ACADEMIC PRESS

Journal of Urban Economics 52 (2002) 582–607

JOURNAL OF Urban Economics

www.academicpress.com

Young workers, learning, and agglomerations

Giovanni Peri

Department of Economics, University of California Davis, One Shields Avenue, Davis, CA 95616, USA

Received 18 January 2001; revised 18 July 2002

Abstract

Since the 90s densely populated locations, such as urban areas, have attracted a disproportionate share of young college-educated workers. We explain this stylized fact as a result of colocation due to learning externalities among educated workers. Workers learn from each other when young, increasing their skills and their productivity. Once they grow mature and their learning decreases, some of them choose to move out of densely populated areas. As skills grew more transferable thanks to computerization and flexibility in the 80s and 90s, urban areas became learning grounds for educated young workers.

JEL classification: R1; R3; R23; J3

Keywords: Learning; Cities; Experience premium; Skills; Agglomerations

1. Introduction

Agglomeration externalities are the centripetal forces that generate and maintain high density of population and production in small areas such as urban regions or cities. Several of these externalities are probably at work and have been analyzed by urban and regional economists. Forward and backward linkages and market proximity generate, in the presence of transport costs, agglomeration externalities (Krugman [20], Krugman and Venables [21], Fujita et al. [11]). Economies of scale and economies of scope generate agglomeration externalities

E-mail address: gperi@ucdavis.edu.

0094-1190/02/$ – see front matter
PII: S0094-1190(02)00510-7

(Abdel Raman and Fujita [1]). Pooling of different workers and firms (Dumais et al. [8]) or search and matching in local labor markets (Helsley and Strange [15]) generate agglomeration externalities. In this article we focus on local learning as a source of agglomeration externalities and we derive some qualitative and quantitative implications from a model that provides interesting insight and explains some stylized empirical facts.

We have several reasons to focus on "local learning" as a source of agglomeration externalities. First and foremost, learning by doing and learning from interactions have been considered important determinants of workers' skills and productivity by several economists in their theoretical (such as Arrow [3], Jovanovic and Rob [17], and Lucas [23]) as well as in their empirical analyses (such as Audretsch and Feldman [2], Foster and Rosenzweig [10], Glaeser and Mare [14], and Jaffe et al. [16]). Working in a local environment that promotes interactions with educated and skilled people generates valuable learning opportunities for workers. "Great are the advantages which people... get from near neighborhood to one another" Marshall[1] says, and Lucas[2] asks: "What can people be paying Manhattan or downtown Chicago rents for, if not for being near other people?"

However, in spite of the acknowledged importance of learning externalities, few regional or urban economists have formally modeled them in relation to location of workers and migration into and out of densely populated areas. While there are some theoretical models analyzing local human capital externalities (such as Black and Henderson [4]) and some empirical papers trying to measure them (such as Rauch [25]) none of them addresses explicitly the issue of workers' learning over time. The only model addressing exactly this issue is Glaeser [13]. While the model he uses is rather different from ours, several implications of that model are similar to ours. Crucially, though, that model does not analyze the role of skills' transferability in determining the localization of mature workers. Duranton and Puga [9] also consider the role of cities in promoting learning, but they focus on learning about new processes to produce goods. They emphasize the role of cities as "nursery" for new firms.

Finally, a reason to focus on learning externalities is that stylized empirical analysis shows that densely populated locations, such as urban areas, appeal more to young educated workers than to anybody else. Young educated workers receive a lower wage premium than their older colleagues to work in urban areas but, in spite of this, they are largely over-represented there. It is reasonable to think that young workers are trading off the current disadvantages of being in cities (high rent, lower real wages) for the advantages of learning and increasing their future productivity and income.

---

[1] The quote is from A. Marshall, Principles of Economics, London, 1890.

[2] The quote is from Lucas [22].

Our model focuses on learning externalities as the source of agglomeration economies in densely populated areas. We explicitly analyze the location decision of young and mature educated workers to predict where they would work. The other sources of potential agglomeration externalities (specialization, transport costs, increasing returns) are eliminated from the model to keep it simple and to focus on learning externalities alone. While learning externalities are simply assumed in our model, we derive several interesting implications on wages, experience premia, migration behavior and density of educated workers that match the stylized facts from data on US cities. The model is kept as parsimonious as possible. Its solution still relies on simulation techniques (just as most recent economic geography models à la Krugman [20]) but we derive analytically several intermediate results. Using the model we are also able to fully understand the channels and mechanisms through which centripetal and centrifugal forces operate.

The key insight of our model concerns the choice of location of educated workers. To keep their choice simple we consider only two locations, similar to what most of the "new economic geography" models (à la Krugman [20]) do. The model is general enough that we can interpret the high-density location as a city, or a metropolitan area or simply as a highly populated region and the low density area as the countryside, the hinterland of a city or simply as a scarcely populated region.

The model has three different types of equilibria that we may suggestively associate to three stages of urban development. Each equilibrium is the only locally stable one within a certain range of values of two key parameters: the intensity of learning and the transferability of skills. Assuming that learning from interactions exists, the first crucial issue is how intense it is. We can think of "the intensity of learning" as the frequency of interactions, or as the intensity of learning from each interaction. Such intensity has to be large enough to even generate an equilibrium with agglomeration. As productive tasks become more complicated, the potentiality and the intensity of learning from interactions increases. Weak learning externalities generate a disperse equilibrium with both locations equally populated: when benefits from learning are too small, the negative crowding effects of living in a dense location cause maximum dispersion of workers. We call this outcome, that represents a dispersed, little urbanized economy, type I equilibrium.

An increase in the "intensity of learning" generates at some point benefits from colocating as educated people learn from each other. Young workers would then work in the dense location. The decision of where to work as they turn mature, though, depends on the second key parameter: the transferability of their skills. If skills learned in a location are not easily transferred because they are specific to the location (i.e., to the establishment or to the firm or to the city) where they have been accumulated, then moving as a mature worker causes a large loss of skills and of productivity. In this case, even when the learning period is over, mature workers stay in dense areas. In spite of crowding effects and no benefits from learning they would be worse off if they moved because of the loss in productivity. We call this equilibrium, featuring high intensity of learning and low transferability of skills, type II equilibrium. Companies with specific technologies, firm-specific machinery, no standardization, no common language, no largely available computer software generate very specific skills. In this situation type II equilibrium is likely to prevail. Up to the 1960s such was the general structure of production. As a consequence urban areas had high concentration of young and mature workers alike.

Finally, as skills become more transferrable because of increased generality of technology, availability of computers, development of common procedures and languages, mature workers move out of dense locations as soon as the learning period is over. Young workers are still attracted to dense locations. Some mature workers though move to less dense areas where, thanks to the skills learned, they enjoy higher wages and lower rents. We call this last scenario, characterized by high intensity of learning and high transferability of skills, type III equilibrium. This picture broadly fits the characteristics of post 1980s production, based on smaller, largely computerized, firms that emphasize general purpose skills. In line with the predictions of our model, since the eighties, urban areas have been the learning laboratories for young workers, while mature workers have moved increasingly to less dense areas outside cities or in the hinterland.

The paper is organized as follows. Section 2 provides stylized facts on wages, experience premia, concentration of educated workers and migration in 1970 and 1990. In Section 3 we describe the model. In Section 4 we analyze the conditions, characteristics and properties of the three types of steady state equilibria. Section 5 discusses the results, provides some measures of skills' transferability and concludes.

2. Stylized facts on productivity and learning

It is useful to present some stylized facts about college-educated workers, their location and their wages. We derive them by analyzing the Public Use Microdata Samples of the US Census for year 1970 and 1990. While some of these facts were already pointed out in previous studies, it is useful to summarize all of them here. We present in Table 1 the qualitative features of the data that our model should match.

Table 1 describes facts relative to college-educated workers only. We focus on them because they are the workers who have highest learning potentialities from interactions and they are the more mobile workers. Stylized facts are relative to 1970 and 1990. We report in regular fonts those facts that are unchanged between 1970 and 1990 and in bold fonts those that changed between the two years. Our classification distinguishes between densely populated areas (High density) and less dense areas (Low density). We also consider migration between them

Table 1
Stylized facts on college-educated workers

Area	1970	1990
High density	Higher concentration Higher experience premia Lower real wage for young **Balanced young-mature ratio**	Higher concentration Higher experience premia Lower real wage for young **High young-mature ratio**
Low density	Lower concentration Lower experience premia Higher real wage for young **Balanced young-mature ratio**	Lower concentration Lower experience premia Higher real wage for young **Low young-mature ratio**
Migrations	**Low**	**High**

The stylized facts are obtained from author's elaboration on PUMS 1970 and 1990 data. Details in the text and in Appendix A.

(Migrations). Here we discuss the stylized facts identifying urban areas as "high density" and non-urban areas as "low density." In Appendix A we show that the same stylized facts also apply if we choose a different comparison such as Megalopolis[3] versus all the other cities, or cities with high population density versus cities with low population density.

The first stylized fact is simply that in urban areas the percentage of college-educated workers is larger than in non-urban areas (25% vs. 20% in 1990 and 14% vs. 11% in 1970). This fact was already pointed out by Glaeser [13] and Glaeser and Mare [14] and is not specific to our analysis. The second stylized fact is that experience premia for the college-educated are larger in urban areas. Considering white males with college education, the increase in their wage between 10 and 30 years of experience was around 47% in urban areas and 41% in non-urban areas in 1990. Similarly that increase was 23 and 13% respectively in urban and non-urban areas in 1970.

The high-experience premium paid in dense areas implies (third stylized fact) that young workers receive in urban areas a smaller wage premium than required by their older colleagues. In 1990 a college-educated white male received a premium of $2 per hour in urban areas over the wage of a non-urban worker, while a mature worker received a premium of $4–5 over her non-urban colleague. Similar differences existed in 1970. This, together with the fact that prices are higher in cities implies that young educated workers paid a real wage premium to be in urban areas.

The fourth stylized fact is that urban areas in 1970 had the same young-mature composition of educated workers as non-urban areas. Conversely, in 1990 urban

[3] We define Megalopolis those Metropolitan Areas with more than two million employees in 1990. They are Los Angeles, New York, Chicago, Philadelphia, Boston, Detroit, and Washington.

areas exhibited significantly higher young-mature ratio of educated workers relative to non-urban areas. Specifically, the ratio of college-educated workers with less than 20 years of experience (young) to college-educated workers with more than 20 years of experience (mature) was 1.5 for urban and non-urban areas alike in 1970. That ratio had grown to 1.85 for non-urban areas by 1990, but for urban areas it had reached 2.12. This last fact was also pointed out by a recent study by Costa and Kahn [7]. They characterize this tendency as stronger for married couples than for singles and consider the "marriage market" for the young educated worker as the main reason to be in cities. That explanation is an interesting and valid alternative to ours. However, let us point out that even restricting the sample to unmarried white males, who should all be equally attracted to dense areas for marriage reasons, young workers are still over-represented in dense areas compared to mature workers (see Appendix A for details).

Finally, we consider migrations of educated workers as they begin their mature career (between 35 and 45 years old). While such group had low migration rates in 1965–1970, the rate was much higher in the 1985–1990 period. In our statistics we only consider white single males head of family. As we do not know if they migrated between a high- and low-density area, we simply count those who moved from a county to another. In order to assess if mobility of 35–45 year old college-educated workers has increased we compare it with mobility of high-school educated white males, 35 to 45 years old. While between 1965–1970 only 20% of the college-educated changed county of residence (same percentage as high-school workers who moved), between 1985 and 1990 51% of college-educated moved across counties (vs. only 33% of the high school-educated). Mobility of educated workers in the pre-maturity phase of their career, therefore, increased significantly between 1970 and 1990.

3. Model

3.1. *The production function*

We consider an economy with two locations A and B. They can be two regions or, alternatively, a localized area and its surroundings. A perfectly tradable homogeneous good is produced in each location using two types of workers, denoted as highly educated H and less educated L, and two generations, denoted as young Y and mature (old) O. The model has an overlapping generations (OLG) structure: each period a generation enters the labor market, works for two consecutive periods, first as young and then as mature, and after two periods exits. Highly educated and less educated workers are assumed to be different types of workers. Our attention is concentrated on the highly educated while we simplify much the treatment of the less educated. The two types of workers differ in two important characteristics. First, and most important, highly educated workers are mobile between locations, while less educated are not. This captures a realistic feature of labor markets and also ensures that there would be some production in each location as not all workers can move to one location. Second, highly educated workers contribute more than less educated to the generation of skills through the learning externality. We can think of this assumption as capturing the idea that there are more skills to be learned by observing or interacting with highly educated workers than with less educated ones.

Mobility of the highly educated is not equal in each period of their life. When young, they can choose where to locate without any cost. When mature, if they decide to move they have a cost. This cost is the loss of a fraction (θ) of the skills accumulated on the job as young.

In the economy there are also unproductive house owners who collect rents from the land they own, consume the tradable good and the land services, and are not mobile between cities. They do not enter the labor market and will be introduced in Section 3.3, when we consider the market for housing (land). No physical capital is used in production.[4] Tradable output is produced in each period t in each location c according to the following CES production function:

$$Y_t^c = \Omega_t \left[\left(\sum_{j \in L_c} e_j^{Lc} l_j^c \right)_t^{\gamma} + \Lambda \left(\sum_{j \in H_c} e_j^{Hc} h_j^c \right)_t^{\gamma} \right]^{1/\gamma},$$
$$c = A, B, \ 0 < \gamma < 1, \ \Lambda > 1, \tag{1}$$

where Ω_t denotes total factor productivity at time t. We assume that it grows at exogenous rate and, for simplicity, we set this rate to 0 and standardize the level of Ω_t to one. We are not interested in analyzing the aggregate growth rate of productivity in the economy, but rather its relative level in different locations and the migration decisions of workers. The first term in brackets represents the contribution from less educated workers. Worker j's supply of labor in location c, l_j^c is scaled up by average effectiveness e_j^{Lc}, which is a measure of skills per worker accumulated on the job. Workers of the same type with different experience levels are perfect substitutes. The second term in brackets represents the contribution of highly educated workers. Again, their personal supply in location c, h_j^c is multiplied by their average effectiveness e_j^{Hc}. We allow for a skill biased technological component Λ which is equal across cities and larger than one. This allows different productivity between highly and less educated workers. The popular form chosen for the production function allows us to use existing estimates of the elasticity of substitution, $1/(1-\gamma)$, when we simulate the model in equilibrium.[5]

[4] As long as capital is mobile between cities none of the results obtained is affected by this simplification.

[5] Katz and Murphy [18] estimate an elasticity between college and high school-educated workers to be 1.4, Ciccone and Peri [6] estimate it to be around 4.

Each of the less educated workers L, supplies one unit of labor and their total number is standardized to one in each location per each generation. Each of the highly educated workers H, supplies one unit of labor and their total number in each generation is standardized to one.[6] Throughout the paper we use the term "education" as referred to the type of workers (H or L) and the term "skills" as referred to workers' effectiveness (e_j^c).

We denote with h_Y^A and h_o^A the share, respectively, of young and mature highly educated workers in location A. Therefore $h_Y^B = (1 - h_Y^A)$ and $h_o^B = (1 - h_o^A)$ are the shares of young- and mature-educated workers in location B. We standardize the effectiveness of workers, when young, to one ($e_Y^{Hc} = e_Y^{Lc} = 1$) and we simplify our notation so that e^{Lc} and e^{Hc} denote the effectiveness of less and more educated mature workers (respectively) in location c. Eq. (1) simplifies to

$$Y_t^c = \left[\left(1 + e^{Lc}\right)_t^\gamma + \Lambda\left(h_Y^c + h_o^c e^{Hc}\right)_t^\gamma\right]^{1/\gamma}, \quad c = A, B,\ 0 < \gamma < 1. \tag{2}$$

The term $(1 + e^{Lc})_t^\gamma$ is the contribution of less educated workers while $\Lambda(h_Y^c + h_o^c e^{Hc})_t^\gamma$ is the contribution of highly educated workers.

3.2. Learning externality

During their youth, due to formal and informal interactions with other workers in the same location, workers build up their "effectiveness." The interactions are thought of as face-to-face interactions (or at least as face-to-face initiated interactions[7]) and therefore require physical proximity, i.e., working in the same location. This type of "learning" is a pure externality; it is a by-product of the interactions and workers are not compensated for it. The dynamic effect on accumulation of personal skills depends on the total presence of workers in the location. More educated and more skilled co-workers help faster learning in young workers. Alternatively, we can think of this externality as generated from "matching" with other workers. Workers accumulate progressively their ability by successive matching with different co-workers and the higher the education and the skills of the co-workers the stronger the effect on their ability. This last interpretation would better fit the evidence of frequent job changes (re-matching) for young workers and of their consequent rapid wage increase in the early phase of their career, pointed out by Topel and Ward [27].

To capture the phenomenon of learning from interactions, we introduce a function that describes the accumulation of workers' effectiveness. We do not describe the microfoundations of the process of matching and learning as is done for instance in Glaeser [13] and Jovanovic and Rob [17]. We simply specify an "ad hoc" function describing the outcome of learning. We can think of the

[6] For simplicity we are assuming no population growth.

[7] Differently from Gaspar and Glaeser [12] here these interactions do not involve costs.

function as a reduced form of a process of learning from matching or learning from observing. Our goal is to capture with this function two main features of the learning externality. First, accumulation of skills takes place only when the worker is young. Second, for given personal characteristics, a worker who is in a location with higher density of educated and skilled workers learns more. To keep things simple, we assume that the accumulation of personal skills is identical for more and less educated so that their skill levels would also be equal. Their contribution to young workers' learning, though, is different as the highly educated generate larger externalities. The evolution of the effectiveness of an agent with type of education I, in location c is

$$\left(e^{Ic}\right)_{t+1} = 1 + \phi\left(h_Y^c\right)_t^{\alpha_1}\left(h_o^c e^{Hc}\right)_t^{\alpha_2}\left(e^{Lc}\right)_t^{\alpha_3}, \quad I = L, H,\ c = A, B. \tag{3}$$

This expression implies that the average young worker in location c, either highly or less educated, accumulates effectiveness over her initial level (one), as a function of the total skills in the location while she is young. The contribution of each group to the learning externality in location c depends on the total skills of the group, namely the product of average skills times the size of the group. Total skills are 1, e^{Lc}, h_Y^c, and $h_o^c e^{Hc}$ for the young less educated, old less educated, young more educated and old more educated, respectively. These levels are combined in a Cobb–Douglas function with elasticities reflecting the importance of each group in generating the externality. The contribution to the externality is stronger from highly educated than from less educated implying $\alpha_1 > \alpha_3$, $\alpha_2 > \alpha_3$, and $\alpha_1 + \alpha_2 \geqslant 0.5$. We also assume that the contribution to the externality is stronger from young educated than from old educated ($\alpha_1 > \alpha_2$). This captures the fact that newly educated incorporate more up-to-date skills. Therefore, the accumulation of skills in location c depends on the average level of skills of mature workers (e^{Lc}, e^{Hc}), and on the share of highly educated workers (h_Y^c and h_o^c). $\phi > 0$ is the parameter capturing the intensity of learning. Improvements in the technology for learning from interactions or improvement in the scope of learning increase this parameter.

For simplicity we assume that the function of accumulation of skills is identical for more (H) and less (L) educated. We could easily assume that the less educated do not accumulate skills, and are simply a fixed factor in the model. We think, though, that accumulation of skills of less educated workers is a realistic channel that magnifies the learning externality and therefore we maintain such assumption. Our notation simplifies to $(e^{Hc})_t = (e^{Lc})_t = e_t^c$. Averaging expression (3) across individuals for each location, we obtain the following dynamic system that describes the evolution of skills of mature workers in location A and B:

$$e_{t+1}^A - e_t^A = 1 - e_t^A + \phi\left(h_Y^A\right)_t^{\alpha_1}\left(h_o^A\right)_t^{\alpha_3}\left(e^A\right)_t^{\alpha_2+\alpha_3}, \tag{4}$$

$$e_{t+1}^B - e_t^B = 1 - e_t^B + \phi\left(h_Y^B\right)_t^{\alpha_1}\left(h_o^B\right)_t^{\alpha_3}\left(e^B\right)_t^{\alpha_2+\alpha_3}. \tag{5}$$

It is easy to find the steady state of this dynamic system, setting to 0 the left-hand side of both Eqs. (4) and (5). Assuming $\alpha_2 + \alpha_3 < 1$, there is a unique, globally stable steady state for e_t^A and e_t^B, once h_Y^A, h_o^A reach constant values. The values of accumulated effectiveness (e^A and e^B) in steady state, in location A and B are given by the solution to the following two conditions:

$$e^A = 1 + \phi\left(h_Y^A\right)^{\alpha_1}\left(h_o^A\right)^{\alpha_3}\left(e^A\right)^{\alpha_2+\alpha_3}, \tag{6}$$

$$e^B = 1 + \phi\left(h_Y^B\right)^{\alpha_1}\left(h_o^B\right)^{\alpha_3}\left(e^B\right)^{\alpha_2+\alpha_3}. \tag{7}$$

Effectiveness accumulated by young workers in location c depends positively on the density of more educated workers in the location h_Y^c and h_o^c.[8] This feature of the model generates the agglomeration economies, explains why experience premia are larger in the densely populated location and provides an important reason for young workers to move into it.

3.3. *Wages and utility*

In the rest of the paper we will consider the steady state equilibria of this economy. In such equilibria, the number of young and old educated workers in a location is constant over time (while young workers in a location can be more or less than old workers in that location implying migration during their life). Wages equal marginal productivity for each worker in location c[9]

$$\begin{aligned}
w_{LY}^c &= \frac{\partial Y^c}{\partial l_Y^c} = \left(Y^c\right)^{(1-\gamma)}\left(1+e^c\right)^{\gamma-1};\\
w_{LO}^c &= \frac{\partial Y^c}{\partial l_O^c} = \left(Y^c\right)^{(1-\gamma)}\left(1+e^c\right)^{\gamma-1}e^c;\\
w_{HY}^c &= \frac{\partial Y^c}{\partial l_Y^c} = \left(Y^c\right)^{(1-\gamma)}\Lambda\left(h_Y^c+h_o^c e^c\right)^{\gamma-1};\\
w_{HO}^c &= \frac{\partial Y^c}{\partial l_Y^c} = \left(Y^c\right)^{(1-\gamma)}\Lambda\left(h_Y^c+h_o^c e^c\right)^{\gamma-1}e^c.
\end{aligned} \tag{8}$$

Notice that, because of the constant return to scale in production, the sum of wages paid to all workers exhausts the total product in a location: $w_{LY}^c + w_{LO}^c + h_Y^c w_{HY}^c + h_o^c w_{HO}^c = Y^c$.

The production function exhibits decreasing returns to educated workers alone. This means that increasing their relative supply in a location ($h_Y^c + h_o^c e^c$), decreases their return relative to the wage of less educated.[10] The "education

[8] This is trivially shown applying the implicit function theorem to Eqs. (6) and (7) and recalling that $e^c \geqslant 1$.

[9] We used the identity: $(Y^c)^{1-\gamma} = [(1+e^c)^\gamma + \Lambda(h_Y^c + h_o^c e^c)^\gamma]^{(1-\gamma)/\gamma}$.

[10] This is consistent with findings in Ciccone and Peri [5].

premium” for young and old alike is: $(w^c_{HY}/w^c_{LY}) = (w^c_{HO}/w^c_{LO}) = [\Lambda(h^c_Y + h^c_o e^c)^{\gamma-1}]/[(1+e^c)^{\gamma-1}]$.[11] The expression is increasing in Λ, and decreasing in h^c_Y and h^c_o. Finally, the accumulated effectiveness determines the “experience premium” that is equal to $(w^c_{LO}/w^c_{LY}) = (w^c_{HO}/w^c_{HY}) = e^c$ for both types of workers.

Agents maximize their inter-temporal utility derived from consumption. Each worker consumes during each period of life a composite bundle, G, which is a Cobb–Douglas combination of the tradable good C (the numeraire) and of housing services T, whose price in location c is denoted as p^c_T. The lifetime utility of a worker is separable and logarithmic:

$$\log(G_t) + \frac{1}{1+\beta}\log(G_{t+1}), \quad \text{where } G_t = C_t^{\delta} T_t^{1-\delta},\ 0<\delta<1,\ \beta \geqslant 0, \tag{9}$$

where β is the inter-temporal discount rate, which we assume equal for all workers in both cities and equal to 0. We assume the existence of a bond market whose interest rate, in terms of the numeraire is r. The rent for the housing services in location c, p^c_T, is paid to non-working local landowners, who spend their income to maximize their inter-temporal utility which is also as in (9). Without loss of generality, the amount of land (and number of landowners) is assumed to be the same in each location and it is normalized to one. The total income of land owners in location c is, therefore, equal to p^c_T.

Let E_{IJ} denote the total expenditure in steady state of a worker of type I $(=L, H)$ in period J $(=Y, O)$ of her life. The optimal allocation between C and T in each period implies that $C_{IJ} = \delta E_{IJ}$ and $T_{IJ} = (1-\delta)E_{IJ}$. The optimal inter-temporal allocation of expenditure between the two periods of life, in steady state is as follows:

$$\left(E^c_{IY}\right)_t = \frac{1}{2+r}\left[w^c_{IY} + \frac{w^c_{IO}}{1+r}\right], \tag{10}$$

$$\left(E^c_{IO}\right)_t = \frac{1+r}{2+r}\left[w^c_{IY} + \frac{w^c_{IO}}{1+r}\right], \tag{11}$$

where w^c_{IY} and w^c_{IO} are the wages for young and old workers of type I in location c defined in Eq. (8). The term in square brackets represents the present discounted value of nominal lifetime income for a worker of schooling I in location c.

We can calculate the total workers’ expenditure in steady state for each period in location c, (E^c_{tot}), by adding the expenditure of each group of workers, multiplied by the number of that type of workers in the location. In particular,

[11] Notice that, as $\Lambda > 1$ the wage of the highly educated is always larger than the wage of the less educated.

substituting (8) into expressions (10) and (11), collecting terms and adding we obtain

$$E_{\mathrm{tot}}^{c}=\left[\left(1+\frac{e^{c}}{1+r}\right)\left(\left(1+e^{c}\right)^{\gamma-1}+\left(\frac{1}{2+r}h_{Y}^{c}+\frac{1+r}{2+r}h_{o}^{c}\right)\Lambda\left(h_{Y}^{c}+h_{o}^{c}e^{c}\right)^{\gamma-1}\right)\right]*\left(Y^{c}\right)^{(1-\gamma)}. \quad (12)$$

The total expenditure of workers in the above expression (12), has two components. The term $(Y^c)^{(1-\gamma)}$ is common to wages of all groups and depends on total production in location c. The term in square brackets is the sum of the specific components of expenditures for each group. The term $(1+e^c/(1+r))$ captures the present discounted value of lifetime income as given by first period wage income plus the second period income discounted at rate $1/(1+r)$. The second term within the large brackets, captures the specific components of income for each of the four groups, multiplied by the share spent in the period.

3.4. *Housing market and goods market*

The market-clearing price of housing, p_T^c, is a function of E_{tot}^c, and can be obtained by equating the total expenditure on land services of people living in location c to the total income accruing to land owners in the same location. Recall that the total quantity of land in each location is one and that each worker and each land owner spends a fraction $(1-\delta)$ of their total expenditure (respectively E_{tot}^c and p_T^c) on housing services. The equilibrium condition is therefore

$$(1-\delta)E_{\mathrm{tot}}^{c}+(1-\delta)p_{T}^{c}=p_{T}^{c}. \quad (13)$$

Solving for p_T^c we obtain: $p_T^c=((1-\delta)/\delta)E_{\mathrm{tot}}^c$. The price of land in a location is proportional to the total expenditure of workers in the location. Larger expenditure on land services, whose supply is fixed, generates an increase in their price (a typical crowding effect).

Similarly, the market-clearing condition for the tradable good implies that total demand for the tradable good should be equal to the total production of the good in the two locations. Each group spends a fraction δ of their income on the tradable good. Adding the income of workers and landowners we get:[12]

$$E_{\mathrm{tot}}^{A}+E_{\mathrm{tot}}^{B}=Y^{A}+Y^{B}. \quad (14)$$

The above equation is the market-clearing condition for the tradable good. The left-hand side captures the demand for the tradable good and the right-hand side represents its supply. As in any overlapping generation economy without capital, condition (14) determines the equilibrium value of the real interest rate r.

[12] The total expenditure in the tradable good for the whole economy is obtained by simplifying the following expression: $\delta E_{\mathrm{tot}}^A+\delta((1-\delta)/\delta)E_{\mathrm{tot}}^A+\delta E_{\mathrm{tot}}^B+\delta((1-\delta)/\delta)E_{\mathrm{tot}}^B$.

The “natural” price index for each location in the economy is the price to purchase one unit of the composite consumption bundle G. This index is specific to each location, as the price of housing is different across cities. Using the conditions derived in this section, the location-specific price index is: $p^c = \zeta(p_T^c)^{1-\delta} = \kappa(E_{\text{tot}}^c)^{1-\delta}$, where ζ, κ are constants, independent from the location.[13]

4. Equilibrium analysis

4.1. Lifetime income and location decisions

First of all let us summarize what the centripetal and the centrifugal forces in this model are. This helps us understand what conditions generate colocation of educated workers rather than their dispersion. Concentration of highly educated workers generates crowding externalities (centrifugal forces) for two reasons. First, it lowers nominal wages of educated workers (Eq. (8)) due to decreasing returns to factor H. Secondly, it increases the local price index p^c as a result of higher demand for housing (Eq. (13)). However, concentration of highly educated workers also generates agglomeration externalities (centripetal forces) due to the increased learning possibilities provided to other young educated workers in the location (Eq. (3)).

Highly educated workers are mobile. Each period they choose to locate in the area that gives them the highest lifetime utility. We need therefore to solve their problem backwards. They decide where to locate when mature, in order to maximize their second period real income and, given this decision, where to locate when young in order to maximize their real lifetime income.[14] As there are only two locations, the choice is made by comparing the real income in the second period in location A and B. Given the choice for the second period agents then compare lifetime income from locating in A and B as young. Before presenting the three relevant equilibria arising from this analysis, let us recall some results and introduce some simplifying notation. From (8) we have that wage[15] received by mature educated workers in location c, if they were in c as young, is: $w_{OH}^c = (w_{YH}^c)e^c$. Vice-versa if they move from location c to location c' they lose a fraction θ of acquired skills and their mature wage would be: $(w_{YH}^{c'})[1 + (1-\theta)(e^c - 1)]$.

We denote by RW_c the real wage per unit of efficiency (which is the real wage as young) of educated workers in location c. It is defined as the ratio of nominal wage of the young in location c (defined in (8)) divided by the local price index in

[13] $\zeta = \delta^\delta/(1-\delta)^{(1-\delta)}$, $\kappa = \delta^{2\delta-1}$.

[14] Utility is a linear function of real lifetime income.

[15] We call simply wage the wage in units of tradable good, while we call real wage, the wage in units of the composite consumption good.

location c defined at the end of Section 3.3: $RW_c = w_{YH}^c/(E_{\text{tot}}^c)^{1-\delta}$. Finally, we define RI_c as the "real lifetime income" from living in location c. It is given by the present discounted value of wages earned by workers who spend their whole life in location c: $RI_c = RW_c(1 + e^c/(1+r))$.

4.2. Conditions for equilibria

For the sake of clarity we present in this section the formal conditions for three possible types of equilibria. In the next two sections, with the help of figures and simulations, we describe the characteristics of these equilibria and the combination of parameter values that give rise to each one of them.

First of all, each stationary equilibrium has to satisfy Eqs. (6) and (7) that determine the level of skills accumulated in each location by mature educated workers. Also they all have to satisfy Eqs. (13) and (14) that ensure equilibrium on the housing and on the goods market. The first kind of equilibrium, that we call type I equilibrium, is the simplest and least interesting. It arises when learning externalities are not too strong (i.e., ϕ is small). This is a perfectly symmetric equilibrium with highly educated workers exactly split between location A and location B. Given the symmetry of the model this implies identical real wages in the two locations for both young and mature workers, and therefore no migration of educated workers when they turn mature as long as there is even a minimal loss of skills when they move ($\theta > 0$). Formally the conditions defining such an equilibrium are following.

Equilibrium type I

$$h_Y^A = h_O^A = 0.5, \tag{15}$$

$$\big[RW_c * e^c > RW_{c'}\big[1 + (1-\theta)\big(e^c - 1\big)\big]\big]_{\text{at } h_Y^A = h_O^A = 0.5} \quad \text{for } c = A, B,\ c' = B, A, \tag{16}$$

$$\underbrace{RW_A\left(1 + \frac{e^A}{1+r}\right)}_{RI_A} = \underbrace{RW_B\left(1 + \frac{e^B}{1+r}\right)}_{RI_B}. \tag{17}$$

Due to the symmetric allocation of workers between A and B (Eq. (15)) we have $RW_A = RW_B$ and $e^A = e^B$. Moreover, Eq. (17) is always satisfied and Eq. (16) is satisfied as long as $\theta > 0$.

The second kind of equilibrium, that we call type II equilibrium, arises when learning externalities are strong (i.e., ϕ is large) and the skills learned via local interactions are highly location specific (i.e., θ is large). In this case educated workers colocate to benefit from learning externalities. While such type of equilibrium could generate concentration of educated workers in either location (multiplicity of equilibria, typical of these models) we choose location A to be the one with high concentration of educated workers. Perfectly symmetrical conditions would hold if we chose B as the densely populated location. Formally we have the following equilibrium type.

Equilibrium type II

$$h_Y^A = h_O^A > 0.5, \tag{18}$$

$$\left[RW_A * e^A \geqslant RW_B\left[1 + (1-\theta)\left(e^A - 1\right)\right]\right]_{\text{at } h_Y^A = h_O^A}, \tag{19}$$

$$\underbrace{RW_A\left(1 + \frac{e^A}{1+r}\right)}_{RI_A} = \underbrace{RW_B\left(1 + \frac{e^B}{1+r}\right)}_{RI_B}. \tag{20}$$

Considering workers in location A, their real income in the second period is larger if they remain in A because of the large loss of skills if they move to B (Eq. (19)).[16] Therefore they stay in location A as mature workers. As young they compare the lifetime income from choosing A (and staying there as mature) and from choosing B (and staying there as mature). Educated workers move to A up to the point in which the real lifetime income in the two locations is equated (Eq. (20)). The prevalence of crowding externalities over the learning externality at high levels of density ensures a non-degenerate allocation with educated workers both in A and in B. Together, the equilibrium conditions determine the equilibrium values of h_Y^A and h_O^A and they define the range of values of the parameter θ for which this equilibrium prevails.

For low values of θ Eq. (19) of equilibrium type II is violated. As skills become transferrable, workers become more and more willing to move out of location A as they grow mature. This would eventually unravel equilibrium type II and give rise to equilibrium type III. Such equilibrium, obtained for large learning intensity (ϕ) and low specificity of skills (θ) features workers moving from location A to location B as they turn mature. Educated workers keep migrating from A to B up to the point where real income for mature educated workers is equated between locations. As some workers move to B when mature and ease the crowding externalities, all young educated workers in this case prefer location A. In A they can benefit from learning and ensure themselves higher lifetime income. Formally equilibrium type III satisfies the following conditions.

Equilibrium type III

$$1 = h_Y^A > h_O^A, \tag{21}$$

$$RW_A * e^A = RW_B\left[1 + (1-\theta)\left(e^A - 1\right)\right], \tag{22}$$

$$\underbrace{RW_A\left(1 + \frac{e^A}{1+r}\right)}_{RI_A} > \underbrace{RW_B\left(1 + \frac{e^B}{1+r}\right)}_{RI_B}. \tag{23}$$

[16] It is always true, as A is the crowded location, that no worker wants to move from B to A when she turns mature.

Equation (21) implies that the equilibrium is a "corner solution" as all young workers choose to locate in A. This is due to the fact that now location A maximizes learning for them, as the young educated are the most valuable group to learn from. At the same time, as part of the old educated leave the location as they grow mature, location A is not so badly plagued by the crowding externalities. Condition (23) ensures that the choice of location A is optimal for young workers. As workers grow mature, and do not benefit any longer from learning, some of them move to location B. In equilibrium, enough people move so as to equate real wage for mature workers in A and B (Eq. (22)). The equilibrium conditions define the values of h_Y^A and h_O^A and determine the range of θ for which this equilibrium prevails.

4.3. Parameter space and characteristics of equilibria

The type of equilibrium prevailing in the economy depends on the value of θ and ϕ. In this section we describe what type of equilibrium prevails in each portion of the θ, ϕ plane (Fig. 1). With the help of some simulations we also describe in Table 2 the characteristic features and the behavior of some key variables in each equilibrium.

While we analyze the behavior of the model as θ and ϕ vary (comparative statics) we choose the other parameters' values so as to match observed statistics for the US between 1970 and 1990. The parameters in the production function are chosen as follows: $\gamma = 0.75$ gives elasticity of substitution between high- and low-educated equal to 4, matching that obtained in Ciccone and Peri [6] and compatible with the estimates in Katz and Murphy [18, pp. 72]. $\Lambda = 1.55$ is chosen to

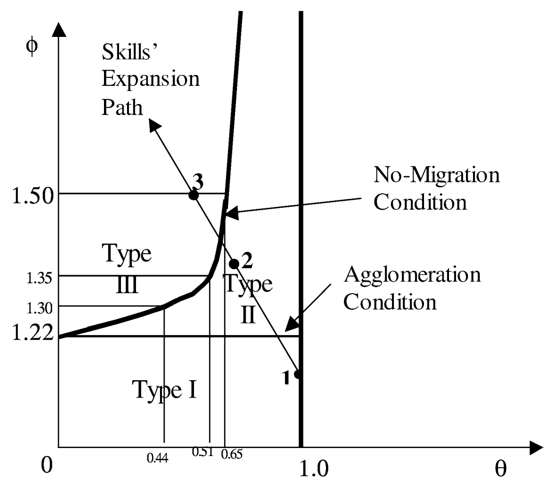

Fig. 1. Parameters' space and equilibria.

Table 2
Characteristics of equilibria

	Type I **1**: ($\theta=1, \phi=1$)			Type II **2**: ($\theta=0.7, \phi=1.4$)			Type III **3**: ($\theta=0.6, \phi=1.5$)		
	h^c_Y/h^c_O	w^c_{HO}/w^c_{HY}	$(h^c_Y+h^c_O)/2$	h^c_Y/h^c_O	w^c_{HO}/w^c_{HY}	$(h^c_Y+h^c_O)/2$	h^c_Y/h^c_O	w^c_{HO}/w^c_{HY}	$(h^c_Y+h^c_O)/2$
Location *A*	1	1.85	0.5	1	3.12	0.82	1.47	3.3	0.84
Location *B*	1	1.85	0.5	1	1.52	0.18	0	1	0.16
Migrants (%, mature)		0%			0%			32%	

Other parameters: $\gamma=0.75$, $\Lambda=1.55$, $\delta=0.75$, $\alpha_1=0.4$, $\alpha_2=0.3$, $\alpha_3=0.2$.

match the average college–high school premium in 1970. From Katz and Murphy [18, Fig. 1], we know that this premium is around 1.7 and in our model we obtain 1.72 as the skilled–unskilled wage ratio in the equilibrium type II. In the utility function we choose $\delta=0.75$ to have a share of housing expenditure equal to 0.25. This value is the average of the shares of family income spent in rents across 1970 (0.19), 1980 (0.26), and 1990 (0.38).[17] Finally, the parameters in the "learning" function are $\alpha_1=0.4$, $\alpha_2=0.3$, $\alpha_3=0.2$. These values capture the larger contribution to learning of educated people and, among them, of young people.[18]

Figure 1 shows the parameter space for θ and ϕ. The vertical lines at $\theta=0$ and $\theta=1$, delimit the relevant area. A horizontal line at $\phi=1.22$ called "Agglomeration Condition" and an upward sloping curve called "No Migration Condition" mark the boundaries between three relevant portions of the plan. Each of the three areas is associated with one type of equilibrium. Within each of the three areas one particular combination of θ and ϕ is marked with a dot and associated to a number ("**1**," "**2**," and "**3**"). For each of these specific parameter combinations, representative of each type of equilibrium, we report in Table 2 the values of the following variables:

- In the first column of each type of equilibrium we have the relative density of young and mature educated workers in each location, h^c_Y/h^c_O.
- In the second column of each type of equilibrium we have the wage of mature workers relative to wage of young workers w^c_{HO}/w^c_{HY} that equals one plus the experience premium. The experience premium is equal to accumulated skills e_c.
- In the third column of each type of equilibrium we have the share of overall (young and mature) educated workers in each location $h^c_O+h^c_Y/2$.
- In the last row we report the percentage of mature workers who migrate at the end of their youth as percentage of total mature workers. This value is given by $(h^A_Y-h^A_O)$.

[17] These values are obtained from our calculations on PUMS of US Census 1970, 1980 and 1990.

[18] Several other simulations and robustness checks have been performed with different values of the parameter. They are available upon request. Their qualitative features are identical to those of the simulation reported in the paper.

In the remainder of this section we describe Fig. 1 and Table 2 that jointly provide a complete characterization of the equilibria and of their characteristics. We analyze their stability and uniqueness properties in the next section.

The area below the "Agglomeration Condition" for any value of θ in $(0, 1]$ is where only equilibrium type I is possible. When the intensity of learning is low ($\phi < 1.22$) there is no incentive for educated workers to colocate as crowding effects are stronger than agglomeration economies. The characteristics of equilibrium in point "**1**" (where $\theta = 1$, $\phi = 1$) are reported in the first three columns of Table 2. Educated workers are evenly spread between the two locations (50% in each location) and do not move as they get mature (so that Migrants are 0 and the young–mature ratio is 1 in each location). Wages are identical in the two locations both for young and mature workers and so are experience premia (equal to 85%).

As the learning intensity ϕ increases above the threshold of 1.22, learning externalities prevail over crowding externalities at the symmetric equilibrium. Above the "Agglomeration Condition" the symmetric equilibrium becomes unstable and two locally stable asymmetric equilibria arise endogenously, one with higher concentration of educated workers in A and one with higher concentration of educated workers in B. We choose (w.l.o.g.) location A as the one attracting more educated workers. Agglomeration equilibria are now possible and a key role is played by θ, the parameter capturing specificity of skills. For θ large (to the right of the "no migration condition" in Fig. 1) type II equilibrium prevails. The "no migration condition" represents the combination of ϕ and θ such that educated workers are indifferent between migrating as they turn mature or staying where they are. Mathematically it is given by Eq. (19) taken with equality or equivalently by Eq. (22). It is upward sloping because as the benefits from learning (ϕ) increase higher specificity of skills (θ) is needed to prevent migrations of workers as they turn mature.

We consider as representative equilibrium for this portion of the plan point "**2**" in Fig. 1 (characterized by $\theta = 0.7$, $\phi = 1.4$). Table 2 shows that educated workers (82% of them) are attracted to location A where they benefit from intense learning and earn large experience premia ($w^A_{HO}/w^A_{HY} = 3.12$ vs. $w^B_{HO}/w^B_{HY} = 1.52$). However, due to the large specificity of their skills (they would lose 70% of their productivity if they moved as mature) all educated workers remain in the same location when they turn mature. This implies the same number of young and old in each location, $h^c_Y/h^c_O = 1$, and no migration of mature workers. Young educated workers locate into the denser area A up to the point where the negative effect of crowding externalities on their lifetime income RI_A balances the positive effect of learning in location A. In equilibrium real lifetime income in each location is the same.

Finally, in the portion of the plan to the left of the "no migration condition" type III equilibrium prevails. As transferability of skills becomes large enough (θ low), mature workers have incentives to move out of the dense location in order to enjoy larger wages and lower prices. Considering point "**3**" ($\theta = 0.6$,

$\phi = 1.5$) in Fig. 1 as representative of this portion of the plan, workers can now transfer 40% ($\theta = 0.6$) of their learned skills. As a consequence many of them migrate as they turn mature. The out-migration of mature workers from A eases the crowding problems of that location. All educated workers choose A as preferred location when young (thus the 0 young–old ratio in B), they learn their skills (experience premium 3.30) and transfer to B as mature. Not all educated workers transfer to B as mature, because the increasing density of that location would eventually drive real wages of mature workers down. The percentage of migrants (33% in our case) is determined by the equality of real wage for mature workers in the two locations (condition (22)). As there are no young educated workers in location B, no learning takes place there and therefore wage of people who are there as young does not increase (hence the reported experience premium $w^B_{HO}/w^B_{HY} = 1$). However, in equilibrium no young worker locates in B. Looking at the experience premium of mature workers in B (who moved from A) relative to the wage that an educated young would earn in B, the "imputed" experience premium w^B_{HO}/w^B_{HY} is 2.01. It is still lower than the premium gained by workers in A due to the loss of some of their skills.

4.4. Stability and uniqueness

We define stability of an equilibrium in the usual way adopted by the "new economic geography models" (see Fujita et al. [11]). An internal equilibrium (i.e., with young educated workers in both locations A and B) is stable if, after a perturbation that moves some workers from A to B, the lifetime income of workers in A increases above the lifetime income of workers in B ($RI_A - RI_B < 0$). Conversely, if the perturbation moves workers from B to A, the lifetime income of workers in A decreases below the lifetime income of workers in B ($RI_A - RI_B < 0$). This condition implies stability because if workers locate in response to real income differentials, they would migrate to undo the perturbation. The stability of the "corner" equilibria (i.e., with educated workers only in one of the two locations) requires that the lifetime income in the location where all workers are, is larger than the lifetime income in the other location, specifically: $RI_A - RI_B > 0$ if $h^A_Y = 1$, $RI_A - RI_B < 0$ if $h^A_Y = 0$.

Therefore, in order to analyze the stability properties of the equilibria, we represent ($RI_A - RI_B$) as a function of h^A_Y (the share of young educated workers located in A). If this function crosses the zero line and is negatively sloped then the corresponding value of h^A_Y is a stable equilibrium.[19] With the help of the simulated values of $RI_A - RI_B$ we analyze the stability and uniqueness of equilibrium for the three representative cases. Figures 2–4 represent the function $RI_A - RI_B$ (dotted line) as h^A_Y increases. As the model is perfectly symmetric we only represent the

[19] Clearly we always impose conditions (6), (7), (13), and (14).

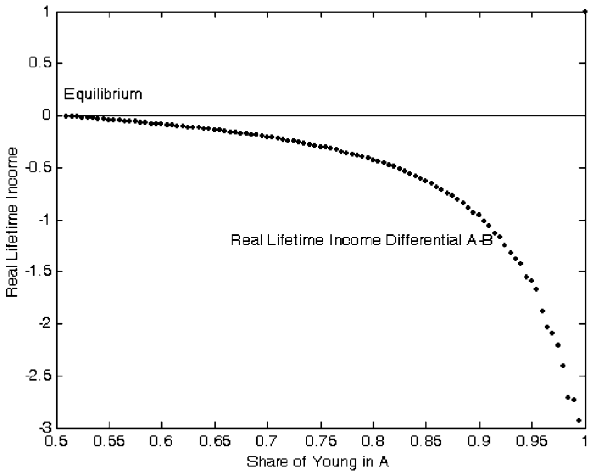

Fig. 2. Equilibrium type I: $\phi = 1$, $\theta = 1$.

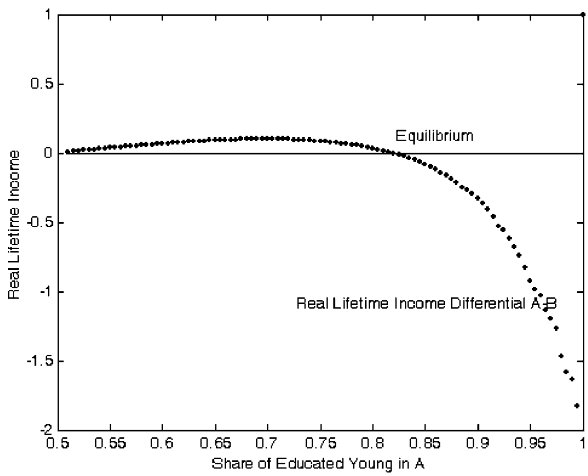

Fig. 3. Equilibrium type II: $\phi = 1.4$, $\theta = 0.7$.

range $h_Y^A \in [0.5, 1]$. The left corner value, $h_Y^A = 0.5$, represents the symmetric equilibrium, with the same share of young educated workers in both locations. The right-corner value, $h_Y^A = 1$, represents full concentration of young educated workers in location A.

Figure 2 is drawn for the parameter values as in point "**1**" of Fig. 1. Such behavior of the function $RI_A - RI_B$ is representative of type I equilibria. The function crosses the zero value at $h_Y^A = 0.5$ and it is globally downward sloping. Clearly in this case the symmetric allocation of young educated workers is the

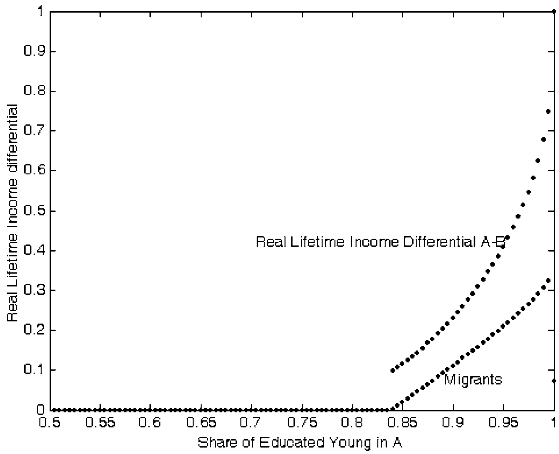

Fig. 4. Equilibrium type III: $\phi = 1.5$, $\theta = 0.6$.

unique stable equilibrium. If externalities are weak, any perturbation from the symmetric allocation of young workers would cause them to relocate in order to re-establish the symmetric equilibrium.

Figure 3 is drawn for the parameter values identified by point "**2**" in Fig. 1. Such behavior of the function $RI_A - RI_B$ is representative of Type II equilibria. The function is obtained imposing condition (18). Notice that now, due to larger learning externalities, the perfectly symmetric equilibrium is not stable any longer. $RI_A - RI_B$ is upward sloping at $h_Y^A = 0.5$ and therefore any small perturbation of the symmetric equilibrium would move young workers away from the symmetric distribution. This is due to the fact that agglomeration externalities are stronger than crowding effects at the symmetric equilibrium and therefore $RI_A - RI_B$ is upward sloping. As h_Y^A increases, though, the crowding externalities gain strength. At the point marked as "Equilibrium" in Fig. 3 the benefits and costs of agglomeration are exactly equal and $RI_A = RI_B$. This point satisfies Eq. (20). Also, for values of θ anywhere in the region denoted as "Type II" in Fig. 1, condition (19) is satisfied as well. This equilibrium is locally stable as $RI_A - RI_B$ is downward sloping.

Finally, we consider the stability property of type III equilibrium. As usual we consider the simulated $RI_A - RI_B$ function for values of the parameters described by point "**3**" in Fig. 1. Such function, drawn in Fig. 4, is representative of the parameter combinations in the whole area denoted in Fig. 1 as "Type III."

Equilibrium of type II cannot exist in this area because, due to the low values of θ, condition (19) is violated. Therefore only equilibrium with positive migration of mature workers from A to B may exist (Eq. (21)). Figure 4 shows, after imposing the equilibrium condition (22), the share of mature workers who migrate from A to B (migrants) as well as, on the same scale, the function $RI_A - RI_B$, in the interval where migration from A to B is positive. We can represent both variables on the same scale because migration is expressed as share of mature workers and income differential are expressed as share of one-period utility. The function $RI_A - RI_B$ is upward sloping and always above the zero line in the interval in which equilibrium Type III exists. Therefore only the corner solution with all young workers in location A is a stable outcome. 33% of the educated workers move to B when they turn mature. This equilibrium is the only locally stable one for the parameters' combination in the area denoted as "Type III" in Fig. 1.

5. Discussion and conclusions

The simple model developed in this paper focuses on the importance of learning externalities in determining the location and migration decision of highly educated workers. We find that, as learning externalities become more important and skills become more transferrable, the distribution of educated workers in dense areas change. Learning externalities are only one among the several agglomeration externalities that cause educated workers to crowd urban areas. They seem, though, the most likely candidate to explain why densely populated areas have been overcrowded with young educated workers since the 90s. In Fig. 1 we show a line connecting points "**1**," "**2**," and "**3**" which we label "Skills Expansion Path." Such a line is drawn because some stylized facts and anecdotal evidence suggest an increase in the intensity of learning and an increase in transferability of skills, during the 70s, 80s, and 90s.

As production tasks became more complex and articulated, the scope for learning increased and the learning intensity from interactions increased. Certainly, moving from a disperse distribution of educated workers to their concentration in urban areas, several agglomeration economies besides learning are at work. However, the sizeable increase of experience premia of college-educated workers in the 70s and 80s (see Katz and Murphy [18]) together with the increased concentration of college educated in urban areas in the same period (see Appendix A) are signs that skills accumulation and concentration of educated workers went hand in hand.

Our attention is focused on the comparison between equilibrium type II and equilibrium type III. We think that our model has a simple and consistent explanation for why densely populated areas have switched from a balanced composition of young and mature educated workers in the 70s to a strong prevalence of young workers in the 90s. This has been driven by a switch from a point like "**2**" to a point like "**3**" in Fig. 1. The key change has been the decrease of the parameter θ, i.e., the increase in transferability of skills. The seventies and eighties were a period of increased flexibility in the US labor market: small firms thrived as large firms went through a slowdown, flexible technology became important vis-a-vis that of firm-specific production (see Piore and Sabel [24]). Flexibility and generality of skills, as well as the importance of learning became key ingredients for success (see Kochan et al. [19] for a survey on change in production during the 80s). Technology promoted general-purpose skills, standard computational languages were adopted, computers were introduced (1980). As a consequence of these changes young educated workers were attracted to densely populated areas as they could learn from interactions there. Moreover, due to higher flexibility and transferability of skills they could retain the higher productivity while moving out of densely populated areas as mature workers.

From a few studies we can put an estimate on the value of the parameter θ during the 80s. As $(1-\theta)$ is the share of productivity lost when workers change location we may use some estimates of wage loss, as workers are displaced from firms, as a lower bound for that value. Topel [26], using 1984–1986 data finds that between 40 and 60% of the tenure premium is lost when workers move out of firms (Tables 5 and 7 in Topel [26]). This implies an estimate of θ between 0.4 and 0.6.

Glaeser and Mare [14] analyze migrants between rural and urban areas, in the two directions. They use data from the National Longitudinal Survey of Youth for the period 1983–1993. Using their most accurate estimates (panel with individual fixed effects, in Column 2 of their Table 5) migrants from urban to rural areas lose on average 5–7% of their wage. Compared to the 12% increase in wage that they accumulate while working in urban areas, this implies a loss of productivity equal to 40 to 60% of the skills learned in cities. Again these estimates put θ between 0.4 and 0.6. Values of θ between 0.4 and 0.6 are exactly in the range that marks the boundary between equilibrium Type II and equilibrium Type III regions in our simulations shown in Fig. 1.

While not the only possible explanation for the increased density of educated young workers in urban areas, our model provides a coherent and reasonable story of what happened. The story hinges crucially on the opportunity of learning skills through interactions and on the possibility of transferring them. Only when the two mechanisms work together the result is that young educated workers choose urban areas and mature workers move out of them. As a consequence, better understanding the process of learning and better measures of transferability of skills, promises to help understanding location and inter-regional migration of educated workers.

Acknowledgments

I thank Graziella Bertocchi, Giuseppe Bertola, Jan K. Brueckner, Andrea Ichino, two anonymous referees, and participants of the CEPR conference, Barcelona 2001, for very useful comments and suggestions. Errors are mine.

Appendix A.

Table A1 reports the complete set of statistics on which the stylized facts described in Section 2 are based. This table reports the comparison between

Table A1
Detailed stylized facts 1970–1990

	1990				1970			
	Comparison 1		Comparison 2		Comparison 1		Comparison 2	
	Mega	Other cities	Urban areas	Non-urban areas	Mega	Other cities	Urban areas	Non-urban areas
Share of college-educated workers	0.29	0.23	0.25	0.20	0.16	0.14	0.135	0.11
W_Y^H = Wage[a] of white, college-educated young* male	20	19	19	17	22.5	19.7	19	16.8
W_O^H = wage[a] of white mature** college-educated male	31	26	28	24	28.5	25.0	23.5	19
Experience premium (W_O^H / W_Y^H)	1.55	1.36	1.47	1.41	1.26	1.26	1.23	1.13
h_Y = share of young* college-educated among all workers	0.20	0.15	0.17	0.13	0.10	0.09	0.08	0.06
h_O = share of mature** college-educated among workers	0.09	0.08	0.08	0.07	0.06	0.05	0.055	0.04
Ratio of young/mature college-educated (h_Y / h_O).	2.22	1.87	2.12	1.85	1.66	1.8	1.5	1.5

* Experience < 20 years. ** Experience $\geqslant 20$ years.
[a] Hourly wage in 1990 US $.
[b] Mega = Megalopolis, the seven largest urban areas in the US: Los Angeles, New York, Chicago, Philadelphia, Boston, Detroit, and Washington (source: Author's Calculations on US Census 1970, 1990).

Table A2
Regressing characteristic of cities on their size measured by employment

	1990		1970	
	ln(Empl)	R^2	ln(Enipl)	R^2
Share of college-educated workers	0.036[a]	0.28	0.01[a]	0.05
	(0.004)		(0.0045)	
Share of young–share of mature**	0.09[a]	0.05	0.01	0.01
among white male college-educated	(0.02)		(0.02)	
Share of young*–share of mature**	0.047[a]	0.04	−0.01	0.01
among single white male college-educated	(0.018)		(0.02)	
Log(experience premium) of white	0.02[a]	0.02	0.06[a]	0.07
males college-educated	(0.01)		(0.02)	

* Experience < 20 years. ** Experience $\geqslant 20$ years. (Source: OLS regressions on PUMS 1970, 1990 data).
[a] White Robust standard error in parenthesis, significant at 99%.

urban and non-urban areas (considered in the paper) as well as the one between Megalopolis and other cities (mega–other cities).

Proceeding from the top row we see that the share of college-educated workers, as well as their experience premium is always larger in the densely populated areas. Conversely, the share of young college-educated workers in densely populated areas became particularly large only in 1990. Table A2 confirms these facts by considering 167 SMSA. We calculate the correlation coefficients of the share of college-educated, experience premia and young–old differentials in SMSA with the size of the SMSA.

The regression coefficients confirm that larger cities had larger share of college-educated workers and paid larger experience premia (both in 1970 and 1990). Again, though, the increased presence of the young college-educated only appears in 1990. In the third row we control for the marital status of workers to avoid that marriage considerations would affect the location decisions. We still find for single males a strong correlation of young workers with larger cities.

References

[1] H.M. Abdel Raman, M. Fujita, Specialization and diversification in a system of cities, Journal of Urban Economics 33 (1993) 189–222.
[2] D. Audretsch, M.P. Feldman, R&D spillovers and geography of innovation and production, American Economic Review 86 (1996) 630–640.
[3] K. Arrow, The economic implications of learning by doing, Review of Economic Studies XXXIX (1962) 155–173.
[4] D. Black, V. Henderson, A theory of urban growth, Journal of Political Economy 107 (2) (1999) 252–284.
[5] A. Ciccone, G. Peri, Human capital and externalities in cities, Discussion paper #2599, CEPR, 2000.

[6] A. Ciccone, G. Peri, Skill substitutability and technological shocks, Mimeo, Bocconi University, 2001.
[7] D.L. Costa, M.E. Kahn, Power couples: Changes in the locational choice of the college educated: 1940–1990, Quarterly Journal of Economics XXX (2000) 1287–1315.
[8] G. Dumais, G. Ellison, E.L. Glaeser, Geographic concentration as a dynamic process, Working paper #6270, NBER, 1997.
[9] G. Duranton, D. Puga, Nursery cities: Urban diversity, process innovation, and the life cycle of products, American Economic Review 91 (5) (2000) 1454–1477.
[10] A.D. Foster, D. Rosenzweig, Learning by doing and learning from others: Human capital and technical change in agriculture, Journal of Political Economy 103 (6) (1995) 1176–1209.
[11] M. Fujita, P. Krugman, A.J. Venables, The Spatial Economy, MIT Press, Cambridge, London, 1999.
[12] J. Gaspar, E.J. Glaeser, Information technology and the future of cities, Journal of Urban Economics 43 (1998) 136–156.
[13] E.J. Glaeser, Learning in cities, Journal of Urban Economics 46 (1999) 254–277.
[14] E.J. Glaeser, P. Mare, Cities and skills, Journal of Labor Economics 19 (2) (2001) 316–342.
[15] R.W. Helsley, W.C. Strange, Matching and agglomeration economies in a system of cities, Regional Science and Urban Economics 20 (1990) 189–212.
[16] A. Jaffe, M. Trajtenberg, R. Henderson, Geographic localization and knowledge spillovers: Evidence from patent citation, Quarterly Journal of Economics 108 (1992) 577–598.
[17] B. Jovanovic, R. Rob, The growth and diffusion of knowledge, Review of Economic Studies 56 (4) (1989) 569–582.
[18] L.F. Katz, K.M. Murphy, Changes in relative wages 1963–1987: Supply and demand factors, The Quarterly Journal of Economics 107 (1992) 35–78.
[19] T.A. Kochan, H. Katz, R.B. McKersie, The Transformation of American Industrial Relations, Basic Books, New York, 1986.
[20] P. Krugman, Increasing returns and economic geography, Journal of Political Economics 99 (1991) 483–499.
[21] P. Krugman, A.J. Venables, Globalization and the inequality of nations, Quarterly Journal of Economics CX (1995) 857–880.
[22] R.E. Lucas, On the mechanics of economic development, Journal of Monetary Economics 22 (1) (1988) 3–42.
[23] R.E. Lucas, Making a miracle, Econometrica 61 (1993) 251–272.
[24] M. Piore, C. Sabel, The Second Industrial Divide: Possibilities for Prosperity, Basic Books, New York, 1984.
[25] J. Rauch, Productivity gains from geographic concentration in cities, Journal of Urban Economics 34 (1993) 380–400.
[26] R.H. Topel, Specific capital mobility and wages: Wage rise with job seniority, Journal of Political Economy 99 (1991) 145–176.
[27] R.H. Topel, M.P. Ward, Job mobility and the careers of young men, Quarterly Journal of Economics 107 (2) (1992) 439–479.