CENTER FOR LABOR ECONOMICS
UNIVERSITY OF CALIFORNIA, BERKELEY
WORKING PAPER NO. 9

# Social Returns to Education and Human Capital Externalities: Evidence from Cities

Enrico Moretti *
Department of Economics
UC Berkeley
enrico@econ.berkeley.edu

November 1998

**Abstract**

Private and social returns to education may differ in the presence of externalities. In this paper, I estimate the external return to education by comparing wages for otherwise similar individuals who work in cities with higher and lower average levels of education. A key issue in this comparison is the presence of unobservable factors that may raise wages and attract more highly educated workers to different cities. I use changes in wages across the 1980 and 1990 Censuses to abstract from any permanent sources of unobserved heterogeneity across cities. To further control for the potential endogeneity of the growth in education across cities, I use an instrumental variable scheme motivated by the observation that younger cohorts have higher education than older cohorts. Specifically, I use the demographic structure of different cities in 1970 as an instrument for changes in education over the 1980s. The results suggest that a one year increase in average education raises the wage of an average worker by about 14%. Consistent with a model that includes both conventional demand and supply factors and externalities, a rise in the portion of better-educated workers has a larger positive effect on less-educated workers. Even for college graduates, however, the external effect is large enough to generate a net positive gain to working in a better-educated city.

*JEL classification*: I20, J31, R1, R50, H52.
*Key words*: Returns to education, human capital externalities, metropolitan areas.

---

*I am indebted to David Card for helpful late-night conversations and insightful advice. I am grateful to Ken Chay, Hilary Hoynes, David Levine, Jeff Perloff, Giovanni Peri and the participants to the UC Berkeley Labor Lunch for useful comments on a previous version of this paper. Michael Greenstone generously provided the computer code to match MSAs across Censuses. I thank Antoine St-Pierre for correcting my English. All errors are naturally my own. Financial support from the UC Berkeley Economics Department Scholarship for Continuing Students is gratefully acknowledged.

# 1 Introduction

Although the magnitude of the return to education is one of the most widely-studied topics in labor economics, the literature has focused almost entirely on *private* rather than *social* returns. This is particularly true of the most recent research, that has explicitly focused on identifying the causal effect of higher education, as compared to any spurious correlation between education and wages due to unobserved ability (Card 1998).

Nevertheless, economists have recognized for at least a century that the social return to education may differ from the private return. Marshall (1890), for example, argued that social interactions among workers in the same industry and location created learning opportunities that enhance productivity. In high technology industries, the link between human capital and innovation is particularly clear. The timing of entry and location of new biotechnology firms are primarily explained by the presence of academic scientists actively contributing to *basic* science research published in academic journals (Zucker, Darby & Brewer 1998).[1] Similarly, the timing and location of patent citations in many industries suggest knowledge spillovers between proximate firms (Jaffe, Trajtenberg & Henderson 1993). Human capital externalities in cities are viewed by some growth theorists as a key determinant of the economic development of nations (Lucas 1988).

Externalities from education may also arise if human and physical capital are complementary factors of production, and firms and workers are imperfectly matched. Complementarity implies that firms invest more in areas where the labor force is more educated. Imperfect matching implies that some uneducated workers in areas with high average education may work with more physical capital than similar workers in areas with low average education (Acemoglu 1996).

Despite the substantial theoretical literature that posits the existence of spillovers, little is known of their empirical relevance outside high-tech industries. Part of the reason is that externalities are not easily quantified. A series of recent empirical studies has provided some indirect evidence by suggesting that neighborhood effects have

[1] Jaffe (1989) and Anselin, Varga & Acs (1997) provide further evidence of spatial spillovers between university research and high technology innovations.

an important role in determining socioeconomic outcomes. For example, Borjas (1995, 1997) argues that the human capital production function depends not only on parental inputs, but also on the average human capital of the relevant ethnic group. Data from the National Longitudinal Survey of Youth confirm the existence of strong human capital externalities both within and across ethnic groups.[2] The most direct evidence of externalities from education is offered by Rauch (1992).[3] He finds that wages are higher in those cities where average education is higher. However, Rauch's assumption that city average education is historically predetermined is problematic, if better-educated workers tend to move to cities with higher wages.

In this paper social returns to education are estimated using a methodology that is robust to unobserved spatial heterogeneity. The hypothesis that the benefits of education are fully reflected in salaries of the educated is tested against one where other agents benefit from spillovers. This is done by comparing wages of otherwise identical individuals living in cities with different education distributions. Using data from the 1980 and 1990 Census, the variation in average education across 282 metropolitan areas is used to measure the external effect of education on wages, after controlling for private returns.

The dependent variable throughout is the regression-adjusted mean wage per education group. This is obtained by regressing individual log wages against standard covariates and city-education-time specific dummies. The city-education-time specific dummies are then regressed on the average education in the city. Under ideal conditions, the coefficient on average education is the social return to education. Although the regression-adjusted city average wages are purged of individual differences in observable characteristics, the results may still suffer from many of the same problems encountered in estimating private returns to education if there are unobserved

[2] Case & Katz (1994) find that neighborhood characteristics appear to affect substantially youth behavior, in a manner suggestive of contagion models of neighborhood effects. See also Corcoran, Gordon, Laren & Solon (1992), Evans, Oates & Schwab (1992), Cutler & L.Glaeser (1995).

[3] Indirect evidence of externalities from education is found in Glaeser, Kallal, Scheinkman & Shleifer (1992). They show that income per capita has grown faster in cities with high human capital in the post-war period. Findings in Glaeser & Mare (1994) are consistent with a model where individuals acquire skills by interacting with one another, and dense urban areas increase the probability of interaction.

differences in ability across workers that happen to be correlated with education.

Like the private return to education, the OLS estimator of social returns is inconsistent in the presence of unobserved heterogeneity and/or measurement error. Contrary the private return case, however, unobserved heterogeneity and measurement error do not necessarily bias OLS in opposite directions. A simple general equilibrium framework that allows for unobserved heterogeneity in labor demand and supply across cities illustrates the reasons. The effect of endogeneity depends on the relative importance of unobserved heterogeneity in labor demand and supply. If variation in average education across cities is driven by unobserved *demand* shocks, OLS is biased up. On the other hand, if variation in average education across cities is driven by unobserved *supply* shocks, OLS is biased down. A city with desirable amenities, for example, may attract skilled workers despite lower wages. Whether the IV estimator is larger or smaller than the OLS one depends on whether supply heterogeneity and measurement error dominate demand heterogeneity.

Information on average education from the Current Population Survey (CPS) is used to assess the extent of measurement error. Measurement error in average education levels of different cities from the CPS is uncorrelated with measurement error in the Census. As might be expected given the large sample size from the Census, the attenuation bias caused by measurement errors in individual education exerts a negligible effect on cross-city comparison.

The use of panel data from 1980 and 1990 Census files permits city-specific fixed effects, which remove all permanent city specific factors that might bias a simpler cross-sectional analysis. However, this differenced regression may still be biased due to transitory factors that affect both wages and mean education. To correct for such biases, an instrumental variable based on exogenous differences in the age structure of cities is used. The US labor force is characterized by a long-run trend of increasing education, since the younger cohorts entering the labor force are better educated than the older ones. This national trend affects cities differently, depending on their age structure. During the 80's, some age groups experienced larger increases in mean education than others. Everything else equal, cities where these age groups were overrepresented experienced larger increases in mean education.

Both the 1980 and the 1970 age structures are used to predict changes in mean education between 1980 and 1990. To serve as valid instrument, age structure must be uncorrelated with wages, except through the endogenous variable in the wage equation. While it is theoretically possible that age structure in 1980 may have affected geographical mobility between 1980 and 1990, labor market changes between 1980 and 1990 are arguably uncorrelated with the 1970 demographic structure, conditional on city and education specific fixed effects. Since geographical mobility across cities is recored in the Census, this assumption may be tested.

The resulting first-differenced IV findings suggest that a 1% increase in the labor force share of college educated increases the wages of high-school drop-outs, high-school graduates and workers with some college by 2.2%, 1.3% and 1.2%, respectively. Surprisingly, it also increases wages of college graduates by 1.1% Consistent with a model that includes both conventional demand and supply factors and externalities, a rise in the portion of better-educated workers has a larger positive effect on less-educated workers. Even for college graduates, however, the external effect is large enough to generate a net positive gain to working in a better-educated city.

The remainder of this paper is organized as follows. The next section provides a general equilibrium framework which illustrates how changes in average education of a city may affect wages there. The first part considers a simple case with 2 identical cities, and a second part introduces unobserved heterogeneity across cities to show how it can bias OLS estimates. Section 3 describes the econometric specification and section 4 outlines the instrumental variables method. Data and results are described in section 5. Section 6 contains concluding remarks.

# 2 General Equilibrium Framework

## 2.1 Homogeneous Cities

In the standard wage equation, productivity depends only on one's own education, and not on other workers' characteristics. However, jobs are rarely performed in isolation: workers within and across firms interact with each other on a daily basis, exchange knowledge, cooperate to solve problems, and learn from each other. If learning from social and professional interactions take place, each worker's education may affect others' productivity and externalities arise. [4]

Alternatively, spillovers from education may take the form of search externalities. If firms and workers are find each other via random matching and breaking the match is costly, equilibrium wages will increase with the average education of the workforce even without learning or technological externalities (Acemoglu 1996). More generally, with imperfect matching, wages are no longer equal to marginal product, and one agent's decision to increase her own education may generate positive spillovers for other agents in the economy. The intuition is simple. With a Cobb-Douglas production function, the privately optimal amount of schooling depends on the amount of physical capital a worker expects to use. The privately optimal amount of physical capital depends on the education of the workforce. If a group of workers in city A increases its level of education, firms in that city, expecting to employ these workers, would invest more. Since search is costly, some of the workers who have not increased their education work with more physical capital and earn more than similar workers in city B. As a consequence, wages increase with average education of the workforce.

A simple example provides more intuition and shows that equilibrium is possible in the presence of externalities. Consider two identical cities and two factors of production, educated and uneducated workers. Each city is a competitive economy that produces an homogeneous good with a Cobb-Douglas technology: $y_c = (\theta_{0c}L_{0c})^{\alpha}(\theta_{1c}L_{1c})^{1-\alpha}$, where $L_{0c}$ and $L_{1c}$ are the number of uneducated and educated workers in city c, respectively, and the $\theta_{jc}$ are productivity shifters. Productivity of education group j in city c is assumed to depend not only on the group's own education but also on the share of educated workers in the city labor force:

$$\theta_{jc} = \theta_j + \gamma \frac{L_{1c}}{L_{0c}} \tag{1}$$

The second term on the right hand side of equation 1 introduces the

[4] The diffusion of knowledge that takes place as a result of social interactions is explicitly modeled in Jovanovic & Rob (1989) and Glaeser (1997). Individuals are assumed to augment their knowledge through pairwise meetings in which they exchange ideas. The higher the level of human capital of the agents, the more rapid is the diffusion and growth of knowledge. Becker, Murphy & Tamura (1990) argue that rates of return on human capital rise rather than decline as the stock of human capital increases. The reason is that education uses educated and other skilled inputs more intensively than sectors that produce consumption goods and physical capital.

externality.

The nominal wage of uneducated workers in city c equals marginal product: $lnw_{0c} = ln\alpha + (\alpha - 1)lnL_{0c} + (1-\alpha)lnL_{1c} + \alpha ln\theta_{0c} + (1-\alpha)ln\theta_{1c}$. Similarly, the nominal wage of educated workers in city c is $lnw_{1c} = ln(1-\alpha) + \alpha lnL_{0c} - \alpha lnL_{1c} + \alpha ln\theta_{0c} + (1-\alpha)ln\theta_{1c}$. If workers are perfectly mobile, real wages, $\frac{w_{jc}}{p_{jc}}$, must be equal across cities. The cost of living for group j in city c, $p_{jc}$, depends on residential rent. Since the consumption good is traded, its price is the same in different cities. For simplicity, educated and uneducated workers are assumed to buy housing services in two separate markets. Housing supply is assumed fixed in both markets. If each worker consumes one housing unit, land prices are proportional to the size of group j in city c: $p_{jc} = \rho L_{jc}$.[5]

Consider what happens when labor force share of educated workers increases in a city. In the absence of any externality ($\gamma = 0$), wages of uneducated workers increase and wages of educated workers decrease. [6] In the presence of external effects ($\gamma > 0$), complementarity and externality both increase the nominal wage of uneducated workers. The impact of an increase in the supply of educated workers on their own wages, however, is determined by two competing forces: the first is the conventional supply effect which makes the economy move along a downward sloping demand curve. The second is the externality that raises productivity. Formally, the change in the educated workers' nominal wage equals a constant times $(\gamma - \alpha)/(1+\eta)$. If the externality is strong enough ($\gamma > \alpha$), the wage of educated workers increases.

Since workers are free to migrate, why are wages not driven to equality? Migration to high wage cities leads to higher housing costs, making real wages equal across cities. Also, migration may entail significant moving costs. Higher nominal wages in a city imply greater productivity in both cases. If workers weren't more productive, firms would leave high wage cities and relocate to low wage cities.

[5]If housing supply is not fixed, results still hold, if we assume that cost of commuting increases with distance.

[6]This is the standard result. Because of complementarity, uneducated workers are now more productive.

## 2.2 Demand and Supply Heterogeneity

A more realistic model allows for heterogeneity across cities. Unobserved heterogeneity may bias OLS estimates if it is correlated with average education in the city. The framework developed in this section identifies potential sources of bias and suggests ways to account for it. Labor demand and supply are assumed to differ across cities because of exogenous factors such as natural resources and amenities. Labor can be divided in J groups according to educational attainment. Workers belonging to the same group are assumed to be perfect substitutes and share the same tastes. Each city is a competitive economy that produces good $y$ using a CES technology. Good $y$ can be consumed locally and its price, $q_{ct}$, may vary across cities.

The production function in city c at time t is as before: $y_{ct} = K_{ct}^{\alpha} L_{ct}^{1-\alpha}$. However, L now has the more general CES specification

$$L_{ct} = [\sum_{j} (\theta_{jct} N_{jct})^{\frac{\sigma-1}{\sigma}}]^{\frac{\sigma}{\sigma-1}} \tag{2}$$

where $N_{jct}$ is the quantity of labor in skill group j in city c at time t and $\theta_{jct}$ is a productivity shifter. Demand for group j labor in city c at time t is

$$lnw_{jct} = \mu_{ct} - \frac{1}{\sigma} lnN_{jct} + \frac{\sigma - 1}{\sigma} ln\theta_{jct} \tag{3}$$

where $\mu_{ct} = \alpha lnK_{ct} + ln(q_{ct} y_{ct}^{1/\sigma})$ is a city-time specific component. Productivity of group j in city c at time t depends on a group specific effect, the externality and an error term:

$$ln\theta_{jct} = \theta_j + \gamma_j S_{ct} + v_{jct} \tag{4}$$

where $S_{ct}$ is the average education in city c at time t. The magnitude of the externality depends on the average education in the city and may have different effects on different education groups. This specification is consistent with both technological and search externalities. The term $v_{jct}$ allows for unobserved geographic heterogeneity in labor demand, using an unrestrictive error component specification:

$$v_{jct} = u_j + u_c + u_t + u_{jc} + u_{jt} + u_{ct} + u_{jct} \tag{5}$$

Geographic heterogeneity in labor demand is due to permanent and transitory differences in technology, natural resources endowment and institutions. For example, if city c is endowed with unique natural resources that makes labor demand for all skill groups permanently higher than in other cities, then $u_c > 0$. If labor demand for skill group j is higher in city c, then $u_{jc} > 0$. Similarly, if returns to unobserved ability are higher in city c and group j is endowed with more ability, $u_{jc} > 0$. Non-neutral technical change, that raises group j's wage across all cities, is captured by $u_{jt}$.

Because of geographic mobility and variable labor force participation, labor supply is assumed to increase with real wages:

$$lnN_{jct} = \eta ln\frac{w_{jct}}{p_{jct}} + v'_{jct} \tag{6}$$

where $p_{jct}$ is the cost of living and the elasticity of labor supply is assumed to be constant across groups. (The assumption of equal elasticities is not critical). The error term $v'_{jct}$ allows for unrestricted supply heterogeneity:

$$v'_{jct} = u'_j + u'_c + u'_t + u'_{jc} + u'_{jt} + u'_{ct} + u'_{jct} \tag{7}$$

Geographic heterogeneity in labor supply comes from different sources. First, it may come from differences in local amenities, such as weather conditions, coastal location, cultural services like museums, cinemas, concerts, etc. (Roback 1982). For example, if weather in city c is particularly favorable, making labor supply of all skill groups permanently higher, then $u'_c > 0$. If only group j individuals are attracted by such weather, then $u'_{cj} > 0$ and $u'_c = 0$.

Second, jobs in the same industry may be very different in different cities. For example, being a teacher may be very different in Los Angeles than Ann Arbor. Parents involvement, security, quality of students are important job characteristics that may be correlated with average education. Antos & Rosen (1975) provide evidence that compensating differentials for teachers may be substantial, particularly in urban areas. The same may be true for social workers, police officers, and many other professions. Finally, heterogeneity in the labor supply of educated workers may also be due to differences in the cost of education.

Two goods are purchased in this economy: the consumption good $y$ and housing. Housing is not produced and its supply is fixed. An upward sloping housing supply can be used without qualitatively changing the results. The cost of living is a weighted average of the price of housing, $r_{ct}$, and the price of the consumption good, $q_{ct}$: $p_{jct} = \lambda_j q_{ct} + (1 - \lambda_j) r_{ct}$. The weight $\lambda_j$ is the share of the budget spent on the consumption good [7]. Since different groups may have different tastes, $\lambda_j$ and $p_{jct}$ are allowed to vary across groups.

The reduced form wage equation can be written as

$$lnw_{jct} = d_j + d_c + d_t + d_{jc} + d_{jt} + \pi_j S_{ct} + \epsilon_{jct} \tag{8}$$

where the d's are dummies that capture any effect that is education, city, time, education-city and education-time specific; and $\pi_j = \frac{(\sigma-1)\gamma_j}{\sigma+\eta}$ is the social return to education. Note that $S_{ct}$ is an equilibrium outcome itself. In fact, $S_{ct}$ is the weighted sum of the education of each group, $S_{jct}$, with weights equal to the share of that group in the city labor force: $S_{ct} = \sum_j \frac{N_{jct}}{N_{ct}} S_{jct}$. Group shares are function of wages and are therefore endogenous.

The residual $\epsilon_{jct}$ is the sum of unobserved heterogeneity in labor supply, labor demand, capital endowment, housing market and output market:

$$\epsilon_{jct} = \frac{1}{\sigma + \eta}(\epsilon^1_{jct} + \epsilon^2_{jct} + \epsilon^3_{jct} + \epsilon^4_{jct} + \epsilon^5_{jct}) \tag{9}$$

where $\epsilon^1_{jct} = \sigma(u_{ct} + u_{jct})$ represents heterogeneity in labor demand; $\epsilon^2_{jct} = -(u'_{ct} + u'_{jct})$, in labor supply; $\epsilon^3_{jct} = \alpha ln K_{ct}$, in capital endowment; $\epsilon^4_{jct} = (\sigma ln q_{ct} + \eta \lambda_j q_{ct} + ln y_{ct})$, in demand for consumption good; and $\epsilon^5_{jct} = \eta(1 - \lambda_j) r_{ct}$ in the housing market.

Equation 8 relates equilibrium wages of group j in city c at time t to the average education in city c at time t. The goal of the empirical work in this paper is to consistently estimate the coefficient $\pi_j$ on average education. The probability limit of the OLS estimate is given by:

$$plim(\hat{\pi}_j) = \pi_j + \frac{1}{\sigma + \eta}\left(\frac{cov(\epsilon^1_{jct}, S_{ct}|\mathbf{d})}{var(S_{ct}|\mathbf{d})} + \frac{cov(\epsilon^2_{jct}, S_{ct}|\mathbf{d})}{var(S_{ct}|\mathbf{d})} + \frac{cov(\epsilon^3_{jct}, S_{ct}|\mathbf{d})}{var(S_{ct}|\mathbf{d})} + \frac{cov(\epsilon^4_{jct}, S_{ct}|\mathbf{d})}{var(S_{ct}|\mathbf{d})} + \frac{cov(\epsilon^5_{jct}, S_{ct}|\mathbf{d})}{var(S_{ct}|\mathbf{d})}\right) \tag{10}$$

[7] This assumption requires Leontieff preferences. Different assumptions on preferences would complicate the model without significantly affecting the results. For simplicity, $\lambda_j$ does not vary by city or time, but this could be relaxed without affecting identification.

where $\mathbf{d} = d_j + d_c + d_t + d_{jc} + d_{jt}$. As in the case of estimating the private return to education, OLS estimates of the social return to education are biased by the presence of various factors that affect workers productivity and supply of education.

The two main sources of endogeneity are unobserved *demand* heterogeneity ($\epsilon^1_{jct}$) and *supply* heterogeneity ($\epsilon^2_{jct}$). Consider first unobserved *demand* heterogeneity. If demand for educated workers increases in city c at time t ($v_{jct} > 0$), then both the wage of educated workers and mean education in the city rise. Formally, this means that $cov[\epsilon^1_{jct}, S_{ct}|\mathbf{d}] > 0$, implying that OLS estimates are biased up. In San Jose, for example, the computer industry boom has increased wages of skilled workers and consequently attracted a well educated labor force.

Secondly, consider unobserved *supply* heterogeneity. If the supply of educated workers increases in city c at time t ($v'_{jct} > 0$), then wages of educated workers decrease and mean education rises. Formally $cov[\epsilon^2_{jct}, S_{ct}|\mathbf{d}] < 0$, making OLS biased down. For example, San Francisco amenities are considered so desirable that educated workers are willing to accept a lower wage to live there.

The existing literature has generally ignored endogeneity issues by assuming that mean education in cities is historically predetermined (Rauch 1992). However, the present model shows that OLS is likely to be biased, although the direction of the bias cannot be determined a priori. If heterogeneity in demand dominates heterogeneity in supply, $\pi_j^{OLS}$ is overestimated. If heterogeneity in supply dominates heterogeneity in demand, $\pi_j^{OLS}$ is underestimated.

The simple supply-demand framework developed here has two advantages. First, and most importantly, it identifies potential sources of bias. Second, it helps identifying, through the relative magnitude of IV and OLS estimates, whether supply or demand heterogeneity drives differences in mean education across cities. Other, more subtle forms of endogeneity arise in the form of general equilibrium effects of changes in the distribution of education. First, if human and physical capital are complements, increases in a city mean education stimulate investment by firms. In this case, $cov[\epsilon^4_{jct}, S_{ct}|\mathbf{d}] > 0$. Second, if some of the output is locally consumed and if tastes differ across education groups, then a change in the group composition of the city alters the demand for city output and hence the demand functions for educated and uneducated labor (Altonji & Card 1991).

Suppose for example that educated workers spend a larger proportion of their income on the consumption good ($\lambda_{educ} > \lambda_{uneduc}$). Even if tastes are similar across groups, income effects might affect output demand. If the share of educated workers increases, demand for $y$ increases. As a consequence, demand for both educated and uneducated labor changes. [8] Therefore $cov[\epsilon^5_{jct}, S_{ct}|\mathbf{d}] > 0$ or $< 0$ depending on whether price and quantity of y increase or decrease as mean education in the city increases.

Finally, a change in the group composition of the city may shift the demand for housing. Although housing is not produced in this model, the price of housing affects labor supply (equation 6). Hence, $cov[\epsilon^4_{jct}, S_{ct}|\mathbf{d}] \neq 0$. Again, this may happen because of differences in tastes or income across groups. [9]

## 3 Econometric Specification

In view of the theoretical discussion in section 2, this section develops a simple estimation method that may be used to identify the externalities from higher education in local labor markets, while recognizing the potential bias that affect a simple cross-sectional analysis. The model is estimated to city-level data using information from the 1980 and 1990 Censuses.

Specifically, I use the Metropolitan Statistical Area (MSA) as a local labor market. MSA's include entire local economic regions and usually contain more than one county. The Census definition of a metropolitan area is a core urban area 'containing a large population nucleus, together with adjacent communities having a high degree of economic and social integration with that core'. A total of 282 MSA's can be identified and matched in 1980 and 1990.[10]

A two-stage econometric specification is adopted. In the first stage, the regression-adjusted mean wage of group j in city c at

[8] As the share of educated workers increases, supply may increase as well. The net effect on labor demand depends on whether changes in demand are larger or smaller than changes in supply.

[9] A limitation of this theoretical framework comes from two assumptions implicit in the CES production function: (1) workers from different education groups have the same degree of complementarity; (2) workers in the same education group are perfect substitutes. In reality, the degree of complementarity varies and workers within the same group are imperfect substitutes.

[10] The Data Appendix describes the procedures used to identify the MSAs in each year and match MSAs across them.

time t, $\hat{\alpha}_{jct}$, is obtained from the following regression:

$$lnw_{ijct} = \alpha_{jct} + X_{ijct}\beta_t + v_{ijct} \quad (11)$$

where $w_{ijct}$ is the wage of individual i in education group j living in city c at time t; $X_{ijct}$ is a vector of individual characteristics including sex, race, Hispanic origin, U.S. citizenship and a quadratic term in potential experience; and $\alpha_{jct}$ is a set of education-city-time specific dummies that represent mean wages after individual differences in observable characteristics have been purged. This equation is estimated separately for 1980 and 1990. In the second stage, $\hat{\alpha}_{jct}$ is regressed on mean education in the city, controlling for city-education fixed effects and city-level covariates:

$$\hat{\alpha}_{jct} = d_j + d_c + d_t + d_{jc} + d_{jt} + \pi_j S_{ct} + \phi Z_{ct} + \epsilon_{jct} \quad (12)$$

where $Z_{ct}$ includes the proportion of blacks, Hispanic, females, US citizens and dummies for New England and Southern States (in certain specification Z includes a richer set of covariates). Equation 12 is analogous to equation 8 and is estimated in first-differences. This specification compares changes in wage of group j in city c over time to changes in mean education in city c, therefore controlling for city-group-specific fixed factors that could bias a cross-sectional analysis.

An alternative one-step estimation strategy could be obtained by substituting equation 12 into 11. The one-step procedure is equivalent to an appropriately weighted two-step procedure (Hanushek 1974).

# 4 Using the Age Distribution to Predict Changes in Average Education

The estimation the coefficient $\pi_j$ in equation 12, which measures the *social* return to education, presents statistical problems similar those encountered in estimating *private* returns to education. Regressing individual wage on own education to obtain estimates of the *private* return yields an estimator that is likely to be biased upward by unobserved heterogeneity and biased downward by measurement error. If unobserved heterogeneity dominates measurement error, IV estimates are lower than OLS ones. Otherwise, IV estimates are larger than OLS ones. Similarly, the OLS estimator of social return in equation 12 is biased.

Measurement error induces attenuation bias, as for private returns. But the heterogeneity bias in equation 12 may be positive or negative. The framework presented in section 2.2 makes the reason clear. If variation in mean education across cities depends on unobserved labor *demand* heterogeneity, there is an upward bias. We observe many educated workers and high wages in San Jose because of the high demand for skilled labor in hi-tech firms. If variation in mean education depends on unobserved labor *supply* heterogeneity, because of differences in amenities, cost of education or job characteristics, there is a downward bias. Therefore, it is not possible to predict a priori the direction of the OLS bias.

An instrumental variable, correlated with changes in mean education across cities and uncorrelated with changes in the residuals, is needed. The demographic structure of a city offers such an instrument. Changes in average education in a city, $\Delta S_c$, can be decomposed in three parts. The first part is a long run trend of increases in education at the national level, due to younger, more educated cohorts entering the labor force, and older, less educated ones leaving it. The second part is the difference between the national trend and the education attainment of the residents. The third part is constituted by changes in the skill composition of the labor force induced by migration. While the second and the third components are likely to depend on local labor market conditions, the first is arguably exogenous.

The secular trend increases the average education of all demographic groups over time. The increase, however, is relatively larger for some groups and provides useful city level variation, since cities have different age structures. For example, the national average in years of education for men aged 55 to 57 increased by 0.75 years from 1980 to 1990. The same variable for men 58 to 60 increased by 0.85. Everything else equal, cities with a larger proportion of 58-60 years old men experienced a larger increase in mean education.

Formally, the instrument for changes in mean years of education in city c that I use in this paper is defined as the weighted sum of changes in mean education by age-gender group. Mean education is calculated at the national level and shares of each age-gender group in city c in 1980 are used as weights:

$$IV_{80} = \sum_m \omega_{mc}[S_{m90} - S_{m80}] \qquad (13)$$

where m indicates age-gender groups (for example: men 58-60); $S_{mt}$ is the mean in years of education for group m at time t at the national level and $\omega_{mc}$ is the share of group m in city c in 1980.[11]

A similar instrument, $IV_{70}$, is obtained from the 1970 age structure instead of the 1980 one. When 1970 data are used, $\omega_{mc}$ is the proportion of people living in city c in 1970 who, in 1980, would eventually belong to age group m (for example, a man who in 1970 is 48 is assigned to the group 'males 58-60'). If there was no mobility, $\omega_{mc}$ estimated with 1970 data would be on average equal to $\omega_{mc}$ estimated using 1980 data (the youngest age group is 16-18). In the presence of mobility, $IV_{70}$ and $IV_{80}$ will differ. Using the 1970 age structure to predict 1980-1990 changes in education has the advantage of independence from potentially endogenous mobility patterns between 1970 and 1980. The disadvantage is that only 115 MSAs are identified in the 1970 Census. Results from both specification are presented.

As will be shown later, both instruments are good predictors of changes in mean education. The instrument $IV_{70}$ ($IV_{80}$) is exogenous if the 1970 (1980) demographic structure is orthogonal to changes in the unobservables between 1980 and 1990. This condition does not require that wages at time t be uncorrelated to the age distribution at the same time. To the extent that younger workers are more mobile, one would expect the proportion of young workers to be low in cities hit by local recessions. Similarly, one would expect the same proportion to be large in cities experiencing economic expansion. Instead, the condition requires that unobserved shocks experienced by a city between 1980 and 1990 not be associated with age distribution in 1970. For example, it is unlikely that the economic boom that Silicon Valley and Seattle are experiencing has much to do with the age distribution in the 1970s.

In terms of equation 12, the instrument is exogenous if

$$\begin{aligned} &cov[iv_2, (\epsilon_{jc90} - \epsilon_{jc80})|\mathbf{d}, X, Z] \\ &= \textstyle\sum_m (S_{m90} - S_{m80}) E[\omega_{mc}(\epsilon_{jc90} - \epsilon_{jc80})|\mathbf{d}, X, Z] = 0 \end{aligned} \qquad (14)$$

[11]Weights $\omega_{mc}$ are estimated using data not only from the labor force, but from the entire population. The age structure of the labor force may be endogenous.

where expectation is taken over j and c. Equation 14 suggests that a sufficient condition for exogeneity is the absence of correlation between the share of age-gender groups in 1970 and changes in unobserved heterogeneity between 1980 and 1990, conditional on observables: $E[\omega_{mc}(\epsilon_{jc90} - \epsilon jc80)|\mathbf{d}, X, Z] = 0$, for each m.

Since conditioning is on city-education specific fixed effects, the residual is purged of all characteristics of the local labor market that do not vary over time. Conditioning on X and Z eliminates any observable individual and city characteristics from the residual.

The potential for correlation of demographic structure with mobility is a concern. This would happen if, for example, cities with a disproportionate share of foreign immigrants were also younger. A tendency of newly-arrived immigrants to move to enclaves established by earlier immigrants (Bartel 1989), implies that the instrument would predict immigrant inflows.

To assess the relation between demographic structure and mobility, table 1 reports the correlation between the share of 3 age groups and the net inflow of domestic and international immigrants. The correlation with percentage change in the population and labor force are also reported.[12] Entries in the first column are +, 0 or - indicating whether the regression of percentage change in population on the share of a demographic group yields a positive, insignificant or negative coefficient. Entries in the remaining 3 columns are obtained similarly.[13]

As expected, the demographic structure in 1970 is in general uncorrelated with changes between 1980 and 1990. The only exception regards the inflow of foreign immigrants. Surprisingly, a younger population in 1970 implies a smaller immigration inflow during the 1980s. The opposite is true for 1980, as the immigrant enclave hypothesis would predict. Net immigration from other US cities and changes in the labor force size are positively correlated with the share of older people in 1980.

[12]Each individual in the 1990 Census was asked to identify the PUMA (Public Use Micro Area) of residence in 1985. By matching the PUMA codes to MSA codes, it is possible to calculate the proportion of the population which moved in, or left, each metropolitan area during the period 1985-90.

[13]The other regressors are the changes in the proportion of blacks, Hispanics, females, US citizens, dummies for New England and for Southern States. Young individuals are aged 16-27; middle-age ones 28-57 and old ones 58-70.

# 5 Data and Results

## 5.1 Average Effect

The empirical analysis of this paper uses individual micro data from the 1970, 1980 and 1990 Census. I use total annual earnings information together with data on weeks worked and hours per week over the year to construct an hourly wage measure and a simple indicator for employment status based on reporting positive earnings and hours. For those employed, I restrict attention to men and women between the ages of 16 and 70, with non negative potential labor market experience. A total of 282 Metropolitan Statistical Areas (MSA) are identified in the Census in 1980 and 1990, while 115 MSAs are identified in 1970. The Data Appendix provides more detailed information on the sample and describes the procedures used to identify the MSAs in each year and match MSAs across them. Two sets of results are presented. In the first one, based on 1980 and 1990 data, the sample consists of 282 cities. In the second one, which uses 1970, 1980 and 1990 data, the sample is made of 115 cities. All cities in the second sample appear in the first one.

Table 2 provides the regression-adjusted mean wage, $\hat{\alpha}_{ct}$, for 15 cities with the highest average education in 1990 as well as for the 15 cities with the lowest such measure. The metropolitan area with both the highest education and adjusted mean wage is Stamford, CT. The one with the lowest education is Mcallen, TX.

Results from a specification with all education groups pooled together are presented first. Equation 12 becomes $\hat{\alpha}_{ct} = d_c + d_t + \pi S_{ct} + \epsilon_{ct}$. This specification provides an estimate of average externality across education groups. This specification provides a parsimonious summary of the external effect. A potential drawback is the imposition of unwarranted structure on the data. Later, the externality is allowed to differ across groups.

Figure 1 plots mean years of education against adjusted mean wage, for 1980 and 1990. The sample includes observations from 282 cities. Figure 2 plots *changes* in mean education against *changes* in adjusted mean wage. OLS regression lines are superimposed. The corresponding coefficient estimates are shown in column 1 of table 3. The top panel of table 3 reports the results for the entire sample. The bottom panel reports results for the subsample of 115 cities for which data for 1970 is available. Because there is wide variation across cities in the number of observations, all the regressions are weighted to account for differences in the precision of the first stage estimates. The weights are the square-root of the number of observations per city.

The cross-sectional OLS coefficients shown in column 1 of table 3 suggest that a one year increase in mean education in a city would have raised wages by 10.9% in 1990 and 5.8% in 1980. The only previous similar study found an estimate of 5.1 % in 1980 (Rauch 1992). The first-differenced OLS estimate in the third row is 7.7%. By first-differencing, the parameter is identified using only within-city variation over time, so that permanent unobserved heterogeneity is controlled for. New York is permanently different than Brownsville in term of size, industry composition, land prices, climate, cultural amenities, etc. However, differencing could exacerbate the bias induced by unobserved transitory heterogeneity since it greatly reduces the variation in mean education.

In column 2, the vector Z of city-level variables is included in the regression.[14] The cross-sectional coefficients decrease slightly, while the first-differenced OLS estimate increases to 12.6 %, suggesting that some of the variables in Z are correlated with transitory factors that affect both wages and average education.

As in the case of private returns to education, OLS estimates of social returns may be biased by unobserved heterogeneity and measurement error. Measurement error may arise because the population mean education in a city is unobserved, and we must rely on an estimate. Using another data source may indicate the extent of measurement error. The 1989 March Current Population Survey (CPS) is used to calculate average education by city. Since measurement error in average education levels of different cities from the CPS is uncorrelated with measurement error in the Census, CPS averages are used as an instrument. This does not solve the simultaneity problem, but does correct for classical measurement error. The estimated coefficient for the 1990 cross-section, shown in column 3 of table 3, is 0.093, slightly larger than the OLS one, but

[14] For now, Z includes the changes in the proportion of blacks, Hispanics, females, US citizens, dummies for New England and for Southern States. A regional dummy for New England is included because the 80's were a period of astonishing wage growth for the region. Among the 20 cities with the largest increase in adjusted mean wage, 13 are in Connecticut or Massachusetts. In section 5.2, more covariates are included to test the robustness of the results.

not statistically different.[15] This suggests that the attenuation bias caused by measurement error exerts a negligible effect on estimates of social returns to education, because the large sample size from the Census provides precise estimates of average education.

Unobserved heterogeneity that varies over time is not controlled for by first-differencing. In the general equilibrium framework of section 2.2, time varying heterogeneity reflects unobserved shocks to labor demand and supply. If the demand for educated workers in a city received a positive shock between 1980 and 1990 ($v_{jct} > 0$), their number and wage would have increased, introducing a positive bias. In this case, OLS estimates are larger than IV estimates (see equation 10). The opposite bias occurs if the number of educated workers increased in a city after a supply shock ($v'_{jct} > 0$). In this case, OLS estimates are smaller than IV estimates.

The instrumental variables, $IV_{70}$ and $IV_{80}$, are based on 1970 and 1980 demographic structure and are discussed in section 4. Table 4 reports results from the first stage regression of changes in average education on the instruments and all the exogenous variables. Both instruments are good predictors of changes in average education, although $IV_{80}$ is more precise, as one might expect. Instrumental variables estimates of social returns to education are in columns 4, 8, 9 and 10 of table 3. The instrumental variables estimate in columns 4 is based on the 1980 age structure. The IV estimate based on 1970 age structure is in column 9 (bottom panel). For comparison purposes, the bottom panel of table 3 includes the estimate obtained using $IV_{80}$ on the subset of 115 MSAs identified in 1970 data.

Instrumental variable estimates in columns 4 and 9 suggest that a one year increase in city mean education raises the adjusted mean wage by 14.8% or 22.6 % respectively, if 1970 or 1980 demographic structure is used. Since the 1970 one is probably more exogenous, as table 1 seems to suggest, more confidence should be put on the lower estimate. The $IV_{80}$ estimate obtained from the smaller sample (column 8) is 0.193, suggesting that the discrepancy between $IV_{80}$ and $IV_{70}$ is due partly to differences in the sample and partly to differences in the instruments. The last column of the table shows the estimate obtained by using both $IV_{70}$ and $IV_{80}$. An over-identification test fails to reject the hypothesis of exogeneity, suggesting that the

[15] The OLS estimate on the subsample of 219 cities for which CPS data are available is 0.087 (0.014).

$IV_{70}$ and $IV_{80}$ estimates are not statistically different (p-value is 0.09). The $IV_{70}$ estimate is almost identical to the corresponding OLS estimate. This suggests that the positive bias introduced by demand heterogeneity is offset by the negative bias of supply heterogeneity. This remains true in most of the specifications presented in this paper.

Visual IV are shown in figure 3 and 4. The top panel shows the correlation between the instrument and wages. The bottom panel shows the correlation between the instrument and education. The IV coefficient is the ratio of the slope in the top panel to the slope in the bottom one.[16]

An estimate of social return to education of 14.8% is large if compared with estimates of private returns, which are usually of the same size, or smaller. There are two reasons why direct comparison of private and social returns is misleading. First, only 2 cities out of 282 experienced a one year increase, or more, in mean education between 1980 and 1990.[17] The median increase was only 0.33 years, i.e. about 3.9 months (the 25th and 75th percentile were 2.0 and 5.5 months, respectively). A 3.9 months increase in average education would imply an overall wage increase by 4.8% over a ten years period. More importantly, the reduced form estimate of social returns to education is conceptually different from private return ones. Increases in average human capital in a city generate general equilibrium effects not observed if only one individual increases her own education. If, for example, human and physical capital are complementary, increases in mean education in a city may stimulate firm investment (see section 2.2). Obviously this does not happen if one individual increases her own education.

## 5.2 Alternative Specifications

Table 5 presents results from alternative specifications designed to probe the robustness of the model. At the top is the base case, the IV estimates from table 3, columns 4 and 9. Rows 2 to 9 show estimates obtained by including a richer set of covariates. One possible explanation for the results is the presence of inter-industry wage differentials. Including 29 industry dummies in the first-stage indi-

[16]This ratio differs from the one reported in table 3 because other regressors are not controlled for in the figures.

[17]They are Lawrence, MA and Huntsville, AL.

vidual level regression takes any industry composition effect out of the adjusted mean wage. Estimates in row 2 suggest that industry composition does not drive the results.

The occupation composition could induce spurious correlation between wages and education if, for example, many small cities were headquarters of large corporations. When four occupation dummies are included in the first-stage regression, results do not change appreciably (row 3).

College towns, where population is younger and more educated, may pose a problem since demand for skill rose in the 1980s. To address this issue, a college town dummy variable is included (row 4). An MSA is defined to be a college town if more then 50% of the age group 18-22 is enrolled in school. According to this definition, there are 32 college towns in the sample. Endogeneity may also arise with retirement communities. If the number and income of retires increase nationally, retirement communities may experience an economic boom. Since location of retirement communities is usually associated with good weather, estimation was repeated including weather characteristics and coastal location (row 5).[18] Estimates in rows 4 and 5 suggest that results of the basic specification are not driven by unobserved characteristics of college towns and retirements communities. This is not surprising, since cities considered here are large metropolitan areas where labor markets rarely depend on a single industry.

Including the percentage change in population yields larger estimates (row 6). This could be due to the endogeneity of population changes.

In row 7, 8 regional dummies corresponding to Census divisions are included to control for unobservables regional shocks that may bias the results, and the estimates drop to 0. This is not surprising, however, because the inclusion of a rich set of regional dummies takes away all the interregional variation needed for identification of the instrument, leaving only within-region variation. The latter represents only a small portion of the total signal and appears to be insufficient for identification.

A similar problem arises in row 8, where I have included among

[18]Weather characteristics include average temperature in January and July, minimum temperature in January, maximum temperature in July, number of cooling days, number of heating days, inches of rain. Coastal location is a dummy equal one if the city is on the ocean or the Great Lakes.

the regressors the 1980 share of individuals aged 61-70. Older and childless voters may be less likely to support public-school spending than younger voters with children. Although the link between population age structure and the level of public education is ambiguous on theoretical grounds, some empirical evidence suggests that per-pupil expenditure are lower in states and districts with a higher proportion of older persons (Cutler, Emendorf & Zeckhauser (1993), Poterba (1998)). With the specification of row 8, the estimated $IV_{70}$ coefficient decreases, and become more imprecise. The likely explanation of the decrease in precision is that the inclusion of the proportion of old individuals reduces the variation necessary to identify the instrument. This becomes evident in row 9, when the proportion of young is also included. $IV_{70}$ ceases to be informative. (Surprisingly, the coefficient on $IV_{80}$ is still precisely estimated).

The reason for the results in row 8 and 9 is that most of the variation that identifies the instrumental variables comes from young and old individual, since the instruments are identified by movements in and out of the labor force, mostly in age groups on the tails of the age distribution. If there was no mobility, people entered the labor force only when they are young and exited only when old, identification of the instrument would come only from the tails.[19] The estimates in row 10 test whether the presence in the instrument of age groups in the middle of the age distribution affects the results. The coefficients in row 10 are obtained by using an instrument that drops age groups in the middle of the distribution, using only the 16-30 and 55-70 ranges. The correlation of the new instrument with the old one is 0.96 and estimated social returns are almost identical.

In this paper individuals are assigned a city on the basis of their residence. MSAs are large enough that for most the city of residence coincides with that of work. To test whether results are sensitive to commuters who live and work in different cities, I repeat the analysis assigning individuals on the basis of the MSA where they work (row 11). Results do not change .

Finally, the residuals are allowed to be spatially autocorrelated. Spatial autocorrelation may arise because cities that are in the same state are subject to similar unobservable shocks. Standard errors increase by about 35%.

[19] In a less extreme case, age groups in the middle of the distribution do provide variation that identifies the instrument.

### 5.3 Heterogeneous Effect

Consider now the richer specification described in equation 12, where social returns to education vary by education group. There are four education groups: less than high-school, high-school, some college and college or more. The sample now includes observations for 282 cities (115 in 1970), 4 education groups and 2 periods. Table 6 presents estimates of the effect of changes in city average education on changes in group regression-adjusted mean wage. Columns 1 and 2 refer to the full sample. Columns 3 and 4 refer to the subset of MSAs identified in 1970. Estimation is performed separately for each education group.

Weighted OLS coefficients vary between 13.2% and 18.5% with the smaller sample, and are somewhat lower for the full sample. Instrumental variable coefficients obtained using the 1970 demographic structure (column 4) suggest that increasing the city mean education by one year increases high-school drop-outs wages by 25.1%, high-school graduates wages by 14.7%, the wage of workers with some college by 13.4% and those of college graduates by 11.9%. For all education groups, except high-school drop-outs, OLS estimates and $IV_{70}$ estimates are similar. To aid in interpreting the magnitude of the coefficients, consider cities like Fargo, ND, or Rochester, NY, which experienced changes in education equal to the median (3.9 months). An increase in average education equal to the median change would imply wages increases for the four education groups by 8.2%, 4.8% , 4.4% and 3.9% over a ten year period, respectively.

Examining the effect of changes in the proportion of college educated workers makes the externality clear. Table 7 shows the effect of an increase in the proportion of college graduates on wages of the four education groups. The instrument here is a weighted sum of changes in the proportion of college graduates in several gender-age cells. As before, the changes in the proportion of college graduates are national averages, and the weights are obtained from the cities' age-gender distribution (section 4). $IV_{70}$ estimates in table 7 suggest that a 1% increase in the share of college educated workers raises high-school drop-outs wages by 2.2%, high-school graduates wages by 1.3%, the wage of workers with some college by 1.2% and those of college graduates by 1.1%. $IV_{70}$ and $IV_{80}$ estimates for high-school drop-outs and college graduate are very similar. To put coefficient estimates in perspective, consider cities like Bakersfield, CA, or Lancaster, PA, which experienced increases in the share of college graduates between 1980 and 1990 close to the median. Such increases raised wages for the four education groups by 4.4%, 2.6%, 2.4% and 2.2%, respectively.

An important feature to note in tables 6 and 7 is that the estimated coefficient decreases with the level of education. This is consistent with the simple model presented in section 2.1 that analyzes the consequences of increases in the proportion of educated workers. Both the externality and complementarity increase the wage of uneducated workers. The impact on the wage of educated workers, however, is determined by two competing forces: the conventional supply effect which makes the wage move along a downward sloping demand curve, and the externality that raises productivity.

Even for college graduates, the external effect seems to be large enough to generate a positive gain to working in a better-educated city. The positive coefficient for college graduates in table 7 implies that the existence of human capital externalities cannot be rejected. Standard demand and supply theory would predict that, without any externality, an increase in the supply of college educate workers would decrease their wage. Here in fact we observe the opposite: an increase in the proportion of college educated workers raises their wage.

Table 8 shows the effect of an increase in the proportion of high-school drop-outs on wages of the four education groups.[20] An increase in the supply of high-school drop-outs hurt uneducated workers in two ways. First, high-school drop-outs wage decreases as it moves along a downward sloping demand curve. Second, as city level human capital decreases, externalities from education decrease. Educated workers suffer only from the latter effect. Results in table 8 support the model, since the estimated coefficient rises monotonically from -3.3% for high-school drop-outs to -1.5% for college graduates.

[20]Similar tables for changes in the proportion of high-school graduates and workers with some college are not reported. It is unclear whether variation in the share of workers with high-school degrees or some college is associated with more or less education in the city. A larger share of high-school graduates may come at the expense of the proportion of high-school drop-outs or workers with some college.

### 5.4 Structural Estimation

The effect of an increase in the proportion of college graduates on their own wage is plotted in figure 5. Initial equilibrium is at point 1. An exogenous increase in the supply of educated workers shifts the supply curve from S1 to S2. Since the city human capital has increased, the marginal product curve shifts from D1 to D2 as a result of the externality.[21]

A test of the reliability of the estimates may be obtained by separately identifying the direct effect of an increase in supply (1 to 2 in figure 5) from the externality (2 to 3). The first one is expected to be negative and the second one to be positive. Let $N_{ct}$ be the number of workers in city c at time t. Adding and subtracting $(1/\sigma)lnN_{ct}$ from equation 3, marginal productivity of group j in city c at time t can be rewritten as

$$lnw_{jct} = \mu'_{ct} - \frac{1}{\sigma}ln\frac{N_{jct}}{N_{ct}} + \frac{\sigma-1}{\sigma}ln\theta_{jct} \quad (15)$$

where $\mu'_{ct} = \mu_{ct} + (1/\sigma)lnN_{ct}$ is a new city-time specific component. The term $ln\frac{N_{jct}}{N_{ct}}$ is the logarithm of the proportion of the labor force of city c in skill group j, denoted $f_{jct}$. Equation 8 becomes

$$lnw_{jct} = d_j + d_c + d_t + d_{jc} + d_{jt} - \frac{1}{\sigma}f_{jct} + \frac{(\sigma-1)\gamma_j}{\sigma}S_{ct} + \epsilon'_{jct} \quad (16)$$

Equation 16 allows separate identification of supply and externality effects. Instrumental variables estimates are shown in table 9. Two instruments are used. The first predicts changes in average education, while the second one predicts changes in the proportion of workers belonging to each group. Separate estimates by education groups are not shown, as there is not enough variation to separately identify different social returns for different groups. Results are consistent with figure 5. Increases in the supply of a given education group, keeping the average education constant, decreases that group's wages, although the coefficient is imprecisely estimated. The $IV_{70}$ estimate of the coefficient on $f_{jct}$ implies a value for the elasticity of substitution, $\sigma$, of about 6.5.[22]

[21] If human and physical capital are complementary, firms may react to the increase in human capital by investing more in physical capital. A larger stock of physical capital would shift D2 further to the right.

[22] Most plausible empirical estimates of the aggregate elasticity of substitution between

## 6 Conclusions

The effect of unobserved factors that affect both wages and education on estimates of social returns to education depends on the relative importance of unobserved heterogeneity in labor demand and supply. If the former dominates, high demand for educated workers increases their wages in cities where mean education is high. If the latter does, wages of educated workers are low in cities where mean education is high, because of the high supply of educated workers.

Although a smaller number of cities is available in 1970, estimation with $IV_{70}$ probably offers the most reliable point estimates. It seems unlikely that age distribution in 1970 affected wages in 1980 and 1990, if city-education specific fixed effects are controlled for. Empirically, the 1970 age structure seems uncorrelated with changes in population, labor force and domestic immigration. The $IV_{70}$ estimates are generally similar to OLS estimates, suggesting that in most cases heterogeneity in labor supply and demand offset one another. The $IV_{80}$ estimates are generally larger than OLS ones, although not statistically different from $IV_{70}$ ones.

An important result of the paper is that a one year increase in average education raises wages by 14.8%. Compare a city like El Paso, TX, a poor border community, with San Jose, CA, which lies in the heart of Silicon Valley; the former with the sixth lowest average education in the US (12 years), the latter with an average education among the highest (13.7 years). Findings in this paper suggest that everything else constant, wages in San Jose are 25% higher because of human capital externalities.

This result may not come as a surprise to El Paso residents. When NAFTA took effect, El Paso was expecting to attract many corporations, with a large young work force and a low cost of living. The city was expected to become a new banking and commerce center for Mexico. But lack of local human capital has prevented this to happen until now. Nathan Christan, the chairman of the Greater El Paso Chamber of Commerce was recently quoted in the New York Times (6/23/98) saying: 'These companies say if you can give us a labor base with adequate skills, we would make you

highly educated and less educated workers in the United States fall between 0.5 and 2.5 (Freeman 1986). Notice, however, that the division in education groups used in this paper is finer than in most studies, where only two categories of workers are considered. With a finer subdivision, the elasticity of substitution increases.

first choice to locate'. As a consequence, 'business leaders this year identified retraining workers as by far the region's top need'.

Since workers are free to migrate, why are wages not driven to equality by workers leaving El Paso and moving to San Jose? First, migration to high-wage cities leads to higher housing costs, making real wages equal across cities. Second, migration may entail significant moving costs. The important thing to note, however, is that higher nominal wages in a city imply greater productivity. If Silicon Valley workers weren't especially productive, firms would leave high wage San Jose and relocate to low-wage El Paso.

This paper's main contribution is the evidence that workers in most jobs, not only in high-tech industries, capture only a part of the benefits of their own education, i.e. that a significant part accrues to others. The focus is on finding a credible methodology of identifying and measuring social returns to education. Identifying the nature and causes of the externality, although important for policy implications, is beyond the scope of this paper. However, proving the existence of externalities from education has in itself important policy implications. First, if social returns to education exceed private returns, then market provision of education is below the efficient level. In this case *efficiency* considerations provide motivation for subsidizing provision of education.

A second implication touches on economic development in depressed areas. A popular recipe among local governments is to attract outside investment. If externalities from education are important, local governments may want to attract human, instead of physical, capital. Anecdotal evidence, like the one relative to El Paso, suggests that some depressed areas are already doing so.

## References

Acemoglu, D. (1996), 'A microfundation for social increasing returns in human capital acumulation', *Quarterly Journal of Economics* pp. 779–804.

Altonji, J. & Card, D. (1991), The effects of immigration on the labor market outcomes of less skilled natives, *in* J. Abowd & R. Freeman, eds, 'Immigration, Trade and Labor', University of Chicago Press.

Anselin, L., Varga, A. & Acs, Z. (1997), 'Local geographic spillovers between university research and high technology innovations', *Journal of Urban Economics* **42**, 422–448.

Antos, J. R. & Rosen, S. (1975), 'Discrimination in the market for public school teachers', *Journal of Econometrics* **3**, 123–150.

Bartel, A. P. (1989), 'Where do the new u.s. immigrants live?', *Journal of Labor Economics* **7**, 371–391.

Becker, G. S., Murphy, K. M. & Tamura, R. (1990), 'Human capital, fertility and economic growth', *Journal of Political Economy* **98**(5), S12–S37.

Borjas, G. J. (1995), 'Ethinicity, neighborhoods and human-capital externalities', *American Economic Review* **85**(3), 365–389.

Borjas, G. J. (1997), 'To ghetto or not to ghetto: Ethnicity and residential segregation', *NBER Working Paper* **6176**.

Card, D. (1998), 'Education in the labor market'.

Case, A. C. & Katz, L. F. (1994), 'The company you keep: the effects of family and neighborhood on disadvanteg youths', *NBER Working Paper* .

Corcoran, M., Gordon, R., Laren, D. & Solon, G. (1992), 'The association between men's economic status and their family and community of origins', *The Journal of Human Resources* **27**(4), 575–601.

Cutler, D. M., Emendorf, D. W. & Zeckhauser, R. J. (1993), 'Demographic characteristics and the public bundle', *Public Finance / Finances Publiques* **48**, 178–198.

Cutler, D. M. & L.Glaeser, E. (1995), 'Are ghettos good or bad?', *NBER Working Paper* **5163**.

Evans, W. N., Oates, W. E. & Schwab, R. M. (1992), 'Measuring peer group effects: a study of teenage behavior', *Journal of political Economy* **100**(5), 966–991.

Freeman, R. B. (1986), 'The demand for education'.

Glaeser, E. L. (1997), 'Learning in cities', *mimeographed* .

Glaeser, E. L., Kallal, H. D., Scheinkman, J. A. & Shleifer, A. (1992), 'Growth in cities', *Journal of Political Economy* **100**(6), 1126–1151.

Glaeser, E. L. & Mare, D. C. (1994), 'Cities and skills', *NBER Working Paper* **4728**.

Greenstone, M. (1998), 'The wage consequences of enclave residence: Evidence from the 1990 and 1980 censuses', *Manuscript, UC Berkeley* .

Hanushek, E. A. (1974), 'Efficient estimators for regressing regression coefficients', *The American Statistician* **28**(2), 66–67.

Jaffe, A. B. (1989), 'Real effects of academic research', *American Economic Review* **79**(5), 957–970.

Jaffe, A., Trajtenberg, M. & Henderson, R. (1993), 'Geographic localization of knowledge spillovers as evidenced by patent citation', *Quaterly Journal of Economics* **108**(3), 577–598.

Jovanovic, B. & Rob, R. (1989), 'The growth and diffusion of knowledge', *Review of Economic Studies* **56**(4), 569–582.

Kominsky & Siegel (1994), 'Measuring educational attainement in the cps', *Population Division, US Bureau of Census* .

Lucas, R. E. (1988), 'On the mechanics of economic development', *Journal of Monetary Economics* **22**, 3–42.

Marshall, A. (1890), *Principles of Economics*, Macmillan.

Poterba, J. M. (1998), 'Demographic change, intergenerational linkages and public education', *America Economic Review* **88**(2), 315–320.

Rauch, J. E. (1992), 'Productivity gains from geographic concentration of human capiatl: Evidence from the cities', *Journal of Urban Economics* **34**, 380–400.

Roback, J. (1982), 'Wages, rents and the quality of life', *Journal of Political Economy* **90**(6), 1257–1278.

Zucker, L. G., Darby, M. R. & Brewer, M. B. (1998), 'Intellectual human capital and the birth of u.s. biothecnology enterprises', *American Economic Review* **88**(1), 290–306.

# 7 Data Appendix

I use individual data from the 1970, 1980 and 1990 Census. For 1980 and 1990, data are from the 5% sample. The labor market information refers to 1979 and 1989. I first randomly select one in two observations, for computational ease. I then assign individuals a metropolitan area on the basis of two geographical identifiers, Public Use Microdata Areas (PUMAs) and metropolitan areas code. The finest geographic unit identified in the 5% samples are PUMAs, which are arbitrary geographic divisions that contains no less than 100,000 people. Most individuals who live in a metropolitan areas are also assigned a metropolitan area identifier (i.e. a MSA or CMSA code). However, some PUMA's straddle the boundary of one or more MSA's and in these 'mixed' PUMA's an MSA code is not assigned. These 'mixed' PUMA's were assigned a MSA code on the basis of the County Group Equivalency files.[23] If over 50% of the PUMA population was attributable to a single MSA, I then assigned all individuals in that PUMA to the majority MSA. The computer code for this assignment is available on request.

Since the MSA definition was changed after the 1980 Census, I redefined 1990 SMSAs to match the 1980 boundaries. The County Group Equivalency files were used to identify PUMAs that contained the affected counties in the 1990 Census. If the counties in question comprised more than half of the PUMA's population, all respondents were assigned to the pertinent MSA. If greater than 10% of a MSA's 1990 population was affected by the boundary changes and was unrecoverable from the County Equivalency files, I dropped the city from the analysis. Dayton and Springfield, Ohio are the only such cities. 282 MSAs are identified in 1980 and 1990.

All the observations from the resulting 1980 sample were used to estimate the age structure. Only individuals in the labor force were used in the estimation of the 1980 and 1990 wage equations. Workers employed in agriculture or in the military were also excluded. Wages rates less than $1.00 per hour, or greater than $400 per hour, were set to missing. Years of education are assigned to education codes used in 1990 Census following table 1 in Kominsky & Siegel (1994). MSA size in 1990 ranges from 935 to 99,371 observations. The

[23] The methodology used to assign MSA codes and to match MSA across Censuses is identical to the one in Greenstone (1998), who generously provided the computer code.

average size is 7288.8. MSA size in 1980 ranges from 667 to 94,343 observations. The average size is 6355.8.

Data from the 1970 Census are used only to estimate age structures. I use the 15% form state sample. The sample universe consist of all individuals residing in one of the 115 MSAs. The finest geographical identifier in the 1970 Census is the county group code. The 1970 county group code was matched to the 1980 MSA code using information in the Census Bureau publication Geographic Identification Code Scheme (1983, 11-17). I followed the same procedure used by Altonji & Card (1991). Of the 282 MSAs identified in 1980 and 1990, only 115 were identified in 1970.

1980

| | Change Populat. | Change Labor Force | Domestic Inflow | Internat. Inflow |
|---|---|---|---|---|
| Share of Young, 1980 | 0 | 0 | - | + |
| Share of Middle-Aged, 1980 | 0 | - | 0 | 0 |
| Share of Old, 1980 | 0 | + | + | - |

1970

| | Change Populat. | Change Labor Force | Domestic Inflow | Internat. Inflow |
|---|---|---|---|---|
| Share of Young, 1970 | 0 | 0 | 0 | - |
| Share of Middle-Aged, 1970 | 0 | 0 | 0 | + |
| Share of Old, 1970 | 0 | 0 | 0 | 0 |

Table 1: Correlation Between 1980 and 1970 Demographic Structure and Changes in Population, Labor Force, Inflows of Domestic Workers, Inflow of Foreign Workers.

NOTES: Entries in the first column are +, 0 or - indicating whether the regression of percentage change in population on the share of a demographic group yields a positive, insignificant or negative coefficient. Entries in the remaining 3 columns are obtained similarly. The other regressors are the changes in the proportion of blacks, Hispanics, females, US citizens, dummies for New England and for Southern States. Young individuals are aged 16-27; middle-age ones 28-57 and old ones 58-70. Models are weighted by the square-root of city size.

| rank | | mean education | adjusted mean wage, $\alpha_{ct}$ |
|---|---|---|---|
| 1 | STAMFORD, CT | 14.5 | 1.10 |
| 2 | ANN ARBOR, MI | 14.4 | 0.77 |
| 3 | NORWALK, CT | 14.2 | 1.08 |
| 4 | GAINESVILLE, FL | 14.1 | 0.56 |
| 5 | TALLAHASSEE, FL | 14.0 | 0.61 |
| 6 | WASHINGTON, DC-MD-VA | 14.0 | 0.91 |
| 7 | BOSTON, MA | 13.9 | 0.89 |
| 8 | CHAMPAIGN-URBANA-RANTOUL, IL | 13.9 | 0.58 |
| 9 | RALEIGH-DURHAM, NC | 13.9 | 0.71 |
| 10 | FORT COLLINS-LOVELAND, CO | 13.8 | 0.55 |
| 11 | MADISON, WI | 13.8 | 0.65 |
| 12 | SAN FRANCISCO, CA | 13.8 | 0.90 |
| 13 | COLUMBIA, MO | 13.7 | 0.52 |
| 14 | DENVER, CO | 13.7 | 0.68 |
| 15 | LINCOLN, NE | 13.7 | 0.52 |
| 267 | STEUBENVILLE-WEIRTON, OH-WV | 12.3 | 0.55 |
| 268 | VISALIA-TULARE-PORTERVILLE, CA | 12.3 | 0.66 |
| 269 | FORT SMITH, AR-OK | 12.2 | 0.48 |
| 271 | YAKIMA, WA | 12.2 | 0.64 |
| 272 | SAN ANGELO, TX | 12.1 | 0.49 |
| 273 | VINELAND-MILLVILLE-BRIDGETON, NJ | 12.1 | 0.77 |
| 274 | DANVILLE, VA | 12.0 | 0.59 |
| 275 | EL PASO, TX | 12.0 | 0.57 |
| 276 | HICKORY, NC | 12.0 | 0.59 |
| 277 | ODESSA, TX | 12.0 | 0.59 |
| 278 | NEW BEDFORD, MA | 11.9 | 0.80 |
| 279 | BROWNSVILLE-HARLINGEN, TX | 11.2 | 0.54 |
| 280 | FALL RIVER, MA-RI | 11.2 | 0.84 |
| 281 | LAREDO, TX | 11.1 | 0.57 |
| 282 | MCALLEN-EDINBURG-MISSION,TX | 10.9 | 0.58 |

Table 2: Cities with the Highest and Lowest Mean Education in 1990.

| | OLS | OLS | $IV_{CPS}$ | $IV_{80}$ |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| 1990 | 0.109 | 0.083 | 0.093 | |
| | (0.015) | (0.012) | (0.021) | |
| 1980 | 0.058 | 0.033 | | |
| | (0.009) | (0.008) | | |
| Changes 90-80 | 0.077 | 0.126 | | 0.226 |
| | (0.023) | (0.020) | | (0.046) |
| Other covariates | no | yes | yes | yes |
| N | 288 | 288 | 219 | 288 |

| | OLS | OLS | $IV_{CPS}$ | $IV_{80}$ | $IV_{70}$ | $IV_{70-80}$ |
|---|---|---|---|---|---|---|
| | (5) | (6) | (7) | (8) | (9) | (10) |
| 1990 | 0.116 | 0.106 | 0.098 | | | |
| | (0.026) | (0.022) | (0.032) | | | |
| 1980 | 0.059 | 0.045 | | | | |
| | (0.016) | (0.015) | | | | |
| Changes 90-80 | 0.028 | 0.151 | | 0.193 | 0.148 | 0.197 |
| | (0.038) | (0.030) | | (0.061) | (0.074) | (0.061) |
| Other covariates | no | yes | yes | yes | yes | yes |
| N | 115 | 115 | 115 | 115 | 115 | 115 |

Table 3: The Effect of Changes in Mean Education on Mean Wages

NOTES: The top panel reports results for the entire sample. The bottom panel report results for the subset of MSAs identified in 1970. Each entry is a separate regression. The dependent variable is the city regression-adjusted mean wage. Entries are the coefficients on mean years of education. $IV_{80}$ is based on 1980 demographic structure. $IV_{70}$ is based on 1970 demographic structure. $IV_{70-80}$ refers to estimates obtained using both $IV_{80}$ and $IV_{70}$. $IV_{CPS}$ is the average education from 1989 March CPS. The other regressors are the changes in the proportion of blacks, Hispanics, females, US citizens, dummies for New England and for Southern States. Models are weighted by the square-root of city size. The OLS estimate for the 1990 cross-section on the subset of 219 cities for which CPS data are available is 0.087 (0.014).

| | $IV_{80}$ (1) | $IV_{70}$ (2) |
|---|---|---|
| | 2.43 | 2.42 |
| | (0.29) | (0.50) |
| $R^2$ | 0.47 | 0.52 |
| N | 282 | 115 |

Table 4: First-Stage: Correlation Between Changes in Mean Years of Education and the Instrumental Variables.

NOTES: The dependent variable is the change in mean years of education. The first entry in column 1 is the coefficient on $IV_{80}$ and the first entry in column 2 is the coefficient on $IV_{70}$. Also included in the regression are proportion of blacks, Hispanics, females and US citizens, dummies for New England and Southern States. Models are weighted by the square-root of city size.

| | $IV_{80}$ (1) | $IV_{70}$ (2) |
|---|---|---|
| (1) Basic specification | 0.226 | 0.148 |
| | (0.046) | (0.074) |
| (2) Industry dummies | 0.228 | 0.157 |
| | (0.044) | (0.069) |
| (3) Occupation dummies | 0.242 | 0.161 |
| | (0.045) | (0.072) |
| (4) College town dummy | 0.239 | 0.125 |
| | (0.056) | (0.078) |
| (5) Weather and coastal location | 0.320 | 0.184 |
| | (0.068) | (0.139) |
| (6) Population change | 0.407 | 0.525 |
| | (0.168) | (0.262) |
| (7) 8 regional dummies | -0.001 | -0.105 |
| | (0.063) | (0.136) |
| (8) Proportion of old | 0.192 | 0.074 |
| | (0.046) | (0.115) |
| (9) Proportion of old and young | 0.120 | -0.732 |
| | (0.060) | (1.067) |
| (10) IV includes only tails of age distr. | 0.237 | 0.142 |
| | (0.058) | (0.075) |
| (11) MSA of Work | 0.207 | 0.145 |
| | (0.051) | (0.074) |
| (12) Spatial autocorrelation | 0.226 | 0.148 |
| | (0.076) | (0.107) |

Table 5: Robustness checks

NOTES: (1) The base case is taken from column 4 and 9 of table 3;
(2) 29 industry dummies are included in the first-stage;
(3) 4 occupation dummies are included in the first-stage;
(4) a college town dummy is included in the second stage;
(5) 8 weather variables and coastal location are included in the second stage;
(6) percentage population change is included in the second stage;
(7) regional dummies correspond to Census divisions;
(8) proportion of 61-70 years old is included in the second stage;
(9) proportion of 16-27 and 61-70 years old is included in the second stage;
(10) age groups 28-30 to 58-60 are not included in the IV;
(11) observations are assigned to MSA of work, not residence;
(12) residuals for cities in the same state are correlated.

| | OLS | $IV_{80}$ | OLS | $IV_{70}$ |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| *less than high-school* | 0.134 | 0.266 | 0.185 | 0.251 |
| | (0.023) | (0.055) | (0.034) | (0.083) |
| *high-school* | 0.120 | 0.253 | 0.142 | 0.147 |
| | (0.023) | (0.054) | (0.035) | (0.085) |
| *some college* | 0.141 | 0.283 | 0.146 | 0.134 |
| | (0.022) | (0.053) | (0.034) | (0.083) |
| *college* | 0.111 | 0.146 | 0.132 | 0.119 |
| | (0.017) | (0.040) | (0.025) | (0.060) |
| N | 282 | 282 | 115 | 115 |

Table 6: The Effect of Changes in Mean Education on Education Group Wage

NOTES: Columns 1 and 2 refer to the entire sample. Columns 3 and 4 refer to the subset of MSAs identified in 1970. Each entry is a separate regression. Equation 12 is estimated. The dependent variable is the change in an education group adjusted mean wage. Entries are the coefficients on changes in mean years of education. The other regressors are the changes in the proportion of blacks, Hispanics, females, US citizens, dummies for New England and for Southern States. Models are weighted by the square-root of city size

| | $IV_{80}$ (1) | $IV_{70}$ (2) |
|---|---|---|
| *less than high-school* | 3.144 | 2.267 |
| | (0.710) | (0.857) |
| *high-school* | 3.049 | 1.314 |
| | (0.668) | (0.814) |
| *some college* | 3.395 | 1.275 |
| | (0.687) | (0.803) |
| *college* | 1.722 | 1.138 |
| | (0.466) | (0.601) |
| N | 282 | 115 |

Table 7: The Effect of Changes in the Proportion of College Educated on Education Group Wage

NOTES: Columns 1 refers to the entire sample. Columns 2 refers to the subset of MSAs identified in 1970. Each entry is a separate regression. The dependent variable is the change in an education group adjusted mean wage. Entries are the coefficients on the share of college educated. The other regressors are the changes in the proportion of blacks, Hispanics, females, US citizens, dummies for New England and for Southern States. Models are weighted by the square-root of city size

| | $IV_{80}$ | $IV_{70}$ |
|---|---|---|
| | (1) | (2) |
| *less than high-school* | -3.337 | -3.361 |
| | (1.105) | (1.152) |
| *high-school* | -3.503 | -2.311 |
| | (1.101) | (1.052) |
| *some college* | -4.048 | -2.190 |
| | (1.151) | (1.008) |
| *college* | -1.855 | -1.570 |
| | (0.706) | (0.704) |
| N | 282 | 115 |

Table 8: The Effect of Changes in the Proportion of High-School Drop-Outs on Education Group Wage.

NOTES: Columns 1 refers to the entire sample. Columns 2 refers to the subset of MSAs identified in 1970. Each entry is a separate regression. The dependent variable is the change in an education group adjusted mean wage. Entries are the coefficients on the change in the share of high-school drop-outs. The other regressors are the changes in the proportion of blacks, Hispanics, females, US citizens, dummies for New England and for Southern States. Models are weighted by the square-root of city size

| | $IV_{80}$ (1) | $IV_{70}$ (2) |
|---|---|---|
| Share of education group, $f_{jct}$ | -0.096 | -0.153 |
| | (0.082) | (0.123) |
| Average education, $S_{ct}$ | 0.251 | 0.180 |
| | (0.056) | (0.071) |
| N | 1128 | 460 |

Table 9: Structural Estimates of the Effect of Changes in City Mean Education and Share of Education Group on an Education Group Wage

NOTES: Columns 1 refers to the entire sample. Columns 2 refers to the subset of MSAs identified in 1970. Equation 16 is estimated. The dependent variable is the change in adjusted mean wage. The other regressors are the changes in the proportion of blacks, Hispanics, females, US citizens, dummies for New England and for Southern States. Standard error account for correlation of residuals at the city level. Models are weighted by the square-root of city size.

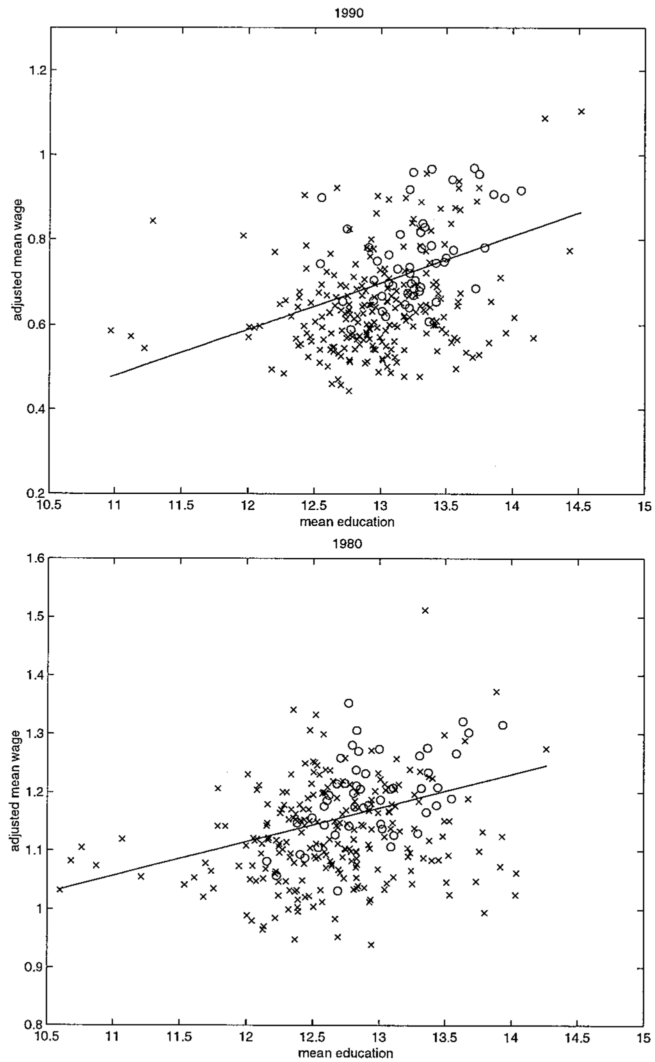


Figure 1: Correlation Between Adjusted Mean Wage and Mean Education in 282 Cities, in 1990 and 1980. Circles Are 50 Largest Cities.

NOTES: Weighted OLS fit superimposed. Outliers (mean education, mean wage):
1990
MCALLEN-EDINBURG-MISSION,TX (10.96 0.585); LAREDO, TX (11.11 0.572) ; BROWNSVILLE-HARLINGEN, TX (11.22 0.543); FALL RIVER, MA-RI (11.28 0.843);
NORWALK, CT (14.24 1.08) ; ANN ARBOR, MI (14.43 0.77) ; STAMFORD, CT (14.51 1.10)
1980
BROWNSVILLE-HARLINGEN, TX (10.59 1.03); MCALLEN-EDINBURG-MISSION,TX (10.68 1.08 ): FALL RIVER, MA-RI (10.76 1.10);
ANCHORAGE, AK (13.34 1.51); ANN ARBOR, MI (14.26 1.27)

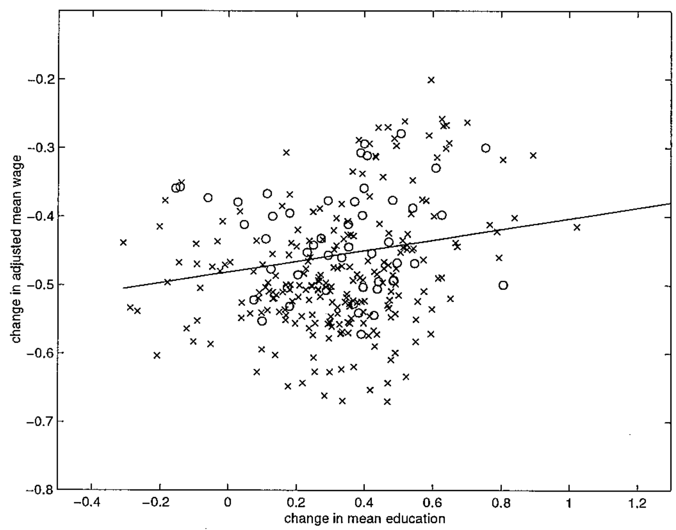


Figure 2: Correlation Between Changes in Adjusted Mean Wage and Mean Education in 282 Cities. Circles are 50 Largest Cities.

NOTES: Weighted OLS fit superimposed.

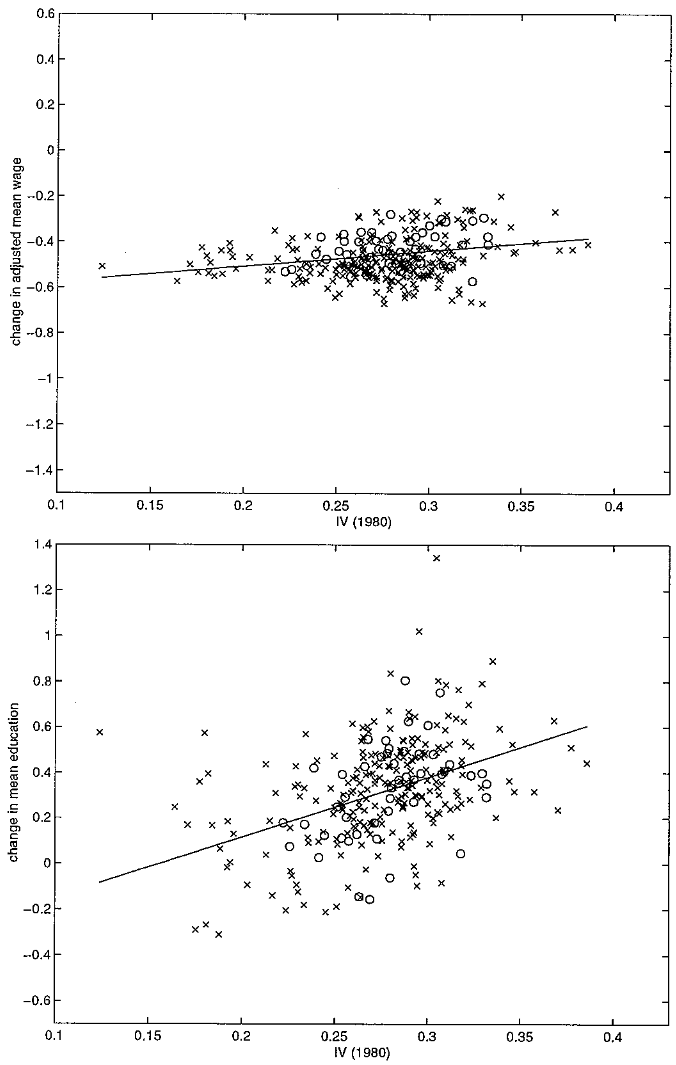


Figure 3: Visual IV. Instrument is $IV_{80}$. Circles Are 50 Largest Cities.

NOTES: Weighted OLS fit superimposed. Scale is the same for both pictures. The weighted IV coefficient is the ratio of the slope in top panel to the one in the bottom panel.

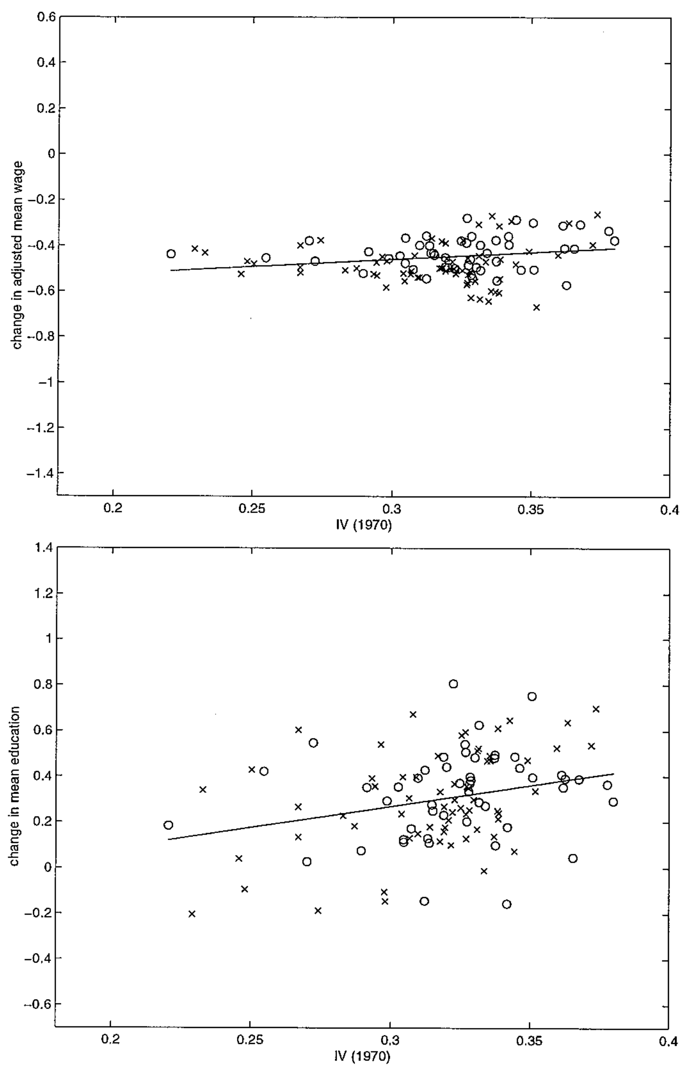


Figure 4: Visual IV. Instrument is $IV_{70}$. Circles Are 50 Largest Cities.

NOTES: Weighted OLS fit superimposed. Scale is the same for both pictures. The weighted IV coefficient is the ratio of the slope in top panel to the one in the bottom panel.

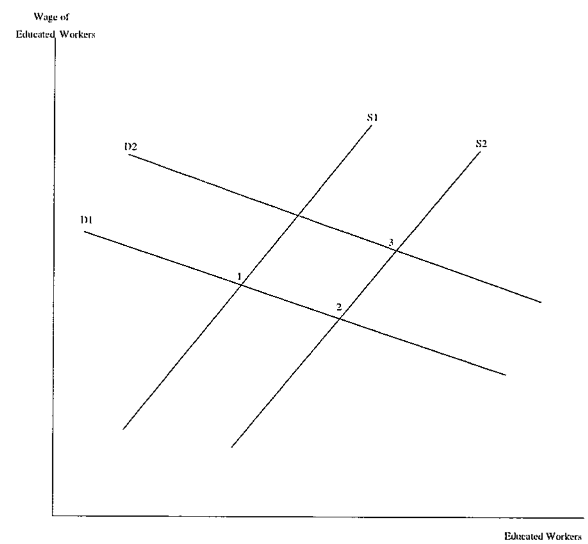


Figure 5: The Effect of an Increase in Supply of College Graduates on College Graduates Wage.