Centre de Referència en Economia Analítica

Barcelona Economics Working Paper Series

Working Paper nº 6

Identifying Human Capital Externalities:
Theory with Applications

Antonio Ciccone and Giovanni Peri

July, 2005

Centre de Referència en Economia Analítica

Identifying Human Capital Externalities:

Theory with Applications

Antonio Ciccone*

and

Giovanni Peri**

Barcelona Economics WP nº 6

July 2005

Abstract: The identification of aggregate human capital externalities is still not fully understood. The existing (*Mincerian*) approach confounds positive externalities with wage changes due to a downward sloping demand curve for human capital. As a result, it yields positive externalities even when wages equal marginal social products. We propose an approach that identifies human capital externalities whether or not aggregate demand for human capital slopes downward. Another advantage of our approach is that it does not require estimates of the individual return to human capital. Applications to US cities and states between 1970 and 1990 yield no evidence of significant average-schooling externalities.

Key Words: Human capital, Externalities, Wages, Downward Sloping Labor Demand, Imperfect substitutability.
JEL Codes: O0, O4, R0, J3.

_____

* ICREA and Universitat Pompeu Fabra;

** University of California, Davis and NBER.

1   Introduction

Depending on their strength, aggregate human capital externalities can help explain cross-country differences in economic development, the lack of capital flows to poor countries, the effects of agglomeration on economic growth, and other macroeconomic phenomena (e.g. Lucas (1988, 1990), Azariadis and Drazen (1990), Benabou (1996), Black and Henderson (1999)). Human capital externalities also determine to what extent human capital accumulation should be subsidized (e.g. Gemmell (1997), Heckman and Klenow (1998), Heckman (2000)). Assessing the strength of human capital externalities is therefore important for applied economic theory as well as economic policy, and empirical research has responded with a variety of different approaches and estimates (e.g. Rauch (1993), Rudd (2000), Acemoglu and Angrist (2001), Conley, Flier, and Tsang (2003), Moretti (2004a, 2004b); see Moretti (2004c) for a survey). Existing work using wages achieves identification by assuming that all effects of the supply of human capital on individual wages are due to externalities. The strength of externalities can therefore be obtained as the effect of the aggregate supply of human capital on individual wages in an otherwise standard Mincerian wage regression (e.g. Rauch (1993), Rudd (2000), Acemoglu and Angrist (2001), Conley, Flier, and Tsang (2003), Moretti (2004a)). This is what we refer to as the *Mincerian* approach to the identification of aggregate human capital externalities.

   In principle, wages may respond to the aggregate supply of human capital because of externalities or because of a downward sloping demand curve for human capital. For example, Fallon and Layard (1975), Katz and Murphy (1992), Angrist (1995), Johnson (1997), Topel (1997), Autor, Katz, and Krueger (1998), Card and Lemieux (2001), and Borjas (2003) show that changes in the education wage premium can be partly explained by supply driven movements along a downward sloping relative demand curve for more educated workers. We therefore analyze the identification of externalities in a framework where the demand for human capital falls as its cost rises. Following the empirical literature, the slope of the demand curve is linked to the substitutability between different levels of human capital in production. In this framework it can be shown that the Mincerian approach to the identification of human capital externalities yields positive externalities even when wages equal marginal social products. Using estimates of the elasticity of substitution between more and less educated workers in the empirical literature, we find an upward bias of the Mincerian approach of between 60 and 70 percent of the individual return to schooling in a first-order approximation and somewhat larger in simulations.

   We propose an alternative approach to the identification of human capital externalities. The theoretical basis is that, under general conditions, the strength of human capital externalities equals the average earnings-weighted effect of human capital on wages, which in turn equals the effect of human capital on the average wage when holding the labor-force skill-

1

composition constant. This result is easiest to explain in the case with two production factors only, S more educated workers and $U = 1 - S$ less educated workers (total employment will be held constant and is normalized to unity). If more educated workers have a positive externality of strength EXT on output Y, their marginal social product $\partial Y / \partial S$ exceeds the difference between the wage of a more and a less educated worker $w_S - w_U$ by EXT, $\partial Y / \partial S = EXT + (w_S - w_U)$. Our approach to identification can now be readily derived from the equality between output and labor income, $Y = w_U (1 - S) + w_S S$. Differentiating both sides with respect to the supply of more educated workers implies $\partial Y / \partial S = (w_S - w_U) + (1 - S)(\partial w_U / \partial S) + S(\partial w_S / \partial S)$ where $\partial w_i / \partial S$ denotes first-order effects of supply on wages. Combining this last expression for the marginal social product of more educated workers with $\partial Y / \partial S = EXT + (w_S - w_U)$ yields the externality in function of the response of wages to the supply of more educated workers: $EXT = (1 - S)(\partial w_U / \partial S) + S(\partial w_S / \partial S)$. Dividing both sides by Y so that externalities are measured in percentage points of output and using $\boldsymbol{b}$ to denote the share of more educated workers in earnings $\boldsymbol{b} = Sw_S / Y$, results in our key equation:

$$\frac{EXT}{Y} = [1 - \boldsymbol{b}]\frac{\partial w_U / \partial S}{w_U} + [\boldsymbol{b}]\frac{\partial w_S / \partial S}{w_S} = \frac{\partial}{\partial S}\ln\left([1 - \bar{S}]w_U + [\bar{S}]w_S\right), \qquad (1)$$

where upper bars denote values that are held constant. Externalities can therefore be identified as the earnings-weighted average percentage-change in wages (the first equality). Or, alternatively, they can be identified as the log-change in the average wage holding skill-composition constant (the second equality). This is what we refer to as the *constant-composition* approach to the identification of aggregate human capital externalities. We show that this approach is easily modified to account for higher-order effects of the supply of human capital on wages.

The approach to identification emerging from this theoretical argument can be used to estimate externalities at the city, region, or country level over any time period in two steps. The first step requires obtaining wages w_x and labor-force shares l_x by skill type x in each city, region, or country at the beginning and the end of the relevant time period to calculate the log-change in the average wage holding skill-composition constant, $\ln\left(\sum \bar{l}_x \tilde{w}_x\right) - \ln\left(\sum \bar{l}_x \bar{w}_x\right)$ where upper bars denote beginning-of-period values and tildes end-of-period values. The second step consists of regressing the log-change in constant-composition average wages on (exogenous) changes in the supply of human capital and other determinants of wages. If the change in the supply of human capital enters positively and significantly, this indicates positive externalities. Higher-order effects of human-capital supply on wages can be dealt with by either including higher-order changes in supply among the regressors or by using an average of the beginning-of-period and end-of-period skill-composition to calculate the constant-composition log-wage change.

We show that ultimately the Mincerian approach identifies human capital externalities as the employment-weighted average percentage-change in wages in response to a greater supply

of human capital, $[1-S](\partial w_U/\partial S)/w_U + [S](\partial w_S/\partial S)/w_S$ in the example above. When the relative demand curve for more educated workers is downward sloping, this approach yields externalities even when wages equal marginal social products. To see this note that (1) implies that earnings-weighted wage changes must average to zero when wages equal marginal social products: $0 = EXT = [1-\boldsymbol{b}](\partial w_U/\partial S)/w_U + [\boldsymbol{b}](\partial w_S/\partial S)/w_S$. The slope of the relative demand curve for more educated workers is key in how wage changes average out. When the relative demand curve is flat, wages do not change in response to relative supply, and wage changes therefore average to zero trivially. But when the relative demand curve for more educated workers is downward sloping, earnings-weighted wage changes average to zero because the strictly positive effect of a greater relative supply of more educated workers on wages of less educated workers $(\partial w_U/\partial S)/w_U > 0$ is offset by the strictly negative effect on wages of more educated workers $(\partial w_S/\partial S)/w_S < 0$. Weighting these wage changes by employment, as in the Mincerian approach, instead of earnings amounts to putting more weight on the rising wage of less educated workers and less weight on the falling wage of more educated workers because less educated workers earn a lower wage than more educated workers. As a result, the Mincerian approach yields positive externalities when there are none.[1]

Another advantage of the constant-composition approach to human capital externalities compared to the Mincerian approach is that it does not require estimating individual returns to human capital. The constant-composition approach can therefore be used even when instruments for individual schooling are unavailable. The Mincerian approach is implemented by estimating a Mincerian wage regression and therefore requires instruments for aggregate schooling as well as individual schooling (Acemoglu and Angrist (2001)).

Our discussion of the identification of human capital externalities when production requires many different, imperfectly substitutable levels of human capital is based on the aggregate human capital framework.[2] We show that this framework yields a parsimonious way of capturing imperfect substitutability. An additional advantage is that it encompasses the Mincerian approach to externalities. The defining feature of the aggregate human capital framework is that the distribution of human capital affects wages only through average human capital. The approach to identification emerging from the human capital framework carries over with minimal variations to any framework where the distribution of skill types affects wages only through a single measure of supply (like, for example, the framework of Katz and Murphy (1992)).[3] Assuming that the supply of human capital can be summarized in a single measure is inevitable in empirical applications because of the difficulties in finding

---

[1] The next section shows how to modify the Mincerian approach to ensure weighting by earnings.

[2] E.g. Lucas (1988), Mankiw, Romer, and Weil (1992), Benhabib and Spiegel (1994), Klenow and Rodriguez-Clare (1997), Topel (1999), Bils and Klenow (2001), and Krueger and Lindahl (2001).

[3] We show this in Appendix A.6. The KM (pp. 67-69) framework does not encompass the Mincerian approach to human capital externalities however, which makes comparisons between the constant-composition and the Mincerian approach less straightforward than in the human capital framework.

instruments for multiple measures of (endogenous) supply.[4] Using a theoretical framework where the use of a single supply measure is justified is therefore practical. It is not necessary for the theoretical validity of the constant-composition approach however, which we show can be used to identify human capital externalities even when the whole human capital distribution matters for wages.

Our main theoretical result in the aggregate human capital framework is that the elasticity of the average wage holding skill-composition constant with respect to average human capital is equal to the strength of average human capital externalities. This result holds whether or not the demand curve for human capital slopes downward. We also analyze second-order effects of average human capital on average wages holding skill-composition constant and prove that these are positive. Moreover, we show that the constant-composition approach can be used to identify human capital externalities even when externalities are biased towards workers with high or low levels of human capital.

As an application of the constant-composition approach, we assess the strength of average-schooling externalities in US cities and US states between 1970 and 1990 using instrumental-variable estimation methods to account for endogenous average schooling. Our results yield no evidence of statistically significant average-schooling externalities. Constant-composition point estimates of the external return to a one-year increase in average schooling are around zero at the city level and not much higher at the state level. Using the Mincerian approach to estimate average-schooling externalities over the same period yields a statistically significant external return around 8 percent at the city level and around 10 percent at the state level. Hence, Mincerian estimates of external returns to schooling are of a similar magnitude as private returns to schooling (e.g. Card (1999)) while constant-composition estimates are relatively small and statistically insignificant. We show using calibration and simulations that this difference in results between the Mincerian and the constant-composition approach is consistent with the degree of imperfect substitutability between more and less educated workers found in the literature (e.g. Johnson (1997), Fallon and Layard (1975), Katz and Murphy (1992), and Ciccone and Peri (forthcoming)).

The variables used as instruments for city level changes in average schooling between 1970 and 1990 are the demographic structure of the labor-force and population as well as the population share of African-Americans in 1970. These variables have predictive power for the change in average schooling at the city level because younger individuals entering the labor-force during this time period had higher levels of schooling than workers going into retirement and because African-Americans were catching up in schooling with the rest of the population. Our identifying hypothesis is that the variables used as instruments affect wage growth of white workers between 1970 and 1990 at the city level only through the supply of human capital. We check this hypothesis by testing the implied overidentifying restrictions and find it cannot be rejected at standard significance levels. Our instruments for changes in average schooling at the state level are either the compulsory-schooling and child-labor law

---

[4] Existing empirical studies use either average years of schooling of the share of workers with schooling above a certain level as a summary measure of the aggregate supply of human capital.

indicators of Acemoglu and Angrist (2001) or the same instruments used at the city level (the two sets of instruments yield basically identical estimates).

The Mincerian approach to human capital externalities was introduced by Rauch (1993) to estimate average-schooling externalities in a cross-section of US cities in 1980. Acemoglu and Angrist (1999, 2001) extend the approach to a panel of US states and account for state fixed effects as well as for the endogeneity of average and individual schooling. Their approach yields no evidence of significant schooling externalities between 1960 and 1980 (the period they focus on). Acemoglu and Angrist are also the first to show that human capital externalities as identified by the Mincerian approach subsume imperfect substitutability between skills (in Appendix A.2 of Acemoglu and Angrist (1999)). Another application of the Mincerian approach to human capital externalities at the US state level is Ruud (2000). Conley, Flier, and Tsang (2003) employ the Mincerian approach to estimate human capital externalities in Malaysian regions. Moretti (2004a) employs the Mincerian approach to estimate externalities associated with increases in the share of college-graduates in US cities between 1980 and 1990. Moretti also proposes an alternative to the Mincerian approach, which consists of testing whether a greater share of college-graduates in cities leads to an increase in their wages. He finds evidence that this was the case between 1980 and 1990.

The remainder of the paper is organized in the following way. Section 2 compares the constant-composition and Mincerian approach to the identification of human capital externalities in the simplest possible case. Section 3 derives the constant-composition approach in the aggregate human capital framework. Section 4 presents the estimating equations and explains the estimation methods. Section 5 describes the data. Section 6 discusses our empirical results and Section 7 summarizes.

2   The Case of Two Skill Types

Before turning to the framework with many different levels of human capital it is useful to elaborate on the differences between the constant-composition and Mincerian approach in the simplest possible case. We therefore return to the case with two production factors only, a number of more educated workers S and a number of less educated workers $U = 1 - S$, and also assume that production is subject to constant returns to scale. In this setting, a competitive equilibrium can be characterized by the following two conditions. First, that the cost-minimizing demand of firms for more relative to less educated workers equals the relative supply of the two types of workers. Second, that full-employment output is equal to aggregate income $[1 - S]w_U + [S]w_S$.

In this framework, the relative supply of more educated workers may affect wages because of externalities or because of a downward sloping relative demand curve for more educated workers (we will refer to wage changes due to downward sloping demand as *neoclassical supply effects*). If each one-point increase in the share of more educated workers has an external effect $EXT = qY$ on output, (1) implies that the strength of the externality q satisfies

5

$$q = [1-b]\frac{\partial w_U / \partial S}{w_U} + [b]\frac{\partial w_S / \partial S}{w_S} = \frac{\partial}{\partial S}\ln\left([1-\overline{S}]w_U + [\overline{S}]w_S\right), \qquad (2)$$

where b continues to denote the share of more educated workers in earnings and upper bars continue to denote values that are held constant. Hence, the strength of externalities can be identified either as the earnings weighted percentage-change in wages due to an increase in the supply of more educated workers or as the effect of an increase in the supply of more educated workers on the constant-composition log-wage.

A useful alternative perspective on the constant-composition approach can be obtained by subtracting the strength of the human capital externality q from both sides of the first equality in (2). This yields

$$0 = [1-b]\left(\frac{\partial w_U / \partial S}{w_U} - q\right) + [b]\left(\frac{\partial w_S / \partial S}{w_S} - q\right) = [1-b]\underset{[+]}{NCSP_U} + [b]\underset{[-]}{NCSP_S}, \qquad (3)$$

where *NCSP* denotes neoclassical supply effects and [+] and [-] denote whether the effect is positive or negative. $NCSP_U = (\partial w_U / \partial S)/w_U - q \geq 0$ is the increase in the wage of less educated workers net of externalities and $NCSP_S = (\partial w_S / \partial S)/w_S - q \leq 0$ the decrease in the wage of more educated workers net of externalities. Hence, the constant-composition approach exploits that neoclassical supply effects offset each other when weighted by earnings.

The Mincerian approach to human capital externalities obtains the strength of externalities q^M as the marginal effect of human capital on the intercept of a Mincerian wage regression (e.g. Rauch (1993)). With two types of labor only, the Mincerian approach is based on the following model for wages: $\ln w_i = q^M S + a + bD_i$ where w_i is the wage of worker i and D_i is 1 if the worker is of the high-education type and 0 otherwise ($q^M S + a$ is the log-wage of less educated workers and b the education log-wage premium). Summing across individuals yields $[1-S]\ln w_U + [S]\ln w_S = q^M S + a + bS$ and hence

$$q^M = \frac{\partial}{\partial S}\left([1-S]\ln w_U + [S]\ln w_S\right) - b. \qquad (4)$$

The Mincerian approach therefore identifies human capital externalities as the wedge between the marginal effect of the supply of more educated workers on the average log-wage (the first term) and the individual log-wage premium b (Acemoglu and Angrist (1999)). Differentiating the first term on the right-hand-side of (4) and making use of $b = \ln w_S - \ln w_U$ yields

$$q^M = [1-S]\frac{\partial w_U / \partial S}{w_U} + [S]\frac{\partial w_S / \partial S}{w_S} = q + \left([1-S]\underset{[+]}{NCSP_U} + [S]\underset{[-]}{NCSP_S}\right), \qquad (5)$$

which can be easily compared to the constant-composition approach in (2). (*NCSP* continues to denote neoclassical supply effects: $NCSP_i = (\partial w_i / \partial S)/w_i - q$.) The first equality makes

clear that the constant-composition and the Mincerian approach ultimately differ only in the weights applied to wage changes. The second equality shows that, unsurprisingly, both approaches yield a consistent estimate of the strength of externalities when there are no neoclassical supply effects, $NCSP_i = 0$. But in the presence of neoclassical supply effects, $\boldsymbol{q}^M$ is strictly greater than the externality $\boldsymbol{q}$, $\boldsymbol{q}^M - \boldsymbol{q} = [1-S]\underset{[+]}{NCSP^U} + [S]\underset{[-]}{NCSP^S} > 0$. Hence, the Mincerian approach overstates human capital externalities. To see this notice that neoclassical supply effects on wages of less and more educated workers offset when weighted by *earnings* shares (see (3)) and that the employment share of less educated workers exceeds their earning share (because they earn a lower wage than more educated workers). The positive neoclassical supply effect on wages of less educated workers therefore more than offsets the negative effect on wages of more educated workers when weighted by *employment* shares.[5] The Mincerian approach therefore confounds positive externalities with neoclassical supply effects (wage changes due to a downward sloping demand curve for more relative to less educated workers).

There is another advantage of the constant-composition compared to the Mincerian approach to human capital externalities. The Mincerian approach is implemented by estimating a Mincerian wage regression and therefore requires instruments for aggregate and individual schooling (Acemoglu and Angrist (2001)). The constant-composition approach is implemented by regressing log-changes in the average wage holding composition constant over the relevant period, $\ln\left([1-\overline{S}]\tilde{w}_U + [\overline{S}]\tilde{w}_S\right) - \ln\left([1-\overline{S}]\overline{w}_U + [\overline{S}]\overline{w}_S\right)$ where upper bars denote beginning-of-period values and tildes end-of-period values, on changes in schooling $\tilde{S} - \overline{S}$. This only requires instruments for aggregate schooling, and the constant-composition approach can therefore be implemented even when instruments for individual schooling are unavailable.

It is straightforward to extend the constant-composition approach to account for higher-order effects of the supply of more educated workers on wages. The simplest way is to include higher-order changes in the supply of more educated worker as regressors in the empirical analysis. An alternative is to use the constant-composition approach to put upper and lower bounds on the strength of externalities. The only additional assumption required is that production is subject to constant or decreasing returns to more educated workers net of externalities. In this case it can be shown that the following inequalities hold: $\Delta\ln\left([1-\overline{S}]w_U + [\overline{S}]w_S\right)/\Delta S \geq \boldsymbol{q} \geq \Delta\ln\left([1-\tilde{S}]w_U + [\tilde{S}]w_S\right)/\Delta S$ where Δ denotes the difference

---

[5] In this framework, there is a simple way to modify the Mincerian approach that results in a consistent estimate of the strength of aggregate human capital externalities. This change consists in weighting individual observations by earnings. The starting point of the earnings-weighted Mincerian approach is $w_i \ln w_i = w_i \boldsymbol{q}^M S + w_i a + w_i b D_i$. Summing across individuals and dividing by total earnings $Uw_U + Sw_S$ yields earnings, where $\boldsymbol{b}$ continues to denote the earnings share of more educated workers. Differentiating with respect to the share of more educated workers S, $\boldsymbol{q}^M = (1-\boldsymbol{b})(\partial w_U/\partial S)/w_U + \boldsymbol{b}(\partial w_S/\partial S)/w_S + \partial\boldsymbol{b}/\partial S\left((\ln w_S - \ln w_U) - b\right)$. Using $b = \ln w_S - \ln w_U$ yields $\boldsymbol{q}^M = (1-\boldsymbol{b})(\partial w_U/\partial S)/w_U + \boldsymbol{b}(\partial w_S/\partial S)/w_S$, which is identical to the constant-composition approach (see (2)). The constant-composition approach and the earnings-weighted Mincerian approach are therefore equivalent.

between end-of-period values and beginning-of-period values.[6] Hence, the constant-composition approach yields a lower bound on externalities when the (constant) skill-composition used corresponds to the end of the period and an upper bound when the beginning-of-period skill-composition is used instead. Or, to put it differently, the constant-composition approach yields the exact strength of externalities when the (constant) skill-composition used is an appropriately weighted average of beginning-of-period and end-of-period values:

$$q = \frac{\Delta \ln \left([1 - S^*] w_U + [S^*] w_S \right)}{\Delta S}, \tag{6}$$

for $\bar{S} \leq S^* \leq \tilde{S}$. Hence, a third way to handle higher-order effects of the supply of more educated workers on wages is to employ an average of the beginning-of-period and end-of-period skill-composition to calculate the constant-composition log-wage change.

A   Quantifying the Bias of the Mincerian Approach

The bias of the Mincerian approach to human capital externalities depends on the slope of the demand curve for more relative to less educated workers. There are several estimates of this slope in the literature and we now consider a framework that allows us to draw on these estimates to quantify the bias of the Mincerian approach.

Calibrating the Bias of the Mincerian Approach

Suppose there are no externalities and that firms produce output Y using a constant-elasticity-of-substitution production function,

$$Y = \left(U^{\frac{s-1}{s}} + (BS)^{\frac{s-1}{s}} \right)^{\frac{s}{s-1}}, \tag{7}$$

where U, S continue to denote the number of less and more educated workers, B captures skill-biased technology, and s denotes the elasticity of substitution between more and less educated workers. In this case, the bias of the Mincerian approach to human capital externalities is

$$\text{Bias of Mincerian Approach} = \frac{1}{s} \left(\frac{w_S - w_U}{w} \right). \tag{8}$$

---

[6] If production is concave in more educated workers net externalities, the following two inequalities hold $\quad \ln \left([1 - \bar{S}] \tilde{w}_U + [\bar{S}] \tilde{w}_S \right) \geq \ln \left([1 - \bar{S}] \bar{w}_U + [\bar{S}] \bar{w}_S \right) + q(\Delta S) \quad$ and $\quad \ln \left([1 - \tilde{S}] \bar{w}_U + [\tilde{S}] \bar{w}_S \right) \geq \ln \left([1 - \tilde{S}] \tilde{w}_U + [\tilde{S}] \tilde{w}_S \right) - q(\Delta S)$. Hence, $\Delta \ln \left([1 - \bar{S}] w_U + [\bar{S}] w_S \right) \geq q(\Delta S) \geq \Delta \ln \left([1 - \tilde{S}] w_U + [\tilde{S}] w_S \right)$. Note that production must be concave in more educated workers net externalities for a competitive equilibrium to exist.

8

Hence, for a given education wage premium the bias is decreasing in the elasticity of substitution between more and less educated workers. This is intuitive because higher values of s imply a flatter relative demand curve for more educated workers and therefore a weaker response of wages to supply-driven movements along the demand curve.

What is the size of the bias of the Mincerian approach for reasonable values for the elasticity of substitution between more and less educated workers and for reasonable values for the education wage premium? When we take more educated workers to be people with two or more years of college and less educated workers to be everybody else, the education premium $(w_H - w_L)/w$ averages to 40 percent in 1970 and 1990 for white males 40-49 (US Bureau of Census (1970, 1990)). (By focusing on white males 40-49 we sidestep the estimation of wage differentials associated with gender, race, and experience.) Almost all available estimates of the elasticity of substitution between college and high-school workers in the US point to values around 1.5 (e.g. Katz and Murphy (1992), Ciccone and Peri (forthcoming)). Combined, these numbers imply that imperfect substitutability may add 27 percent (0.4/1.5) to the Mincerian estimate of the external return to more educated workers. To facilitate the interpretation of this estimate, we make use of the fact that average schooling of white males 40-49 with two or more years of college exceeded average schooling of white males with less than two years of college in the same age group by 4.2 years in 1970 and 1990. Hence, an increase in the share of more educated workers by 24 percent (1/4.2) amounts to a one-year increase in average schooling. Imperfect substitutability may therefore add around 6.5 percent (0.27*0.24) to the Mincerian estimate of the external return of a one-year increase in average schooling.

Simulating the Bias of the Mincerian Approach

The 6.5-percent estimate of the bias of the Mincerian approach to human capital externalities is based on small changes in the relative supply of more educated workers. This may not reflect the bias in actual applications. To get an alternative estimate of the bias we therefore apply the Mincerian approach to city-level wage data that is generated by combining the constant-elasticity-of-substitution production function above with several features of the distribution of wages and human capital across US cities in 1970 and 1990.

The model underlying our simulations is built on the production function in (7). We continue to assume that there are no externalities and that the elasticity of substitution between more and less educated workers is $s = 1.5$. Assuming competitive labor markets implies that the wage of both types of workers is equal to their marginal productivity

$$w_{U,ct} = \left(1 + \left(\frac{B_{ct} S_{ct}}{1 - S_{ct}}\right)^{\frac{s-1}{s}}\right)^{\frac{1}{s-1}} \tag{9}$$

and

$$w_{S,ct} = B_{ct} \left[1 + \left(\frac{1 - S_{ct}}{B_{ct} S_{ct}} \right)^{\frac{s-1}{s}} \right]^{\frac{1}{s-1}}, \tag{10}$$

where c, t are subscripts for city and year respectively.

Our simulations assume 163 cities, because this is the number of cities in our empirical application. The initial shares of more educated workers are chosen to match the share of workers with two or more years of college in each city in the 1970 US Census, which will be denoted by S_{c1970} (Section 5 contains a description of the Census data used). The initial levels of skill-biased technology, which will be denoted by B_{c1970}, are chosen to match the level of skill-biased technology implicit in the data assuming that (9) and (10) hold. More precisely, (9) and (10) imply that B_{c1970} is linked to relative wages and relative supplies of more educated workers by

$$\ln B_{c1970} = \frac{s}{s-1} \ln \left(\frac{w_{S,c1970}}{w_{U,c1970}} \right) + \frac{1}{s-1} \left(\frac{S_{c1970}}{1 - S_{c1970}} \right). \tag{11}$$

We will measure $w_{S,c1970}$ ($w_{U,c1970}$) as the average wage of white male workers aged 40-49 with 2 or more years of college (less than 2 years of college) in city c in 1970.

Starting from this calibration of the initial values for the relative supply of more educated workers and levels of skill-biased technology for each city we generate 163 city-specific human capital shocks, ΔS_c, and skill-biased technology shocks, ΔB_c. These shocks are drawn from identical and independent normal distributions. The standard deviation is chosen to match the standard deviation of $S_{c1990} - S_{c1970}$ (S_{c1990} is obtained analogously to S_{c1970}) and the mean is set to zero. The mean and standard deviation of the distribution for the skill-biased technology shock are chosen to match the mean and standard deviation of $B_{c1990} - B_{c1970}$ (where B_{c1990} is obtained analogously to B_{c1970}).

We take a total of 5000 draws from the distribution for the shocks to human capital and skill-biased technology (a draw consists of 163 human capital shocks and 163 skill-biased technology shocks). Starting from the calibrated values for S_{c1970} and B_{c1970}, each draw results in 163 values for S_{c1990} and B_{c1990}, which substituted in (9) and (10) yield 163 values for $w_{S,c1990}$ and $w_{U,c1990}$. Combining this data on wages with relative supplies of more and less educated workers allows us to calculate the city-specific intercept of a Mincerian wage regression $\ln a_{c1990}$ for each draw as

$$\ln a_{c1990} = \left(S_{c1990} \ln(w_{S,c1990}) + (1 - S_{c1990}) \ln(w_{U,c1990}) \right) - S_{ct} \overline{\overline{(\ln w_{S,c1990} - \ln w_{U,c1990})}}, \tag{12}$$

where the double upper bar denotes the average across cities. Repeating this calculation using the 1970 values yields $\ln a_{c1970}$. The Mincerian approach identifies the strength of aggregate

human capital externalities as the effect of aggregate human capital on the intercept of a Mincerian wage regression. Hence, the estimating equation for the strength of human capital externalities using the Mincerian approach q^M is

$$\ln \boldsymbol{a}_{c1990} - \ln \boldsymbol{a}_{c1970} = \text{constant} + \boldsymbol{q}^M \Delta S_c. \tag{13}$$

Estimating this equation using least squares yields a point estimate $\hat{\boldsymbol{q}}_d^M$ and a standard error $\hat{\boldsymbol{r}}_d^M$, where d denotes a specific draw for the 163 city-specific human capital shocks and skill-biased technology shocks, for each of the 5000 draws.

The constant-composition estimate of the strength of the human capital externality is obtained in the following way. For each of the 5000 draws for city-specific human capital shocks and skill-biased technology shocks we calculate the log average wage in 1990 using the 1970 labor-force composition,

$$\ln w_{c1990}^F = \ln \left(S_{c1970} w_{S,c1990} + (1 - S_{c1970}) w_{U,c1990} \right) \tag{14}$$

and then estimate the equation

$$\ln w_{c1990}^F - \ln w_{c1970} = \text{constant} + \boldsymbol{q}^{CC} \Delta S_c, \tag{15}$$

using least squares. This yields a point estimate $\hat{\boldsymbol{q}}_d^{CC}$ and a standard error $\hat{\boldsymbol{r}}_d^{CC}$ for each of the 5000 draws.

Table 1 summarizes the results of our simulations. Panel A contains the answer to the following question. Suppose we use the point estimate $\hat{\boldsymbol{q}}_d^M$ and standard error $\hat{\boldsymbol{r}}_d^M$ obtained with the Mincerian approach to test the hypothesis that the strength of human capital externalities is equal to zero against the alternative of positive externalities at some standard significance level (assuming asymptotic normality of the estimator). What fraction of the 5000 draws would result in rejection of the null hypothesis? The entries in the lower right-hand cell of the table for example indicate that we would reject the hypothesis of no human capital externalities for 52 percent (75 percent) of the draws when tests are performed at the 5-percent (10-percent) significance level. Results in this particular cell are based on simulations assuming a standard deviation of the human capital shock ($StdDev(\Delta S)$) equal to 0.04 and a standard deviation of the skill-biased technology shock ($StdDev(\Delta A)$) equal to 0.08, which are the values implied by the calibration described above. Hence, the hypothesis of no human capital externalities is rejected far too frequently given the nominal size of the test and the fact that there are no human capital externalities in the model underlying the simulations. Other cells in the table contain analogous results for different values of $StdDev(\Delta S)$ and $StdDev(\Delta A)$. It can be seen that a reduction of $StdDev(\Delta A)$ implies that the hypothesis of no externalities is rejected even more frequently and that the frequency of rejection reaches 100 percent when $StdDev(\Delta A)$ is equal to 0.

Panel B of Table 1 answers exactly the same question for the constant-composition-approach simulation results. That is, suppose we use the point estimate $\hat{q}_d^{CC}$ and standard error $\hat{r}_d^{CC}$ obtained with the constant-composition approach to test the hypothesis that the strength of human capital externalities is equal to zero at some standard significance level (assuming asymptotic normality of the estimator). What is the frequency of rejection? The entries in the lower right hand corner of the table indicate that we would reject the hypothesis of no human capital externalities in 1.1 percent (4.8 percent) of the draws when tests are performed at the 5-percent (10-percent) significance level. Hence, compared to the Mincerian approach, the hypothesis of no human capital externalities is rejected far less frequently (rejection frequencies using the Mincerian approach were 47 times greater in the case of the 5-percent test and 16 times greater in the case of the 10-percent test). This is desirable as the underlying simulations assume that there are no human capital externalities. The results in Panel A and B are virtually unchanged when we increase the number of draws from 5000 to 10000.

Table 2 contains the average estimates of the strength of human capital externalities in our simulations using the Mincerian as well as constant-composition approach. The values in the table should be read as the simulated bias of the two different approaches as the model underlying the simulations assumes that there is no human capital externality. Panel A contains the bias of the Mincerian approach to human capital externalities. The range of values is between 0.33 and 0.35, depending on the standard deviation of the human capital shock and the skill-biased technology shock. Hence, according to our simulations the Mincerian estimate of human capital externalities is biased upward. The size of the bias is somewhat greater than suggested by the calibrations based on first-order effects in the previous section (which were between 0.24 and 0.29). To facilitate the interpretation of the range of estimated values, we again make use of average schooling of white males aged 40-49 with two or more years of college exceeding average schooling of those with less than two years of college by 4.2 years in 1970 and 1990. Hence, an increase in the share of more educated workers by 24 percent amounts to a one-year increase in average schooling. The range of estimates obtained using the Mincerian approach therefore implies an upward bias of average-schooling-externality estimates between 8 and 8.4 percent. The numbers in square brackets below the estimate are the fraction of simulations yielding positive Mincerian estimates of the strength of human capital externalities.

Panel B in Table 2 contains the simulated average bias of the constant-composition approach to human capital externalities. It can be seen that the bias is rather small (between -0.003 and -0.018). Using the same approach as in the Mincerian case to facilitate the interpretation of these values (that an increase in the share of more educated workers by 24 percent amounts to a one-year increase in average schooling), the constant-composition estimates imply a bias of average-schooling-externality estimates between -0.07 percent and -0.4 percent. Hence, the bias is very small relative to the Mincerian bias (the absolute value of the Mincerian bias is between 11 and 24 times greater) and it does not imply any economically significant human capital externality. The numbers in square brackets below the estimates are

the fraction of simulations yielding positive constant-composition estimates of the strength of human capital externalities.

Summarizing, our simulations matching city-level data on wages and human capital most closely yield a Mincerian estimate of the city-level external return to a one-year increase in average schooling of around 8 percent, which is somewhat greater than the first-order bias. The constant-composition estimate of the external effect closely reflects the absence of a human capital externality in the model underlying our simulations. The simulations have also shown that an econometrician using the Mincerian approach to human capital externalities would reject the null hypothesis of no human capital externalities far too often in favor of positive externalities.

3   The Human Capital Framework with Externalities

We now turn to the identification of human capital externalities in the aggregate human capital framework. A key feature of this framework is that the supply of different levels of human capital affects individual wages only through average human capital. Another important feature is that the framework captures imperfect substitutability among workers with many different levels of human capital in a parsimonious way.[7]

Suppose that output Y of cities (or other spatial units) depends on the aggregate amount of labor L and human capital H employed according to the following production function

$$Y = AF(L,H), \tag{16}$$

where A denotes the level of total factor productivity (TFP) in the city and

$$H \equiv \sum_x xL(x), \tag{17}$$

where $L(x)$ is the number of workers with human capital x in the city (using this notation the aggregate amount of labor in the city is $L \equiv \sum_x L(x)$). Assume also that the aggregate production function is twice continuously differentiable and subject to constant returns to scale to labor $L(x)$ for all x (or, alternatively, subject to constant returns to scale to L,H) as well as constant or decreasing returns to human capital, $F_{22}(L,H) \leq 0$.

Firms in each city produce according to (16) and maximize profits taking the city specific levels of TFP as given. Suppose also that product and labor markets are perfectly competitive and that output is tradable. Under these assumptions the equilibrium product wage of workers with human capital x in a city with a supply of human capital relative to labor $h \equiv H/L$ can be written as

$$w(x,h) \equiv A\boldsymbol{w}_L(h) + A\boldsymbol{w}_H(h)x$$

---

[7] In Appendix A.7 we show that the constant-composition approach can be used to identify externalities even if the whole distribution of human capital matters for individual wages, there are no restrictions on the pattern of substitutability among different types of workers, and externalities are driven by the whole distribution of worker types. However, data requirements necessary to implement the constant-composition approach are formidable in this case.

$$\boldsymbol{w}_L(h) \equiv F_1(1,h) \text{ and } \boldsymbol{w}_H(h) \equiv F_2(1,h),$$

The wage of workers with human capital x is therefore the sum of two components: the price of labor, $A\boldsymbol{w}_L \equiv w_L$, (multiplied by the one unit supplied) and the price of human capital, $A\boldsymbol{w}_H \equiv w_H$, multiplied by the quantity of human capital supplied (x). A higher level of TFP translates into a higher price of labor and human capital. An increase in the relative supply of human capital h raises the price of labor but lowers the price of human capital in the case of strictly decreasing returns to human capital, $F_{22}(L,H) < 0$, and leaves them unchanged if $F_{22}(L,H) = 0$.

We will allow for the possibility that the marginal social product of workers with above-average (below-average) human capital is greater (smaller) than their equilibrium wage. This is accomplished by assuming that TFP in each city may be increasing in the average level of human capital h in the city

$$A = h^q, \tag{19}$$

where q captures the strength of average human capital externalities. This setup yields $A = 1$ if $q = 0$, which combined with (18) allows us to interpret $w(x,h) \equiv \boldsymbol{w}_L(h) + \boldsymbol{w}_H(h)x$ as the wage of workers with human capital x in the absence of externalities, and $\boldsymbol{w}_L$ and $\boldsymbol{w}_H$ as the price of labor and human capital in the absence of human capital externalities. While our discussion of the identification of aggregate human capital externalities focuses on non-pecuniary externalities, the same issues arise when externalities have a pecuniary origin as in Acemoglu (1996) for example.

Whether product wages of identical workers in different cities will be equalized or not depends on the motivations for inter-city migration. Identical workers in different cities will earn different product wages in equilibrium if cities differ in characteristics that are relevant for workers' utility. Examples of such characteristics are the cost of housing, the quality of local public schools, local tax rates, the degree of air pollution, the crime rate, climate, and recreational opportunities.

The model presented so far is the simplest framework that allows us to discuss identification of human capital externalities when workers with many different levels of human capital may be imperfect substitutes. It can be extended in several dimensions without affecting our theoretical results on identification or our empirical approach. The most basic extension would include physical capital and land as factors of production and distinguish between tradable and non-tradable goods. Allowing for physical capital as a production factor does not alter our approach at all when physical capital moves to equalize its rate of return across cities. The main implication of extending the theoretical analysis to allow for land as a production factor is that our approach identifies externalities net of congestion effects. The main insight of allowing for non-tradable goods is that only externalities in the tradable goods sector are identified. All these extensions are discussed in the Appendix of Ciccone and Peri (2002). It may be worthwhile to point out that the model with land and physical capital has

many similarities with the theoretical work of Roback (1982). The constant-composition approach can also be used to identify human capital externalities at the aggregate level when physical capital is not perfectly mobile across the geographic units of analysis (the relevant case for human capital externalities at the country level), see Appendix A.3.

Substitutability and Returns to Human Capital

The framework described so far is flexible enough to allow workers with different levels of human capital to be perfect or imperfect substitutes in production. It is straightforward to show that assuming perfect substitutability is equivalent to assuming constant marginal returns to human capital given TFP, $F_{22}(1,h) = 0$, or to assuming that the production function in (16) simplifies to

$$Y = A(L + BH) \tag{20}$$

where B determines the marginal rate of substitution between labor and human capital. In this case, wages of workers with a given level of human capital and the return to human capital will be independent of the average level of human capital in the city for a given level of TFP. Hence, all effects of the average level of human capital on the equilibrium wage curve must arise through TFP and can be interpreted as externalities.

Imperfect substitutability among different types of workers in production on the other hand is equivalent to decreasing marginal returns to human capital, $F_{22}(1,h) < 0$. To see this suppose that the supply of workers with low human capital x^l in a city decreases while the supply of workers with high human capital x^h ($x^h > x^l$) increases so as to keep the total number of workers constant. It can be shown that the implied change in the relative wage of low human capital workers $w(x^l)/w(x^h)$ is proportional to $-F_{22}(1,h)(x^h - x^l)^2$ (this result is derived in Appendix A.4).[8] Hence, the decrease in the supply of low human capital workers and increase in the supply of high human capital workers will increase the relative wage of low human capital workers if and only if there are decreasing returns to human capital. Moreover, the implied increase in the relative wage is smaller the closer x^l to x^h. This is because the closer the levels of human capital of the two types of workers, the better they substitute for one another.

A   Identifying Human Capital Externalities

The constant-composition approach is based on the theoretical result that the elasticity of average wages holding labor-force skill-composition weights constant with respect to average human capital is equal to the strength of average human capital externalities. To state and proof this result it is useful to note that the average wage w can be written as

---

[8] If there are only two types of labor, the production function in (16) implies that the elasticity of substitution between the two types is inversely proportional to $-F_{22}(1,h)(x^h - x^l)^2$.

$$w \equiv \sum_{x \in X} w(x,h)\, l(x) \equiv w(h, l(x): x \in X)\,, \tag{21}$$

where $l(x) = L(x)/L$. This notation emphasizes that the average wage depends on individual wages of workers with human capital x as well as labor-force skill-composition weights $l(x)$ and that individual wages depend on average human capital in the city h.

Proposition 1: The elasticity of the average wage when labor-force skill-composition weights $l(x)$ are held constant with respect to the average level of human capital yields the strength of average human capital externalities,

$$\frac{\partial \ln w(h, l(x): x \in X)}{\partial h}\, h = q\,. \tag{22}$$

Proof: To proof this result it is useful to write the average wage in the absence of human capital externalities $\sum_x w(x,h)\, l(x)$ as a function of the average level of human capital and the labor-force composition

$$w(h, l(x): x \in X) \equiv \sum_{x \in X} w(x,h)\, l(x). \tag{23}$$

Using this notation, $\ln w(h, l(x): x \in X) = \ln w(h, l(x): x \in X) + \ln A$, which implies that

$$\frac{\partial \ln w(h, l(x): x \in X)}{\partial h}\, h = \frac{\partial \ln A}{\partial h}\, h + \frac{\partial \ln w(h, l(x): x \in X)}{\partial h}\, h$$

$$= q + \frac{\partial \ln w(h, l(x): x \in X)}{\partial h}\, h$$

where we have made use of (19). Hence, (22) follows if the elasticity of the average wage with respect to average human capital holding skill-composition weights constant is equal to zero when there are no human capital externalities. To see that this is the case suppose that the shares of workers with different human capital go from $\{l(x): x \in X\}$ with an average level of human capital $h = \sum_x x l(x)$ to $\{l^*(x): x \in X\}$ with an average level of human capital $h^* = \sum_x x l^*(x)$ (by definition $\sum_x l^*(x) = \sum_x l(x) = 1$). In this case the change in the average wage holding labor-force skill-composition constant at $\{l(x): x \in X\}$ is

$$\Delta \equiv \sum_{x \in X} w(x,h^*)\, l(x) - \sum_{x \in X} w(x,h)\, l(x)\,.$$

To prove that $\partial \ln w(h, l(x): x \in X)/\partial h = 0$ we need to show that $\Delta/(h^*-h) \to 0$ as $h^* \to h$. Adding and subtracting $\sum_x w(x,h^*)\, l^*(x)$ and rearranging terms yields

$$\Delta = \left(\sum_{x \in X} w(x,h^*)\, l^*(x) - \sum_{x \in X} w(x,h)\, l(x) \right) - \left(\sum_{x \in X} w(x,h^*)\, l^*(x) - \sum_{x \in X} w(x,h^*)\, l(x) \right)$$

The first term in brackets is the change in the average wage that is implied by shares of workers with different human capital going from $\{l(x): x \in X\}$ to $\{l^*(x): x \in X\}$. Constant

16

returns to scale and perfect competition imply that this term is equal to the change in average labor productivity $F(1,h^*)-F(1,h)$. Hence,

$$\Delta = \left(F(1,h^*)-F(1,h)\right) - \left(\sum_{x\in X} \mathbf{w}(x,h^*)l^*(x) - \sum_{x\in X} \mathbf{w}(x,h^*)l(x)\right).$$

Making use of $\mathbf{w}(x,h^*) \equiv \mathbf{w}_L(h^*) + \mathbf{w}_H(h^*)x$ and collecting terms, the second term in brackets can be rewritten as the price of human capital $\mathbf{w}_H(h^*)$ multiplied by the change in average human capital $h^*-h = \sum_x x(l^*(x)-l(x))$,

$$\sum_{x\in X} \mathbf{w}(x,h^*)l^*(x) - \sum_{x\in X} \mathbf{w}(x,h^*)l(x) = \mathbf{w}_H(h^*)\sum_{x\in X} x(l^*(x)-l(x)) = \mathbf{w}_H(h^*)(h^*-h) .$$

Hence,

$$\Delta = F(1,h^*)-F(1,h)-\mathbf{w}_H(h^*)(h^*-h)$$

$$= \left(\frac{F(1,h^*)-F(1,h)}{h^*-h} - \mathbf{w}_H(h^*)\right)(h^*-h)^,$$

which implies that $\Delta/(h^*-h) \to F_2(1,h)-\mathbf{w}_H(h)$ as $h^* \to h$. Combined with $\mathbf{w}_H(h) \equiv F_2(1,h)$ in (18) this proves that $\partial \ln \mathbf{w}(h,l(x): x\in X)/\partial h = 0$.   **Q.E.D.**

The proof of Proposition 1 and the intuition behind the result is closely related to the dual approach to TFP accounting. The main difference is that dual TFP accounting identifies the change in TFP associated with the passing of time while the constant-composition approach identifies the change in TFP associated with an increase in the aggregate supply of human capital.[9]

Proposition 1 suggests that the strength of average-schooling externalities between t and T can be estimated in two steps. First obtain the average wage at time T in city c using wages at T but the labor-force composition at t, $w_{cT}^t = \sum_x w_{cT}(x)l_{ct}(x)$. Second regress the log-change in wages holding labor-force skill-composition constant, $\ln w_{cT}^t - \ln w_{ct}$, on the increase in average schooling between t and T (and other variables that may affect wages).

So far we have concentrated on first-order effects of the average level of human capital on average wages holding labor-force skill-composition constant. It is straightforward to show

---

[9] One way to see the relationship between dual TFP accounting and the constant-composition approach to human capital externalities is to derive the result in Proposition 1 in a way that is analogous to the derivation of the main result of dual TFP accounting. Note that $w(x,h) = w_L(h) + w_H(h)x$ yields that the left-hand side of (22) is a weighted average of the effect of human capital on the price of labor $d\ln w_L(h)/d\ln h = \mathbf{e}_L$ and on the price of human capital $d\ln w_L(h)/d\ln h = \mathbf{e}_H$, $(1-\mathbf{b})\mathbf{e}_L + \mathbf{b}\mathbf{e}_H$, where $\mathbf{b} = w_H(h)h/w(h)$ is the share of human capital in wages. Log-differentiating both sides of the equality between output and labor income $y = w_L(h) + w_H(h)h$ yields $\mathbf{q} + \mathbf{b} = (1-\mathbf{b})\mathbf{e}_L + \mathbf{b}(\mathbf{e}_H + 1)$ and hence $(1-\mathbf{b})\mathbf{e}_L + \mathbf{b}\mathbf{e}_H = \mathbf{q}$.

that second-order effects are always positive. The intuition is easiest to explain in the case without human capital externalities (when the marginal social product of human capital is equal to the price of human capital). Suppose that returns to human capital are constant. In this case the marginal social product of human capital does not depend on the average level of human capital used in production. Hence, equality between the price and the marginal social product of human capital implies that the price of human capital is also equal to the intra-marginal social product of human capital. Even a large increase in average human capital will therefore not result in an increase in average wages holding skill-composition constant. When the marginal social product of human capital is strictly decreasing in the average level of human capital however, the intra-marginal social product of human capital exceeds the marginal social product and the price of human capital. Hence, a large increase in average human capital will result in an increase in average wages holding labor-force skill-composition constant. This result is proven in Appendix A.5. Empirically, higher-order effects of human capital on wages can be dealt with just like in the model with two skill types only. Either by including higher-order changes in human-capital supply among the regressors or by using an average of the beginning-of-period and end-of-period skill-composition to calculate the constant-composition log-wage change.

B   Identifying Biased Human Capital Externalities

Our analysis so far has maintained that human capital externalities enter production in a Hicks-neutral way. We now turn to the case where human capital externalities at the aggregate level may be biased towards workers with high levels of human capital or workers with low levels of human capital. To do so we return to the aggregate human capital framework and replace the aggregate production function in (16) by

$$Y = F(A_L L, A_H H), \tag{24}$$

where

$$A_L = h^{q_L} \text{ and } A_H = h^{q_H}; \tag{25}$$

q_L, q_H capture externalities of average human capital at the city level (H and L are defined as in the baseline model, $H \equiv \sum_x x L(x)$ and $L \equiv \sum_x L(x)$). We also assume that the production function is twice continuously differentiable and subject to constant returns to scale given A_L, A_H as well as constant or decreasing returns to human capital given A_L, A_H, $F_{22}(A_L L, A_H H) \leq 0$. The specification in (24) and (25) implies that human capital externalities affect relative wages of workers with different human capital if $q_L \neq q_H$ and the elasticity of substitution between L and H is different from unity.

   To determine the strength of aggregate human capital externalities implied by (24) and (25) suppose that average human capital increases by one percent. The resulting increase in average labor productivity is $q_L(1-b)+(1+q_H)b$ where b is the share of human capital in the average wage, $b = w_H h / w$. Of this total increase, $q_L(1-b)+q_H b$ is due to human capital

18

externalities and will be referred to as the strength of aggregate human capital externalities at the aggregate level. The next proposition proves that the strength of aggregate human capital externalities can be identified with the constant-composition approach.

Proposition 2: Suppose that the aggregate production function is given by (24). Then the elasticity of the average wage holding labor-force skill-composition weights $l(x)$ constant with respect to average human capital is equal to $q_L(1-b)+q_H b$.

Proof: The aggregate production function implies that the equilibrium wage schedule is given by $w(x,h) = F_1(h^{q_L}L, h^{q_H}H)h^{q_L} + F_2(h^{q_L}L, h^{q_H}H)h^{q_H}x$. This equilibrium wage schedule implies

$$\frac{\partial \ln \sum_x w(x,h)\, l(x)}{\partial h}\, h = \left.\frac{\partial \log\left(F_1(h^{q_L}L, h^{q_H}H)h^{q_L} + F_2(h^{q_L}L, h^{q_H}H)h^{q_H}h_0 \right)}{\partial \log h}\right|_{h=h_0}.$$

Constant returns to scale of the aggregate production function given A_L and A_H yields that the marginal product of human capital is homogenous of degree zero. The right-hand-side of the equation can therefore be written as

$$\left.\frac{\partial \log\left(F_1(1, h^{1+q_H-q_L})h^{q_L} + F_2(1, h^{1+q_H-q_L})h^{q_H}h_0 \right)}{\partial \log h}\right|_{h=h_0}$$

$$= (1+q_H-q_L)h^{q_H-q_L}\left. \left(F_{12}(1, h^{1+q_H-q_L})h^{q_L} + F_{22}(1, h^{1+q_H-q_L})h^{q_H}h_0 \right)\frac{h}{w}\right|_{h=h_0}$$

$$+ \left.\left(\frac{q_L w^L}{h} + \frac{q_H w^H h_0}{h} \right)\frac{h}{w}\right|_{h=h_0}.$$

Homogeneity of degree zero of the marginal product of human capital combined with the aggregate production function being twice continuously differentiable implies that $F_{12}(1, h^{1+q_H-q_L})h^{q_L} + F_{22}(1, h^{1+q_H-q_L})h^{q_H}h_0 = F_{21}(1, h^{1+q_H-q_L})h^{q_L} + F_{22}(1, h^{1+q_H-q_L})h^{q_H}h_0 = 0$ for $h = h_0$. Hence,

$$\frac{\partial \ln w(h, l(x): x \in X)}{\partial h}\, h = q_L(1-b)+q_H b \, . \qquad \textbf{Q.E.D.}$$

While the constant-composition approach identifies the aggregate strength of human capital externalities, it cannot identify the parameters q_L and q_H separately.

4 Estimation

We now turn to how the constant-composition and the Mincerian approach can be used to estimate average-schooling externalities.

A   The Constant-Composition Approach

The first step is to eliminate gender, marital status, and race effects from the data on individual wages. This is done by estimating

$$\log w_{ict} = \log \boldsymbol{w}_{ct}(s,e) + \boldsymbol{l}_{pt} X_{ipt} + v_{ict}, \tag{26}$$

where w denotes the hourly wage; ic denotes individual i in city c; t stands for either 1970 or 1990; s,e are individual schooling and experience; X stands for dummies for gender, race, and marital status; and v captures other factors determining wages.[10] The regression is set up so that the city-time specific intercept, $\log \boldsymbol{w}_{ct}(s,e)$, corresponds to the log-wage of married white males. Gender, marital status, and race effects are allowed to differ across macro-regions p (the regions we use are South, East, Midwest, Mountain, and West). As an alternative approach we will estimate $\log \boldsymbol{w}_{ct}(s,e)$ using data on white males only, following Acemoglu and Angrist (2001). The method used to estimate (26) is least squares.

Once we have estimated city-time specific wages of workers with given levels of schooling and potential experience, $\hat{\boldsymbol{w}}_{ct}(s,e)$, we can construct average wages necessary for implementation of the constant-composition approach. The average wage in 1970 is defined as $\hat{w}_{c1970} \equiv \sum \hat{\boldsymbol{w}}_{c1970}(s,e) l_{c1970}(s,e)$, where $l_{c1970}(s,e)$ is the fraction of workers with individual schooling and potential experience s,e in city c in 1970. The average wage in 1990 using the 1970 education-experience labor-force composition is

$$\hat{w}_{c1990}^{1970} \equiv \sum_{s,e} \hat{\boldsymbol{w}}_{c1990}(s,e) l_{c1970}(s,e). \tag{27}$$

The strength of average-schooling externalities in cities between 1970 and 1990, $\boldsymbol{a}^{CC}$, can now be estimated by regressing the log-change of average constant-composition wages, $\Delta \log \hat{w}_{c1970-90}^{1970} \equiv \log \hat{w}_{c1990}^{1970} - \log \hat{w}_{c1970}$, on the change in average schooling, $\Delta S_{c1970-90}$, and other controls

$$\Delta \log \hat{w}_{c1970-90}^{1970} = \text{Controls} + \boldsymbol{a}^{CC} \Delta S_{c1970-90} + u_c. \tag{28}$$

By focusing on the determinants of average wage growth our approach eliminates city-specific fixed effects. The control variables considered are the log-change in city employment, to capture aggregate scale effects (e.g. Sveikauskas (1975), Moomaw (1981), and Henderson (1986), the change in average years of potential experience of workers in the city, and four dummies for the macro-regions described above.

Equation (28) will be estimated using two-stage least squares with the demographic structure and the share of African-Americans in 1970 (as well as various interaction effects) as instruments.

B   The Mincerian Approach

---

[10] Our estimates of the strength of average-schooling externalities do not change when we add controls for industry or occupation to the individual wage regressions.

The first step of the Mincerian approach consists of adding city-time specific fixed effects a_{ct} to an otherwise standard least-squares wage regression

$$\log w_{ict} = a_{ct} + b_t s_{ict} + c_t e_{ict} + d_t e_{ict}^2 + \boldsymbol{m}_{pt} X_{ipt} + v_{ict}, \tag{29}$$

where s, e are individual schooling and experience. X continues to stand for dummies for gender, race, and marital status (gender, marital status, and race effects are again allowed to differ across macro-regions p). The strength of average-schooling externalities in cities between 1970 and 1990, $\boldsymbol{a}^M$, is then obtained by regressing the growth of the estimated city-time specific intercept, $\Delta \hat{a}_{c1970-90} = \hat{a}_{c1990} - \hat{a}_{c1970}$, on the change in average schooling and other controls

$$\Delta \hat{a}_{c1970-90} = \text{Controls} + \boldsymbol{a}^M \Delta S_{c1970-90} + u_c. \tag{30}$$

The methods of estimation, instruments, and control variables used to obtain the Mincerian estimate of average-schooling externalities are identical to those used to obtain the constant-composition estimate.

5 Data and Instruments

Our constant-composition and Mincerian estimates of average-schooling externalities at the city level for the period 1970-1990 are based on data for approximately 2 million individuals in 163 cities in 1970 and 1990. The data comes from the public use micro samples (PUMS) of the US Census (US Bureau of Census (1970, 1990)). Individual wages are measured per hour worked. Experience is measured as potential experience (age minus years of schooling minus six). The control variables X included in the individual wage regressions are marital status, gender, and race (White; Black; Hispanic; Indian or Eskimo; Japanese, Chinese, or Filipino; and Pacific Islander or Hawaiian). To estimate (26) potential experience is partitioned in five intervals and schooling in six intervals, which yields a total of thirty schooling-experience combinations (listed in Appendix A.2).

Our definition of cities corresponds with some exceptions to the US Bureau of Census definition of standard metropolitan statistical areas (SMSAs) in 1990 and is explained in detail in Appendix A.2. City level employment in 1970 and 1990 is obtained by summing employment of all counties that were contained in the city in 1990. County-employment is the number of people with part-time or full-time jobs and comes from the U.S. Department of Commerce (US Department of Commerce (1992)). We only consider employment in the private sector and exclude agriculture and mining.

Average years of schooling (experience) at the city level are obtained by aggregating years of schooling (potential experience) of workers in the city. Average schooling across cities rose by 1.12 years during the 20-year period 1970-1990. The standard deviation of the increase in average schooling was 0.56 and the maximal increase 2.1 years. Average potential experience across cities fell by 5.3 years.

Table 3 contains the results of regressing the 1970-1990 increase in average schooling across cities on the 1970 instruments using the specification that fits the data best. The R^2 of the average schooling regression is 48 percent without macro-region dummies and 57 percent with macro-region dummies.[11] The coefficient estimates of the average-schooling regressions in columns (1) and (2) combined with the sample values of the explanatory variables yield the following three main results (the non-linear specification implies that coefficient estimates must be combined with the sample values of the explanatory variables to assess the effect of changes in the explanatory variables on average schooling). First, cities with a larger share of workers older than 50 in 1970 (AGE50P70) experienced a greater increase in average schooling between 1970 and 1990. This is because workers who retired in this period had levels of education below the labor-force average. Second, cities with a larger number of people younger than 18 per adult in 1970 (YOUNG70) experienced a greater increase in average schooling between 1970 and 1990. This is because young people entering the labor force in this period had levels of education above the labor-force average. The quadratic specification implies that the marginal effect of YOUNG70 on the increase in average schooling was larger in cities with a larger number of people younger than 18 per adult in 1970 (and also that the marginal effect would be negative for small values of YOUNG70; for sample values the effect is always positive however). When we add macro-region dummies in column (2), YOUNG70 and YOUNG70 squared are no longer individually significant but remain jointly significant at the 5-percent level. Third, cities with a larger population share of African-Americans in 1970 experienced a greater increase in average schooling between 1970 and 1990. This is because African-Americans were catching up in schooling levels with the rest of the population over this time period.

Our constant-composition and Mincerian estimates of average-schooling externalities at the state level for the 1970-1990 period are based on data for white males aged 40-49 collected by Acemoglu and Angrist (2001), (the original data sources is US Bureau of Census (1970, 1990)). Following Acemoglu and Angrist we instrument for the change in average schooling at the state level between 1970 and 1990 using data on compulsory-schooling and child-labor laws. The basic information is summarized in eight dummies, CL6-CL9 and CA8-CA11, associated with each individual in our sample. For example the dummy CL7 is equal to one, and all other child-labor law dummies are equal to zero, if the state where the individual is likely to have lived when aged 14 had child-labor laws imposing a minimum of 7 years of schooling. And the dummy CA8 is equal to one, and all other compulsory attendance law dummies are equal to zero, if the state where the individual is likely to have lived when aged 14 had compulsory attendance laws imposing a minimum of 8 years of schooling. The eight dummies are used to calculate the share of individuals for whom each of the CL6-CL9 and

---

[11] The instruments (without macro-region dummies) explain 38 percent of the increase in the share of workers with a high school education or more, 31 percent of the increase in the share of workers with some college or more, 32 percent of the increase in the share of workers with a college education or more, 25 percent of the increase in the share of workers with a high school education only, and 37 percent of the decrease in the share of high school dropouts.

CA8-CA11 dummies is equal to one in each state. Six out of these eight shares (we omit CL6 and CA8 as both sets of variables add up to one) are used as instruments for the relative supply of more educated workers. The data does not include precise information on where individuals lived when aged 14, which is why we follow AA in assuming that at age 14 individuals either all lived in the current state of residence (state-of-residence approach) or in the state where they were born (state-of-birth approach). Both the state-of-residence and the state-of-birth instruments predict more than 40 percent of the change in average schooling 1970-1990 (not in table). As an alternative to the Acemoglu and Angrist instruments we also use the instruments of Table 3 at the state level. These instruments predict just above 50 percent of the change in average schooling 1970-1990 at the state level (not in table).

6   Results

We first discuss the results using the constant-composition approach to average-schooling externalities and then compare the constant-composition results with those of the Mincerian approach.

A   The Constant-Composition Approach

Table 4 contains the results of estimating (28) at the city level using two-stage least squares (2SLS) with the instruments discussed in the previous section.[12] Column (1) uses the constant and four (of the five) macro-region dummies as controls. The estimate of the strength of average-schooling externalities is 0.014 with a standard error of 0.03 and hence highly insignificant. Column (2) eliminates the (individually and jointly) insignificant macro-region dummies SOUTH and WEST. The estimate of the strength of average-schooling externalities is now –0.004 with a standard error of 0.017. Column (3) uses the constant and four macro-region dummies as well as the change in average potential experience 1970-1990 as controls. The estimate of the strength of average-schooling externalities does not change much compared to the specification without average experience in column (1). Changes in average potential experience have a significantly negative effect on average wages holding labor-force skill-composition constant. Hence, cities where the average age of the labor force fell more than average saw an above-average increase of average wages holding labor-force skill-composition constant, which suggests that these cities experienced an inflow of workers with high wages due to unobservable characteristics. The P-value of the test of overidentifying restrictions in the last row (0.53) indicates that these restrictions cannot be rejected at standard significance levels. Column (4) eliminates the (individually and jointly) insignificant macro-region dummies SOUTH and WEST. The estimate of average-schooling externalities is

---

[12] Least-squares (LS) estimation of (28) is likely to yield biased estimates because the increase in average schooling is endogenous and measured with error. Still, in practice LS estimates of the strength of average-schooling externalities are very similar to 2SLS estimates (the difference is at most half a percentage point) and highly insignificant. For example, the LS estimate of the average-schooling coefficient in the specification of column (1) of Table 4 is 0.011 with a standard error of 0.026. This suggests that the different biases present in least-squares estimation tend to offset each other in this particular application.

now –0.01 with a standard error of 0.018. The P-value of the test of overidentifying restrictions in the last row (0.41) indicates that these restrictions cannot be rejected at standard significance levels. Re-estimating the specifications in columns (1) to (4) for full-time workers only yields slightly larger average-schooling externalities, although the difference never exceeds half a percentage point. Columns (5) to (10) estimate equation (28) using selected instruments as additional control variables. The direct effect of the instruments on average wages holding labor-force skill-composition constant is in all cases small and statistically insignificant. For example, when adding the population share of African-Americans in 1970 as a control variable in column (7), we find that a 5 percentage points increase in this share lowers average wages holding labor-force skill-composition constant by only 0.2 percent (the maximum variation in the share of African-Americans across cities in 1970 is 25 percentage points) and that this effect is highly insignificant. Moreover, estimates of the strength of average-schooling externalities in columns (5) to (10) remain close to zero and insignificant.

Table 5 contains the results of estimating (28) using data on white males only to construct constant-composition average wages using (26) and (27). The method of estimation is 2SLS with the usual instruments. The results are very similar to those obtained using all workers once the (individually and jointly) insignificant macro-region dummies SOUTH and WEST are eliminated. For example the strength of average-schooling externalities in column (2) is –0.001 with a standard error of 0.021. The P-value of the test of overidentifying restrictions in the last row (0.73) indicates that these restrictions cannot be rejected at standard significance levels. Estimating (28) using data on white males aged 40-49 only to construct constant-composition average wages in (27) also yields results that are similar to those obtained with all workers (not in the table).

Estimates of the strength of aggregate scale effects in Tables 2 and 3 are very imprecise and larger than the 4 to 10 percent reported in the literature (e.g. Henderson (1986), Ciccone and Hall (1996)). To see whether our results are sensitive to the strength of aggregate scale effects we estimate (28) restricting aggregate scale effects to values between 4 and 10 percent. The results are reported in Table 6. Estimates of the strength of average-schooling externalities are in all cases close to the values obtained earlier.

Table 7 contains constant-composition-approach estimates of average-schooling externalities between 1970 and 1990 at the US *state level*. The method of estimation is 2SLS using either the instruments of Acemoglu and Angrist (2001) or our instruments at the state level. Column (1) contains the result of estimating $\Delta \log \hat{w}^{1970}_{state\,1970-90} = \text{Constant} + a^{CC} \Delta S_{state\,1970-90}$, where state level constant-composition average wages are constructed using data on white males aged 40-49 only, with the Acemoglu and Angrist state-of-residence instruments (like them, we do not include the change in log-employment or other control variables in the estimating equation). Using the state-of-birth instruments instead yields very similar results. Column (2) contains the result of estimating the same equation using 2SLS with our instruments at the state level. It can be seen that estimates of the strength of average-schooling externalities obtained with both sets of instruments are nearly identical. Constant-composition estimates of

average schooling externalities between 1960 and 1980, obtained using the same methods and procedures as described above, are reported in column (3) and (4) and are very similar to those obtained for the 1970-1990 period.

Re-estimating all specifications using the 1990 education-experience labor-force composition to calculate the log-change in the constant-composition wage yields estimates of average-schooling externalities that are within one percentage point of the results using the 1970 composition. Higher-order effects of average schooling on wages do therefore not appear to play an important role in these applications.

B The Mincerian Approach

Table 8 contains estimates of average-schooling externalities at the city level for the period 1970-1990 using the Mincerian approach in (30). The control variables, estimation method, and instruments are the same ones used for the constant-composition approach. Columns (1) and (2) contain the results when the underlying Mincerian wage regression in (29) is estimated using all workers. The external effect on productivity of an additional year of average schooling is around 8 percent and statistically significant at the 5-percent level. Columns (1) and (2) differ in that the latter eliminates the (individually and jointly) insignificant SOUTH and WEST macro-region dummies.[13] Column (3) contains the average-schooling-externality estimate when the Mincerian intercept is estimated using data on white males only. Results are very similar to the case where all workers are used. The P-value of the test of overidentifying restrictions in the last row of columns (1) to (3) indicates that these restrictions cannot be rejected at standard significance levels. Estimating the Mincerian intercept with data on white males aged 40-49 only does not change the results significantly (not in the table).

Estimating average-schooling externalities at the state level for the period 1970-1990 using the Mincerian approach also yields statistically significant externalities whether we use our instruments or the Acemoglu and Angrist (2001) state schooling law instruments. Using our instruments, the estimates external effect on productivity of an additional year of average schooling is 10.2 percent with a standard error of 3.6 percent (statistically significant at the 5-percent level). Using the Acemoglu and Angrist instruments yields average-schooling externalities of 9.8 percent with a standard error of 3.8 percent (also statistically significant at the 5-percent level).

When we apply the Mincerian approach at the state level to the 1960-1980 period, we find average-schooling externalities of around 2 percent whether we use our instruments or the Acemoglu and Angrist (2001) state schooling law instruments. Hence, the Mincerian approach yields different results for the 1970-1990 period than for the 1960-1980 period. Simulation results of Acemoglu and Angrist (2001) show that this difference can be partly explained by the US Census recording individual schooling as a categorical variable starting with the 1990 Census. They show that this change in measurement can explain a spurious average-schooling

---

[13] The LS estimate of average-schooling externalities in the specification of column (1) of Table 8 is 0.092 with a standard error of 0.029. LS estimates of average-schooling externalities in columns (2) and (3) are 0.076 and 0.082 respectively.

externality of 1.7-2.1 percent. This leaves about three quarters of the difference between the Mincerian schooling-externality estimates for the 1960-1980 and 1970-1990 period unexplained.

7   Conclusions

Wages may respond to the aggregate supply of human capital because of a downward sloping demand curve for human capital. The existing (*Mincerian*) approach to human capital externalities confounds such wage changes with externalities. As a result, it yields positive human capital externalities when wages equal marginal social products. Using estimates of the elasticity of substitution between more and less educated workers in the empirical literature, we find that the upward bias of the Mincerian approach is between 60 and 70 percent of the individual return to schooling in a first-order approximation and somewhat larger in simulations.

We propose an alternative approach to the identification of human capital externalities that yields consistent estimates whether or not the demand curve for human capital is downward sloping. The theoretical basis is that, under general conditions, the strength of human capital externalities equals the effect of human capital on the average wage when holding the labor-force skill-composition constant. Another advantage of our (constant-composition) approach compared to the Mincerian approach is that it does not require estimating individual returns to human capital and can therefore be used to estimate average-schooling externalities even when instruments for individual schooling are unavailable.

The approach to identification emerging from this theoretical argument can be used to estimate externalities at the city, region, or country level over any time period in two steps. The first step requires obtaining wages w_x and labor-force shares l_x by skill type x in each city, region, or country at the beginning and the end of the relevant time period to calculate the log-change in the average wage holding skill-composition constant, $\ln\left(\sum \bar{l}_x \tilde{w}_x\right) - \ln\left(\sum \bar{l}_x \bar{w}_x\right)$ where upper bars denote beginning-of-period values and tildes end-of-period values. The second step consists of regressing the log-change in constant-composition average wages on (exogenous) changes in the supply of human capital and other determinants of wages. If the change in the supply of human capital enters positively and significantly, this indicates positive externalities. Higher-order effects of human-capital supply on wages can be dealt with by either including higher-order changes in supply among the regressors or by using an average of the beginning-of-period and end-of-period skill-composition to calculate the constant-composition log-wage change.

As an application of the constant-composition approach we assess the strength of average-schooling externalities in US cities and states between 1970 and 1990 using instrumental-variable estimation methods to account for endogenous average schooling. Our empirical results yield no evidence of statistically significant average-schooling externalities at the city level or the state level. Point estimates of the external return to a one-year increase in average schooling are around zero at the city level and around 2 percent at the state level.

References

Acemoglu, Daron (1996) "A Microfoundation for Social Increasing Returns in Human Capital Accumulation," *Quarterly Journal of Economics* 111 (1), pp. 779-804.

Acemoglu, Daron and Joshua Angrist (1999) "How Large are the Social Returns to Education: Evidence from Compulsory Schooling Laws," NBER Working Paper No. W7444.

Acemoglu, Daron and Joshua Angrist (2001) "How Large are the Social Returns to Education: Evidence from Compulsory Schooling Laws," in Ben Bernanke and Kenneth Rogoff (Editors), *NBER Macroeconomic Annual 2000*, pp 9-59.

Angrist Joshua, "The Economic Returns to Schooling in the West Bank and Gaza Strip," *American Economic Review* 85 (1995), pp. 1065-1087.

Autor, David, Larry Katz and Alan Krueger (1998) "Computing Inequality: Have Computers Changed the Labor Market?" *Quarterly Journal of Economics* 113 (4), pp. 1169-1213.

Azariadis, Costas and Allen Drazen (1990) "Threshold Externalities in Economic Development," *Quarterly Journal of Economics* 105, pp, 501-26.

Benhabib, Jess and Mark Spiegel (1994) "The Role of Human Capital in Economic Development, Evidence from Aggregate Cross-Country Data," *Journal of Monetary Economics* 34 (2), pp. 143-173.

Benabou Roland (1996) "Equity and Efficiency in Human Capital Investment: The Local Connection" *Review of Economic Studies*, Vol. 63, pp. 237-264

Bils, Mark and Peter Klenow (2000) "Does Schooling Cause Growth?" *American Economic Review*, 90, pp.1160-1193.

Black, Duncan and Vernon Henderson (1999) "A Theory of Urban Growth," *Journal of Political Economy* 107 (2), pp. 252-284.

Borjas George (2003) "The Labor Demand Curve is Downward Sloping: Reexamining the Impact of Immigration on the Labor Market" *Quarterly Journal of Economics*, Vol CXVIII (4), pp. 1335-1374.

Card, David, (1999) "The Causal Effect of Education on Earnings," in *Handbook of Labor Economics*, Orley Ashenfelter and David Card (Editors), North Holland, Amsterdam.

Card, David and Thomas Lemieux (2001) "Can Falling Supply Explain the Rising Return to College for Younger Men? A Cohort Based Analysis," *Quarterly Journal of Economics*, 116 (2), pp. 705-746.

Ciccone, Antonio and Robert Hall (1996) "Productivity and the Density of Economic Activity," *American Economic Review* 86 (1), pp. 54-70.

Ciccone, Antonio and Giovanni Peri (2002) "Identifying Human Capital Externalities: Theory with an Application to U.S. Cities," IZA Discussion Paper No. 488.

Ciccone, Antonio and Giovanni Peri (forthcoming) "Long-Run Substitutability Between More and Less Educated Workers: Evidence from U.S. States 1950-1990," *Review of Economics and Statistics*.

Conley, Timothy, Frederick Flier, and Grace Tsiang (2003) "Spillovers from Local Market Human Capital and the Spatial Distribution of Productivity in Malaysia," *Advances in Economic Analysis & Policy*, Vol. 3, No. 1 (2003), Article 5. http://www.bepress.com/bejeap/advances/vol3/iss1/art5

Fallon, P.R. and P.R.G. Layard (1975) "Capital-Skill Complementarity, Income Distribution and Output Accounting" *Journal of Political Economy*, 83(2), pp. 279-302.

Gemmell, Norman (1997) "Externalities to Higher Education," Report to the National Committee of Inquiry into Higher Education, HMSO Copyright Unit, St Clements House, Norwich.

Heckman, James and Peter Klenow (1998) "Human Capital Policy," in Michael Boskin (Editor) *Policies to Promote Capital Formation,* Hoover Institution.

Heckman, James (2000) "Policies to Foster Human Capital," *Research in Economics* 51, No 1 March, pp. 3-56.

Henderson, Vernon (1986) "Efficiency of Resource Usage and City Size," *Journal of Urban Economics* 18, pp. 23-56.

Johnson, George (1997) "Changes in Earnings Inequality: The Role of Demand Shifts," *Journal of Economic Perspectives* 11 (2), pp. 41-54.

Katz, Larry and Kevin Murphy (1992) "Changes in Relative Wages 1963-1987- Supply and Demand Factors," *Quarterly Journal of Economics* 107 (1), pp. 35-78.

Klenow, Peter and Andres Rodriguez-Clare (1997) "The Neoclassical Revival in Growth Economics: Has it Gone too Far?" in Ben Bernanke and Julio Rotemberg (Editors). *NBER Macroeconomic Annual* 1997, pp.73-103, MIT Press, Cambridge (MA).

Krueger, Alan and Mikael Lindahl (2001) "Education for Growth: Why and for Whom?" *Journal of Economic Literature* 39 (4), pp. 1101-1136.

Lucas, Robert E. (1988) "On the Mechanics of Economic Development," *Journal of Monetary Economics* 22, pp. 3-41.

Lucas, Robert E. (1990) "Why Doesn't Capital Flow From Rich To Poor Countries?" *American Economic Review Papers and Proceeding*, 80(3), pp. 92-96.

Mankiw, Gregory, David Romer, and David Weil (1992) "A Contribution to the Empirics of Economic Growth" *Quarterly Journal of Economics*, 107 (2), pp. 407-438.

Moomaw, Ronald (1981) "Productivity and City Size: A Critique of the Evidence," *Quarterly Journal of Economics* 95, pp. 194-223.

Moretti, Enrico (2004a) "Estimating the Social Return to Higher Education: Evidence from Longitudinal and Repeated Cross-Sectional Data," *Journal of Econometrics* 121 (1), pp. 175-212.

Moretti, Enrico (2004b) "Workers' Education, Spillovers and Productivity: Evidence from Plant-Level Production Functions," *American Economic Review*, 94(3), pp.656-690.

Moretti, Enrico (2004c) "Human Capital Externalities in Cities" in Handbook of Urban and Regional Economics, Vol. 4, Geography and Cities, J.V. Henderson and J-F Thisse (eds.)

Rauch, James (1993) "Productivity Gains from Geographic Concentration in Cities," *Journal of Urban Economics* 34, pp. 380-400.

Roback, Jennifer (1982) "Wages, Rents, and the Quality of Life," *Journal of Political Economy* 90, pp. 1257-1278.

Rudd, Jeremy (2000) "Empirical Evidence on Human Capital Spillovers," Federal Reserve Board, Finance and Economics Discussion Paper 2000-46.

Topel, Robert (1997) "Factor Proportions and Relative Wages," *Journal of Economic Perspectives*, 11(2), pp. 55-74.

Topel, Robert (1999) "Labor Markets and Economic Growth," in *Handbook of Labor Economics*, Orley Ashenfelter and David Card (Editors). North Holland, Amsterdam.

U.S. Census (1970), *Census of Population and Housing, United States*, Washington D.C.

U.S. Census (1990), *Census of Population and Housing, United States*, Washington D.C.

US Department of Commerce (1992), Economics and Statistics Administration, Bureau of Economic Analysis, Regional Economic Measurement Division, *REIS CD-ROM*.

Appendix

A.1 Tables and Figures

Table 1: Rejection frequencies of the null of no human capital externalities in a simulated model without human capital externalities

Panel A: Mincerian approach

Confidence Level: 5% 10%	$StdDev(\Delta A)$ =0	$StdDev(\Delta A)$ =0.02	$StdDev(\Delta A)$ =0.04	$StdDev(\Delta A)$ =0.06	$StdDev(\Delta A)$ =0.08
$StdDev(\Delta S)$=0.02	100% 100%	94% 98%	49% 73%	27% 51%	17% 37%
$StdDev(\Delta S)$=0.03	100% 100%	99.7% 99.9%	81% 93%	51% 76%	32% 58%
$StdDev(\Delta S)$=0.04	100% 100%	100% 100%	96% 99%	75% 90%	52% 75%

Notes: The reported values are the percentages of draws in which an econometrician rejects the null of no externalities against an alternative of positive externalities at the 5% or 10% confidence level respectively using the Mincerian approach. Each test is based on the t-statistic for the parameter q_d^M estimated as explained in Section 2. The percentages are based on 5000 random draws of city-specific human capital shocks and skill-biased technology shocks.

Panel B: Constant-composition approach

Confidence Level: 5% 10%	$StdDev(\Delta A)$ =0	$StdDev(\Delta A)$ =0.02	$StdDev(\Delta A)$ =0.04	$StdDev(\Delta A)$ =0.06	$StdDev(\Delta A)$ =0.08
$StdDev(\Delta S)$=0.02	0.6% 3.0%	1.6% 7.5%	2.0% 9.1%	2.1% 8.3%	2.1% 9.1%
$StdDev(\Delta S)$=0.03	0.3% 0.9%	0.8% 4.6%	1.5% 6.4%	1.7% 7.5%	2.0% 7.4%
$StdDev(\Delta S)$=0.04	0.08% 0.2%	0.2% 1.2%	0.6% 2.9%	1.0% 4.3%	1.1% 4.8%

Notes: The reported values are the percentages of draws in which an econometrician rejects the null of no externalities against an alternative of positive externalities at the 5% or 10% confidence level respectively using the constant-composition approach. Each test is based on the t-statistic for the parameter q_d^{CC} estimated as explained in Section 2. The percentages are based on 5000 random draws of city-specific human capital shocks and skill-biased technology shocks.

Table 2: Average estimates of human capital externalities in a simulated model without human capital externalities

Panel A: Mincerian approach

Average $\hat{q}_d^M$ [Share of $\hat{q}_d^M > 0$]	*StdDev*(Δ**A**) =0	*StdDev*(Δ**A**) =0.02	*StdDev*(Δ**A**) =0.04	*StdDev*(Δ**A**) =0.06	*StdDev*(Δ**A**) =0.08
StdDev(Δ**S**)=0.02	0.339 [0.999]	0.340 [0.997]	0.338 [0.950]	0.343 [0.871]	0.336 [0.810]
StdDev(Δ**S**)=0.03	0.340 [0.999]	0.346 [0.999]	0.347 [0.992]	0.349 [0.950]	0.344 [0.900]
StdDev(Δ**S**)=0.04	0.351 [0.999]	0.353 [0.999]	0.352 [0.999]	0.351 [0.998]	0.351 [0.963]

Notes: The reported values in each cell are the average value of $\hat{q}_d^M$ (estimated using (13)) and the fraction of values larger than zero based on 5000 random draws of city-specific human capital shocks and skill-biased technology shocks.

Panel B: Constant-composition approach

Average $\hat{q}_d^{CC}$ [Share of $\hat{q}_d^{CC} > 0$]	*StdDev*(Δ**A**) =0	*StdDev*(Δ**A**) =0.02	*StdDev*(Δ**A**) =0.04	*StdDev*(Δ**A**) =0.06	*StdDev*(Δ**A**) =0.08
StdDev(Δ**S**)=0.02	-0.003 [0.300]	-0.003 [0.445]	-0.003 [0.488]	-0.002 [0.495]	-0.004 [0.493]
StdDev(Δ**S**)=0.03	-0.008 [0.110]	-0.008 [0.334]	-0.008 [0.399]	-0.008 [0.433]	-0.008 [0.448]
StdDev(Δ**S**)=0.04	-0.017 [0.051]	-0.018 [0.166]	-0.018 [0.277]	-0.017 [0.332]	-0.018 [0.364]

Notes: The reported values in each cell are the average value of $\hat{q}_d^{CC}$ (estimated using (15)) and the fraction of values larger than zero based on 5000 random draws of city-specific human capital shocks and skill-biased technology shocks.

Table 3: Quality of the 1970 instruments for the change in average schooling and average experience 1970-1990

	Change in average schooling 1970-1990 (ΔS)	
	(1)	(2)
Share of the city labor force older than 50 in 1970 (50PLUS70)	5.7**	3.6**
	(1.1)	(1.1)
Share of African-Americans in the city-population in 1970 (AA70)	11.6**	6.8**
	(3.1)	(2.9)
People in the city younger than 18 per adult in 1970 (YOUNG70)	-3.7*	-3.0
	(2.2)	(2.2)
YOUNG70*YOUNG70	2.7**	2.8
	(1.2)	(2.1)
YOUNG70*AA70	-5.9**	-2.9
	(2.6)	(2.5)
50PLUS70*AA70	-16.7**	-8.8
	(8.2)	(8.2)
MIDWEST		-0.17**
		(0.06)
MOUNTAIN		-0.29**
		(0.14)
WEST		-0.47**
		(0.07)
SOUTH		-0.14**
		(0.07)
R^2	0.48	0.57

Notes: Results of regressing the increase in average years of schooling and average years of potential experience 1970-1990 at the city level on a constant and the variables in the leftmost column using least squares with robust standard errors. The number of observations is 163. YOUNG70 and YOUNG70 squared are always jointly significant at the 5-percent level. * and ** denote estimates that are significantly different from zero at the 10 and 5-percent level. The quadratic specification for YOUNG70 implies that the marginal effect of YOUNG70 on the increase in average schooling (average experience) would be negative (positive) for small values of YOUNG70; for sample values the effect is always positive (negative) however.

Table 4: Average-schooling externalities at the city level: constant-composition approach 1970-1990

	All workers									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Change in average schooling 1970-1990 (ΔS)	0.014 (0.03)	-0.004 (0.017)	0.005 (0.034)	-0.01 (0.018)	0.003 (0.029)	-0.001 (0.018)	-0.006 (0.032)	-0.005 (0.028)	-0.006 (0.018)	0.001 (0.033)
Change in average experience 1970-1990 (ΔE)			-0.018** (0.008)	-0.017** (0.008)	-0.013 (0.009)	-0.017** (0.008)	-0.014 (0.010)	-0.014 (0.010)	-0.012 (0.017)	-0.015 (0.010)
Log-change in aggregate employment ($\Delta \log L$)	0.16** (0.06)	0.081** (0.027)	0.2** (0.07)	0.11** (0.04)	0.11** (0.05)	0.11** (0.05)	0.12** (0.04)	0.12** (0.05)	0.12** (0.04)	0.11** (0.04)
MOUNTAIN	-0.11** (0.02)	-0.09** (0.01)	-0.11** (0.03)	-0.09** (0.008)	-0.09** (0.008)	-0.09** (0.008)	-0.09** (0.009)	-0.09** (0.008)	-0.09** (0.01)	-0.09** (0.008)
MIDWEST	-0.07** (0.013)	-0.06** (0.01)	-0.06** (0.015)	-0.05** (0.011)	-0.05** (0.01)	-0.05** (0.011)	-0.053** (0.011)	-0.053** (0.011)	-0.051** (0.011)	-0.05** (0.011)
SOUTH	-0.03 (0.02)		-0.036 (0.026)							
WEST	-0.027 (0.03)		-0.032 (0.033)							
YOUNG70*AA70					-0.06 (0.1)					-0.35 (0.4)
YOUNG70						-0.022 (0.036)				
AA70							-0.04 (0.09)			0.28 (0.33)
AGE50P*AA70								-0.004 (0.34)		
AGE50P									0.13 (0.37)	
P-Value overidentifying restrictions			0.53	0.41						

Notes: 2SLS estimation with robust standard errors of (28) at the city level. All regressions contain a constant. Constant-composition average wages constructed with data on all workers. Right-hand-side variables used are those in the leftmost column. Instruments used are: people in the city younger than 18 per adult in 1970 (YOUNG70), the share of the city labor force older than 50 in 1970 (50PLUS70), the share of African-Americans in the city-population in 1970 (AA70), YOUNG70*YOUNG70, YOUNG70*AA70, YOUNG70*50PLUS70, and four macro-region dummies. * and ** denote estimates that are significantly different from zero at the 10 and 5-percent level

Table 5: Average-schooling externalities at the city level: constant-composition approach 1970-1990

	White males				
	(1)	(2)	(3)	(4)	(5)
Change in average schooling 1970-1990 (ΔS)	0.046 (0.04)	-0.001 (0.021)	-0.017 (0.045)	-0.005 (0.041)	-0.001 (0.023)
Change in average experience 1970-1990 (ΔE)	-0.02* (0.011)	-0.02* (0.011)	-0.02* (0.013)	-0.021* (0.013)	-0.02* (0.011)
Log-change in aggregate employment ($\Delta \log L$)	0.16** (0.08)	0.13** (0.05)	0.14** (0.05)	0.14** (0.05)	0.14** (0.05)
MOUNTAIN	-0.11** (0.03)	-0.12** (0.013)	-0.12** (0.012)	-0.12** (0.013)	-0.12** (0.014)
MIDWEST	-0.029 (0.019)	-0.035** (0.014)	-0.03** (0.01)	-0.035** (0.014)	-0.035** (0.014)
SOUTH	-0.018 (0.028)				
WEST	0.026 (0.052)				
YOUNG70*AA70 YOUNG70					-0.005 (0.05)
AA70 AGE50P*AA70			-0.055 (0.14)		
AGE50P				0.017 (0.14)	
P-Value overidentifying restrictions	0.87	0.73			

Notes: 2SLS estimation with robust standard errors of (28) at the city level. All regressions contain a constant. Constant-composition average wages constructed with data on white males only. Right-hand-side variables used are those in the leftmost column. Instruments used are: people in the city younger than 18 per adult in 1970 (YOUNG70), the share of the city labor force older than 50 in 1970 (50PLUS70), the share of African-Americans in the city-population in 1970 (AA70), YOUNG70*YOUNG70, YOUNG70*AA70, YOUNG70*50PLUS70, and four macro-region dummies. The number of observations is 163. * and ** denote estimates that are significantly different from zero at the 10 and 5-percent level.

Table 6: Average-schooling externalities: constant-composition approach 1970-1990 with restricted scale effects

	(1)	(2)	(3)
	$\hat{a}$	$\hat{a}$	$\hat{a}$
$d = 0.04$	-0.008	-0.017	-0.022
	(0.026)	(0.027)	(0.016)
$d = 0.06$	-0.004	-0.014	-0.019
	(0.025)	(0.027)	(0.016)
$d = 0.08$	0.0002	-0.01	-0.015
	(0.025)	(0.028)	(0.016)
$d = 0.1$	0.0004	-0.006	-0.011
	(0.026)	(0.028)	(0.016)

Notes: Method of estimation is 2SLS with robust standard errors. Estimated parameter is the strength of average-schooling externalities at the city level using (28) and restricting the strength of aggregate scale effects d to the values in the leftmost column. Instruments used are: people in the city younger than 18 per adult in 1970 (YOUNG70), the share of the city-labor force older than 50 in 1970 (50PLUS70), the share of African-Americans in the city-population in 1970 (AA70), YOUNG70*YOUNG70, YOUNG70*AA70, YOUNG70*50PLUS70, and four macro-region dummies. The number of observations is 163. * and ** denote estimates that are significantly different from zero at the 10 and 5-percent level. The control variables used are:
- Column (1): a constant and four macro-region dummies. The P-value of the hypothesis that the macro-region dummies SOUTH and WEST can be excluded from the estimating equation is 0.43.
- Column (2): same as in (5) plus the increase in average experience. The P-value of the hypothesis that the macro-region dummies SOUTH and WEST can be excluded from the estimating equation is 0.57.
- Column (3): a constant and two macro-region dummies (MOUNTAIN, MIDWEST) plus increase in average experience. The P-values of the test of overidentifying restrictions (not in the table) indicate that these restrictions cannot be rejected at standard significance levels for the values of d in the table.

Table 7: Average-schooling externalities at the *state level*: constant-composition approach 1970-1990 (and 1960-80)

	1970-1990		1960-1980	
	(1)	(2)	(3)	(4)
Change in average schooling at the US state level 1970-1990	0.029	0.027	0.018	0.016
	(0.028)	(0.026)	(0.024)	(0.023)
P-Value overidentifying restrictions	0.86	0.66	0.75	0.69
Comments	AAIV	OURIV	AAIV	OURIV

Notes: 2SLS estimation at the US state level of $\Delta \log \hat{w}_s^F = \text{Constant} + a\Delta S_s$ for 1970-1990 and 1960-1980, where state level constant-composition average wages are constructed using white males aged 40-49 only. AAIV refers to the instruments used by Acemoglu and Angrist (2001) and OURIV refers to our instruments used at the state level. The number of observations is 49. See the main text for additional explanations.

Table 8: Average-schooling externalities: Mincerian approach 1970-1990

	All workers		White males
	(1)	(2)	(3)
Change in average schooling 1970-1990 (ΔS)	0.085** (0.036)	0.071** (0.017)	0.079** (0.038)
Change in average experience 1970-1990 (ΔE)	-0.017* (0.01)	-0.013* (0.008)	-0.016* (0.01)
Log-change in aggregate employment ($\Delta \log L$)	0.28** (0.08)	0.16** (0.04)	0.22** (0.09)
MOUNTAIN	-0.09** (0.026)	-0.053** (0.013)	-0.16** (0.03)
MIDWEST	0.066** (0.016)	-0.055** (0.012)	-0.054** (0.017)
SOUTH	-0.047 (0.035)		
WEST	-0.047 (0.036)		
P-Value overidentifying restrictions	0.64	0.21	0.23

Notes: 2SLS estimation with robust standard errors of (30) at the city level. All regressions contain a constant. Mincerian intercepts are estimated using data on all workers in columns (1) and (2) and white males only in column (3). Right-hand-side variables used are those in the leftmost column. Instruments used are: people in the city younger than 18 per adult in 1970 (YOUNG70), the share of the city labor force older than 50 in 1970 (50PLUS70), the share of African-Americans in the city-population in 1970 (AA70), YOUNG70*YOUNG70, YOUNG70*AA70, YOUNG70*50PLUS70, and four macro-region dummies. The number of observations is 163. * and ** denote estimates that are significantly different from zero at the 10 and 5-percent level.

A.2   Data and Some Statistics

The data used in the empirical analysis comes from the "Census of Population and Housing" Public-Use Microdata Samples (PUMS) files. For 1970 we have used the 5-percent sample modifying the extraction code kindly provided by David Card. The geographic identifier used for 1970 is the "County Group Code". For 1980 and 1990 we have used the "Card and Chay" extracts of the 5-percent PUMS (available at ftp://elsa.berkeley.edu/pub/census/), which include the standard metropolitan statistical area (SMSA) code as a geographic identifier.

Construction of Cities

The definition of cities that we use corresponds, with some exceptions, to the US Bureau of Census definition of SMSAs in 1990. The PUMS of the 1980 and 1990 US Census have (FIPS) codes identifying the SMSA where individuals live. With this information we can assign individuals in 1980 and 1990 to one of 236 cities. The 1970 US Census does not identify the SMSAs where individuals live, only whether they live in a SMSA or not. Individuals are instead assigned to so-called county groups. County groups can be related to SMSAs by using the so-called county group map (attached to the PUMS in 1970). We match individuals to SMSAs in the following way. When one or more county groups are contained in one SMSA, we assign individuals located in one of the county groups to the SMSA that contains them. When a county group contained more than one SMSA, we merged the different SMSAs into one (13 of our 163 cities are obtained this way) applying the same criterion to SMSAs in 1980 and 1990 (to ensure that cities are defined in the same way in 1970, 1980, and 1990). Finally, when a county group was contained partly in a SMSA and partly in a non-SMSA area, we assigned all individuals located in the county group who lived in a SMSA to the SMSA that contained part of the county group. This procedure resulted in 163 cities for 1970, 1980, and 1990. The code to perform the identification and merge of cities is available from us upon request.

Definition of Individual Wages and Schooling

Hourly wages in a given year have been calculated as yearly salary and wage divided by weeks worked times average hours per week worked in the year. All regressions are run using only individuals with positive hours worked and non-negative potential experience. All variables refer to the year previous to the census. Top-codes differ across years. Individual years of schooling have been obtained in the following way. For 1970 and 1980, we use the variables "Highest Grade Attended" and "Grade" which yields nineteen levels of schooling. For 1990, we use the variable "Yearsch" which yields eleven levels of schooling. When we only observe an interval for years of schooling, we use the midpoint of the interval as years of schooling in the Mincerian wage regressions in (29). The Variable "years of experience" used in (29) is potential experience, i.e. age minus years of schooling minus six. For the constant-composition approach in (26) we partition years of schooling in six intervals [0-9), [9-12), [12-14), [14-16), [16-17), and [17 and more) and years of experience in five intervals [0-10), [10-20), [20-30), [30-40), [40-more).

A.3 The Constant-Composition Approach to Human Capital Externalities at the Country level

The constant-composition approach as developed so far cannot be applied to the identification of human capital externalities at the country level because it would be unreasonable to assume that all countries have access to a perfectly competitive international market for physical capital. This raises the question of how the strength of average human capital externalities can be estimated at the country level. To answer this question suppose that the production function at the country level is

$$Y = AF(L, H, K) \tag{A1}$$

where K is the physical capital stock used in production and the level of TFP is $A = Bh^q$ where B captures exogenous differences in TFP and q the strength of average human capital externalities at the country level; L, H are defined as usual. Assume that the aggregate production function is twice continuously differentiable and subject to constant returns to scale to L, H, K as well as constant or decreasing returns to human capital and to physical capital. Suppose also that labor markets and the market for physical capital at the country level are perfectly competitive and that firms maximize profits taking the level of TFP as given. Denote the rental price of physical capital at the country level by r and define "factor income per worker" by $\sum w(x, h) l(x) dx + rk$, where k is the physical capital intensity. Then the following proposition holds.

Proposition A1: The elasticity of factor income per worker with respect to the average level of human capital yields the strength of average human capital externalities when labor-force skill-composition weights $l(x)$ and the physical capital intensity k are held constant

$$\frac{\partial\left(\sum w(x,h)l(x)dx + rk\right)}{\partial h} \frac{h}{\sum w(x,h)l(x)dx + rk}\Bigg|_{l(x) \forall x \text{ and } k \text{ constant}} = q . \tag{A2}$$

Proof: The argument is very similar to the proof of Proposition 2, which is why we will only sketch the main elements. Competitive factor markets at the country level, profit-maximization, and the aggregate production function imply that factor income per worker can be written as $\sum w(x,h)l(x)dx + rk = AF_1 + AF_2 h + AF_3 k$, where F_i denotes the partial derivative of $F(L, H, K)$ with respect to the i-th argument. Hence, (A2) follows if $\partial(F_1 + F_2 h + F_3 k)/\partial h = F_{12} + F_{22} h + F_{32} k = 0$. To demonstrate this last equality, notice that constant returns to scale to L, H, K and twice continuous differentiability of the production function imply $F_{21}(1,h,k) + F_{22}(1,h,k)h + F_{23}(1,h,k)k = F_{12}(1,h,k) + F_{22}(1,h,k)h + F_{32}(1,h,k)k = 0$. **Q.E.D.**

A.4 Effects of Labor Supply on Relative Wages

Notice that equation (18) can be written as $w(x) = AF(1,h) + AF_2(1,h)(x-h)$ using constant returns to scale given TFP of (16). Hence,

$$\frac{\partial \frac{w(x^l)}{w(x^h)}}{\partial h} = -\frac{F(1,h)F_{22}(1,h)(x^h - x^l)}{\left(F(1,h) + F_2(1,h)(x^h - h)\right)^2} \qquad (A3)$$

or, denoting the amount of labor with human capital x used in production by $L(x)$,

$$\left.\frac{\partial \frac{w(x^l)}{w(x^h)}}{\partial L(x^l)}\right|_{L(x^h)+L(x^l)=constant} = \frac{F(1,h)}{\left(F(1,h) + F_2(1,h)(x^h - h)\right)^2 L} F_{22}(1,h)(x^h - x^l)^2, \quad (A4)$$

using $L \equiv L(x^l) + L(x^h)$ and $h \equiv (L(x^l)x^l + L(x^h)x^h)/L$. The increase in the relative wage of low human capital workers is therefore proportional to $-F_{22}(1,h)(x^h - x^l)^2$ when the supply of low human capital workers decreases and the supply of high human capital workers increases by the same amount.

A.5   Second-Order Effects of the Constant-Composition Approach

Proposition A2: Suppose that the aggregate production function in (16) is three times continuously differentiable. Then the second-order effect of the log of average human capital on the log of average wages when holding labor-force skill-composition weights constant is $\mathbf{h} \equiv -F_{22}(1,h)h^2 / F(1,h)\big|_{h=h_0} \geq 0$, where $h_0 = \sum l(x)x$. The quadratic approximation of the relationship between the log-change of average wages holding labor-force skill-composition weights constant and the log-change of average human capital is therefore $\Delta \log w(h, l(x) : x \geq 0)\big|_{l(x) \text{ constant} \forall x} = \mathbf{q}(\Delta \log h) + \mathbf{h}(\Delta \log h)^2$.

Proof: The first-order effect is $\mathbf{q} + \partial \log(F_1(1,h) + F_2(1,h)h_0)/\partial \log h = \mathbf{q} + (F_{12}(1,h) + F_{22}(1,h)h_0) h/(F_1(1,h) + F_2(1,h)h_0)$ evaluated at $h = h_0$. The second-order effect can be obtained by differentiating the expression above with respect to $\log h$ and evaluating it at $h = h_0$. Differentiation yields $h(F_{12}(1,h) + F_{22}(1,h)h_0)\partial\big(h/(F_1(1,h) + F_2(1,h)h_0)\big)/\partial h + (F_{122}(1,h) + F_{222}(1,h)h_0)\big(h^2/(F_1(1,h) + F_2(1,h)h_0)\big)$. The first term evaluated at $h = h_0$ is zero because constant returns to scale of the production function given TFP implies that $F_2(L,H)$ is homogenous of degree zero, which combined with twice continuous differentiability of the production function yields $F_{12}(1,h) + F_{22}(1,h)h_0 = F_{21}(1,h) + F_{22}(1,h)h_0 = 0$ for $h = h_0$. To simplify the second term notice that constant returns to scale given TFP also imply that $F_1(1,h_0) + F_2(1,h_0)h_0 = F(1,h_0)$ and that $F_{22}(L,H)$ is homogenous of degree minus one. The latter combined with three times continuous differentiability of the production function yields $-F_{22}(1,h_0) = F_{221}(1,h_0) + F_{222}(1,h_0)h_0 = F_{122}(1,h_0) + F_{222}(1,h_0)h_0$. Hence, the second term evaluated at $h = h_0$ becomes $-F_{22}(1,h_0)(h_0)^2 / F(1,h_0)$.   **Q.E.D.**

A.6   Identifying Human Capital Externalities in the Katz and Murphy (1992) Framework

It is straightforward to show that the constant-composition approach to the identification of human capital externalities carries over to the Katz and Murphy (1992) framework. To see this suppose following Katz and Murphy that total output Y_{ct} in city c at time t is produced according to a constant returns to scale production function $Y_{ct} = h_{ct}^q F(L_{ct}, H_{ct})$, where $h = H/L$ and H and L are related to the hours supplied by high school dropouts N_1, high school graduates N_2, college dropouts N_3, and college graduates N_4 through $H \equiv \boldsymbol{b}N_1 + bN_3 + N_4$ and $L \equiv \boldsymbol{a}N_1 + N_2 + aN_3$.[14] Assume also that all markets are competitive and do not internalize the effect of human capital on output captured by $\boldsymbol{q}$ ($\boldsymbol{q}$ measures the strength of human capital externalities).

It is straightforward to show that the supply of workers of different types affects the wage of type-i workers only through the average level of human capital h. Hence, the average wage in a city can be written as

$$w(h, n_i : i = 1,2,3,4) \equiv \sum_i w_i(h) n_i . \tag{A5}$$

where $w_i(h)$ is the wage of type-i workers as a function of average human capital and $n_i = N_i / (\sum N_i)$ their share in the labor force. The definition in (A5) allows us to state the main proposition.

Proposition A3: The elasticity of the average wage with respect to the average level of human capital yields the strength of average human capital externalities when labor-force skill-composition weights n_i are held constant, $\partial \ln w(h, n_i : i = 1,2,3,4) / \partial \ln h = \boldsymbol{q}$.
Proof: Notice that

$$\frac{\partial \ln w(h, n_i : i = 1,2,3,4)}{\partial \ln h} = \sum_i \left(\frac{n_i w_i}{w} \right) \left(\frac{\partial w_i}{\partial h} \frac{h}{w_i} \right) = \sum_i \left(\frac{n_i w_i}{w} \right) \boldsymbol{q} + \left[\sum_i \left(\frac{n_i w_i}{w} \right) \frac{\partial \left(\frac{w_i}{h^q} \right)}{\partial h} \frac{h}{w_i} \right] .$$

Hence, the result follows if the term in square brackets is equal to zero, i.e. if the first-order effect of the average level of human capital on the average wage holding labor-force skill-composition weights n_i are held constant is equal to zero when human capital externalities are absent. In competitive equilibrium, the wage of type-i workers is linked to the production function by $w_1(h) = \boldsymbol{a}h^q F_1(1,h) + \boldsymbol{b}h^q F_2(1,h)$, $w_2(h) = h^q F_1(1,h)$, $w_3(h) = ah^q F_1(1,h) + bh^q F_2(1,h)$, and $w_4(h) = h^q F_2(1,h)$. Making use of the relationship between wages and the production function and continuous differentiability of $F(L,H)$, it can be shown that the term in square

---

[14] Katz and Murphy (pp. 67-69) consider a framework with two labor inputs, high school equivalents and college equivalents. High school graduates are treated as pure high school equivalents and college graduates as pure college equivalents. High school and college dropouts are seen as combinations of high school graduates and college graduates. The weights are determined by regressing wages of high school (college) dropouts on the wages of high school graduates and college graduates. Our assumptions on production are necessary and sufficient to obtain the KM framework.

brackets is equal to $\left(LF_{21}(1,h)+HF_{22}(1,h)\right)/\sum_i N_i$. Constant returns to scale of $F(L,H)$ imply $LF_{21}(1,h)+HF_{22}(1,h)=0$.   **Q.E.D.**

A.7   Human Capital Externalities Through Higher Moments

We now consider the case where the whole distribution of human capital matters for productivity. This amounts to relaxing the (implicit) assumptions regarding the pattern of substitutability among different levels of human capital in the human capital framework and allowing externalities to be driven by the whole distribution of human capital.

Assume that there are $X \equiv \{1,...,Q\}$ types of workers and that output Y is determined by the production function $Y = AF(N(x):x\in X)$,   where A denotes TFP, $N(x)$ denotes employment of workers with human capital x. The only assumptions imposed on the production function are weak concavity, twice continuously differentiability, and constant returns to scale. Suppose that externalities may be driven by the whole distribution of human capital $A = g(n(x):x\in X \setminus q)$, where $n(x) \equiv N(x)/\sum_x N(x)$ denotes the share of workers with human capital x and q is the reference level of human capital. This specification captures the external effect of workers with human capital x through $\partial \ln g(n(x): x\in X \setminus q)/\partial n(x)$, the percentage increase in TFP caused by a one percentage point increase in $n(x)$ accompanied by a one percentage point decrease in $n(q)$, the share of workers with the reference level of human capital q. The average wage can now be written as

$$w = g(n(x):x\in X \setminus q)\left(\sum_{x\in X} \mathbf{w}(n(x):x\in X \setminus q)l(x)\right)\Bigg|_{l(x)=n(x)\ \forall x\in X}, \qquad (A6)$$

where the wage of workers with human capital x in the absence of externalities, $\mathbf{w}(n(x):x\in X \setminus q)$, is equal to

$$\mathbf{w}(n(x):x\in X \setminus q) \equiv F_{n(x)}(n(1),...,n(q-1),1-\sum_{x\in X \setminus q} n(x),n(q+1),...,n(Q)) \qquad (A7)$$

under perfect competition ($F_{n(x)}$ denotes the partial derivative of $F(N(x):x\in X)$ with respect to $n(x)$).

Proposition A4: The elasticity of the average wage with respect to $n(x)$ holding labor-force skill-composition $l(x)$ constant is equal to the external effect of workers with human capital x,

$$\frac{\partial \ln\left(g(n(x):x\in X \setminus q)\left(\sum_{x\in X} \mathbf{w}(n(x):x\in X \setminus q)l(x)\right)\right)}{\partial n(i)}\Bigg|_{n(x)=l(x)\ \forall x\in X} = \frac{\partial \ln g(n(x):x\in X \setminus q)}{\partial n(i)}. \quad (A8)$$

Proof: The left-hand-side of (A8) can be written as

$$\frac{\partial \ln g(n(x):x\in X\ \backslash q)}{\partial n(i)} + \frac{\partial \ln\left(\sum_{x\in X} w(n(x):x\in X\ \backslash q)l(x)\right)}{\partial n(i)}\Bigg|_{n(x)=l(x)\ \forall x\in X} \qquad . \qquad (A9)$$

Making use of (A7) the second term is equal to

$$\sum_{x\in X}\Big(F_{n(x)n(i)}(n(x):x\in X) - F_{n(x)n(q)}(n(x):x\in X)\Big)l(x)$$

$$= \sum_{x\in X} F_{n(i)n(x)}(n(x):x\in X)l(x) - \sum_{x\in X} F_{n(q)n(x)}(n(x):x\in X)l(x)$$

where the equality uses that $F(N(x):x\in X)$ is twice continuously differentiable ($F_{n(x)n(x)}$ denotes the partial derivative of $F_{n(x)}(N(x):x\in X)$ with respect to $n(x)$). Constant returns to scale implies that $\sum_{x\in X} F_{n(j)n(x)}(n(x):x\in X)n(x) = 0$ for all $j\in X$, which yields that the second term in (A9) is zero. **Q.E.D.**