of college graduates, and β_2, the coefficient on the interaction of college graduates with median annual earnings of college graduates. Their sign will guide us on what theory seems to prevail in the data. While models that focus on the role of social networks and identity predict the marginal effect of X_{ijt} on E_{ijt} to be positive and independent of income: $\frac{\partial E_{ijt}}{\partial X_{ijt}} = \beta_1 > 0$ and $\beta_2 = 0$, the social learning mechanism proposed in this work predicts that the marginal effect of X_{ijt} on E_{ijt} can be positive or negative depending on the level of income: $\frac{\partial E_{ijt}}{\partial X_{ijt}} = \beta_1 + \beta_2 Y_{ijt}$ with $\beta_1 < 0$ and $\beta_2 > 0$ such that there is threshold of Y_{ijt}, $Y^* = \frac{\beta_1}{\beta_2}$, below which the marginal effect is negative.

3.3 Results

I start by presenting a parsimonious model where I only include as covariates in equation 15 city and year fixed effects. I then go on introducing demographic characteristics of the school-district and characteristics of the 12th grade. The results from these models are reported in table 2 column 1 and 2, respectively. The coefficient on the variable *college graduates* is positive and statistically significant: an increase in the share of college graduates by 10 percentage points increases college enrollment by 4.06 percentage points. Even though I control for the cohort ability using the average ACT score of the class, part of this effect can be masked by the effects of school quality. As suggested by Bénabou [1996a], Bénabou [1996b], Férnandez and Rogerson [1996] and Durlauf [1996], it could be that in school-districts with more college graduates, local funding is higher, children attend to better schools hence are more likely to enroll in college. I address this issue by estimating a version of the model in equation 15 where I include three proxies for school quality: expenditures, local revenue and teachers per student. Column 3 presents estimation results of this model. Not only the coefficient estimates of school quality are always small and statistically insignificant, the coefficient estimate of *college graduates* is also very similar to the one reported in column (2). It appears then that the role played by neighborhoods goes beyond the school resource mechanism.

In column 4, I introduce the interaction term. The coefficient on the variable *college graduates* becomes negative while the coefficient on the interaction is positive and both coefficients are statistically significant. This means that the marginal effect of *college graduates* is increasing in Y_{ijt}, hence there is a median earnings threshold, Y^*, above which the the marginal effect of college graduates on college enrollment is positive and below which its negative. These results are consistent my model's prediction. The role of skilled neighbors in education decisions depends on their wage distribution: only if earnings of college graduates are sufficiently high do we observe a positive relation between the share of college graduates in the neighborhood and education investment. To the best of my knowledge, this threshold effect is novel and suggests that the effect of neighborhood composition is heterogeneous, depending namely on the earnings distribution of individuals.[17]

Robustness. As a robustness check, I also estimate equation 15 with two different proxies

[17] Appendix table C.1 shows estimated coefficients on all included variables.

Table 2: College enrollment and college graduates: Regression Evidence

	(1)	(2)	(3)	(4)	(5)
college graduates	0.776***	0.389***	0.399***	-4.675***	-4.643***
	(0.029)	(0.050)	(0.055)	(1.224)	(1.220)
expenditure per student			0.00367	0.00279	0.00235
			(0.005)	(0.005)	(0.005)
local revenue per student			0.00204	0.00410	0.00529
			(0.005)	(0.005)	(0.005)
teachers to student ratio			-0.00178	-0.00423	-0.00384
			(0.006)	(0.006)	(0.006)
college graduates × median earnings, bachelor				0.470***	0.466***
				(0.113)	(0.113)
Observations	1838	1824	1548	1548	1548
R-squared	0.708	0.801	0.803	0.811	0.825
Demographic controls	N	Y	Y	Y	Y
Cohort controls	N	Y	Y	Y	Y
Year FE	Y	Y	Y	Y	N
MSA FE	Y	Y	Y	Y	N
MSA-year FE	N	N	N	N	Y

Notes: Each column 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. Column 2 to 5 control for demographic characteristics of the school-district (the share of black and white residents, the unemployment rate, the median family income, school-district size and earnings of high-school graduates) and characteristics of the graduating class (the share of females among the high-school graduates and the average ACT score). Column 3 to 5 also include school quality controls. All columns include MSA and year fixed effects. Significance: *** p<0.01, ** p<0.05, * p<0.1. Source: CEPI, NCES-EDGE and NCES-CCD. (2008-2014)

for the earnings of college graduates and two alternative samples. Using instead the median annual earnings of individuals with a post-graduate degree (column 1 in Appendix table C.2) and the average between this measure and the median annual earnings of individuals with a bachelor degree (column 2 in Appendix table C.2) leaves the coefficients of interest relatively unchanged.

In all specifications in table 2, I use the full sample of urban school-districts. However, in Michigan there is a school choice program under which families can opt to move their children out of the schools they would attend by residency to neighboring districts. To the check the robustness of the results to the inclusion/exclusion of students that do not live in the school-district they attend, I re-estimate equation 15 focusing only on school-districts which have a low share of non-resident students attending the 12th grade (I fix the share threshold at 10%). Appendix table C.2 reports the results. Overall, the results previously found hold when we exclude from the analysis school-districts with a higher share of non-resident students attending 12th grade. More important, the coefficients estimates are relatively similar.

In addition to this, the sample used to estimate equation 15 covers the period 2008-2014, which includes the period of the Great Recession. According to the National Bureau of Economic Research, the recession began in December 2007 and ended in June 2009. To check if my results are driven by the Great Recession, I re-estimate 15 restricting the sample to the years after the Great Recession (2010-2013) and I find that the results are robust to the Great Recession. Appendix table C.2 presents the results.

Discussion. In interpreting these results, I have so far abstracted from the important issue of location choice and the influence that selection might have on the results. Sorting by individuals may mean that individuals with different education choose to locate in systematically different neighborhoods, whose characteristics might lead to different decisions regarding college enrollment of their kids. Even though I control for school quality, there might still be other unobserved characteristics that affect both sorting of individuals and college enrollment. In light of this issue, one should interpret the evidence presented here as suggestive and descriptive of social learning as the channel linking neighborhoods and educational outcomes. Importantly, to be the best of my knowledge, this is the first paper to uncover a threshold effect on neighborhood effects, suggesting that the effect of being around skilled neighbors on human capital formation of high-school graduates depends on their wage distribution.

4 Quantitative analysis

Even though I show evidence consistent with the model's prediction, it remains an open question the importance of social learning as channel through which neighborhoods affect the education decision of young adults, in particular, the decision to enroll in college. To take up this debate, in this section I estimate the model presented in section 2 and use it to quantify the spillover effects from college graduates to high-school graduates through social learning. I calibrate the model in order to match 2013 data regarding college enrollment, defined as the % of high-school graduates that enroll in a 4-year college within 6 months, at the school-district level in Detroit. I choose Detroit as its the largest metropolitan area in Michigan, with 95 school-districts in 2013, which I define to be a neighborhood in the model. I proceed in the following steps. First, I set parameters that have a counterpart in the data using data from the American Community Survey (ACS) for 2008-2013. Second, given data on the distribution of labor force across neighborhoods ($l_{s,n}$ and $l_{u,n}$) and rents paid (R_n) in each neighborhood, I recover values of skilled and unskilled amenities ($\alpha_{s,n}$ and $\alpha_{u,n}$), under the assumption that the current allocation of the labor force across neighborhoods is an equilibrium. Third, I use simulated method of moments to estimate the remaining parameters. Finally, I perform counterfactual analysys.

4.1 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 $C(a_i) = \overline{c} - \phi a_i^{\varphi}$.

4.2 Calibration strategy

To explore the quantitative predictions of the theory presented previously, I calibrate the economy to reproduce some key statistics regarding 4-year college enrollment in Detroit in 2013. Calibration is proceed in two steps. First, I select all parameters that either have a direct counterpart in the data or that have been used in previous literature. Then I estimate the rest of the parameters with the simulated method of moments.

First step. The number of neighborhoods equals the number of school-districts in Detroit in 2013. I set $\overline{a}$ equal to one and $\underline{a}$ to zero, also I consider that skilled individuals only care about skilled amenities and unskilled individuals only care about unskilled amenities: $\beta_s = 1$ and $\beta_u = 0$. As in Babock, Choi and Feinerman (1993), I set the risk aversion parameter, γ, to 0.5. Following Férnandez [2013] and Fogli and Veldkamp [2011], I calibrate beliefs about the skilled wage and the unskilled wage using individual level data from the 2013 ACS. The sample consists of individuals aged between 25 and 55 years working at least 35 hours per week. 48 weeks per year. A skilled worker is defined as a worker with at least a 4-year college degree. All other workers are considered to be unskilled. I use information on wage, namely the log-monthly wage, to calibrate (i) unskilled wage, w_u, (ii) the common component, ω and (iii) prior's mean and variance, $\tilde{\mu}$, $\tilde{\sigma}^2$. In particular, I set the unskilled wage to 7.83, which is the average of the log-monthly wage for unskilled workers in Detroit. Then, ω and $\tilde{\mu}$ are set to 8.46 and 8.54, which are the average log-monthly wage for skilled workers at the national level and in Detroit, respectively. Finally, $\tilde{\sigma}^2$ is equal to the standard deviation of the log-monthly wage for skilled workers. in Detroit.

I recovered the distribution of skilled and unskilled amenities ($\alpha_{s,n}$ and $\alpha_{u,n}$) across the city from the data. Using data from NCES-EDGE, I observe for each school-district, which correspond to a neighborhood in the model, total population (L_n), the number of skilled and unskilled parents, $L_{s,n}$ and $L_{u,n}$, and rents (R_n). 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.[18] For any two neighborhoods n and k, the following holds

$$\frac{L_{j,n}}{L_{j,k}} = \frac{\alpha_{j,n}}{exp(\gamma R_n)} \frac{exp(\gamma R_k)}{\alpha_{j,k}}$$

where $L_{j,n}$ is the number of type j-families that live in n. I set $\alpha_{j,1} = 1$ for Detroit City school-district, and then calculate the level of amenities for the other school-districts.

Second step. The parameters left to be estimated are the following: σ_ϵ, $\bar{c}$, φ and ϕ. To estimate these parameters, I match the fraction of high-school graduates that enroll in 4-year colleges within 6 months at the school-district level, hence I use a total of 95 moments. The method of moments picks the parameter vector θ=(σ_ϵ, $\bar{c}$, φ, ϕ) that minimizes the weighted sum of square deviations between data moments and their model-generated counterpart:

$$\hat{\theta} = argmin\Big((y_i - \hat{y}_i)' W (y_i - \hat{y}_i)\Big)$$

where W is the identity matrix, implying that each moment is equally weighted, y_i represents the observed fraction of high-school graduates that enroll in college in neighborhood i, and $\hat{y}_i$ is the predicted value of this fraction by the model.

Model fit. Table 3 summarizes all parameters. Column 1 in table 4 shows the correlation between the observed and fitted values for college enrollment by high-school graduates. The relatively high correlation indicates a reasonably good fit for the data. Perhaps more critical for counterfactuals is whether the model can fit moments that are not target in the estimation. The remaining columns in table 4 show correlations between the matched variable,

[18] 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, therefore I use the reported median gross rent in the NCES-EDGE dataset.

Table 3: Baseline model parameters

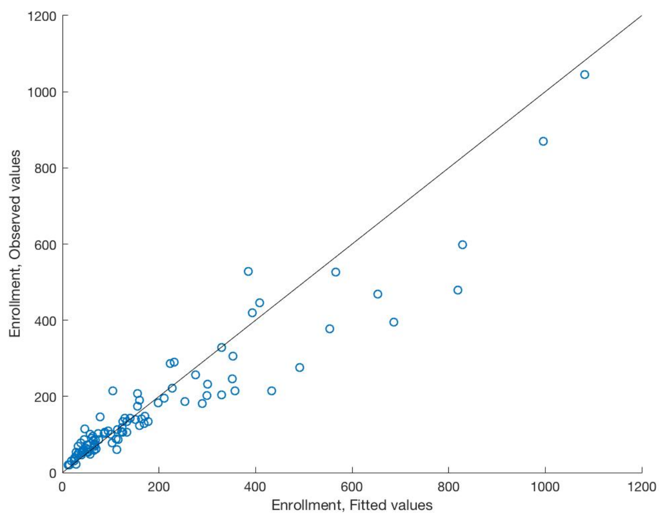
Description	Parameter	Value
Panel A: Independently chosen		
# neighborhoods	N	95
risk-aversion (CARA)	γ	0.5
unskilled wage	w_u	7.83
skilled wage, common component	ω	8.46
prior's mean	$\tilde{\mu}$	8.54
prior's variance	$\tilde{\sigma}^2$	0.64
Panel B: Estimated		
cost function parameter	$\overline{c}$	8.98
cost function parameter	φ	1.34
cost function parameter	ϕ	2.56
variance, idiosyncratic term	σ^2_ϵ	0.03

college enrollment, and other variables, both for the observed and fitted values. The correlations for fitted values, which are moments not targeted in the estimation, resemble the actual correlations reasonably well. Figure 3 depicts the observed and fitted values for the number of high-school graduates that enroll in college and also shows a high correlation between fitted and predicted values. To assess whether these results are sensitive to functional forms or the way I calibrate some parameters, I re-estimated the model using a variety of different specifications. These results are reported in Appendix table C.3. To sum up, the ability of the model to fit the data seems robust to alternative parameters (β_j, $\overline{a}$ and θ) and an alternative distribution for innate ability, in particular I considered that innate ability followed a pareto distribution. Overall, I view the evidence presented here as indicative that the model successfully captures patterns in the data.

Table 4: Goodness of fit: fitted versus observed values

	Enrollment	College grad	Rents	HS dropouts
enrollment, (%, observed data)	1.00	0.84	0.51	-0.54
enrollment, (%, fitted data)	0.78	0.89	0.55	-0.41

Notes: Observations are at the school-district level. Column 1 refers to the correlation between predicted and observed values of college enrollment as measured by the % of high-school graduates that enroll in a 4-ear college within 6 months of graduation. Column 2 to 4 display the correlation between three variables and college enrollment for the observed and fitted values. *College grad* is the % of individuals over 25 years old with at least a 4 year college degree. *Rents* is the median gross rent at the school-district. *HS dropouts* is the % of dropouts in high-school. Observations: 95 school-districts in Detroit.



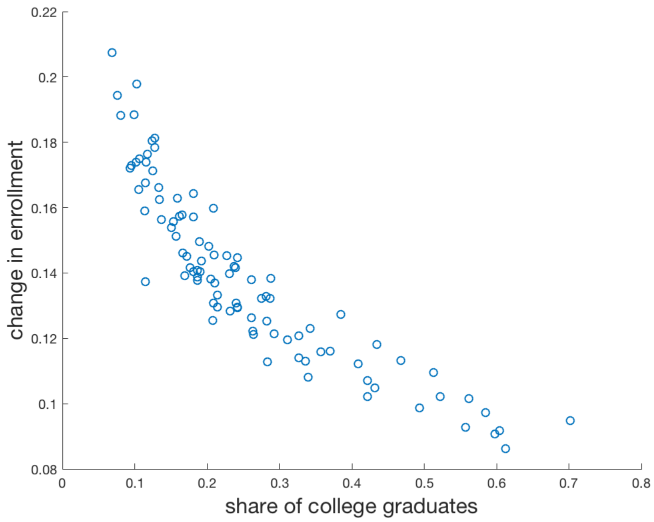
Notes: This graph plots fitted and observed values for the number of high-school graduates that enroll in a 4-year college degree 6 months after graduation. The correlation is . Fitted values are on the horizontal axis; observed values are on the vertical axis.

Figure 3: Goodness of fit: fitted versus observed values

4.3 Counterfactual analysis

Given the calibrated and estimated parameters, I now simulate what would happen in two alternative scenarios to evaluate the impact on college enrollment of (i) information frictions and (ii) neighbors' skills through the social learning mechanism. I first ask: do information frictions play a role in college enrollment? To answer this question, I maintain the baseline estimation but assume there is no uncertainty by making the skilled wage (w_s) known. I find that if there was no uncertainty about the skilled wage, college enrollment would increase by 13% at the city level. Estimated gains are, however, heterogeneous across neighborhoods as depicted in figure 4, ranging between 20.7% and 8.6% depending on the information endowment at the neighborhood level. In locations with a high share of college enrollment, uncertainty about the skilled wage is lower to start with hence information frictions do not play such an important role in determining college enrollment as in neighborhoods where the share of highly educated individuals is low. In these neighborhoods, information about the earnings of skilled workers is limited, hence uncertainty is higher, which translates into a higher impact of information frictions on the % of high-school graduates that enroll in 4-year colleges.

To understand the importance for college enrollment of neighbors' skills through the social learning mechanism, I keep the baseline calibration but shut down entirely the learning channel, by assuming that individuals do not update their prior's mean or variance. I find that while in the model with social learning I am able to explain 91% of the % of college enrollment at the city level, in a model with no social learning I can only explain 78%, it seems then to be an important feature of the model in order to replicate the data. How much of college enrollment does social learning account for? Social learning mitigates underlying uncertainty about the skilled wage by 81%, on average. This translates into an improvement in college enrollment by 14.5% at the city level. Nonetheless, this aggregate impact hides important differences across neighborhoods.

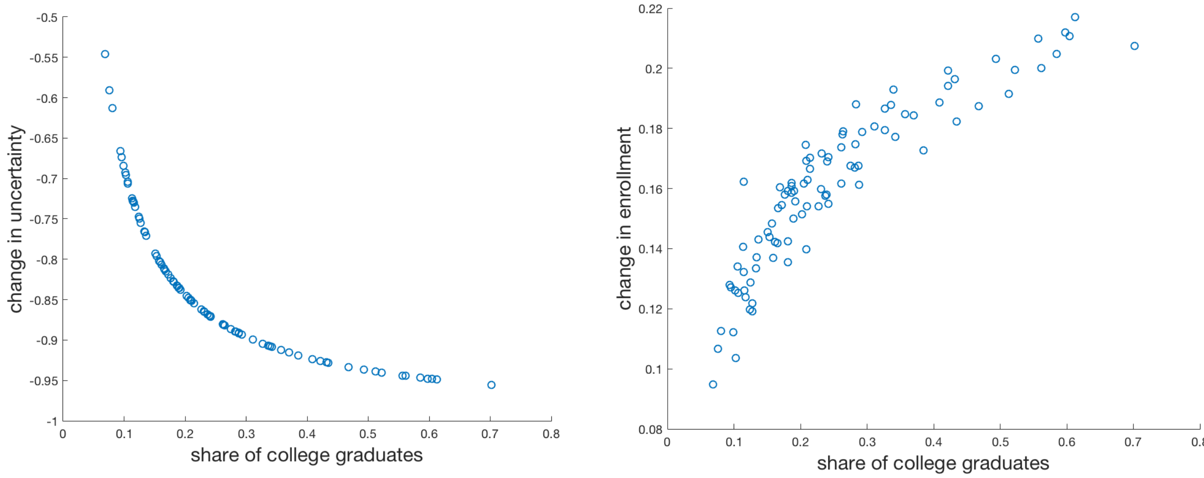


Notes: On the horizontal axis is the share of college graduates living the neighborhood; on the vertical axis is the change on college enrollment, which shows how much college enrollment would change if there was no uncertainty about the skilled wage. Change on college enrollment is measured as the % difference college enrollment in the model without uncertainty about the skilled wage (I maintain the baseline estimation but consider the skilled wage w_s to be known) and college enrollment in the baseline model.

Figure 4: The effect of uncertainty on college enrollment across neighborhoods

The left panel in figure 5 depicts the % change in uncertainty across neighborhoods when I shut down the social learning mechanism. I measure the change in uncertainty as the difference between the posterior variance in the baseline model, $\hat{\sigma}_n^2$, and the posterior variance in the model without social learning, which is equal to the prior's variance, $\tilde{\sigma}_n^2$ $\left(\frac{\Delta\sigma_n^2}{\tilde{\sigma}^2},\right.$ where $\left.\Delta\sigma_n^2 = \hat{\sigma}_n^2 - \tilde{\sigma}^2\right)$. The right panel in figure 5 depicts % change in college enrollment in each neighborhood due to social learning. I measure the % change in college enrollment for each neighborhood as the difference between college enrollment in the baseline model and college enrollment in the model without social learning.

From figure 5, one can see that neighborhoods with a higher share of college graduates experience a higher decrease in the posterior variance relative to the prior variance. This result follows from bayesian updating: skilled neighbors generate more information about the returns to educational investment, decreasing uncertainty and risk associated with col-

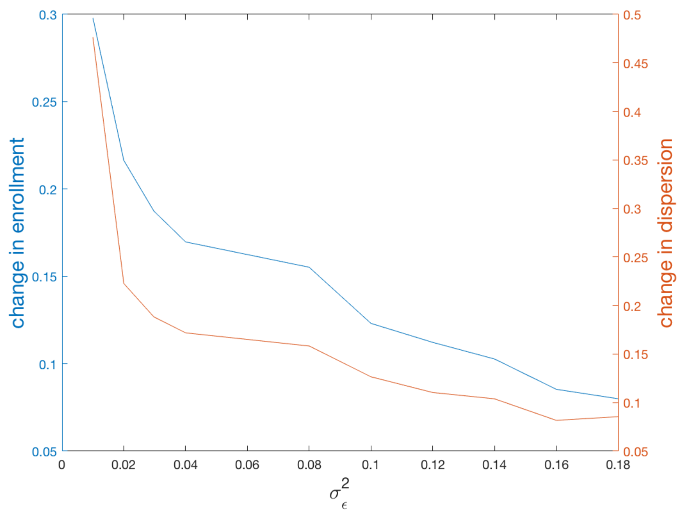


Notes: The left panel describes change in uncertainty across neighborhoods due to social learning. The right panel describes change in college enrollment across neighborhoods due to social learning. Change in uncertainty (college enrollment) is measured as the % difference between uncertainty (college enrollment) in the model baseline model and uncerainty (college enrollment) without social learning (I maintain the baseline estimation but assume that individuals do not update their prior about the skilled wage).

Figure 5: Effects of social learning across neighborhoods

lege enrollment. Lower uncertainty translates into higher college enrollment. Hence, social learning accounts for a larger % of college enrollment in neighborhoods with a larger information endowment, i.e., with a higher share of college graduates. Recall that the model presented in section 2 predicts that a decrease in uncertainty due to a higher share of college graduates may increase or decrease the share of high-school graduates that invest in education. The direction of this effect depends on the level of the signal relative to the prior. In the case of Detroit, it can be shown that for all neighborhoods $\omega_n > \omega^*$, which implies a positive relationship between the share of college graduates and the fraction of high-school graduates that enroll in college. Given this, it comes as no surprise that the neighborhoods more affected by social learning are the ones with a higher information endowment: a higher share of college graduates not only decreases uncertainty but it also increases the weight on the public signal, since the signal is higher then the prior this leads to a higher perception of the education value, hence higher investment in education. These heterogeneous effects across neighborhoods translate into an increase dispersion in college enrollment across neighborhoods by 14.1%. So, even though social learning means less uncertainty about college returns hence more college enrollment at the city level, it also increases inequalities across neighborhoods because the decrease in uncertainty is uneven within the city

The magnitude of the aggregate impact of social learning is sensitive to σ_ϵ. Figure 6 shows how the aggregate effect of social learning changes across different values of σ_ϵ: higher values of σ_ϵ translate into lower decreases in uncertainty, on average, which in turn reflect into lower increases in college enrollment and lower increases in dispersion of college enrollment across neighborhoods.



Notes: On the horizontal axis is the variance of idiosyncratic term; on the vertical axis are the change in college enrollment at the city level (left axis) and the change in dispersion of college enrollment within the city (right axis) due to social learning. Change in college enrollment (dispersion) is measured as the % difference between college enrollment (dispersion) in the model baseline model and college enrollment (dispersion) in the model without social learning (I maintain the baseline estimation but assume that individuals do not update their prior about the skilled wage).

Figure 6: The aggregate effect of social learning

4.4 Extension

I augment the model with uncertainty about the unskilled wage. As the skilled wage, I consider that the unskilled wage is now the sum of a common component (ω_u) and an idiosyncratic term, $\epsilon_{i,u}$ that is normally distributed with zero mean and variance equal to $\sigma^2_{\epsilon_{i,u}}$. Children also share a common prior about the unskilled wage, which follows a normal distribution with mean $\tilde{\mu}_u$ and variance $\tilde{\sigma}_u$. As before, children update their priors using Bayes' rule, but now they observe two public signals one about the skilled wage ($\omega_{n,s}$) and another about the unskilled wage ($\omega_{n,u}$):

$$\omega_{n,s} \equiv \frac{1}{L_{s,n}} \sum_{i=1}^{L_{s,n}} \omega_{i,s} = \omega_s + \varepsilon_{\omega_{n,s}}, \quad \varepsilon_{\omega_{n,s}} \sim \mathcal{N}(0, l_{s,n}^{-1} \sigma^2_{\epsilon_s}) \qquad \text{(signal about } w_{i,s}) \tag{16}$$

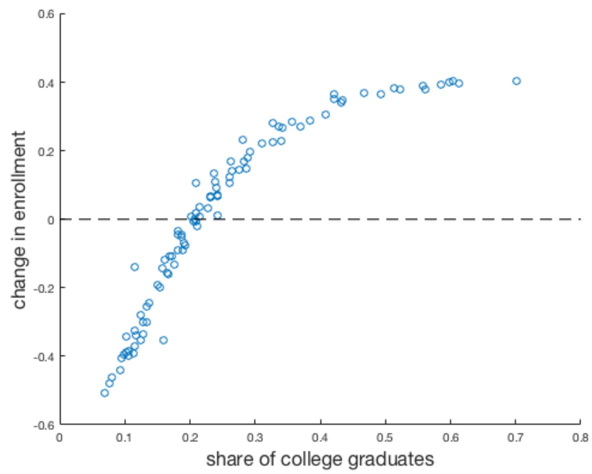
$$\omega_{n,u} \equiv \frac{1}{L_{u,n}} \sum_{i=1}^{L_{u,n}} \omega_{i,u} = \omega_u + \varepsilon_{\omega_{n,u}}, \quad \varepsilon_{\omega_{n,u}} \sim \mathcal{N}(0, l_{u,n}^{-1} \sigma^2_{\epsilon_u}) \qquad \text{(signal about } w_{i,u}) \tag{17}$$

while the precision of the signal about the skilled wage increases in the share of skilled parents living in the neighborhood, the precision of the signal about the unskilled wage increases with the share of unskilled parents in the neighborhood. Children born to low skilled families that live isolated from skilled families will have more information about the unskilled wage, reducing uncertainty and risk associated with no investment in education. On the other hand, children living around skilled families have low uncertainty about the skilled wage, but at the same time have very little information regarding being an unskilled worker.

I calibrate this model in order to match 2013 data about college enrollment in Detroit, defined as the % of high-school graduates that enroll in a 4-year college within 6 months of graduation. I calibrate beliefs about the unskilled wage using the same strategy as the one used to calibrate beliefs about the skilled wage. I also use the simulated method of moments to estimate the remaining parameters, which are the same as before plus ϵ_u. The augmented model also fits the data reasonably well as shown in Appendix table C.4.

Given the baseline parameters, I simulate the model shutting down entirely the social learning mechanism. Social learning decreases, on average, uncertainty about the skilled and the unskilled wage by 30.5% and 23.5%, respectively. This translates into an increase in college enrollment by 8.3% at the city level.[19] As before the the role of neighbors through social learning is heterogeneous across neighborhoods, but now the effect is different not only in magnitude but also in sign. Figure 7 depicts the % change in college enrollment across different neighborhoods. In locations with a sufficiently low share of college graduates, learning about wages from neighbors translates into a decrease in college enrollment while in neighborhoods with a high share of college graduates social learning increases college enrollment. The intuition is the following. In neighborhoods with a low share of college graduates, children learn more about the unskilled wage in comparison to the skilled wage: the information endowment about the unskilled wage is higher, hence children in these neighborhoods are very little uncertain about the unskilled wage but are very uncertainty about the skilled wage which leads to more children choosing to become unskilled workers. As a consequence, dispersion in college enrollment across neighborhoods increases by 14.6%.

[19]Recall that I measure the change in uncertainty as the difference between the posterior variance in the baseline model and the posterior variance in the model without social learning, which is equal to the prior's variance. Also, I measure the % change in college enrollment for each neighborhood as the difference between college enrollment in the baseline model and college enrollment in the model without social learning



Notes: On the horizontal axis is the share of college graduates in the neighborhood; on the vertical axis is the change in college enrollment due to social learning. Change in college enrollment is measured as the % difference between college enrollment in the extended model and college enrollment in the extended model without social learning (I maintain the baseline estimation but assume that individuals do not update their prior about the skilled wage).

Figure 7: Effects of social learning across neighborhoods

5 Conclusion

This paper introduces uncertainty at the investment stage and social learning in a model of human capital formation with local interactions. I show that the effect of the exposure of skilled neighbors in schooling investment within a neighborhood will depend on the labor market information they disclose, in particular, their wage distribution. If the information from skilled neighbors leads to an increase in the perceived college premium, then overall schooling investment increases. Otherwise, it decreases.

I study college enrollment by high-school graduates in Michigan over the period 2008-2013, and find empirical evidence supporting the proposed mechanism: the share of college graduates increases college enrollment if median earnings of these college graduates is sufficiently high. Because in my empirical analysis I abstract completely from the issue of location, I calibrate the model to match facts on college enrollment in Detroit in 2013. In a quantitative exercise, I find that social learning decreases uncertainty about schooling returns, boosting college enrollment. However, it also increases inequalities across neighborhoods.

On on-going work, I augment the present model and introduce other channels proposed in the literature: (i) local financing of education and (ii) social networks. Parents choose where to live, knowing that they will vote in a public good, children decide whether to invest in education or not. The difference here is that the cost of investing in education will depend on the public good, as a proxy for education expenditures, peer effects and innate ability. Within this model, neighbors affect education decision in three ways (i) public good, (ii) peer effects and (iii) social learning. I plan to calibrate the model to a metropolitan area in the US and understand the importance of these three mechanisms. This analysis will allows to better understand the black box of neighborhood effects.

References

T. Allen. Information frictions in trade. *Econometrica*, 82(6):2041–2083, 2014.

P. Arcidiacono, E. Aucejo, A. Maurel, and T. Ransom. College attrition and the dynamics of information revelation. *Working paper*, 2016.

G. Arkelof and R. Kranton. Economics and identity. *Quarterly Journal of Economics*, 115 (3):715–753, 2000.

O. Attanasio and K. Kaufmann. Education choices and returns to schooling: Intrahousehold decision making, gender and subjective expectations. *Journal of Development Economics*, 109:203–216, 2014.

I. Baley, L. Veldkamp, and M. Waugh. Information globalization, risk sharing, and international trade. *Working paper*, 2016.

P. Bayer, F. Ferreira, and R. McMillan. A unified framework for measuring preferences for schools and neighborhoods. *Journal of Political Economy*, 115(4):588–638, 2007.

C. Belfield, T. Boneva, C. Rauh, and J. Shaw. Money or fun? why students want to pursue further education. *Working paper*, 2016.

R. Bénabou. Workings of a city: Location, education and productions. *Quarterly Journal of Economics*, 108(3):619–652, 1993.

R. Bénabou. Heterogeneity, stratification and growth: Macroeconomic implications of community structure and school finance. *American Economic Review*, 86(3):584–609, 1996a.

R. Bénabou. Equity and effiency in human capital investment: The local connection. *Review of Economic Studies*, 63(2):237–264, 1996b.

Z. Bleemer and B. Zafat. Intended college attendance: Evidence from an experiment on college returns and costs. *Working paper*, 2015.

M. Bobba and V. Frisancho. Learning about oneself: The effects of signaling academic ability on school choice. *Working paper*, 2014.

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

P. Carneiro and J. Heckman. The evidence on credit constraints in post-secondary schooling. *The Economic Journal*, 112(482):705–734, 2002.

T. Cavalcanti and G. Gisnnitarou. Growth and human capital: A network approach. *Working paper*, 2013.

R. Chetty, N. Hendren, and L. Katz. The effects of exposure to better neighborhoods on children: New evidence from the moving to opportunity experiment. *American Economic Review*, 106(4):855–902, 2016.

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

R. Diamond. The determinants and welfare implications of us workers' diverging location choices by skill: 1980-2000. *American Economic Review*, 106(3):479–524, 2016.

S. Durlauf. A theory of persistent income inequality. *Journal of Economic Growth*, 1(1): 75–93, 1996.

S. Durlauf. Neighborhood effects. *Handbook of Regional and Urban Economics*, pages 2174–2230, 2004.

J. Eeckhout, R. Pinheiro, and K. Schmidheiny. Spatial sorting. *Journal of Political Economy*, 122(3):554–620, 2014.

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

R. Férnandez. Cultural change as learning: The evolution of female labor force participation over a century. *American Economic Review*, 103(1):472–500, 2013.

R. Férnandez and R. Rogerson. Income distribution, communities and the quality of public education. *Quarterly Journal of Economics*, 111(1):135–164, 1996.

A. Fogli and L. Veldkamp. Nature or nurture? learning and the geography of female labor force participation. *Econometrica*, 79(4):1103–1138, 2011.

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

A. Glitz. Ethnic segregation in germany. *Labour Economics*, 29:28–40, 2014.

C. Goldin and L. Katz. The race between education and technology. *Press for Harvard University Press.*, 2008.

J. Hellerstein and D. Neumark. Workplace segregation in the united states: Race, ethnicity, and skill. *The Review of Economics and Statistics*, 90(3):459–477, 2008.

C. Hoxby and C. Avery. The missing "one-offs": The hidden supply of high-achieving, low-income students. *Brookings Papers on Economic Activity*, 2014.

C. Hoxby and S. Turner. Expanding college opportunities for high-achieving, low income students. *Working paper*, 2014.

C. Hoxby and S. Turner. What high-achieving low-income students know about college. *American Economic Review*, 105(5):514–517, 2015.

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

L. Katz and K. Murphy. Changes in relative wages, 1963-1987: Supply and demand factors.the. *The Quarterly Journal of Economics*, 107(1):35–78, 1992.

K. Kaufmann. Understanding the income gradient in college attendance in mexico: The role of heterogeneity in expected returns. *Quantitative Economics*, forthcoming, 2014.

Y. Kim and G. Loury. Social externalities, overlap and the poverty trap. *Journal of Economic Inequality*, 12(4):535–554, 2013.

J. Kunz and K. Staub. Subjective completion beliefs and the demand for post-secondary education. *ECON - Working Papers 218, Department of Economics - University of Zurich*, 2016.

D. McFadden. Conditional logit analyisis of qualitative choice behavior. *Frontiers in Econometrics*, ed. P. Zarembka:105–142, 1973.

R. Stinebrickner and T. R Stinebrickner. A major in science? initial beliefs and final outcomes for college major and dropout. *The Review of Economic Studies*, 81(1):426–472, 2014.

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

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

F. Yamuchi. Social learning, neighborhood effects, and investment in human capital: Evidence from green-revolution india. *Journal of Development Economics*, 83:37–62, 2007.

Appendix A

From the US census data, I built a dataset that combined population counts for educational attainment, race and high-school dropouts at the census tract level for the years 1980, 1990 and 2000. Using data on educational attainment and race, I measure to what extent individuals with different educational attainment and different race live apart from each other. For this purpose, I use a segregation index used by Glitz [2014] and Hellerstein and Neumark [2008]. This index is based on the exposure index and isolation index. While the latter measures the extent to which members of one particular group are only exposed to each other, the former measures the exposure of members of one particular group to members of another group.

In my analysis, I divided the population over 25 years old into four groups: non-college graduates (which aggregates educational attainment categories "no high-school", "high-school" and "some college"), college graduates, composed by individuals with 4 or more years of college, white and non-white individuals. Apart from defining the relevant population groups, measuring exposure and isolation requires a definition of neighborhood: I computed both indices within a MSA using census tracts, which have on average 4,000 residents, as neighborhoods. Given this, the exposure index in MSA m is

$$E_m = \sum_{i=1}^{N} \frac{x_i}{X_m} \frac{y_i}{p_i}$$

where N is the number of census tracts in MSA m, x_i and y_i are, respectively, the number of non-college (non-white) and college graduates (white individuals) in census tract i, p_i is the number of people living in census tract i and X_m is the number of non-college graduates (non-white individuals) living in MSA m. E_m is the % of college graduates (white individuals) living in the census tract where the average non-college graduate (non-white individual) lives. On the other hand, the isolation index for college graduates and white individuals is given by

$$I_m = \sum_{i=1}^{N} \frac{x_i}{X_m} \frac{x_i}{p_i}$$

where N is the number of census tracts in MSA m, x_i is the number of college graduates (white individuals) in census tract i, p_i is the number of people living in census tract i and X_m is the number of college graduates (white individuals) living in MSA m. I_m is the % of college graduates (white individuals) living in the census tract where the average college graduate (white individual) lives.

I focus on a segregation measure, S_m, which is the difference between I_m and E_m, $S_m = I_m - E_m$. This measure can be interepreted the following way: in comparison to the average non-college graduate, the average college graduate lives in a neighborhood where the % of college graduates is higher by Sm percentage points. If all college graduates live in the same neighborhood, then $I_m = 1$, $E_m = 0$ and $S_m = 1$, and the two groups are fully segregated. In contrast, if the percentage of neighbors that are college-graduates are the same for the average college graduate and the average non-college graduate, then I_m=E_m and $S_m = 0$, and there is no skill segregation.

Using this measure, I find that, in comparison to non-college graduates, college graduates live in neighborhoods where the % of college graduates is 10 percentage points higher, on average.

Appendix B

Proof of Lemma 3. For a child born in neighborhood n, the perceived value of being skilled, V_n^S, is given by

$$
\begin{aligned}
V_n^S = \sum_{k=1}^{N} & \underbrace{\Gamma\left(1+\frac{1}{\theta}\right)\mathbb{E}_{w_{i,s}}[U(c_{i,s},\alpha_{i,k})|\mathcal{I}_n]}_{\text{Expected utility from living in neighborhood } k} \quad \rho_{s,k} = \\
&= \sum_{k=1}^{N}\Gamma\left(1+\frac{1}{\theta}\right)\mathbb{E}_{w_{i,s}}\Big[\frac{-\exp(-\gamma(w_{i,s}-R_k))}{\alpha_{s,k}}|\mathcal{I}_n\Big]\rho_{s,n} = \\
&= \sum_{k=1}^{N}\Gamma\left(1+\frac{1}{\theta}\right)\mathbb{E}_{w_{i,s}}\Big[-\exp(-\gamma w_{i,s})|\mathcal{I}_n\Big]\frac{\exp(\gamma R_k)}{\alpha_{s,k}}\rho_{s,k} = \\
&= \sum_{k=1}^{N}\Gamma\left(1+\frac{1}{\theta}\right)\Big[-\exp\Big(-\gamma\Big(\hat{\mu}_n-\gamma\Big(\hat{\sigma}_n^2/2\Big)\Big)\Big]\frac{\exp(\gamma R_n)}{\alpha_{s,k}}\rho_{s,k} = \\
&= -\exp(-\gamma\Big(\hat{\mu}_n-\gamma(\hat{\sigma}_n^2/2))\left(\sum_{k=1}^{N}\Gamma\left(1+\frac{1}{\theta}\right)\frac{\exp(\gamma R_k)}{\alpha_{s,k}}\rho_{s,k}\right) = \\
&= -\exp(-\gamma\Big(\hat{\mu}_n-\gamma(\hat{\sigma}_n^2/2))\left(\frac{N\Gamma\Big(1+\frac{1}{\theta}\Big)}{\sum_{k=1}^{N}\frac{\alpha_{s,k}}{\exp(\gamma R_k)}}\right)
\end{aligned}
$$

The expected value of being unskilled, V^U, which is equal for all children, is given by

$$
\begin{aligned}
V^U = \sum_{k=1}^{N} & \underbrace{\Gamma\left(1+\frac{1}{\theta}\right)U(c_{i,u},\alpha_{i,k})}_{\text{Expected utility from living in neighborhood } k} \quad \rho_{u,k} = \\
&= -\exp(-\gamma w_u\sum_{k=1}^{N}\Gamma\left(1+\frac{1}{\theta}\right)\Big[\frac{\exp(\gamma R_k)}{\alpha_{u,n}}\Big]\rho_{u,k} = \\
&= -\exp(-\gamma w_u)\left(\frac{N\Gamma\Big(1+\frac{1}{\theta}\Big)}{\sum_{k=1}^{N}\frac{\alpha_{u,k}}{\exp(\gamma R_k)}}\right)
\end{aligned}
$$

Given V_n^S and V^U, the perceived well-being gap for a child born in neighborhood n, $\triangle V_n \equiv$

$V_n^S - V^U$, is given by

$$\triangle V_n = N\Gamma\Big(1+\frac{1}{\theta}\Big)\left(\frac{-\exp(-\gamma\Big(\hat{\mu}_n - \gamma(\hat{\sigma}_n^2/2))}{\sum_{k=1}^{N}\frac{\alpha_{s,k}}{\exp(\gamma R_k)}} - \frac{-\exp(-\gamma w_u)}{\sum_{k=1}^{N}\frac{\alpha_{u,k}}{\exp(\gamma R_k)}}\right)$$

where n indexes the neighborhood where the child lives, N is the number of neighborhoods in the city and term $\Gamma\Big(1+\frac{1}{\theta}\Big)$ is the expected value of the idiosyncratic component of utility with $\Gamma(.)$ being the gamma function. The optimal investment decision takes the form of a cut-off rule. The ability cut-off, a_n^*, is definied by the indiference condition $\triangle V_n = c(a_n^*)$. Defining $\varpi_n = \triangle V_n - c(a_n^*)$, I establish the following:

1. $\frac{\partial s_n}{\partial \hat{\mu}_n} > 0$. The effect of $\hat{\mu}_n$ on the probabilit of becoming a skilled worker, s_n is given by

$$\frac{\partial s_n}{\partial \hat{\mu}_n} = \frac{\frac{\partial s_n}{\partial a_n^*}}{\frac{\partial a_n^*}{\partial \hat{\mu}_n}}$$

   By the implicit function theorem, $\frac{\partial a_n^*}{\partial \hat{\mu}_n} = -\frac{\frac{\partial \varpi}{\partial a_n^*}}{\frac{\partial \varpi}{\partial \hat{\mu}_n}} < 0$. Since $\frac{\partial s_n}{\partial a_n^*} < 0$ and $\frac{\partial a_n^*}{\partial \hat{\mu}_n} < 0$, we can conclude that $\frac{\partial s_n}{\partial \hat{\mu}_n} > 0$.

2. $\frac{\partial s_n}{\partial \hat{\sigma}_n^2} < 0$. The effect of $\hat{\sigma}_n^2$ on the probabilit of becoming a skilled worker, s_n is given by

$$\frac{\partial s_n}{\partial \hat{\sigma}_n^2} = \frac{\frac{\partial s_n}{\partial a_n^*}}{\frac{\partial a_n^*}{\partial \hat{\sigma}_n^2}}$$

   By the implicit function theorem, $\frac{\partial a_n^*}{\partial \hat{\sigma}_n^2} = -\frac{\frac{\partial \varpi}{\partial a_n^*}}{\frac{\partial \varpi}{\partial \hat{\sigma}_n^2}} > 0$. Since $\frac{\partial s_n}{\partial a_n^*} < 0$ and $\frac{\partial a_n^*}{\partial \hat{\sigma}_n^2} > 0$, we can conlude that $\frac{\partial s_n}{\partial \hat{\sigma}_n^2} < 0$.

Appendix C

Table C.1: Regression results: all coeficients

	(1)	(2)	(3)	(4)	(5)
college graduates	0.776***	0.389***	0.399***	-4.675***	-4.643***
	(0.029)	(0.050)	(0.055)	(1.224)	(1.220)
median family income		0.0444	0.0507*	0.0237	0.0223
		(0.028)	(0.030)	(0.029)	(0.030)
black residents		0.207***	0.197**	0.215***	0.205***
		(0.078)	(0.081)	(0.075)	(0.075)
white residents		0.121	0.111	0.147**	0.135*
		(0.075)	(0.078)	(0.074)	(0.074)
unemployment rate		0.437***	0.438***	0.342***	0.328***
		(0.113)	(0.117)	(0.115)	(0.118)
average ACT score		0.0378***	0.0360***	0.0338***	0.0341***
		(0.003)	(0.003)	(0.003)	(0.003)
females		-0.00597	0.0213	0.0232	0.0285
		(0.045)	(0.049)	(0.047)	(0.047)
total population (log)		0.00260	0.00218	0.00443	0.00410
		(0.004)	(0.004)	(0.004)	(0.004)
expenditure per student			0.00367	0.00279	0.00235
			(0.005)	(0.005)	(0.005)
local revenue per student			0.00204	0.00410	0.00529
			(0.005)	(0.005)	(0.005)
teachers to student ratio			-0.00178	-0.00423	-0.00384
			(0.006)	(0.006)	(0.006)
college graduates × median earnings, bachelor				0.470***	0.466***
				(0.113)	(0.113)
median earnings, bachelor				-0.0443*	-0.0467*
				(0.026)	(0.027)
median earnings, high sch graduates				-0.0238	-0.0163
				(0.023)	(0.023)
Observations	1838	1824	1548	1548	1548
R-squared	0.708	0.801	0.803	0.811	0.825

Notes: Each column 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. Columns 1 to 4 include MSA and year fixed effects, column 5 includes MSA-year fixed effects. Significance: *** p<0.01, ** p<0.05, * p<0.1. Source: CEPI, NCES-EDGE and NCES-CCD.

Table C.2: Robustness checks

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
college graduates	-4.715***	-4.415***	0.481***	-4.091**	-5.064***	-4.608**	-4.608**	-5.118***	-4.608**
	(1.408)	(1.502)	(0.072)	(1.578)	(1.872)	(1.949)	(1.949)	(1.621)	(1.949)
college graduates × median earnings, post-graduate	0.460***				0.496***				
	(0.127)				(0.170)				
college graduates × median earnings, average		0.437***				0.462**	0.462**	0.502***	0.462**
		(0.136)				(0.179)	(0.179)	(0.147)	(0.179)
college graduates × median earnings, bachelor				0.420***					
				(0.147)					
Observations	1546	1546	701	701	700	700	700	1153	700
R-squared	0.814	0.814	0.791	0.797	0.800	0.800	0.800	0.791	0.800
Demographic controls	Y	Y	Y	Y	Y	Y	Y	Y	Y
Cohort controls	Y	Y	Y	Y	Y	Y	Y	Y	Y
Year FE	Y	Y	Y	Y	Y	Y	Y	Y	Y
MSA FE	Y	Y	Y	Y	Y	Y	Y	Y	Y

Notes: Each column 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. Column 1 to 2 use different earnings measures. From column 3 to 7, I restrict the sample to school-districts with less than 10% of non-resident students. Column 8 and 9 report estimation results using only the post-Great Recession years (2010-2014). Column 8 uses all the sample of urban school-districts, while column 7 only uses urban school-districts with less than 10% of non-resident students. Demographic controls include the share of black and white residents, the unemployment rate, the median family income, total population and earnings of high-school graduates. Cohort controls include the share of females among the high-school graduates and the average ACT score of the 12th grade. School quality controls include teacher to student ratio, expenditure and local revenue per student. All columns include MSA and year fixed effects. Significance: *** p<0.01, ** p<0.05, * p<0.1. Source: CEPI, NCES-EDGE and NCES-CCD.

Table C.3: Goodness of fit: fitted versus observed values

	Enrollment	College grad	Rents	HS dropouts
Panel A: observed data				
enrollment (%)	1.00	0.84	0.51	-0.54
Panel B: fitted data				
enrollement, baseline	0.78	0.89	0.55	-0.41
enrollement, estimated $\overline{a}$ & $\underline{a}$	0.79	0.90	0.57	-0.44
enrollement, pareto distribution	0.79	0.95	0.62	-0.42
enrollement, θ=0.5	0.51	0.63	0.46	-0.23

Notes: Observations are at the school-district level. Panel A displays observed correlations between college enrollment, measured by the % of high-school graduates that enroll in a 4-year college within 6 months of graduation, and college graduates, rents and high-school dropouts. Panel B displays correlations between these three variables and fitted values for college enrollment under four different calibration strategies. Row 1 in Panel B reports the model fit under the baseline calibration. In Panel B Row 2, instead of exogenously setting the support of the innate ability distribution, I included in the estimator vector $\overline{a}$ and $\underline{a}$. Row 3 in Panel B shows the godness of fit for a model in which innate ability is assumed to be drawn from a Pareto distribution with support $[1, \infty)$, shape parameter k, scale parameter z and c.d.f $I(a) = 1 - (\frac{z}{a})^k$, z and k are parameters estimated using the simulated method of moments. Row 4 used the baseline calibration but sets $\theta = 0.5$. *College grad* is the % of individuals over 25 years old with at least a 4 year college degree. *Rents* is the median gross rent at the school-district. *HS dropouts* is the % of dropouts in high-school. Observations: 95 school-districts in Detroit.

Table C.4: Goodness of fit: fitted versus observed values

Correlations	Enrollment	College grad	Rents	HS dropouts
Panel A: observed data				
enrollment (%)	1.00	0.84	0.51	-0.54
Panel B: fitted data				
enrollment, baseline	0.78	0.89	0.55	-0.41
enrollment, extension	0.75	0.90	0.57	-0.49

Notes: Observations are at the school-district level. Panel A displays observed correlations between college enrollment, measured as the % of high-school graduates that enroll in a 4-year college within 6 months of graduation, and college graduates, rents and high-school dropouts. Panel B displays correlations between these three variables and fitted values for college enrollment under the baseline and the extended model. Row 1 in Panel B reports the model fit under the baseline calibration. Row 2 reports the model fit for the extended model. *College grad* is the % of individuals over 25 years old with at least a 4 year college degree. *Rents* is the median gross rent at the school-district. *HS dropouts* is the % of dropouts in high-school. Observations: 95 school-districts in Detroit.