Chapter 51

HUMAN CAPITAL EXTERNALITIES IN CITIES

ENRICO MORETTI

Contents

Handbook of Regional and Urban Economics, Volume 4. Edited by J.V. Henderson and J.F. Thisse

DOI: 10.1016/S0169-7218(04)07051-0

Abstract

What is the effect of an increase in the overall level of human capital on the economy of a city? Although much is known about the private return to education, much less is known about the more important question of what happens to productivity, wages and land prices when the aggregate stock of human capital in a city increases. Increases in the aggregate stock of human capital can benefit society in ways that are not fully reflected in the private return of education. Human capital spillovers can in theory increase aggregate productivity over and above the direct effect of human capital on individual productivity. Furthermore, increases in education can reduce criminal participation and improve voters' political behavior. In this chapter, I review what we know about social returns to education, with a particular emphasis on those externalities that accrue to local geographic areas. The focus of the chapter is on the empirical issues that arise in identifying these externalities and on the existing empirical evidence on their magnitude.

Keywords

education, social return, spillovers, externalities, cities

JEL classification: J0, I2, H23

1. Introduction

After 40 years of research on the relationship between education and earnings, economists have a good idea of the private benefits of human capital. We know that individuals with more education earn more, and most empirical work suggests that this difference in earnings is in fact a reflection of education per se and not a result of differences in unmeasured worker characteristics.

But despite this general agreement on the private return to education, much less is known about the social return, although economists have speculated about the possibility of externalities for at least a century. In this chapter, I review what we know about the social benefits of human capital, with a particular emphasis on those benefits that accrue to local geographic areas. Although I briefly review theories of human capital externalities, the focus of this chapter is on the empirical issues that arise in estimating these externalities and on the existing empirical evidence on their magnitude.

What is the effect of an increase in the overall level of human capital on a local economy? In the presence of externalities, the effect of aggregate schooling on aggregate earnings is not necessarily the same as the effect of individual schooling on individual earnings. These earnings externalities can be either positive or negative. On one hand, a large theoretical literature in both urban economics and macroeconomics has argued that aggregate human capital has a positive effect on productivity over and above its private effect-making human capital spillovers important factors in explaining the economic growth of cities, regions, and countries.

On the other hand, it is in theory possible that education has little effect on individual productivity, but it is simply a signal of innate ability. In this case education generates negative (pecuniary) externalities, and the effect of increases in aggregate schooling on aggregate earnings is smaller than the effect of increases in individual schooling on individual earnings.

In another branch of research, economists have hypothesized that education may have other social benefits in addition to its effect on earnings. For example, education is often thought to reduce the probability that an individual will engage in activities that generate negative externalities, such as crime. Alternatively, economists from Adam Smith to Milton Friedman have argued for public subsidies to education on the grounds that a better-educated electorate makes better decisions on policy issues that affect the economy.

The possibility that the social return to human capital differs from its private return has tremendous practical importance. For example, the magnitude of the social return to education is a crucial tool for assessing the efficiency of public investment in education, since state and local governments subsidize almost all direct operating costs of primary and secondary educational institutions. In fact, much of the argument for public education comes from the recognition that education not only rewards the educated individual, but also creates a variety of benefits that are shared by society at large.

Furthermore, the magnitude of externalities from education is important for local development policies. Local governments are increasingly interested in fostering economic growth and they have a number of alternative policy options at their disposal: subsidizing new business, changing environmental or labor standards, or developing policies to attract or create an educated labor force. Local governments must strike a balance between these options, as they may involve important trade-offs. For example, lowering environmental standards may result in a reduction in the number of educated workers if demand for environmental quality increases with education. Knowing the magnitude of the social benefits of human capital is therefore a crucial consideration in the choice of an optimal development policy.

In this chapter, I present a unified equilibrium framework with productivity spillovers. The framework indicates that geographically local spillovers can be identified either directly – by comparing the productivity of firms in cities with different overall levels of human capital, holding constant firms' individual characteristics – or indirectly, using factor prices. In the indirect approach, externalities can be identified in two ways, either by comparing the wage of workers in cities with different overall levels of human capital, holding constant workers' individual characteristics; or by comparing housing prices in these cities, holding constant houses' characteristics. The framework also clarifies the precise relationship between the estimates obtained from these three empirical strategies. I use this framework to interpret existing estimates of human capital spillovers.

The issue of endogeneity of aggregate human capital is probably the most important empirical challenge facing researchers in this area. Human capital is not distributed randomly across cities; it tends to be higher in areas with high productivity and good amenities. The reason is that workers endogenously choose where to locate based on wages, cost of living and the match between their taste and city amenities. Similarly, firms also endogenously choose where to locate based on wages, cost of land and the match between their cost function and city characteristics. Empirically, we observe that cities with a well-educated labor force tend to have better amenities, better institutions, better infrastructure, a more modern industry structure, and more technologically advanced firms than cities with a less-educated labor force. In addition, workers in cities with a well-educated labor force are likely to have unobserved characteristics that make them more productive than workers with the same level of schooling in cities with a less-educated labor force.[1]

As a consequence, it is empirically difficult to disentangle the effect of higher overall levels of human capital on productivity, wages, and land prices from the effects of these unobserved characteristics of workers and cities. The framework developed here indicates that unobserved heterogeneity of firms, workers, and cities is likely to bias least squares (OLS) estimates of the externalities, but the direction of the bias is not obvious a priori. Whether the true magnitude of the spillover is larger or smaller than the OLS estimate will depend on whether the unobserved factors that affect the relative demand

[1] For example, a lawyer in New York is likely to be different from a lawyer in El Paso, TX. Similarly, a high-school graduate in a biotech firm in San Francisco is likely to be different from a high-school graduate in an apparel plant in Brownsville, TX.

of skilled labor across cities dominate the unobserved factors that affect the relative supply of skilled labor. I discuss a number of ways to address the potential endogeneity of human capital stocks.

The empirical literature on human capital externalities should arguably have two objectives. First, it should credibly assess the magnitude of spillovers. Given the significant policy implications and a large theoretical literature that assumes the existence of human capital externalities, it is an important first step to quantify the size of such externalities, if they exist at all. A second goal should be to empirically investigate the mechanisms that give rise to externalities.

After reviewing the existing evidence, I conclude that the empirical literature provides some intriguing evidence on the existence of human capital externalities, but we are still far from a consensus on the magnitude of such externalities. The empirical literature on the subject is still very young and the econometric challenges are difficult to overcome. More work is needed before we can draw convincing conclusions about the size of human capital externalities and the mechanisms that drive them.

The remainder of the chapter is organized as follows. In the next section, I describe recent trends in the geographic distribution of education across U.S. cities. In Section 3 I briefly review theories of the social return to human capital. In Section 4, I discuss the empirical challenges that arise in estimating human capital spillovers and the existing empirical evidence on the magnitude of these spillovers. In Section 5, I turn to the social benefits of education that are not reflected in increased earnings. The last section concludes.

2. Recent trends in the geographic distribution of human capital across cities

The distribution of human capital in the United States is geographically uneven. Urban areas typically tend to have a better-educated populace than rural areas. But even among urban areas, there are substantial differences in the number of skilled individuals, as well as in the changes over time in that number. For example, the fraction of college-educated individuals in cities at the top of the education distribution in 2000 (e.g., San Francisco, CA) is four times larger than the fraction of college-educated individuals in cities at the bottom of the distribution (e.g., Danville, VA). In this section, I document recent trends in the distribution of schooling across major U.S. metropolitan areas from 1980 to 2000. I also analyze which characteristics of cities in 1990 are associated with large increases in the stock of human capital between 1990 and 2000.

My findings suggest that virtually all U.S. cities experienced increases in the fraction of educated individuals in the 1990s. The increases were on average similar to those experienced during the 1980s. But the increases were by no means uniform across cities. In particular, cities that had a relatively high fraction of educated individuals in 1990 experienced *larger* increases between 1990 and 2000 than cities that had a relatively smaller fraction of educated individuals that year. As a consequence, the distribution of human capital across cities became more unequal during the 1990s. One reason for the

Table 1
Percent and dispersion of college graduates by year

	(1)	(2)	(3)	(4)
	Mean	Variance	P75–P25	P10–P90
2000	0.234	0.0044	0.088	0.172
1990	0.201	0.0037	0.070	0.152
1980	0.176	0.0028	0.073	0.121

Note. Sample includes 222 metropolitan areas.

increased concentration of human capital in some cities was the high-tech boom of the 1990s, since it benefited a handful of already highly skilled cities. But this tendency of increasing inequality in the distribution of human capital across U.S. cities during the 1990s was not a new phenomenon, as it was already in place during the 1980s.

Table 1 reports summary statistics for 222 metropolitan areas that I was able to link in 1980, 1990, and 2000. The data are from the Census of Population and Housing. Throughout the paper, the unit of analysis is the metropolitan statistical area (MSA). MSAs are defined to include local economic regions with populations of at least 100,000. Most MSAs contain more than one county. The term "city" and "MSA" will be used interchangeably in this paper. A total of 320 MSAs can be identified in the 2000 Census, but only 222 can be matched consistently across censuses.[2]

Historically, the U.S. population is characterized by a long-run trend of increasing education, since the younger cohorts are better-educated than the older ones. Column 1 in Table 1 confirms that the average fraction of college graduates across 222 cities increased from 17% in 1980, to 20% in 1990 to 23% in 2000. The fact that the share of college graduates has been steadily increasing over time is well documented, and should not come as a surprise. More interesting is the fact that the dispersion of human capital across cities also appears to be increasing. I present three measures of dispersion: variance (column 2), interquartile range (column 3) and difference between 90th and the 10th percentile. The variance increased from 0.028 in 1980 to 0.037 in 1990 to 0.088 in 2000. The other two measures of dispersion are generally consistent with such increase.

These trends in the mean and dispersion of the stock of human capital across U.S. cities are depicted graphically in Figures 1 and 2. The top panel of Figure 1 plots the percent of college graduates in 1990 on the x-axis, against the percent of college graduates in 2000 on the y-axis for each city in the sample, and it superimposes the 45 degrees line. The first thing to notice in the Figure is the wide variation in the average stocks of human capital among cities. In both 1990 and in 2000, the fraction of college graduates goes from about 10% in the least-educated cities to above 40% in the

[2] Data for 1980 and 1990 are from the individual-level 5% PUMS. Because individual-level data for the 2000 Census are not yet available, I rely on aggregate statistics provided by the Census for year 2000. One limitation of the aggregate data is that the average years of schooling is not reported. Only the percent of individuals with different level of schooling achievements in each city is reported.

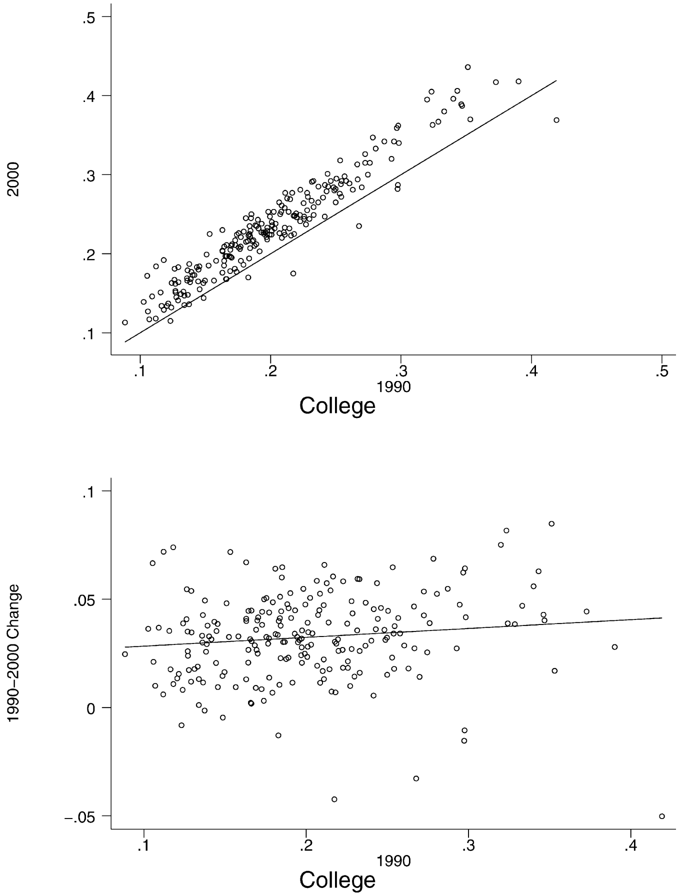

Figure 1. Changes in percent college, by city: 1990–2000.

highest-educated cities. A second feature to notice is that the stock of college graduates has increased almost everywhere. Only an handful of cities have a smaller percentage of college graduates in 2000 than in 1990. Most of the cities lie above the 45 degree line.

Third, the increase between 1990 and 2000 appears to be larger the higher the 1990 level of human capital. To see this last point clearly, the bottom panel of Figure 1 plots the percent of college graduates in 1990 on the x-axis, against the change in the percent of college graduates between 1990 and 2000 on the y-axis. Unlike the top panel, here the superimposed line is the OLS fit. The panel shows that cities that had higher levels of human capital in 1990 experienced larger increases during the 1990s. This finding is consistent with the increase in dispersion documented in Table 1. The slope (standard error) of the OLS fitted line is 0.041 (0.022).

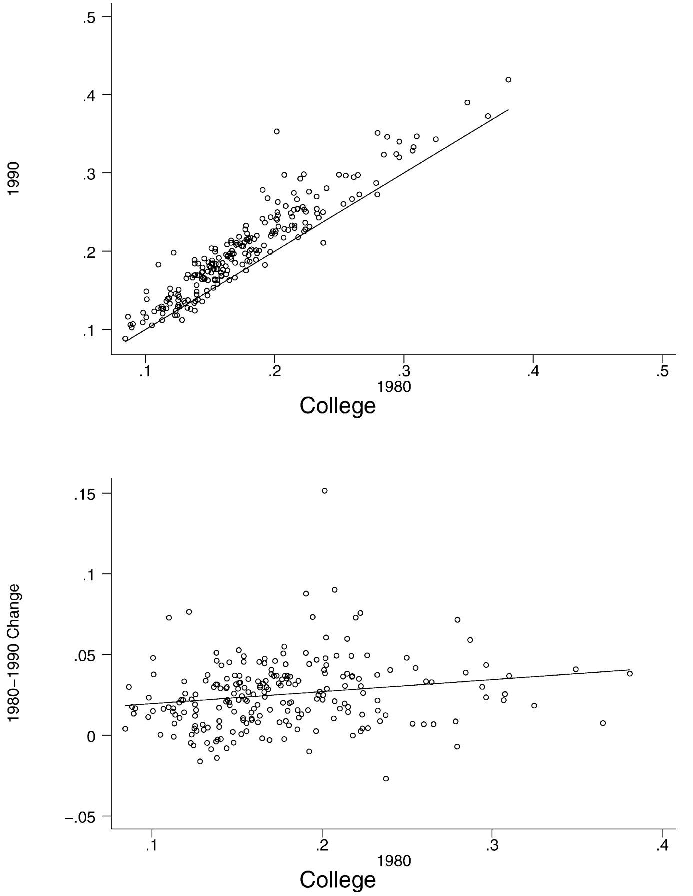

Figure 2. Changes in percent college, by city: 1980–1990.

Figure 2 shows a similar picture for the 1980s. In particular, the figure shows that the three features of the data uncovered in Table 1 are not specific to the 1990s, but have been going on much longer. The increase in dispersion is qualitatively consistent with the increase in dispersion documented in Figure 1, but the slope (standard error) of the OLS fitted line is even steeper: 0.0744 (0.0260).

To give a better sense of the distribution of human capital across cities, Table 2 lists the 15 cities with the largest and smallest per capita number of college-educated residents in 2000. San Francisco, where more than 43% of adults are college educated, appears to be the city with the largest per capita stock of human capital. Medium-sized cities that host one or more large research universities are overrepresented in the top group: examples are Madison, WI; Raleigh, NC; Gainesville, FL; Urbana-Champaign, IL; Austin, TX; College Station, TX; State College, PA; and Santa Cruz, CA. Table 3

Table 2
Cities with large and small percent of college graduates in 2000

Cities with the largest percentage	(1)	*Cities with the largest percentage*	(1)
San Francisco, CA	0.436	Jacksonville, NC	0.148
Washington, DC–MD–VA–WV	0.418	Beaumont–Port Arthur, TX	0.147
Columbia, MO	0.417	Hagerstown, MD	0.146
Madison, WI	0.406	Stockton–Lodi, CA	0.145
San Jose, CA	0.405	Huntington–Ashland, WV–KY–OH	0.144
Bloomington, IN	0.396	Modesto, CA	0.141
Fort Collins–Loveland, CO	0.395	Altoona, PA	0.139
Raleigh–Durham–Chapel Hill, NC	0.389	Ocala, FL	0.137
Gainesville, FL	0.387	Hickory–Morganton–Lenoir, NC	0.136
Champaign–Urbana, IL	0.38	Bakersfield, CA	0.135
Bryan–College Station, TX	0.37	Brownsville–Harlingen–San Benito, TX	0.134
Ann Arbor, MI	0.369	Lima, OH	0.134
Austin–San Marcos, TX	0.367	Yuba City, CA	0.132
State College, PA	0.363	McAllen–Edinburg–Mission, TX	0.129
Bloomington–Normal, IL	0.362	Johnstown, PA	0.127
Seattle–Bellevue–Everett, WA	0.359	Mansfield, OH	0.118
Rochester, MN	0.347	Vineland–Millville–Bridgeton, NJ	0.117
Santa Cruz–Watsonville, CA	0.342	Visalia–Tulare–Porterville, CA	0.115
Denver, CO	0.342	Danville, VA	0.113

Table 3
Cities with large and small changes in percent of college graduates between 1990 and 2000

Cities with the largest increases	(1)	*Cities with the largest increases*	(1)
San Francisco, CA	0.0848341	Corpus Christi, TX	0.0090711
San Jose, CA	0.0816702	Killeen–Temple, TX	0.0084904
Fort Collins–Loveland, CO	0.0750746	Yuba City, CA	0.0081555
Kenosha, WI	0.0739364	Las Cruces, NM	0.0073815
Odessa–Midland, TX	0.0718681	Salinas, CA	0.0070028
Roanoke, VA	0.0717642	Terre Haute, IN	0.0068426
Rochester, MN	0.0686671	Mansfield, OH	0.006075
Waterloo–Cedar Falls, IA	0.0670066	Montgomery, AL	0.0055255
New Bedford, MA	0.0666163	Utica–Rome, NY	0.0031253
Cedar Rapids, IA	0.0658358	Longview–Marshall, TX	0.0021987
Charleston–North Charleston, SC	0.0647602	Fresno, CA	0.001815
Colorado Springs, CO	0.0647194	Bakersfield, CA	0.0011622
Bloomington–Normal, IL	0.064214	Hickory–Morganton–Lenoir, NC	−0.001399
Asheville, NC	0.0640587	Huntington–Ashland, WV–KY–OH	−0.0046227
Madison, WI	0.0628737	Visalia–Tulare–Porterville, CA	−0.0081666
Seattle–Bellevue–Everett, WA	0.0622931	Lexington, KY	−0.0105446
West Palm Beach–Boca Raton, FL	0.0604797	Clarksville–Hopkinsville, TN–KY	−0.0128526
Fort Lauderdale, FL	0.0600323	Lafayette, IN	−0.0153243
Columbus, OH	0.0593833	Kalamazoo–Battle Creek, MI	−0.0327374
Baltimore, MD	0.0592615	Lafayette, LA	−0.0423326

lists the 15 cities with the largest and smallest changes in the per capita number of college-educated individuals between 1990 and 2000. The cities at the heart of the Silicon Valley boom experienced the largest increase. From 1990 to 2000, the share of college graduates in San Francisco and San Jose increased by 8 percentage points, almost three times the national average. The flow of young, highly-educated professionals and technicians attracted to Silicon Valley by the dot com boom in the second half of the 1990s is likely to have been a major reason for this impressive increase. As I show below, the 1990 fraction of high-tech jobs is a key predictor of the increase in the stock of human capital during the 1990s.

I now turn to a more formal analysis of the determinants of changes in the stock of human capital. Table 4 reports the coefficients from a regression of 1990–2000 changes

Table 4
Correlation between 1990 city characteristics and changes in percent of college graduates between 1990 and 2000

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
College		0.066							−0.019
		(0.020							0.031)
Population			0.002						0.000
			(0.001						0.001)
Family income				0.038					0.022
				(0.008					0.014)
Black					−0.003				−0.027
					(0.019				0.020)
Hispanic					−0.030				−0.032
					(0.012				0.018)
Immigrants						0.0002			0.023
						(0.017			0.026)
Agriculture							−0.312		−0.221
							(0.066)		(0.074)
Manufacturing							0.002		−0.029
							(0.024)		(0.028)
Hi tech								0.512	0.376
								(0.131)	(0.149)
Northeast	0.037	0.024	0.007	−0.374	0.038	0.037	0.040	0.035	−0.199
	(0.003)	(0.005)	(0.015)	(0.094)	(0.003)	(0.003)	(0.003)	(0.003)	(0.144)
Midwest	0.036	0.023	0.008	−0.373	0.037	0.036	0.041	0.034	−0.197
	(0.002)	(0.004)	(0.014)	(0.095)	(0.002)	(0.002)	(0.002)	(0.002)	(0.144)
South	0.028	0.015	−0.001	−0.378	0.031	0.028	0.032	0.025	−0.201
	(0.002)	(0.004)	(0.014)	(0.093)	(0.002)	(0.002)	(0.002)	(0.002)	(0.143)
West	0.032	0.017	0.003	−0.377	0.036	0.032	0.043	0.026	−0.201
	(0.002)	(0.005)	(0.014)	(0.093)	(0.002)	(0.002)	(0.002)	(0.002)	(0.144)

Note. Standard errors in parenthesis. Entries are OLS coefficients. The dependent variable is the 1990–2000 change in percent college. Each column is a separate regression. $N = 237$.

in college share on several geographic and socio-economic indicators in 1990. I begin by analyzing whether there are differences across U.S. regions in the change in college share. Column 1 indicates that the average increase over the decade was 3.7 percentage points for Northeastern cities, a slightly smaller increase for Midwestern and Western cities, and only 2.8 percentage points for Southern cities.

In column 2, I include the percentage of college graduates in 1990. Consistent with Figure 1, I find that cities that had a large share of college graduates in 1990 further increased their share over the course of the decade. The coefficient on college share – obtained by conditioning on regional dummies – is now 0.066 larger than the unconditional coefficient reported in Figure 1.

U.S. cities differ widely in size and per capita income and it is well known that in a cross section, the share of college graduates is positively correlated with these variables. What is interesting is that the concentration of human capital in cities that are relatively larger and richer increased further during the 1990s. In particular, columns 3 and 4 indicate that the overall fraction of college graduates grew faster between 1990 and 2000 in cities that were larger and richer in 1990.

Race and ethnic background are also important predictors of cross sectional differences in human capital across cities. But as it turns out, the 1990 percentage of Blacks is not a significant predictor of changes in college share between 1990 and 2000 (column 5). On the other hand, the percentage of Hispanics is negatively correlated with changes in college share. One might expect that the fact that cities with larger Hispanic population in 1990 experienced relatively smaller increases in college share is explained by the inflow of unskilled immigrants. It is well-documented that immigrants tend to migrate to cities with high densities of immigrants. However, in column 6 I find little evidence that the 1990 fraction of immigrants is correlated with 1990–2000 changes in college share.

I now turn to the industrial structure of cities. The 1990 percentage of manufacturing jobs appears to be uncorrelated with changes in college share, while the percentage of agricultural jobs is negatively correlated with changes in college share (column 7). Perhaps the most interesting result on the correlation between industry structure and human capital is in column 8, where I focus on the relationship between the 1990 share of high-tech jobs in a city and the 1990–2000 change in college share.[3] I find that the share of high-tech jobs in a city is a strong predictor of change in college share. This is consistent with the finding in Table 2 that San Jose and San Francisco experienced the largest increase in college share over the 1990s. Finally, column 9 reports results from a specification where all the variables are included.

[3] To classify jobs as high-tech or low-tech, I use the definition of high-tech industries provided by the American Electronics Association (1997) based on 45 4-digit SIC codes. The definition is based on SIC codes, which are not exactly equivalent to the Census industry definition. In my analysis, high-tech industries include: Computers and related equipment; Scientific and controlling instruments; Guided missiles, space vehicles, and parts.

3. Theories of social returns to education

Through the chapter, I use the terms "human capital" and "education" interchangeably. The focus of this chapter is mainly empirical. Although human capital is in theory a broader concept than education, in practice most empirical studies use education to measure human capital.

After four decades of debate, there seems to be a consensus on the magnitude of the private benefits of human capital. Most empirical studies indicate that all else equal, individuals with one extra year of schooling earn about 8–12% more per year. Yet economists have speculated for at least a century that education may have additional benefits that are not reflected in the private return. If this is indeed the case, the social return to education will exceed the private return. By social return to education I mean the sum of all the benefits that accrue to society resulting from an increase in the overall level of education.

The social return to education differ from the private in the presence of externalities. I will consider three type of externalities. First, I consider what I will call productivity spillovers. Productivity spillovers arise if the presence of educated workers makes other workers more productive. In the presence of such spillovers, an increase in aggregate human capital may have an effect on aggregate productivity that is quite different from the effect of an increase in individual education on individual earnings. A large body of theoretical literature in urban and macroeconomics has argued that these types of spillovers are important determinants of economic growth.

On the other hand, it is also possible that education generates negative spillovers. For example, if education functions as a signal of productive ability, rather than enhancing productivity directly, the private return may exceed the social return.[4] In this case, increases in average schooling in a labor market may result in increases in earnings that are smaller than the private return.

Second, education may reduce the probability of engaging in activities that generate negative externalities. The most obvious example is the effect of education on criminal activities. If education reduces an individual's incentive to commit a crime, then cities with a better-educated populace will enjoy lower crime rates. Finally, economists like Adam Smith, Milton Friedman and others have argued that a better-educated electorate makes better decisions on policy issues that affect the collectivity. If this is true, cities and states with a better-educated population will elect better representatives and enact better public policies.

In the remainder of this section, I briefly review the theoretical arguments that have been proposed in support of these three sources of human capital externalities. In Sections 4 and 5 I describe the most recent empirical evidence on the magnitude of each of these externalities.

[4] This is a case where people with higher innate ability signal their higher innate productivity by enduring extra years of schooling.

For policy implications, it is important to keep in mind that not every spillover is necessarily a market failure that requires government intervention. One can think of many spillovers that are internalized. For example, an increase in the number of high skilled workers may generate positive spillovers that benefit productivity of low skill workers in the same firm. One reason for such increase in the productivity of low skilled workers is the imperfect substitution between high skill and low skill workers. Another reason is the presence of learning spillovers, if low skilled workers acquire better skills in the presence of high skilled workers. In either case, these within-firm spillovers are likely to be internalized and will be reflected in higher wages for educated workers. (I will come back to the issue of imperfect substitution in more detail in Section 4.2.1.) In this chapter, most of the analysis focuses on spillovers between firms, which are hard to internalize and therefore are market failures.

Because the geographic scope of externalities does not need to be the same for all types of externalities, the social return for a city does not need to equal the social return for a state or a country. The geographic scope of externalities is important for policy implications. For example, if spillovers have only local effects, one would argue in favor of Pigouvian subsidies to education financed at the local level, similar to those currently in place in the United States. If, however, spillovers from schooling have a broader geographical scope, so that their benefits are realized at a national level, then one would argue in favor of a federal role in public education.

From the point of view of local governments, one problem with subsidizing the production of human capital is that human capital is mobile, so that the link between production and utilization of human capital is not clear a priori. In a recent paper, Bound et al. (in press) study the relationship between production of college graduates, and their geographical distribution. Bound et al. (in press) argue that because college graduates are highly mobile, states and counties that generate large flows of new college graduates are not necessarily the ones where college graduates tend to locate.

On one hand, the production of a large number of college-educated residents in an area may lead to increases in the employment of skilled workers, if industries that are human capital intensive locate there. Examples of this phenomenon include Silicon Valley (electronics), Cambridge, MA (biotech and pharmaceutical), San Diego (medical, biotech, pharmaceutical). On the other hand, given graduates' high mobility, the link between production of college graduates and stock of college graduates may be weak. This clearly has important policy implications for states that invest heavily in public education since it is not obvious a priori what the return on such an investment is. Results in Bound et al. (in press) are not too encouraging for states like Michigan or Ohio that invest heavily in their system of public higher education. Estimates based on Census data indicate that the link between the production and use of BA degree recipients is modest. States awarding relatively large numbers of college degrees do have somewhat higher concentrations of college-educated workers, but the effect is not very large.

3.1. Productivity spillovers

The question of whether education raises a person's productivity and earnings has generated a large body of empirical literature.[5] The consensus that has emerged is that a worker's schooling does in fact raise her productivity and her earnings. For the United States in the 1990s, the private return to schooling is believed to be about 8–12%: each extra year of schooling appears to be associated with an 8–12% increase in earnings.[6]

Much less is known about the more important question of what happens to productivity and wages when the aggregate stock of educated workers increases. The fact that employers pay individual workers 8–12% more for each extra year of schooling does not necessarily imply that raising the average education in a city (or state or nation) by one year would result in a 8–12% increase in aggregate earnings. I will consider three reasons why the social return to schooling – as measured in terms of increased aggregate earnings or aggregate income – may differ from the private return that has received so much attention in the literature.

Technological externalities. While many different explanations have been proposed for positive externalities, these models can be grouped in two broad families, that I will call technological externalities and pecuniary externalities. In the first class of models, externalities are built into aggregate production functions in the form of technological increasing returns. Learning through social interaction is often cited as the mechanism through which externalities arise. Marshall (1890) is the first to argue that social interactions among workers in the same industry and location create learning opportunities that enhance productivity.

Perhaps the most influential example of the class of models where externalities are built into aggregate production functions in the form of technological increasing returns is a paper by Lucas (1988). In that paper, human capital is assumed to have two effects. First, an individual's own human capital has the standard effect of increasing her own productivity. Second, the average aggregate level of human capital contributes to the productivity of all factors of production. This second effect is an externality, because "though all benefit from it, no individual human capital accumulation decision can have an appreciable effect on average human capital, so no one will take it into account" in deciding how much to invest in human capital accumulation. In Lucas' view, human capital externalities may be large enough to explain long-run income differences between rich and poor countries.

What are the mechanisms that generate these human capital externalities? In Lucas' model the externality is simply built into the production function, but Lucas goes on

[5] Although college-educated individuals clearly earn more than high-school graduates, it is possible that college graduates have higher earnings potential because of innate ability, family background, ambition and determination. If these unmeasured workers characteristics are important, college graduate would earn more than high-school graduates even in the absence of a college education.

[6] See Card (1999) for a comprehensive survey of the evidence on the private return to schooling.

to argue that the sharing of knowledge and skills through formal and informal interaction is the mechanism that generates positive externalities across workers.[7] More recent models build on this idea by assuming that individuals augment their human capital through pairwise meetings with more skilled neighbors at which they exchange ideas.[8] See Duranton and Puga (2004) for a detailed survey of this class of models.

Pecuniary externalities. A second class of models explain positive human capital externalities as pecuniary externalities. Labor market pooling externalities were first proposed by Marshall (1890). One recent example is a model where job search is costly, and spillovers from education arise because of the complementarity between physical and human capital [Acemoglu (1996)]. Because of the complementarity between physical and human capital, the privately optimal amount of schooling depends on the amount of physical capital a worker expects to use. The privately optimal amount of physical capital depends on the education of the workforce. If a group of workers in a city increases its level of education, firms in that city, expecting to employ these workers, would invest more in physical capital. Since search is costly, some of the workers who have not increased their education would end up working with more physical capital and hence earn more than similar workers in other cities.

As in Lucas, the presence of skilled workers in a city generates external benefits for other workers there. Both Lucas and Acemoglu agree that the average wage of unskilled workers in a city increases with the average human capital of the labor force. But what

[7] In Lucas' words:

> We know that there are group interactions that are central to individual productivity and that involve groups larger than the immediate family and smaller than the human race as a whole. Most of what we know we learn from other people. We pay tuition to a few of these teachers, either directly or indirectly by accepting lower pay so we can hang around them, but most of it we get for free, and often in ways that are mutual – without a distinction between student and teacher. Certainly in our own profession, the benefits of colleagues from whom we hope to learn are tangible enough to lead us to spend a considerable fraction of our time fighting over who they shall be, and another fraction travelling to talk with those we wish we could have as colleagues but cannot. We know that this kind of external effect is common to all the arts and sciences – the 'creative professions'. All of intellectual history is the history of such effects.

But, Lucas argues, the external effect of human capital in not limited to academia:

> Much of economic life is creative in much the same way as is art and science. New York City's garment district, financial district, diamond district, advertising district and many more are as much intellectual centers as is Columbia or New York University.

[8] See, for example, Glaeser (1999), Peri (2002), Jovanovic and Rob (1989). Black and Henderson (1999) specifically explore human capital spillovers in urban areas. They investigate how urbanization affects the efficiency of the growth process and how growth in turn affects urbanization when human capital generates positive spillovers. Other authors focus on the importance of basic research in fostering technological innovation and productivity, the public good nature of this type of research, and the resulting positive externalities in the form of knowledge spillovers. See, for example, Arrow (1962), Griliches (1986), Jaffe, Trajtenberg and Henderson (1993), and Saxenian (1994).

distinguishes Acemoglu's story from Lucas' story is that this result does not follow from assumptions on the production function, but rather is derived from market interactions. Even though all the production functions of the economy exhibit constant returns to scale in Acemoglu, the complementarity of human capital and physical capital coupled with frictions in the job search process, generates a positive relationship between average wage and average human capital, holding constant workers' individual human capital.[9]

Although differences across cities in their *quantity* of physical capital play a central role in this model, differences in the *quality* of physical capital (technology) could arguably generate similar conclusions. Specifically, if skills and technology are complementary, it is plausible to assume that the privately optimal amount of human capital depends not only on the amount of physical capital a worker expects to use, but also on the technological level that characterizes such capital. Similarly, in models with endogenous skill-biased technical change, an increase in the supply of educated workers increases the size of the market for skill-complementary technologies and stimulates the R&D sector to spend more effort upgrading the productivity of skilled workers [Acemoglu (1998)].

Negative externalities. Another reason why the social return to schooling, as measured in terms of increased aggregate earnings, may differ from the private return is the presence of negative externalities. If education functions as a signal of productive ability, rather than enhancing productivity directly, the private return may exceed the social return. This is a case where people with higher innate ability signal their higher innate productivity by enduring extra years of schooling. If schooling is more difficult for individuals with low innate productivity than individuals with high innate productivity, then, even if schooling itself is worthless in terms of enhancing productivity, it still may be a useful screening device for employers to identify more productive job applicants. This possibility is important because it implies that one extra year in average schooling in a city (or state or nation) would result in less than an 8–12% increase in aggregate earnings.

In the most extreme version of the model, a one-year increase in average schooling in a city would have no effect on earnings. Employers would simply increase their hiring standard, and everyone would end up at the same jobs they would have had without the increase in education. In this extreme case, the private return to schooling would be 8–12%, but the social return would be 0. Although this is certainly possible in theory, this scenario is unlikely to be relevant in practice. The existing empirical evidence on private returns to schooling indicates that education has a causal effect on productivity.

[9] Empirically, manufacturing plants located in cities that have a more skilled labor force do tend to have a larger stock of human capital. This is true both in a cross section of cities and when looking at within cities changes over time. (Author's calculation using firms data from the Census of Manufacturers matched with workers data from the Census of Population.)

3.2. Crime

Besides its effects on productivity and earnings, human capital may also reduce the probability that an individual engage in socially costly activities, such as crime. Crime is a negative externality with enormous social costs. If education reduces crime, then schooling will have social benefits that are not taken into account by individuals, and most of this benefit is likely to be realized at the local level: cities with high levels of education would have lower crime rates. Given the large social costs of crime, even small reductions in crime associated with education may be economically important.

There are a number of reasons to believe that education can reduce criminal activity. First, schooling increases the returns to legitimate work, raising the opportunity costs of illegal behavior. Additionally, punishment for criminal behavior often entails incarceration. By raising wage rates, schooling makes any time spent out of the labor market more costly.

Second, schooling may directly affect the psychic rewards from crime itself. For example, Arrow (1997), discussing the social benefits of education, argues that

> Like everything else interesting about human beings, preferences are a mixture of hereditary and environment. Schools must surely have a major part, if only because they occupy a large part of a child's day. It is a traditional view that not only does education influence values but it ought to do so.

Third, schooling may alter preferences in indirect ways, which may in turn affect decisions to engage in crime. For example, education may increase one's patience (as in Becker and Mulligan (1997)) or risk aversion. A lower discount rate or higher risk aversion will reduce the probability that an individual will engage in criminal activities.

3.3. Voting

Many economists have argued that education provides social benefits through enhanced political behavior. Among many other authors, Hanushek (2002), makes this argument in his survey of public education. Interestingly, the argument that education generates externalities by improving the political behavior of voters resonates both with noted advocates of a limited role for government – such as Adam Smith and Milton Friedman – as well as with liberal proponents of a larger role of government in the economy. For example, a document by the liberal Center on National Education Policy (1996) lists several benefits of public education, including the preparation of "people to become responsible citizens". Forty years ago, Friedman (1962) made exactly the same point:[10]

[10] Even earlier, Adam Smith (1776) emphasized the benefits of increased cognitive capacity among the common people, claiming that: they are more disposed to examine, and more capable of seeing through, the interested complaints of faction and sedition, and they are, upon that account, less apt to be misled into any wanton or unnecessary opposition to the measures of government.

> A stable and democratic society is impossible without a minimum degree of literacy and knowledge on the part of most citizens and without widespread acceptance of some common set of values. Education can contribute to both. In consequence, the gain from education of a child accrues not only to the child or to his parents but also to other members of the society. The education of my child contributes to your welfare by promoting a stable and democratic society. There is therefore a significant "neighborhood effect". [...] Most of us would probably conclude that the gains are sufficiently important to justify some government subsidy.

Why might education affect political behavior? First, and most importantly, more-educated voters may have more information on candidates' and political parties' positions. The fact that better-educated citizens are likely to be more informed voters may be due to active accumulation of information during campaigns (higher newspaper readership, for example), or to a better ability to process a given amount of information (if, for example, education improves cognitive skills). According to this argument, better-educated citizens are in a position to make more informed choices at election time. By choosing better candidates, they create an externality that may benefit all citizens. A second channel through which education might affect political behavior is if education increases civic participation, for example, by raising voter turn-out rates. If increased civic participation improves social decision-making, then education may also affect the quality of political decisions.[11] If enhanced political behavior produces social benefits, then Pigouvian subsidies for education may produce more efficient education acquisition decisions. Unlike the social benefits of schooling that arise because of reductions in criminal activities, the benefits of schooling that accrue because of improved political behavior are not necessarily limited to a local area. State and national elections may benefit residents in other cities and states.

4. Estimating productivity spillovers in cities

In Section 3.1, I described alternative theoretical models based on productivity spillovers. In this section, I discuss the challenges that arise in the estimation of these spillovers, and I summarize the existing empirical evidence on their magnitude.

I begin in Section 4.1 by presenting a simple equilibrium framework that helps identify three possible strategies for estimating human capital spillovers in cities:

(1) by comparing the output (or productivity) of firms located in cities with high and low level of aggregate human capital;
(2) by comparing the wages of workers located in cities with high and low level of aggregate human capital;
(3) by comparing cost of land in cities with high and low levels of aggregate human capital.

[11] Different models have been proposed in which increased civic participation lead to better outcomes. See for example Osborne, Rosenthal and Turner (2000) and Feddersen and Pesendorfer (1996).

I then discuss possible empirical strategies for estimating spillovers using these three models. The fundamental issue in the interpretation of these three models is the presence of unobservable determinants of productivity, wages or rent that are correlated with aggregate human capital across cities. The equilibrium framework suggests that simple OLS models are likely to be biased, but the sign of the bias is not obvious a priori. I discuss alternative identifications strategies to account for unobserved heterogeneity.

A considerable portion of the discussion is devoted to empirical models based on wages (Section 4.2), since these are most prevalent in the existing literature. I discuss in detail the interpretation of wage equations in the presence of externalities, since this interpretation is complicated by the fact that increases in the stock of skilled workers in a city may affect the wage distribution even in the absence of externalities, if skilled and unskilled workers are imperfect substitutes. I review the findings of some of the existing empirical studies based on wages, productivity and land prices. Finally, in Section 4.3 I turn to models based on firm productivity. I discuss the interpretations of these models, and the existing empirical evidence.

4.1. Equilibrium with spillovers

I begin by presenting a simple general equilibrium framework of perfect competition that includes both standard demand and supply factors and spillovers from human capital. The framework identifies the effect of an increase in the relative supply of educated workers in a city on the productivity, land prices, and wages of skilled and unskilled workers. The framework, which is based on models in Roback (1982), Moretti (in press, 2002), aims to make two points. First, it indicates how human capital spillovers can be measured by relating differences across cities in firms' productivity, land prices, or wages to differences in the overall level of human capital. Second, the model identifies potential sources of unobserved heterogeneity that might bias empirical estimates of the spillover.[12]

4.1.1. Framework

The intuition is quite simple. If there are spillovers, firms and workers are more productive in cities with high overall levels of human capital. In equilibrium, firms are indifferent between cities because wages and land prices are higher in cities with high overall levels of human capital, and lower in cities with low overall levels of human capital, making unit costs similar across cities. Similarly, workers are indifferent because housing prices are higher in cities with high overall levels of human capital.

Consider two cities, A and B, and two types of labor, educated and uneducated. Workers and firms are perfectly mobile. The market structure is assumed to be perfectly competitive, so that the profits of firms are assumed to be zero. Assume that there are two

[12] See Gabriel and Rosenthal (2004) for a recent empirical application of the Roback model.

types of goods, a composite good y – nationally traded – and land h – locally traded. Each city is a competitive economy that produces y combining skilled and unskilled workers (N_1 and N_0) and capital: $y = Ag(N_0, N_1, K)$.

To introduce the possibility of human capital spillovers in the model, I allow the productivity of plants in a city to depend on the aggregate level of human capital in the city, S: $A = f(S)$. This specification is consistent with most of the explanations of spillovers suggested in the literature and described in the previous section. Note that this specification assumes that the spillover augments both capital productivity and labor productivity. Alternative assumptions are possible. For example, one could assume that human capital spillovers benefit only labor productivity, or alternatively, that they benefit only capital productivity. Empirically, it is difficult to distinguish between these alternative specifications.

Cities differ in the amenities that they offer. Workers maximize utility subject to a budget constraint by choosing quantities of the composite good and residential land, given the city amenity, v'. Because the composite good, y, is traded nationally, its price is the same everywhere and set equal to 1. Variations in the cost of living between cities depend only on variations in the price of land, p, which is assumed to be the same for all workers in the same city, irrespective of the education group. The quantity of land is fixed. Because of the perfect mobility and perfect competition assumptions, equilibrium is obtained when workers have equal utilities in all cities and firms have equal unit costs across cities.

The equilibrium for the simple case of only two cities, A and B, is described in Figure 3. The upward sloping lines in each panel represent indifference curves for the two education groups. Indirect utility of workers belonging to group j, $V_j(w_j, p, v')$, is a function of the group's nominal wage, w_j, cost of land and the amenity. The indifference curves are upward sloping because workers prefer high wages and low rent. Since workers are free to migrate, utility of workers is equalized across locations:

$$V_1(w_1, p, v') = k_1 \quad \text{and} \quad V_0(w_0, p, v') = k_0$$

for educated and uneducated workers, respectively. The downward sloping lines show combinations of wages and rents which hold constant firms' unit costs: $C_c(w_0, w_1, p)$ =1, where w_0 and w_1 are wages of uneducated and educated workers, respectively; and c indexes city. (If production functions vary across cities, for example because of spillover effects, then the unit cost functions are city-specific.) A zero-profit condition for the firms ensures that production takes place along the downward sloping curve. Thus the model has three equations (unit cost and indirect utility for each skill group) in three unknowns (w_0, w_1 and p). Point 1 in the left panel of Figure 3 represents the equilibrium combination of the educated workers' wage and the price of land in city A. Point 1 in the right panel represents the same combination for uneducated workers.

If the two cities are identical, the equilibrium in city B is the same. However, there are two ways to make the overall level of human capital higher in city B than in city A – either by increasing the relative supply of educated workers in city B, or by increasing the relative demand for educated workers in city B. I begin by considering what happens

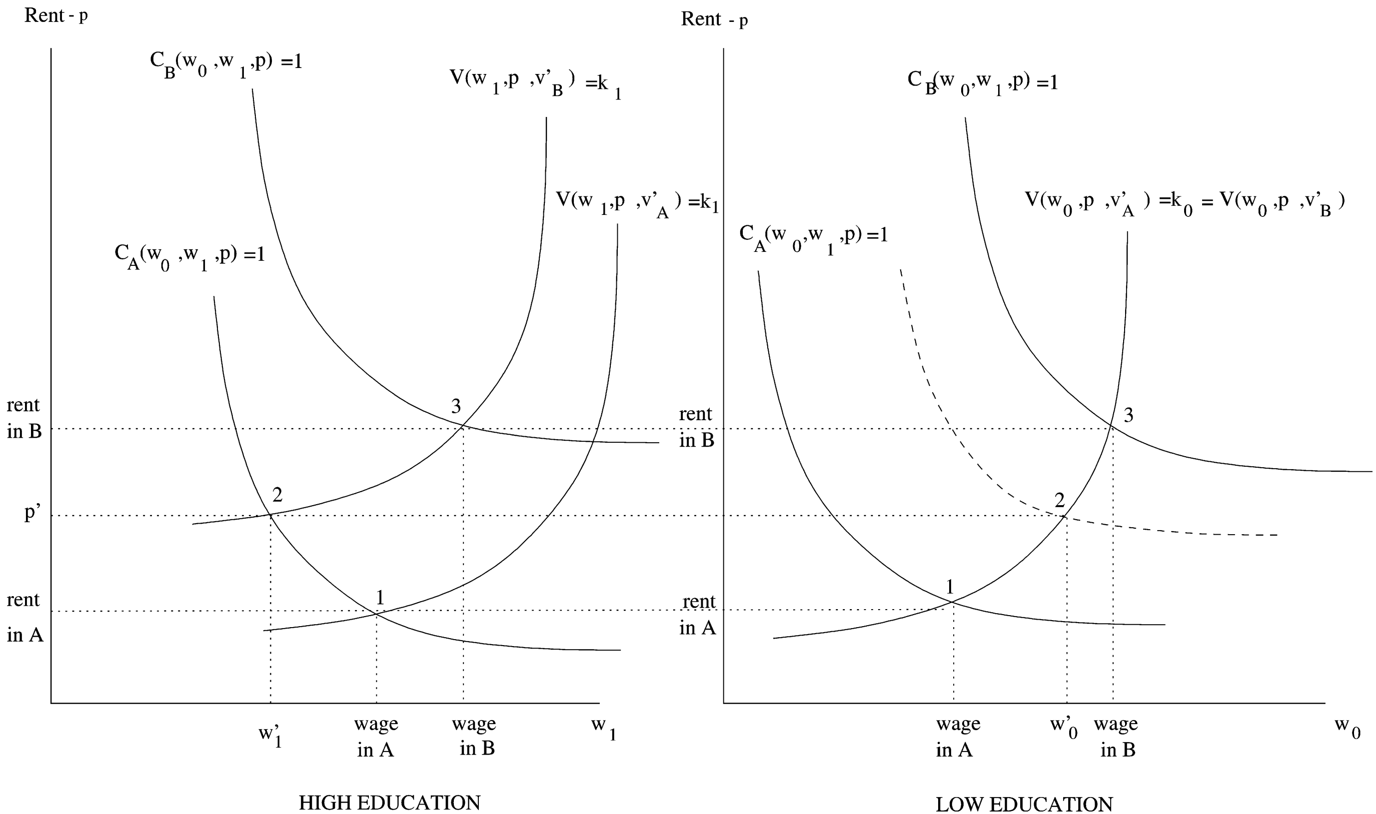

Figure 3. Equilibrium wages and rent when amenities differ across cities. Point 1 is the equilibrium in city A. Point 2 is the equilibrium in city B without spillover. Point 3 is the equilibrium in city B with spillover. The dashed lines in both panels are the isocost curves in city B without spillover. w_1 and w_0 are the nominal wage of educated and uneducated workers, respectively.

to equilibrium wages when the relative supply of educated workers is higher in B than in A.

One way of making the relative supply of educated workers higher in B than in A is to assume that city B has an higher level of the local amenity than city A ($v'_B > v'_A$) and educated workers value the amenity, while uneducated workers do not. It is important to note that, in this general framework, I interpret v' broadly, as any *exogenous* factor that increases the relative supply of educated workers.

As shown in Figure 3, the indifference curve at level k_1 of educated workers in city B is to the left of the corresponding curve in city A, while the indifference curve for uneducated workers does not change. In this context, even without externalities, the wage of the uneducated workers is higher. If there are no spillovers, the increase in the supply of educated workers in city B raises the wage of uneducated workers to w'_0 and lowers the wage of educated ones to w'_1 (point 2 in both panels of Figure 3). This is the standard result. Because of imperfect substitution, uneducated workers are now more productive in city B and because of the amenity, educated workers accept lower wages there.[13]

In the presence of spillovers, however, the combinations of wages and rents that hold firms' costs constant in city B lies to the right of the corresponding combination in city A for both groups (point 3). For educated workers, the shift of the isocost curve is caused by the spillover only; for uneducated workers the shift is caused by both complementarity (movement from 1 to 2) and the spillover (movement from 2 to 3). The distinction between complementarity and spillovers is important both for theoretical reasons as well as for policy implications. (Complementarity is clearly not a market failure). Below, I discuss how it is possible to empirically distinguish between complementarity and spillovers.

So far I have considered the case where differences in the relative number of educated workers in city A and city B are driven by differences in the relative supply. I now turn to the case where differences in the relative number of educated workers are driven by differences in the relative demand for educated workers. In Figure 4 cities are identical in term of amenities, but differ in term of technology, T. I interpret T broadly, as any *exogenous* factor that increases the relative productivity of educated workers and therefore their relative demand. To make technology differences more explicit in Figure 4, T appears in the isocost: $C(w_0, w_1, p, T)$. (Since cities are identical, the amenity is dropped from the indifference curves.) Suppose that, because of technological differences, skilled workers are particularly productive in city B and demand for them is high. Attracted by higher wages, skilled workers move to city B. In so doing, they raise average education there. Point 2 represents the equilibrium in city B if there are no spillovers. The wage of educated workers is higher because technology makes them

[13] For simplicity, we follow Roback (1982) and take the level of utility k_1 and k_0 as parameters for simplicity. Closure of the model would require that the level of utility is made endogenous. This would complicate the model, without changing its implications.

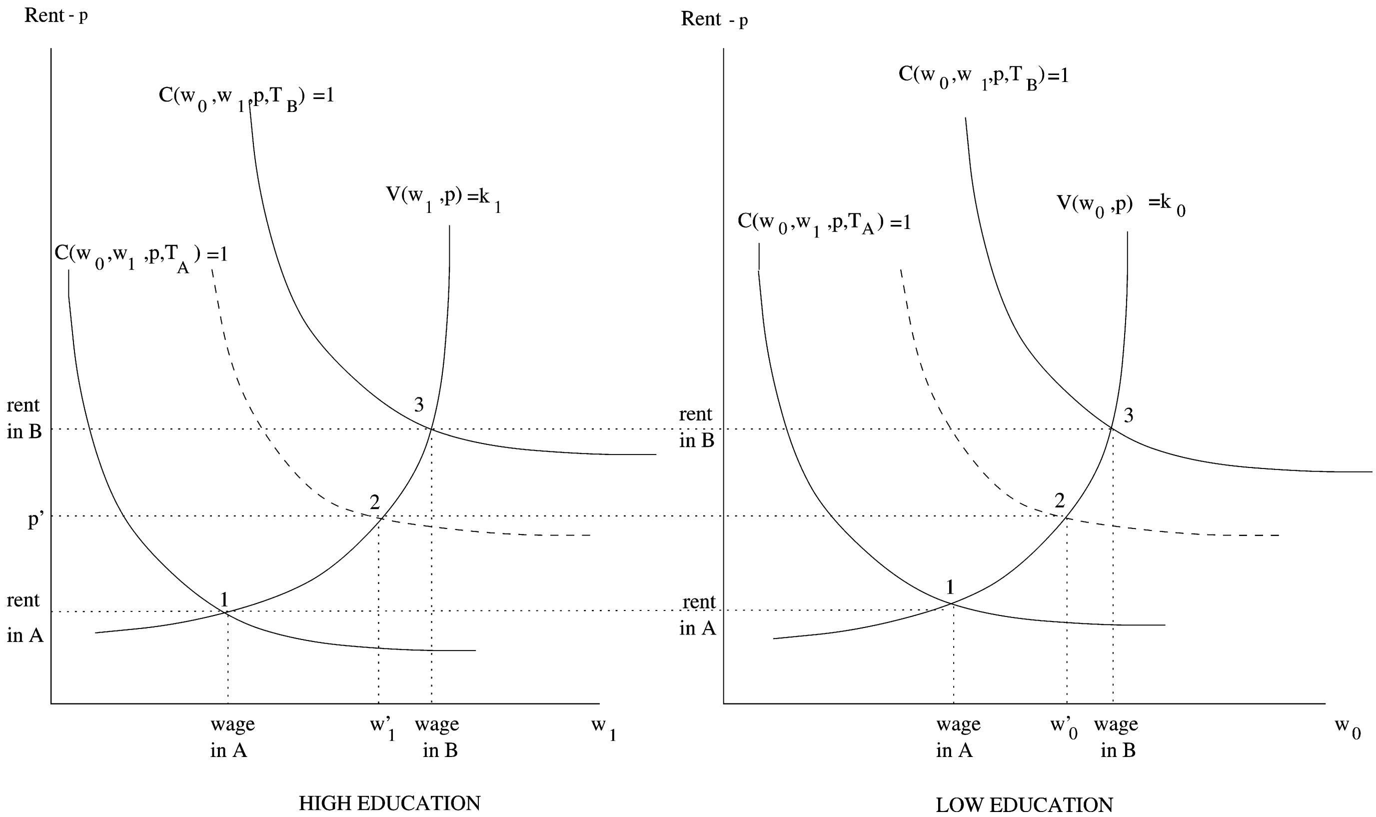

Figure 4. Equilibrium wages and rent when technology differs across cities. Point 1 is the equilibrium in city A. Point 2 is the equilibrium in city B without spillover. Point 3 is the equilibrium in city B with spillover. The dashed line in the right panel is the isocost curve in city B without spillover. w_1 and w_0 are the nominal wage of educated and uneducated workers, respectively.

more productive, while the wage of uneducated workers is higher because of complementarity. In the presence of spillovers effects, the isocost curve shifts further to the right. In this case, the true spillover effect is a shift from 2 to 3, but the observed effect is larger, from 1 to 3.

In equilibrium, both skill groups are present in both cities. Since workers are free to migrate from city A to city B, why are equilibrium wages – net of the compensating differential – not driven to equality? In this model, migration to high-wage cities leads to higher rent, making workers indifferent between cities. Although in equilibrium workers in cities with higher human capital earn higher nominal wages than workers in cities with low human capital, in real terms workers in cities with high human capital are not better off because land is more expensive.[14] A similar intuition holds for firms. Since firms are free to relocate from A to B, why is productivity not driven to equality? Wages and rent are higher in city B, making firms indifferent between cities.

Note that in this context, where cities are small open economies that face a perfectly elastic supply of labor at a fixed utility level, landowners in cities with high levels of human capital are the only real beneficiaries of the spillovers. Because land is the only immobile factor in this model, all the rent generated by the externality in terms of increased productivity is capitalized in land prices. The policy implications are not obvious. On one hand, the common U.S. system of financing public education with local property taxes seems efficient. Since landowners are the beneficiaries of the spillover, taxing land may work to internalize the externality. On the other hand, workers are mobile and heterogeneous in their tastes, and localities differ in their amenities. Under these circumstances, it is possible that municipalities that invest heavily in schooling may retain only some of the benefits. Black and Henderson (1999) present a theoretical discussion of whether local governments can successfully internalize human capital externalities.[15] And Bound et al. (in press) undertake an empirical investigation of the mobility of college graduates and its implications for local and state education policies.

It is important to realize that the model presented here assumes that cities are small relative to the whole economy, and face a perfectly elastic supply of labor at a fixed utility level, so that technology shocks do not affect the average technology for the whole economy. If the number of cities is limited and cities are large relative to the whole economy, conclusions are different, because general equilibrium effects may arise. For example, if a large city experience a large technology shock, this will result in an improvement of the average technology for the whole economy, and utility levels will in general rise.

[14] Other models achieve the same result assuming that quality of life is declining in the size of the city [Glaeser, Scheinkman and Shleifer (1995)].

[15] Black and Henderson (1999) use a dynamic framework that is more general than the one presented here, because it allows spillovers to affect economic growth.

4.1.2. Implications for empirical analysis

The model developed in the previous section predicts that the productivity of firms is higher in cities with higher overall levels of human capital. Because workers are more productive, wages are also higher in cities with higher overall levels of human capital. But for this to be an equilibrium, land prices must adjust to make workers and firms indifferent. A useful implication of this model is that there are three possible empirical strategies to identify the magnitude of human capital spillovers. We can compare productivity, wages or land prices between cities.

Using the first metric, the magnitude of the spillover can be identified by taking the difference in the unit cost functions in city A, the city with low levels of human capital, and city B, the city with high levels of human capital, holding constant the price of factors:

$$\ln c(w_{0\mathrm{A}}, w_{1\mathrm{A}}, p, S_{\mathrm{A}}) - \ln c(w_{0\mathrm{A}}, w_{1\mathrm{A}}, p, S_{\mathrm{B}}). \tag{1}$$

If there are spillovers, unit costs are lower in city B than in city A, holding constant wages and land prices.[16] Although appealing in theory, an estimation strategy based on the comparison of unit costs across cities like the one suggested by Equation (1) is hard to implement in practice because of data limitations. Large scale datasets with information on production costs for many firms in many cities are hard to obtain.

On the other hand, data on inputs and output are more readily available.[17] So, instead of identifying spillovers by comparing unit costs of otherwise identical plants located in cities with high and low levels of human capital and holding input prices fixed, one can more easily identify spillovers by comparing the output of otherwise identical plants located in cities with high and low levels of human capital, holding input quantities fixed. In the notation of the simple example in the previous section, spillovers can be measured by taking the difference in the production functions of city B and city A, holding labor and capital constant:

$$\ln\big[f(S_{\mathrm{A}})g(N_0, N_1, K)\big] - \ln\big[f(S_{\mathrm{B}})g(N_0, N_1, K)\big] = \ln f(S_{\mathrm{A}}) - \ln f(S_{\mathrm{B}}). \tag{2}$$

The second option is to measure the magnitude of the spillovers in term of land prices. The model in the previous section shows that the spillover is capitalized in land prices. If data on property values in different cities are available, estimates of the spillovers can be obtained by simply measuring differences in land prices between cities with high levels of human capital and cities with low levels of human capital. In terms of the example in the previous section, the magnitude of the spillover is simply the difference in housing

[16] Note that I now write the cost function c as a function of human capital in the city. This reduced form representation of the cost function captures the idea that in cities with higher human capital, total factor productivity is higher, so the same amount of output can be produced with less inputs.

[17] The most prominent example is the Census of Manufacturers, which provides longitudinal data on the universe of U.S. manufacturing establishments with 1 employee or more. The Census of Manufacturers has detailed information on output produces as well as capital and labor used in production.

prices between city B and city A: $(p_{\mathrm{B}} - p')$. Graphically, this is the difference in rent between point 2 and point 3 in Figure 3 or 4.

Two caveats need to be considered. First, the model assumes that land is fixed, which may not be always true. Second, there is the issue of how to empirically measure $(p_{\mathrm{B}} - p')$. Because data on land prices are difficult to obtain, researchers often rely on housing prices (adjusted for housing characteristics), which are readily available for most large cities (for example in the Census of Population and Housing). One limitation of using housing prices is that the stock of housing is not necessarily fixed.

Finally, one can use wages to measure spillovers. Most of the existing empirical studies that attempt to quantify the magnitude of human capital spillovers have focused on wages. In theory, one might think of using the difference in the wage of educated workers, $(w_{1\mathrm{B}} - w_1')$, or the difference in the wage of uneducated workers $(w_{0\mathrm{B}} - w_0')$ in the two cities, or a weighted average of the two:

$$\frac{N_1}{N_1 + N_0}\left(w_{1\mathrm{B}} - w_1'\right) + \frac{N_0}{N_1 + N_0}\left(w_{0\mathrm{B}} - w_0'\right).$$

Graphically, the difference in the wage of educated workers is the distance between point 2 and point 3 in the left panel of Figure 3 or 4 and the difference in the wage of uneducated workers is the distance between point 2 and point 3 in the right panel of Figure 3 or 4.

Three points are important here. First, nominal wages should be used in the empirical analysis. Wages adjusted for cost of living are not the correct dependent variable. The reason is that higher nominal wages in a city imply greater productivity. If workers were not more productive, firms producing goods that are traded nationally (such as manufacturing goods) would leave high-wage cities and relocate to low-wage cities. Some workers are employed in industries that produce output that is not traded nationally (for example, local services). But firms producing traded goods face the same price everywhere in the nation, so that, as long as there are some firms producing traded goods in every city, average productivity has to be higher in cities where nominal wages are higher [Acemoglu and Angrist (2000)].

Second, it is important to recognize that wage changes affecting workers in a city not only capture human capital spillovers, but also capture the complementarity (or imperfect substitutability) between skilled and unskilled workers. If skilled and unskilled workers are imperfect substitutes, unskilled workers benefit from an increase in the number of skilled workers even in the absence of any externality. Therefore, the average effect on wages,

$$\frac{N_1}{N_1 + N_0}\left(w_{1\mathrm{B}} - w_1'\right) + \frac{N_0}{N_1 + N_0}\left(w_{0\mathrm{B}} - w_0'\right)$$

reflects both the spillover effect and imperfect substitution between high- and low-education workers. The distinction is important, because, unlike human capital externalities, complementarity is not a market failure. In Section 4.2.1, I formally show the difference between complementarity and spillovers and I suggests two ways to empirically separate the two.

Finally, even controlling for the complementarity effect, the difference in wages between cities with high and low human capital is not exactly equal to the spillover, because land prices also adjust. Only in the case where no land is used in commercial production will the wage difference between cities with high and low human capital equal the spillover.

4.1.3. The consequences of unobserved heterogeneity

The discussion so far has ignored the possible presence of confounding factors that may introduce spurious correlation in the relationship between wages (or productivity or land prices) and aggregate human capital. There are many unobserved characteristics of workers and cities that affect wages and at the same time may be correlated with the overall level of human capital. A goal of the model is to identify potential sources of unobserved heterogeneity that might bias empirical estimates of the human capital spillover.

In the stylized framework developed in Section 4.1.1, unobserved heterogeneity is of two types: demand shocks that affect the relative productivity of workers with high human capital in a city; and supply shocks, that affect the relative attractiveness of a city for high human capital workers. As mentioned above, these demand and supply shocks need to be interpreted broadly, as any factor that affects the relative demand or supply of skilled workers and that is unmeasured by the econometrician.

In the presence of unobserved heterogeneity that is correlated with aggregate human capital, OLS regressions of wages on aggregate human capital can be biased upward or downward depending on the relative magnitude of unobserved demand and supply heterogeneity. To see this, consider first Figure 3, where variation in the relative number of educated workers across cities is driven by *supply* factors. To the extent that the amenity that attracts skilled workers to city B is not observed, this unobserved heterogeneity biases the OLS coefficient in a regression of wages of educated workers on share of educated workers *downward*. In Figure 3 (left panel), the true spillover is the difference between the wage at point 3 and the wage at point 2. The observed effect is instead the difference between the wage in point 3 and the wage in point 1, which is smaller than the true spillover. The intuition is straightforward. The compensating differential that skilled workers implicitly pay for the amenity is unobserved, and enters the wage of skilled workers as a negative city-specific residual. The correlation between this residual and average education is negative, as skilled workers trade some of their wage for the amenity, so that the OLS coefficient on average education is biased down.

The opposite bias arises from heterogeneity in relative labor demand. Consider Figure 4. The size of the spillover is the size of the shift from 2 to 3. But if T is unobserved, the OLS coefficient in a regression of wages of educated workers on share of educated workers assigns all of the observed correlation between wages and average education to the spillover, and yields an estimate of the spillover that is upward biased (the size of the shift from 1 to 3). Again, the intuition is clear. A positive unobserved shock to the demand of skilled workers implies a wage equation residual that is positively correlated with the overall level of human capital.

Overall, whether the true magnitude of the spillover is larger or smaller than the OLS estimate depends on whether supply heterogeneity dominates demand heterogeneity.

In the case of land, the bias is unambiguously positive. The reason is that the compensating differential paid for the amenity in term of housing prices raises prices in city B with respect to city A. In Figure 3, the true spillover is the difference between the rent at point 3 and the rent at point 2. The observed effect is instead the difference between the rent at point 3 and the rent at point 1, which is smaller than the true spillover.

In Section 4.2.3, I survey different approaches that have been used in the empirical literature to try to obtain consistent estimates of human capital spillovers in the presence of unobserved heterogeneity.

4.2. *Empirical models based on wages*

The previous section indicates that there are three potential empirical strategies to identify human capital spillovers: regressing wages, land prices or output on aggregate human capital. In this section I focus on empirical models based on wages. Most of the existing empirical studies that attempt to quantify the magnitude of human capital spillovers have focused on wages. I also briefly mention models based on land prices, although the evidence on these is limited. In Section 4.3, I focus on empirical models based on firm productivity.

I begin this section by discussing the question of whether and how it is possible to differentiate between spillover effects and imperfect substitution between skilled and unskilled workers. Wages are a less straightforward measure of spillovers than production costs, productivity or land prices, because the effect of an increase in the number of educated individuals in a city has different implications for the wages of educated and uneducated workers. I then describe the main specification that wage studies have adopted, and the strategies used to deal with the issue of endogeneity of human capital. Finally, I review the existing estimates.

4.2.1. *Spillovers vs. imperfect substitution*

Increases in the aggregate level of human capital in a city have two distinct effects on the wage distribution. First, the standard neoclassical model with imperfect substitution between educated and uneducated workers indicates that an increase in the number of the educated will lower the wage of the educated and raise the wage of uneducated workers. Second, human capital spillovers will raise the wage of both groups. The distinction between spillover and imperfect substitution is analyzed in great detail in Ciccone and Peri (2002). This section is based on a simple model in Moretti (in press).

Under the assumption of complementarity (imperfect substitutability) between educated and uneducated workers, an increase in the relative number of college graduates is unambiguously positive for the wage of unskilled workers, while for college graduates its sign depends on the size of the spillover. Intuitively, complementarity and spillover both increase wages of uneducated workers, while the impact of an increase in the supply of educated workers on their own wage is determined by two competing forces: the first is the conventional supply effect which makes the economy move along a downward sloping demand curve; the second is the spillover that raises productivity. If the spillover is strong enough, as in Figure 3, the equilibrium wage of educated workers in city B is *higher* than in city A.[18] To see this point in more detail, assume that the technology is Cobb–Douglas:

$$y = (\theta_0 N_0)^{\alpha_0} (\theta_1 N_1)^{\alpha_1} K^{1-\alpha_1-\alpha_0}, \tag{3}$$

where the θ's are productivity shifters. As before, I allow for human capital spillovers by letting workers' productivity depend on the share of educated workers in the city, as well as on their own human capital:

$$\log(\theta_j) = \phi_j + \gamma\left(\frac{N_1}{N_0 + N_1}\right) \quad j = 1, 2, \tag{4}$$

where ϕ_j is a group-specific effect that captures the direct effect of own human capital on productivity ($\phi_1 > \phi_0$);

$$s = \frac{N_1}{N_0 + N_1} < 1$$

is share of college-educated workers in the city. If $\gamma = 0$, the model is the standard Mincerian model of wage determination without spillovers. If there are positive spillovers, $\gamma > 0$. If wages are equal to the marginal product of each type of labor and the spillover is external to individual firms in the city but internal to the city as a whole (so that firms take the θ's as given), the logarithm of wages for educated and uneducated workers respectively are:

$$\log(w_1) = \log(\alpha_1) + \alpha_1 \log(\theta_1) + (1 - \alpha_1 - \alpha_0)\log(K/N) + (\alpha_1 - 1)\log(s) + \alpha_0 \log\bigl(\theta_0(1 - s)\bigr)$$

and

$$\log(w_0) = \log(\alpha_0) + \alpha_0 \log(\theta_0) + (1 - \alpha_1 - \alpha_0)\log(K/N) + (\alpha_0 - 1)\log(1 - s) + \alpha_1 \log(\theta_1 s),$$

where $N = N_0 + N_1$. Consider what happens to the wages when the share of educated workers increases in the city:

$$\frac{\mathrm{d}\log(w_1)}{\mathrm{d}s} = \frac{\alpha_1 - 1}{s} - \frac{\alpha_0}{1 - s} + (\alpha_1 + \alpha_0)\gamma, \tag{5}$$

$$\frac{\mathrm{d}\log(w_0)}{\mathrm{d}s} = \frac{1 - \alpha_0}{1 - s} + \frac{\alpha_1}{s} + (\alpha_1 + \alpha_0)\gamma. \tag{6}$$

[18] Empirical evidence confirms that educated and uneducated workers are imperfect substitutes; see, for example, Katz and Murphy (1992).

The wage of uneducated workers, w_0, benefits for two reasons. First, an increase in the number of educated workers raises uneducated workers' productivity because of imperfect substitution:

$$\frac{1-\alpha_0}{1-s}+\frac{\alpha_1}{s}>0.$$

Second, the spillover further raises their productivity: $(\alpha_1+\alpha_0)\gamma>0$. The impact of an increase in the supply of educated workers on their own wage, w_1, is determined by two competing forces, as I mentioned above: the first is the conventional supply effect which makes the economy move along a downward sloping demand curve:

$$\frac{\alpha_1-1}{s}-\frac{\alpha_0}{1-s}<0.$$

The second is the spillover that raises productivity.

The important feature of Equations (5) and (6) is that unskilled workers benefit from an increase in the share of educated workers in the city even in the absence of any spillovers ($\gamma=0$), but the effect on the wage of skilled workers depends on the magnitude of the spillover. If γ is large enough, the net effect for skilled workers should be positive although smaller than for unskilled workers. If $\gamma=0$, the net effect should be negative.

It is interesting to notice that an increase in the number of educated workers in a city may raise the average wage above the private return to schooling even in the absence of any spillovers ($\gamma=0$). To see this, take the derivative of average wage with respect to s minus the private return β:

$$\frac{\mathrm{d}\log(\overline{w})}{\mathrm{d}s}-\beta=s\frac{\mathrm{d}\beta}{\mathrm{d}s}+\frac{\mathrm{d}\log(w_0)}{\mathrm{d}s}+(\alpha_1+\alpha_0)\gamma, \tag{7}$$

where $\log(\overline{w})$ is the weighted average of log wages of the two groups, $\log(\overline{w})=s\log(w_1)+(1-s)\log(w_0)$; and β is the private return, defined as the difference between the wage of educated and uneducated workers $\beta=\log(w_1)-\log(w_0)$. The first component in Equation (7) is the effect of an increase of educated workers on the private return to education. This effect is negative, because as the supply of educated workers in a city increases, the private return decreases. The second effect captures the imperfect substitution between educated and uneducated workers, and is positive. The third effect reflects the spillover. In the U.S., the share of college-educated workers, s, is approximately 0.25. Therefore, the sum of the first two components,

$$s\frac{\mathrm{d}\beta}{\mathrm{d}s}+\frac{\mathrm{d}\log(w_0)}{\mathrm{d}s}=\frac{(1-s)\alpha_1-s\alpha_0}{s(1-s)}$$

is positive if the share of output that goes to college-educated workers is more than a third of the share of output that goes to less-educated workers: $\alpha_1>0.33\alpha_0$. In this case, the increase in productivity for low-education workers more than offsets the effect of the decrease in the private return to education and an increase in s raises average wages over and above the private return to schooling even in the absence of spillovers.

The distinction between imperfect substitutability and spillovers is important for the interpretation of empirical estimates. Finding that *average* wages are affected by aggregate human capital does not necessarily indicate a spillover effect: rather this finding may indicate imperfect substitution between high- and low-educated workers. This distinction is relevant not only for theoretical reasons, but also for policy reasons. The standard imperfect substitution effect is not itself a market failure. However, if human capital spillovers exist, a market failure may occur. This depends on whether the spillover takes place within or outside the firm. It is in theory possible that within-firm spillovers are reflected in the wages of educated workers, so that no market failure arises. If the spillover has effects outside the firm, however, it is likely to be a pure externality.[19]

If the spillover effect is not constant across groups (γ_j instead of γ), it is not empirically possible to separately identify externalities and imperfect substitutability. However, under the assumption that the spillover effect γ is constant across education groups, there are two ways to empirically distinguish between imperfect substitutability and externalities. First, one can estimate separate effects of changes in the fraction of highly educated workers on wages of different education groups [for example, Moretti (in press)]. By comparing the effect of an increase in the share of college graduates across education groups, it is in theory possible to shed some light on the size of the spillover. Standard demand and supply considerations suggest that the effect of an increase in college share should be positive for low-education groups and that for college graduates its sign should depend on the size of the spillover. If the spillover is strong enough, the effect for skilled workers is positive although lower than the one for unskilled workers.

Second, Ciccone and Peri (2002) propose an alternative approach – called the "constant-composition approach" – to estimate human capital externalities when highly educated workers and less-educated workers are imperfect substitutes. They propose estimating the effect of average schooling on average wages across cities, holding the relative size of each skill group constant through a re-weighting scheme. This is obtained by first estimating a city–year–education group specific conditional average wage, and then regressing these cell averages on average schooling, weighting the regression by the size of the group in a base year. The intuition is that weighting makes it possible to separate complementarity from spillovers by holding the skill distribution of the labor force in the city constant.

4.2.2. Econometric specification

Most of the direct evidence on the magnitude of the spillovers is based on models that regress wage on measures of the aggregate stock of human capital. The basic source

[19] One can think of the "imperfect substitutability effect" as a form of pecuniary externality. However, this type of pecuniary externality is very different from the pecuniary externalities proposed by Acemoglu that I discussed in Section 3.1.

of identification therefore consists of the comparison of wages for otherwise similar individuals who work in cities with different aggregate human capital. Typically, authors have estimated variations of the following equation

$$\log(w_{ict}) = X_{it}\beta_{ct} + \pi P_{ct} + \alpha Z_{ct} + d_c + d_t + u_{ict}, \tag{8}$$

where w_{ict} is wage of individual i living in city c in period t; X_{it} is a vector of individual characteristics, including years of schooling; P_{ct} represents a measure of aggregate human capital in city c in year t; Z_{ct} is a vector of city characteristics which may be correlated with P_{ct}; d_c represents a city fixed effect; and d_t is a year effect.

The coefficient of interest is π, which is the estimate of the effect of aggregate human capital on average wages after controlling for the private return to education. Typically, authors have measured aggregate human capital in a city, P_{ct}, using either average years of schooling or the percent of individuals with a college education. Ciccone and Peri (2002) show the conditions under which Equation (8) can be derived from the standard framework used in theoretical macro economics to model the effect of human capital on economic growth at the aggregate level [see for example, Lucas (1988) or Bils and Klenow (2000)].

A source of confusion in the existing literature has been the issue of whether nominal or real wages would be used in estimating Equation (8). From the model of Section 4.1, it should by now be clear that *nominal* wages (i.e., wages unadjusted for cost of living) are the appropriate dependent variable in Equation (8) and that no control for cost of living need to be included in Z_{ct}. In equilibrium, higher nominal wages in a city should reflect greater productivity. Although workers may be indifferent between high nominal wages and high rents on one hand, and low nominal wages and low rent on the other hand, firms still are willing to pay high nominal wages presumably only if workers are more productive.

The wage equation residual can be thought as the sum of three components:

$$u_{ict} = \mu_c\theta_i + \nu_{ct} + \varepsilon_{ict}, \tag{9}$$

where θ_i is a permanent unobservable component of individual human capital, such as ability or family background; μ_c is a factor loading which represents the return to unobserved skill in city c; ν_{ct} represents time-varying shocks to labor demand and supply in city c in period t; ε_{ict} is the transitory component of log wages which is assumed to be independently and identically distributed over individuals, cities and time.

A first source of omitted variable bias is the presence of time-varying shocks to local labor markets that are correlated with aggregate human capital. Cities differ widely in geographical location, industrial structure, technology, weather and amenities. City fixed effects sweep out the effect of *permanent* city characteristics such as the industrial structure and physical and cultural amenities that might bias a simpler cross-sectional analysis. But first-differenced models may still be biased by the presence of *time-varying* factors that are correlated with changes in human capital and wages across cities – for example, transitory productivity shocks that attract highly educated workers and raise wages: $\mathrm{cov}(\nu_{ct}, S_{ct}) \neq 0$. As argued in Section 4.1.3, the resulting OLS bias is positive (negative) if positive shocks to wages are associated with increases (decreases) in the human capital stock in a city. For example, the San Jose economy experienced an unprecedented economic expansion starting in the second half of the 1980s that was driven by the Silicon Valley computer industry boom. The same boom attracted a highly educated labor force to San Jose. On the other hand, if variation in human capital stock across cities is driven by unobserved *supply* factors, OLS is biased downward.

A second source of omitted variable bias is the presence of unobserved worker characteristics if individuals observed in cities with high human capital are better workers than individuals with the same observable characteristics who live in cities with low human capital. In terms of Equation (8), this implies that $\text{cov}(\theta_i, P_{ct}) > 0$. For example, a high-school graduate working in a biotechnology firm in San Francisco is probably different along some unobservable dimension from a high-school graduate working in a shoe factory in Miami. Similarly, a lawyer working for a Wall-Street firm in New York is likely to differ from a lawyer in El Paso, TX. This type of sorting may take place if a higher overall level of human capital in a city is associated with a higher return to unobserved ability, causing higher quality workers to move to cities with higher college share [Borjas, Bronars and Trejo (1992), Rauch (1993)]. Consider a simple Roy model where different cities reward workers' skills – both observed and unobserved – differently, and mobility decisions are based on comparative advantage. In such a model, workers are not randomly assigned to cities, but choose the city where their skills are most valued and skill–price differentials determine the skill composition of migratory flows. Cities that have an industrial structure that demands more education are also likely to offer a higher price for unobserved ability. In this case, the correlation of high P_{ct} with high wages may simply reflect higher unobserved ability of workers rather than higher productivity.

4.2.3. Accounting for the endogeneity of human capital

In an ideal analysis, the researcher could randomly assign different overall levels of human capital across cities and measure differences in the value of wages before sorting occurs. This experimental design would solve the econometric problems discussed in Section 4.1.3. (Note, however, that the experimental design would not solve the problem of distinguishing between complementarity and externalities discussed in Section 4.2.1). In its absence, three strategies can be used to account for endogeneity of overall levels of human capital.

First, some authors have tried accounting for time-varying shocks by controlling for observable characteristics of cities, such as racial composition or unemployment rate. It is particularly important to fully control for shocks to the relative demand for skilled labor, as they lead to overestimates of the spillover. In an effort to accomplish this goal, some researchers have used an index of demand shifts proposed by Katz and Murphy (1992). The index, a generalization of a widely used measure of between-sector demand shifts, is based on nation-wide employment growth in industries, weighted by the city-specific employment share in those industries. It captures exogenous shifts in the relative demand for different education groups that are predicted by the city industry mix.[20]

One limitation of this approach is that it is hard to argue persuasively that observables can fully account for shocks. For this reason, some studies have turned to instrumental variable techniques. This approach requires an instrument that is correlated with changes in the overall level of human capital in a city and uncorrelated with changes in unobserved factors that affect wages directly. Examples of instrumental variable used are compulsory schooling laws, child labor laws, the entry of the baby boom cohort into the labor market, and the presence of land grant colleges. The advantage of instrumental variable techniques is that a valid instrument isolates the effect of exogenous changes in human capital levels on wages. The disadvantages are that valid, exogenous instruments are rare. Furthermore, if the effect of overall human capital on wages is not homogeneous, IV estimates and OLS estimates may not be directly comparable.

As a third possible identification strategy, individual-level longitudinal data have been used. By observing the same individual over time, one can control for factors that make an individual permanently more productive. But note that if longitudinal data on multiple individuals and cities are available, individual fixed effects models are not the most general model that can be estimated. In particular, the term $\mu_c \theta_i$ in Equation (8) can be absorbed by including a set of individual × city dummies. By controlling for the individual–city match, variation that comes from movers is lost. Identification is based on stayers and comes from changes of P in a city over time. *Conditional on a city–individual match*, the longitudinal model estimates what happens to an individual's wage as aggregate human capital around her increases. The key identifying assumption is that the return to unobserved ability μ_c may vary across cities, but not over time or, if it does change over time, the change is not systematically correlated with the stock of human capital. Under this assumption, differences in the level of unobserved ability and in return to unobserved ability across cities are absorbed by individual × city fixed effects. One limitation of this longitudinal strategy is that stayers are not necessarily a random sample of the population. If stayers are different from other workers, longitudinal estimates may be biased.

[20] Different cities specialize in the production of different goods, so that industry-specific demand shocks at the national level have a differential impact on cities [Bound and Holzer (1996)]. If employment of skilled workers in a given industry increases (decreases) nationally, cities where that industry employs a significant share of the labor force will experience a positive (negative) shock to the labor demand of skilled workers. Formally, the index is

$$\text{shock}_{jc} = \sum_{s=1} \eta_{sc} \Delta E_{js}, \tag{10}$$

where s indexes two-digit industry; shock_{jc} represents the predicted employment change for workers belonging to education group j in city c; η_{sc} is the share of total hours worked in industry s in city c in 1980; ΔE_{js} is the change in the log of total hours worked in the same industry nationally between 1980 and 1990 by workers belonging to education group j. See, for example, Moretti (in press).

4.2.4. Empirical findings of wage and land price models

I now turn to a discussion of some of the empirical evidence on the magnitude of human capital externalities generated by wage models. What do we know about the magnitude of human capital spillovers? There is some indirect evidence that human capital spillover may play a role in increasing income and growth. For example, Glaeser, Scheinkman and Shleifer (1995) report that income per capita has grown faster in cities with high initial human capital in the post-war period. Findings in Glaeser and Mare (2001) are consistent with a model where individuals acquire skills by interacting with one another, and dense urban areas increase the probability of interaction.

Yet, despite these significant policy implications, there is remarkably little systematic empirical evidence on the magnitude of human capital spillovers. Researchers have only recently begun to estimate the size of spillovers from education by comparing the wages of otherwise similar individuals in cities or states with different average levels of education. Most of these wage studies have used variants of the wage equations in Equation (8).

Rauch (1993) often cited study is the first to exploit differences in human capital across cities to identify externalities. Using the 1980 Census, he estimates a cross sectional version of Equation (8) and finds that a one year increase in average education raises wages by 3 to 5 percent in 1980. Rauch is also one of the very few researchers to examine the effect of human capital on the cost of housing. He finds that the cost of housing is higher in cities with a larger stock of human capital (holding constant housing characteristics), and concludes that spillover appear to be capitalized in land prices. A limitation of Rauch's methodology is that he does not directly account for the endogeneity of aggregate human capital. Rauch uses only one cross section and treats average schooling as historically predetermined. A second limitation is that he does not distinguish between externalities and complementarity between skilled and unskilled workers.

The correlation between wages and one measure of the overall level of human capital is shown in Figure 5, where the percentage of college graduates is plotted against the regression-adjusted average wage for 282 cities in 1990.[21] The regression-adjusted average wage is obtained by conditioning on individual education, gender, race, Hispanic origin, U.S. citizenship, and a quadratic term in potential experience. The figure shows that, *after controlling for the private return to education*, wages are higher in cities where the labor force is better educated.[22] Obviously, it is far from clear whether the association between wages and human capital uncovered in Figure 5 reflects human capital spillovers. As argued in detail in Section 4.1.3, this correlation is likely to be biased by the presence of unmeasured characteristics of workers and cities that are

[21] Calculations by the author based on the 1990 Census.

[22] Similar figures can be obtained by plotting housing cost or firms productivity against aggregate human capital.

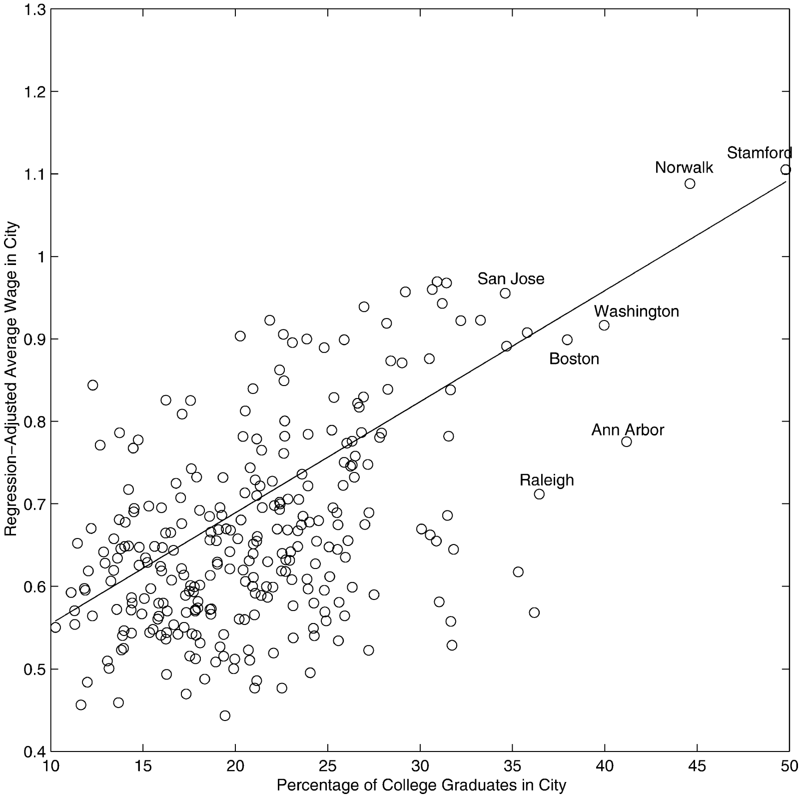

Figure 5. Correlation between regression-adjusted average wage and percentage of college graduates in 282 cities, in 1990. Regression-adjusted average wage is obtained by conditioning on individual education, gender, race, Hispanic origin, U.S. citizenship and work experience. Weighted OLS fit superimposed.

potentially correlated with the fraction of college-educated individuals across cities. Although we know that the correlation is likely to be biased, we do not know the direction of the bias. The OLS coefficient in a regression of wages on share of educated workers can be biased either downward or upward, depending on whether variation in the relative number of college-educated workers across cities is driven by unobserved *supply* factors or unobserved *demand* factors. And even if we could account for the endogeneity of college share, from the simple graph in Figure 5 we cannot distinguish between complementarities and externalities (Section 4.2.1).

Moretti (in press) attempts to address the endogeneity created by city-wide demand shocks using two instrumental variables. The first instrument is based on differences in the age structure of cities. The U.S. labor force is characterized by a long-run trend of increasing education, with younger cohorts better-educated than older ones. The second instrument used is an indicator for the presence of a land-grant college in the city.

Land-grant colleges were established by the federal Morrill Act of 1862.[23] He also tries to account for unobserved individual ability by exploiting the panel structure of the National Longitudinal Survey of Youth (NLSY) to estimate models that condition on individual × city effects.

Moretti finds that a one percentage point increase in college share in a city raises average wages by 0.6–1.2%, above and beyond the private return to education. As argued in previous sections, the finding that *average* wages are affected by the percentage of college graduates in the labor force does not necessarily indicate a spillover effect: rather this finding may indicate imperfect substitution between high- and low-education workers. For this reason, Moretti estimates the effect of changes in the fraction of highly educated workers on wages of different education groups. He finds that a one percentage point increase in the labor force share of college graduates increases the wages of high-school drop-outs and high-school graduates by 1.9 and 1.6%, respectively. It also increases wages of college graduates by 0.4%. This findings are consistent with a model that includes both conventional demand and supply factors as well as spillovers: as expected, an increase in the proportion of better-educated workers has a large positive effect on less-educated workers, and a smaller but still positive effect on the wages of the best-educated group.

Acemoglu and Angrist (2000) use state variation in child labor and compulsory school attendance laws to instrument for average schooling. They show that within state changes in these laws affect the education distribution at the "right" point, by increasing the probability of high-school graduation but not college graduation. Unlike Rauch and Moretti, Acemoglu and Angrist (2000) also address the endogeneity of individual schooling. They point out that inconsistent estimates of the private return to education will lead to inconsistent estimates of the externality, because individual and aggregate schooling are correlated.

To account for the endogeneity of individual schooling, they use quarter of birth as an instrumental variable. While their OLS estimates of the externality are qualitatively consistent with Rauch's and Moretti's OLS estimates, their IV estimates are smaller and in most cases not significantly different from zero.

The difference in findings between Acemoglu and Angrist (2000) and Moretti can be explained in part by the fact that child labor and compulsory attendance laws affect educational attainment in the lower part of the educational distribution, mostly in middle school or high school. On the contrary, Moretti identifies externalities using variation in the number of college graduates, i.e., the upper part of the distribution. It appears that a one year rise in a city's average education resulting from an increase in the number of those who finish high school has a different effect than a similar increase resulting from an increase in the number of those who go to college. A second factor that may account for the difference in estimates is the fact that Acemoglu and Angrist's analysis is at the

[23] Because the program was federal and took place more than one hundred years ago, the presence of a land-grant institution is unlikely to be correlated with local labor market conditions in the 1980s.

state level. When Moretti estimates state-level spillovers, he finds coefficients closer to those of Acemoglu and Angrist.[24]

Ciccone and Peri (2002) propose a new econometric approach – the "constant-composition approach" – to estimate human capital externalities when highly educated workers and less-educated workers are imperfect substitutes. This new approach is a generalization of the approach based on Mincer wage equations like Equation (8), and is to date the most comprehensive attempt to distinguish between complementarity and externalities. The constant-composition approach consists of estimating the effect of average schooling on average wages across cities, holding the relative size of each skill group constant with a re-weighting scheme. The weights are based on the size of each skill group in a base year.

While the re-weighting procedure accounts for the possibility of complementarity between skilled and unskilled workers, Ciccone and Peri (2002) also use a set of instrumental variables to account for the endogeneity of aggregate human capital. When they constrain highly educated workers and less-educated workers to be perfect substitutes, Ciccone and Peri (2002) find significant positive externalities, with magnitudes consistent with estimates in Rauch and Moretti. However, when they allow for imperfect substitutability, they find little evidence of positive human capital spillovers.

In a related paper, Peri (2002) models the location decisions of young and old workers as a function of human capital externalities. Using Census data, he begins by showing that the experience premium is higher in urban areas than in rural areas. For example, in 1990 a college-educated urban White male received a $2 hourly premium over the wage of a similar non-urban worker. The premium for a mature White worker was twice as large. This result indicates that young educated workers receive a lower wage premium in urban areas than their older colleagues, but in spite of this, they are overrepresented in urban areas. Why do urban areas attracts young educated workers? Peri argues that learning externalities are an important explanation. Workers learn from each other when they are young, so living in dense urban areas may raise human capital accumulation more than living in a rural area. The negative compensating differential indicates that young workers value such human capital externalities. As they grow older, the importance of knowledge spillovers diminishes, and some of them move toward non-urban areas.[25]

In most empirical applications, the geographic scope of the human capital spillover is taken as exogenous. In Conley, Flyer and Tsiang (in press), however, the scope of the spillover is allowed to vary across individuals. The stock of aggregate human capital that is relevant for an individual is defined using a measure of economic distance based on

[24] A third difference concerns the period under consideration. Most models in Acemoglu and Angrist (2000) are estimated using 1960–1980 Census data. When they add data from the 1990 Census, they find statistically significant positive estimates for the externality, when child labor laws are used as instruments. Since the private return to education increased during the 1980s, this finding may reflect a change in the social value of human capital.

[25] The evidence in Costa and Kahn (2000) offers an alternative to the learning story.

estimates of travel time between locations. They estimate the number and human capital level of potential trading partners within, say, a two hours trip for each individual. By using this economic distance metric rather than administrative geographic units like cities or provinces, they better characterize the human capital of the set of neighbors that potentially interact with each agent. Furthermore, they can identify the geographic scope of the spillover by varying the definition of the local labor market and testing how far two individuals must be for their human capital spillover to have no effect. They find significant instrumental-variable estimates of human capital externalities. They also find that externalities tend to disappear for distances further than a 90 minute trip. Conley et al. also investigate whether human capital spillovers are capitalized in land prices. They find that land prices are positively correlated with human capital stocks.

In sum, the existing *direct* empirical evidence on the magnitude of human capital spillovers is mixed. While most papers find OLS estimates that are consistent with the pattern shown in Figure 5, estimates that attempt to account for the endogeneity of aggregate human capital rarely agree. The differences across studies appear to be mostly due to differences in the particular empirical strategies adopted to account for unobserved heterogeneity. Clearly, more research is needed before we can be confident in our estimates of the magnitude of human capital spillovers.

4.3. Empirical models based on firm productivity

Having analyzed the empirical evidence based on differences in wages and land prices across cities, I now turn to evidence based on differences in productivity levels. The model in Section 4.1 indicates that if externalities exist, we should find that firms located in cities with high levels of human capital produce more output with the same inputs than otherwise similar firms located in cities with low levels of human capital. Furthermore, the model indicates that these differences between cities should coincide with observed differences in wages of workers and land prices. In equilibrium, if firms really are more productive in cities with high levels of human capital, we would expect to find that these firms incur higher labor and land costs. If this was not the case, firms (at least those producing nationally traded goods) would relocate from cities with low human capital to cities with high human capital.

4.3.1. Econometric specification

To see how spillovers can be identified by comparing the productivity of firms in cities with different level of human capital, assume that technology can be described by the following Cobb–Douglas production function:[26]

$$y_{pjct} = A_{pjct} L1_{pjct}^{\alpha_{1j}} L2_{pjct}^{\alpha_{0j}} K_{pjct}^{\beta_j}, \tag{11}$$

[26] This section is based on Moretti (2002).

where y_{pjct} is output of firm p, belonging to industry j, in city c, and year t; j indexes industry; $L1_{pjct}$ is the number of hours worked by skilled workers in the firm; $L2_{pjct}$ is the number of hours worked by unskilled workers; K_{pjct} is capital. As before, assume that A_{pjct} is a function of aggregate human capital outside the firm in the same city and unobservable productivity shocks:

$$\ln A_{pjct} = \gamma \overline{S}_{ct} + \varepsilon_{pjct}, \tag{12}$$

where $\overline{S}_{ct}$ is some measure of the overall stock of human capital among all workers in city c at time t; and ε represent unobserved heterogeneity in productivity. The coefficient of interest is γ, the external effect of education on productivity. If $\gamma = 0$, the model reduces to a standard production function without externalities.

Empirically, the production function (11) can be either estimated directly or estimated using its total factor productivity (TFP) version. The TFP version can be estimated in two steps. Under the assumption that input prices are equal to their marginal product, a plant-specific measure of TFP is easily calculated by subtracting the sum of each input cost share multiplied by the quantity of that input, from the value of the output. This estimate of TFP can then be regressed on aggregate human capital.

Like for wage models, the main concern is that there may be unobservable productivity shocks that are potentially correlated with aggregate human capital. For example, ε may reflect unmeasured firm characteristics such as the quality of machines, patents, the quality of workers and management, and the culture within the firm. Alternatively, ε may capture city characteristics that make some cities more productive than others. These may include the public infrastructure (ports, highways, or airports), weather conditions, the presence of a research universities, and efficiency of local authorities.

In general, if plants with a positive ε tend to be located in cities with a high overall level of human capital, then OLS estimates of overestimate γ.[27]

4.3.2. Empirical findings of productivity models

Empirical evidence suggests that knowledge spillovers may be particularly important in certain hi-tech industries. One interesting piece of evidence on knowledge externalities is a well-cited paper by Jaffe, Trajtenberg and Henderson (1993) that shows that references to existing patents that inventors include in their patent applications are likely to come from the same state or metropolitan area as the originating patent application.

[27] A similar point is made graphically in Figure 4. This is the case, for example, if unusually productive entrepreneurs are more likely to seek out productive locations; or if unusually skilled individuals are disproportionately recruited to the most productive locations. The true spillover is the difference in productivity between a plant in point 3 and a plant in point 2. But if the technology that raises productivity of educated workers in city B relative to city A is unobserved to the econometrician, a naive estimate of the spillover is the difference in productivity between a plant in point 3 and a plant in point 1, which overestimates the true spillover.

Because human capital spillovers and knowledge spillovers are invisible, most empirical studies resort to indirect evidence to test for the presence of spillovers. The studies based on wage equations described in the previous section test indirect implications of the spillover hypothesis, rather than directly measuring the spillover itself. But Jaffe, Trajtenberg and Henderson (1993) argue that patent citations offer a direct measure of spillovers, an observable paper trail in the form of citations in patents. Jaffe, Trajtenberg and Henderson (1993) use citation patterns to test the extent to which spillovers are geographically localized. Because patents are publicly available, in the absence of localized spillovers, citations would not depend on the location of the inventor.

The key empirical challenge of the paper is to distinguish between geographic patterns of patent citations caused by spillovers from patterns caused by exogenous sources of agglomeration effects. To address this issue, the authors construct "control" samples of patents that have the same temporal and technological distribution as the patent citations. To identify the presence of externalities, they compare these two patterns of geographic concentration under the assumption that the geographic correlation between the controls and originating patents is only due to exogenous agglomeration forces that are independent of spillovers. The proposed test of localization is whether the correlation is significantly greater for the cited patents than the control patents. Their findings suggest that patents citations are indeed geographically localized and that knowledge spillovers appear to be large.

Adams and Jaffe (1996) also study the composition of the knowledge transfers within and across firms. They use a TFP framework that is related to the one presented in Section 4.3.1, but instead of using the stock of human capital as their main independent variable, they focus on R&D performed in formal research labs. In particular, they postulate that a plant has an "effective stock of knowledge" that is generated in several ways: by learning-by-doing at this and other plants in the same city or industry, by informal research activities performed at the plant, by formal research of the plant's parent firm, and by formal research of other firms in the same city or industry. Empirically, they use manufacturing plant-level data to examine the productivity effects of R&D performed in a plant, outside a plant but inside the parent firm that owns the plant, and in external plants in the same geographical area or industry. They find that spillovers of R&D are important, both within and across firms – a result that is consistent with the notion that the social return to research is higher than the private return.

Interestingly, they find that the effect of parent firm R&D on plant-level productivity is diminished by both the geographical distance and the technological distance between the research lab and the plants. They interpret this finding as a reflection of the fact that communications costs rise with distance. They also provide evidence of within-industry spillover effects: R&D of other firms in the same industry does appear to affect a plant's productivity, holding industry constant. The magnitude of these spillovers is surprisingly large. The marginal product of industry R&D is approximately 40% as large as the marginal product of parent firm research.

Another piece of indirect evidence on the role of human capital spillovers on the productivity of high-tech firms is a recent paper by Zucker, Darby and Brewer (1998).

They argue that geographic differences in specialized human capital across cities is the main determinant of where and when American biotechnology industries developed. In particular, they show that the stock of human capital of outstanding scientists in certain cities – measured in terms of the number of publications reporting genetic-sequence discoveries in academic journals – plays a key role in the entry decisions of new biotech firms. This effect seems to reflect, at least in part, human capital externalities, because it is not just a reflection of the presence of universities and government research centers in areas where outstanding scientists are located.[28]

The studies described so far focus on high-tech industries. Moretti (2002) attempts to systematically assess the magnitude of human capital externalities in all industries by estimating production functions similar to those in Equation (11). Using longitudinal data, he estimates establishment-level production functions controlling for establishment-specific permanent heterogeneity, as well as time-varying industry-specific and state-specific heterogeneity. Moretti finds that productivity gains from human capital spillovers appear to be empirically relevant for manufacturing establishments in U.S. cities. However, because the stock of human capital grows slowly over time, the contribution of human capital spillovers to economic growth does not appear to be large. Estimates in the paper indicate that human capital spillovers were responsible for an average of 0.1% increase in output per year during the 1980s.[29] Most of the estimated spillover comes from high-tech plants. For non-high-tech producers, the spillover appears to be virtually zero.

Importantly, the magnitude of spillovers between plants in the same city appear to depend on their level of interaction. If input–output tables are used to measure the interaction between plants in the same city, spillovers between plants that often interact are found to be significant, while spillovers between plants that rarely interact are much smaller. This is consistent with the notion that human capital spillovers decay not only with geographic distance, but also with economic distance.

Consistent with the predictions of the theoretical model presented in Section 4.1.1, the productivity gains generated by human capital spillover appear to be offset by increased labor costs. Findings indicate that the estimated productivity differences between cities with high human capital and low human capital coincide with observed differences in wages of manufacturing workers.

4.3.3. Empirical findings of country-wide models

Although many empirical studies use cities as their unit of analysis, some studies focus on the effect of nation-wide increases in human capital on national income.[30] In gen-

[28] Audretsch and Stephan (1996) use data on IPO of biotech firms to link the location of the biotechnology firm with the location of the university-based scientists affiliated with the firm. They conclude that "while proximity matters in establishing formal ties between university-based scientists and companies, its influence is anything but overwhelming".

[29] For the average manufacturing plant in the U.S., this amounts to about $10,000 per year.

[30] See, for example, Topel (1999), Krueger and Lindahl (1998), and Bils and Klenow (2000).

eral, cities have several advantages over states, regions and countries. First, most of the geographical scope of knowledge spillover is likely to be local. Workers interact mostly at the local level. Although communications technologies like phones, e-mail and the Internet allow low-cost communication across cities, most of the interactions between individuals take place in a limited amount of space. If anything, one may argue that local spillovers are likely to arise at the neighborhood level as well as at the city level.

Second, cities are more specialized and less arbitrary economic units than countries. National boundaries that limit labor and capital mobility and national policies that encourage industrial diversification reduce the gain from factor mobility [Glaeser, Scheinkman and Shleifer (1995)].

Third, countries differ along so many variables that it is hard to interpret cross-country estimates. Legal and political institutions, cultural attitudes and social norms are important factors in determining wages and productivity of countries, and they are likely to be correlated with the aggregate level of human capital. To the extend that it is hard to convincingly control for cross countries differences in these unobservables, estimates of spillovers are hard to interpret. Looking at within countries changes over time is unlikely to solve the problem, as time-varying shocks that affect productivity and wages are likely to be highly uncorrelated with changes in human capital of the labor force. But by looking at cities within the United States, this source of unobserved heterogeneity is minimized. Although cities may have different institutions, the difference are smaller than cross-countries differences.

Finally, the lack of high quality multiple-country datasets makes cross country comparisons difficult. Available educational attainment data for many countries are noisy. Measurement error becomes an even more serious problem when looking not at cross-sectional models, but at models based on changes in education over time. For example, Krueger and Lindahl (1998) find that at least half of the variability of measured changes in schooling across countries is pure noise.[31] They re-analyze the correlation between education and growth, taking into account the measurement problems. They conclude that, on average, economic growth is positively correlated with increases in schooling, and the estimated effect is not too different from the 10 percent rate of private return to schooling. They caution, however, that cross-country evidence is relatively weak, and fraught with problems of non-comparability, measurement error, and most fundamentally, a lack of a credible "research design".

Bils and Klenow (2000) propose an overlapping generations model in which each generation learns from previous generations and human capital creates positive externality in the level of technology adoption. Although they do not estimate the model, they calibrate it using existing evidence based on Mincerian regressions of the type described in Equation (8). They conclude that schooling explains less than one-third of the empirical variation in growth rates.

[31] On the contrary, Census, CPS, and other large scale U.S. government datasets provide high quality, consistent information on schooling levels, wages and productivity for all large U.S. metropolitan areas.

5. Empirical evidence on other social benefits of education: crime and voting

In Section 3, I explained that investment in human capital may generate both private and social benefits. I also argued that two important examples of the social benefits of high levels of human capital are reduced crime rates and improved political participation. In this section I review the existing literature on the link between education on the one hand, and crime and voting on the other.

Crime. Despite the enormous policy implications, little research has been undertaken to evaluate the relationship between schooling and criminal behavior. In Section 3, I outlined several theoretical reasons for why education may reduce the probability of engaging in criminal activities. But is it possible in practice to reduce crime rates by raising the education of potential criminals? Witte (1997) argues that based on the existing empirical studies "... neither years of schooling completed nor receipt of a high-school degree has a significant effect on an individual's level of criminal activity". But, this conclusion is based on only a few of the available studies, including Tauchen, Witte and Griesinger (1994) and Witte and Tauchen (1994), which find no significant link between education and crime after controlling for a number of individual characteristics. While Grogger (1998) estimates a significant negative relationship between wage rates and crime, he finds no relationship between education and crime after controlling for wages. (Of course, increased wages are an important consequence of schooling.)[32]

More recently, Lochner (1999) estimates a significant and important link between high-school graduation and crime using data from the National Longitudinal Survey of Youth (NLSY). Other research relevant to the link between education and crime has examined the correlation between crime and time spent in school [Gottfredson (1985), Farrington (1986), Witte and Tauchen (1994)]. These studies find that time spent in school significantly reduces criminal activity – more so than time spent at work – suggesting a contemporaneous link between school attendance and crime.

The key difficulty in estimating the effect of education on criminal activity is that unobserved characteristics affecting schooling decisions are likely to be correlated with unobservables influencing the decision to engage in crime. For example, individuals with high discount rates or high returns to criminal activity are likely to spend more time on crime rather than work, regardless of their educational background. To the extent that schooling does not raise criminal returns, there is little reward to finishing high school or attending college for these individuals. As a result, we might expect a negative correlation between crime and education even if there is no causal effect of education

[32] Freeman (1996), Gould, Mustard and Weinberg (2000), Grogger (1998), Machin and Meghir (2000), and Viscusi (1986) empirically establish a negative correlation between earnings levels (or wage rates) and criminal activity. The relationship between crime and unemployment has been more tenuous [see Chiricos (1987) or Freeman (1983, 1995)) for excellent surveys]; however, a number of recent studies that better address problems with endogeneity and unobserved correlates [including Gould, Mustard and Weinberg (2000) and Raphael and Winter-Ebmer (2001)] find a sizeable positive effect of unemployment on crime.

on crime. State policies may induce bias with the opposite sign – if increases in state spending for crime prevention and prison construction trade off with spending for public education, a *positive* spurious correlation between education and crime is also possible.

In a recent paper, Lochner and Moretti (2002) analyze the effect of schooling on incarceration, arrests and self-reported criminal activity using changes in state compulsory school attendance laws as an instrument for schooling. Changes in these laws have a significant effect on educational achievement, and the authors reject tests for reverse causality. Moreover, increases in compulsory schooling ages do not appear to be correlated with increases in state resources devoted to fighting crime. Both OLS and IV estimates agree and suggest that additional years of secondary schooling reduce the probability of incarceration with the greatest impact associated with completing high school. Differences in educational attainment between Black and White men can explain as much as 23% of the Black–White gap in male incarceration rates. Education has the largest impact on the prevention of murder, assault, and motor vehicle theft. Lochner and Moretti also find evidence that the estimates for imprisonment and arrest are caused by changes in criminal behavior and not educational differences in the probability of arrest or incarceration conditional on crime.

If these results are correct, cities with higher high-school graduation rates should have lower crime rates, holding everything else constant. The social savings from crime reduction associated with high-school graduation rates appear to be economically important. The externality is about 14–26% of the private return, suggesting that a significant part of the social return to completing high school comes in the form of externalities from crime reduction.

Voting. I now turn to the evidence on the effect of schooling on political behavior. Two channels have been suggested. First, education may improve citizens' ability to choose better candidates. Second, it may improves civic participation. There is virtually no evidence on the first margin due to the fact that it is very difficult to measure "quality" of candidates in an objective, value-free way. On the other hand, a vast body of empirical research in political science focuses on civic participation.[33] The key weakness of the existing evidence lies in the treatment of causality. Since both education acquisition and civic participation are choices made by individuals, these decisions might be jointly caused by some excluded individual characteristic. Lacking a strategy to address this possibility, the available literature offers little firm evidence on the causal nature of the relationship.

Brady, Verba and Schlozman (1995) are the first to address the potential endogeneity of schooling in this literature, although the exclusion restrictions they impose on their estimation are not convincing. More recently, Dee (2002) and Milligan, Moretti and Oreopoulos (2003) use an instrumental variables strategy based on changes in compulsory schooling laws to account for endogeneity. Milligan, Moretti and Oreopoulos

[33] See Verba and Nie (1972), Teixeira (1987), Helliwell and Putnam (1999), Powell (1985), Leighley and Nagler (1992), and Weisberg and Box-Steffensmeier (1999).

(2003) find a strong effect of education on voting in the U.S. The effect appears to stem from differences in voter registration across education groups. Results from the UK, where citizens are legally responsible and actively assisted to register, show no effect of education on voting. They also find strong and persistent effects of education on civic behavior in both the U.S. and the UK. Educated adults are more likely to discuss politics with others, associate with a political group, work on community issues, and follow election campaigns in the media.

6. Conclusion

What is the effect of an increase in the aggregate level of human capital on a city economy? Although much is known about the private returns to human capital, the answer to this question is not straightforward. Increases in the skill level of a city can affect the local economy in ways that are not fully reflected in the private return of education. Human capital spillovers can in theory increase productivity over and above the direct effect of human capital on individual productivity. Furthermore, increases in education can reduce criminal participation and improve voters' political behavior.

The size of the social return to human capital has enormous policy implications. Local governments are increasingly involved in policies aimed at fostering economic growth, so a measure of the magnitude of human capital spillovers is crucial in deciding how many resources to invest in attracting skilled workers. Moreover, the magnitude of the social return to education is an important tool for assessing the efficiency of investment in public education.

Three empirical strategies are available to estimate human capital spillovers. First, human capital spillovers can be identified by comparing the productivity or the unit costs of identical plants located in cities with high and low levels of human capital. Second, spillovers can be identified by comparing land prices in cities with high and low levels of human capital; and finally spillovers can be identified by comparing wages of identical workers located in cities with high and low levels of human capital.

Despite the important policy implications and a large theoretical literature that assumes the existence of human capital externalities, the empirical literature on the magnitude of these externalities is still young. Given the limited number of empirical studies on this subject, it is still too early to draw definitive conclusions on the size of externalities. This is particularly surprising given the huge literature estimating the private return to education that has emerged in labor economics in the past four decades. Nevertheless, "economists are conspicuous by their absence" on the subject of human capital spillovers [Topel (1999)].

More research is needed to overcome the formidable identification issues that the endogeneity of human capital presents. Current research on the topic is now at a stage that is reminiscent of the literature on the private returns to education in its early stages. The empirical challenges are enormous, but the potential rewards are also large. Judging by the creativity and ingenuity that economists have shown in addressing the empirical challenges that arose in the estimation of the private return to human capital, it is not unreasonable to expect to see substantial progress on the issue of social returns in the not-too-distant future.

Acknowledgements

I thank the editors Vernon Henderson and Jacques Thisse for useful comments. I am also grateful to Luis Bertinelli, Gilles Duranton, Gerhard Glomm, Giovanni Peri, Stephanie Riegg and Stuart Rosenthal for many useful suggestions on previous versions of this chapter.

References

Acemoglu, D. (1996). "A microfundation for social increasing returns in human capital accumulation". Quarterly Journal of Economics, 779–804.

Acemoglu, D. (1998). "Why do new technologies complement skills? Directed technical change and wage inequality". Quarterly Journal of Economics 113, 1055–1090.

Acemoglu, D., Angrist, J. (2000). "How large are human capital externalities? Evidence from compulsory schooling laws". NBER Macroeconomic Annual 15.

Adams, J., Jaffe, A.B. (1996). "Bounding the effects of R&D: An investigation using matched establishment-firm data". RAND Journal of Economics 27 (4), 700–721.

American Electronics Association (1997). Area definition of the high-tech sector. Cyberstates.

Arrow, K. (1962). "Economic welfare and the allocation of resources for invention". In: Nelson, N. (Ed.), The Rate and Direction of Inventive Activity. Princeton University Press, Princeton, NJ.

Arrow, K. (1997). "The benefit of education and the formation of preferences". In: Behrman, J., Stacey, N. (Eds.), The Social Benefits of Education. University of Michigan Press.

Audretsch, D., Stephan, P. (1996). "Company–scientists location links: The case of biotechnology". American Economic Review 86 (3), 641–652.

Becker, G., Mulligan, C. (1997). "The endogenous determination of time preference". Quarterly Journal of Economics 112 (3), 729–758.

Bils, M., Klenow, P. (2000). "Does schooling cause growth?" American Economic Review 90 (5), 1160–1183.

Black, D., Henderson, J.V. (1999). "A theory of urban growth". Journal of Political Economy 107, 252–284.

Borjas, G.J., Bronars, S.G., Trejo, S.J. (1992). "Self-selection and internal migration in the United States". Journal of Urban Economics 32, 159–185.

Bound, J., Groen, J., Kezdi, G., Turner, S. (in press). "Trade in university training, cross-state variation in the production and use of college-educated labor". Journal of Econometrics.

Bound, J., Holzer, H.J. (1996). "Demand shifts, population adjustments, and labor market outcomes during the 1980s". NBER Working Paper 5685.

Brady, H.E., Verba, S. (1995). Schlozman, K.K., "Beyond SES: A resource allocation model of political participation". American Political Science Review 89 (2), 271–294.

Card, D. (1999). "Education in the labor market". In: Ashenfelter, O., Card, D. (Eds.), The Handbook of Labor Economics. North-Holland, Amsterdam.

Chiricos, T. (1987). "Rates of crime and unemployment: An analysis of aggregate research". Social Problems 34 (2), 187–211.

Ciccone A., Peri G. (2002). "Identifying human capital externalities: Theory with application to US cities". IZA Discussion Paper No. 488.

Conley, T., Flyer, F., Tsiang, G. (in press). "Local market human capital and the spatial distribution of productivity in Malaysia". Contributions to the Economics and Growth of Developing Areas.
Costa, D.L., Kahn, M.E. (2000). "Power couples: Changes in the locational choice of the college educated, 1940–1990". Quarterly Journal of Economics 115 (4), 1287–1315.
Dee, T.S. (2002). "Are there civic returns to education?" Mimeo. Swathmore College.
Duranton, G., Puga, D. (2004). "Micro-foundations of urban agglomeration economies". In: Henderson, J.V., Thisse, J.-F. (Eds.), Handbook of Regional and Urban Economics, vol. 4. Elsevier, Amsterdam, pp. 2063–2117. This volume.
Farrington, D., et al. (1986). "Unemployment, school leaving and crime". British Journal of Criminology 26, 335–356.
Feddersen, T.J., Pesendorfer, W. (1996). "The swing voter's curse". American Economic Review 86 (3), 408–424.
Freeman, R. (1983). "Crime and unemployment". Chapter 6. In: Wilson, J.Q. (Ed.), Crime and Public Policy. ICS Press, San Francisco.
Freeman, R. (1995). "The labor market". Chapter 8. In: Wilson, J.Q., Petersilia, J. (Eds.), Crime. ICS Press, San Francisco.
Freeman, R. (1996). "Why do so many young American men commit crimes and what might we do about it?" Journal of Economic Perspectives 10 (1), 25–42.
Friedman, M. (1962). Capitalism and Freedom.
Gabriel, S.A., Rosenthal, S.S. (2004). "Quality of the business environment versus quality of life: Do firms and households like the same cities?" The Review of Economics and Statistics 86 (1), 438–444.
Glaeser, E.L. (1999). "Learning in cities". Journal of Urban Economics 46 (2), 254–2334377.
Glaeser, E.L., Mare, D.C. (2001). "Cities and skills". Journal of Labor Economics 19, 316–342.
Glaeser, E.L., Scheinkman, J.A., Shleifer, A. (1995). "Economic growth in a cross-section of cities". Journal of Monetary Economics 36 (1).
Gottfredson, D. (1985). "Youth employment, crime, and schooling". Developmental Psychology 21, 419–432.
Gould, E., Mustard, D., Weinberg, B. (2000). "Crime rates and local labor market opportunities in the United States: 1979–1997". Working paper.
Griliches, Z. (1986). "Economic data issues". In: Griliches, Z., Intrilligator, M.D. (Eds.), Handbook of Econometrics. North-Holland, Amsterdam.
Grogger, J. (1998). "Market wages and youth crime". Journal of Labor Economics 16 (4), 756–791.
Hanushek, E. (2002). "Publicly provided education". In: Auerbach, A., Feldstein, M. (Eds.), The Handbook of Public Economics, vol. 3. North-Holland, Amsterdam.
Helliwell, J., Putnam, J. (1999). "Education and social capital". Working paper 9999. National Bureau of Economic Research.
Jaffe, A., Trajtenberg, M., Henderson, R. (1993). "Geographic localization of knowledge spillovers as evidenced by patent citation". Quarterly Journal of Economics 108 (3), 577–598.
Jovanovic, B., Rob, R. (1989). "The growth and diffusion of knowledge". Review of Economic Studies 56 (4), 569–582.
Katz, L.F., Murphy, K.M. (1992). "Changes in relative wages, 1963–1987: Supply and demand factors". The Quarterly Journal of Economics 107, 35–78.
Krueger, A.B., Lindahl, M. (1998). "Education for growth in Sweden and the World". Industrial Relation Section Working Paper. Princeton, NJ.
Leighley, J.E., Nagler, J. (1992). "Socioeconomic class bias in turnout, 1964–1988, The voters remain the same". American Political Science Review 86 (3), 725–736.
Lochner, L. (1999). "Education, work, and crime: Theory and evidence". Rochester Center for Economic Research Working Paper No. 465.
Lochner, L., Moretti, E. (2002). "The effect of education on crime: Evidence from prison inmates, arrests, and self-reports". American Economic Review. In press.
Lucas, R.E. (1988). "On the mechanics of economic development". Journal of Monetary Economics 22, 3–42.
Machin, S., Meghir, C. (2000). "Crime and economic incentives". Working Paper. Institute for Fiscal Studies.

Marshall, A. (1890). Principles of Economics. Macmillan, New York.
Milligan, K., Moretti, E., Oreopoulos, P. (2003). "Does education improve citizenship? Evidence from the U.S. and the UK". Journal of Public Economics. In press.
Moretti, E. (2002). "Workers' education, spillovers and productivity: Evidence from plant-level production functions". American Economics Review. In press.
Moretti, E. (in press). "Estimating the external return to higher education: Evidence from cross-sectional and longitudinal data". Journal of Econometrics.
Osborne, M.J., Rosenthal, J.S., Turner, M.A. (2000). "Meetings with costly participation". American Economic Review 90 (4), 927–943.
Peri, G. (2002). "Young workers, learning, and agglomeration". Journal of Urban Economics 52 (3), 582–607.
Powell Jr., G. (1985). "American voter turnout in comparative perspective". American Political Science Review 80 (1), 17–43.
Raphael, S., Winter-Ebmer, R. (2001). "Identifying the effect of unemployment on crime". Journal of Law and Economics 44 (1).
Rauch, J.E. (1993). "Productivity gains from geographic concentration of human capital: Evidence from the cities". Journal of Urban Economics 34, 380–400.
Roback, J. (1982). "Wages, rents and the quality of life". Journal of Political Economy 90 (6), 1257–1278.
Saxenian, A. (1994). Regional Advantage: Culture and Competition in Silicon Valley and Route 128. Harvard University Press.
Smith, A. (1776). The Wealth of Nations.
Tauchen, H., Witte, A.D., Griesinger, H. (1994). "Criminal deterrence: Revisiting the issue with a birth cohort". Review of Economics and Statistics 76 (3), 399–412.
Teixeira, R.A. (1987). Why Americans Don't Vote: Turnout Decline in the United States 1960–1984. Greenwood Press.
Topel, R. (1999). "Labor market and economic growth". In: Ashenfelter, O., Card, D. (Eds.), The Handbook of Labor Economics. North-Holland, Amsterdam.
Verba, S., Nie, N.H. (1972). Participation in America: Political Democracy and Social Equality. Harper and Row.
Viscusi, K. (1986). "Market incentives for criminal behavior". Chapter 8. In: Freeman, R., Holzer, H. (Eds.), The Black Youth Employment Crisis. University of Chicago Press, Chicago.
Weisberg, H.F., Box-Steffensmeier, J.M. (1999). Reelection 1996: How Americans Voted. Chatham House Publishers.
Witte, A.D. (1997). "Crime". Chapter 7. In: Behrman, J., Stacey, N. (Eds.), The Social Benefits of Education. University of Michigan Press, Ann Arbor, MI.
Witte, A.D., Tauchen, H. (1994). "Work and crime: An exploration using panel data". NBER Working Paper 4794.
Zucker, L.G., Darby, M.R., Brewer, M.B. (1998). "Intellectual human capital and the birth of U.S. biotechnology enterprises". American Economic Review 88 (1), 290–306.