CENTER FOR LABOR ECONOMICS
UNIVERSITY OF CALIFORNIA, BERKELEY
WORKING PAPER NO. 22

Estimating the Social Return to Education: Evidence From Repeated Cross-Sectional and Longitudinal Data *

Enrico Moretti
UC Berkeley

October 1999

Abstract

The social return to education may exceed the private return if there are externalities. In this paper, I estimate externalities from education by comparing wages for otherwise similar individuals who work in cities with higher or lower shares of college educated workers in the labor force. OLS estimates based on Census data show a large positive relationship between the share of college graduates in a city and individual wages, over and above the private return to education. A key issue in this comparison is the presence of *city–wide* unobservable factors that may raise wages and attract more highly educated workers to different cities. To control for the potential endogeneity of education across cities, I use three instrumental variables: the presence of a land–grant college; the cost of tuition at state colleges and universities; and the city demographic structure. I then investigate the hypothesis that the correlation between college share and wages is due to omitted *individual* characteristics, such as ability. I turn to the National Longitudinal Survey of Youth (NLSY) to build a richer econometric model of non–random selection of workers among cities. By observing the same individual over time and in different cities, I can control for permanent factors that make an individual–city match particularly productive. The model is more general than the standard individual fixed effects model. The results from the NLSY sample are remarkably consistent with those based on Census data. A percentage point increase in the supply of college graduates raises high-school drop-outs' wages by 2.3%, high-school graduates' wages by 1.4%, the wages of college graduates by 1.2%. The effect is larger for less educated groups, as predicted by a conventional demand and supply model. But even for college graduates, an increase in the supply of college graduates increases wages, as predicted by a model that includes both conventional demand and supply factors and externalities. (JEL I2, J31)

*Financial support from the A. P. Sloan Dissertation Scholarship is gratefully acknowledged. I am deeply indebted to David Card for many helpful conversations and insightful advice. I also thank John Bound, Ken Chay, David Levine, and Jeff Perloff for detailed comments, as well as Joshua Angrist, Michael Ash, Richard Blundell, Michael Greenstone, Hillary Hoynes, Darren Lubotsky, Marco Manacorda and the participants in the Econometric, Labor and Public Economics seminars at Berkeley for many useful suggestions on this and an earlier draft. Michael Greenstone generously provided the computer code to match MSAs across Censuses. All errors are naturally my own. A previous version of this paper appeared under the title 'Social Returns to Education and Human Capital Externalities: Evidence From Cities', *Center for Labor Economics*, University of California, Berkeley, Working paper No. 9 (1998). E-mail: enrico@econ.berkeley.edu

1 Introduction

Economists have speculated for at least a century that the social return to education may exceed the private return. Different explanations have been offered for such externalities. For example, the sharing of knowledge and skills through formal and informal interaction may generate positive externalities across workers.[1] Alternatively, spillovers from education may arise through search externalities or endogenous skill–biased technical change.[2] Some have hypothesized that education may even have other economic and non–economic benefits in addition to its effect on earnings. Milton Friedman (1962) argued for public subsidies to education on the grounds that a better–educated electorate makes better decisions over policy choices that affect the economy.

State and local governments subsidize primary and secondary education. Almost all direct operating costs are completely subsidized through high school. The current subsidy of direct costs to students at major public universities in the U.S. is around 80% (Heckman 1999). The magnitude of the social return to education is important for assessing the efficiency of public investment in education. Despite the significant policy implications and a large theoretical literature that assumes the existence of externalities from education, the empirical evidence on the magnitude of any externalities is limited.[3] This is particularly surprising given the huge literature on estimating the *private* return to education that has emerged in labor economics in the past three decades. Nevertheless, "labor economists are conspicuous by their absence" on the subject of social returns to education (Topel 1999). If we are to evaluate the full set of benefits that education generates for society, then the private return to education may not tell us the whole story.

In this paper, I test the hypothesis that the economic returns to education are fully reflected in the earnings of educated workers against the alternative that other individuals in the same labor market benefit from externalities associated with higher overall levels of education. I focus on local labor markets and identify the external effect of human capital by comparing the wages of otherwise similar individuals living in metropolitan areas with different shares of college educated workers in the labor force. If externalities exist, then we should find that workers in cities where the labor force is better educated are more productive and earn more than otherwise similar workers in cities where the labor force is less educated.

OLS estimates show a large positive relationship between the share of college graduates in

[1] See for example Marshall (1890), Lucas (1988), Glaeser (1997), Jovanovic & Rob (1989).

[2] See Acemoglu (1996) and Acemoglu (1998).

[3] Borjas (1995, 1997) shows that economic success of children of immigrants depends not only on parental inputs, but also on the average human capital of the relevant ethnic group. Glaeser, Scheinkman & Shleifer (1995) report that income per capita has grown faster in cities with high initial human capital in the post-war period. Findings in Glaeser & 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. In high technology industries, innovation has been shown to be linked to human capital (Zucker, Darby & Brewer 1998, Jaffe 1989, Anselin, Varga & Acs 1997, Jaffe, Trajtenberg & Henderson 1993). The only studies that adopt an approach similar to the one used here are Rauch (1993) and Acemoglu & Angrist (1999), and are discussed below.

a city and individual wages, even after controlling for the direct effect of individual education on wages. This correlation is shown in figure 1, where the percentage of college graduates in 282 cities is plotted against the regression-adjusted average wage.[4] The figure shows that, after controlling for the private return to education, wages are higher in cities where labor force is better educated. However, it is not clear whether this documented association is causal.

A fundamental issue in the interpretation of this simple specification is the presence of unobservable factors that are correlated with college share and wages across cities. First, there are city–specific unobserved characteristics that are correlated with college share. If cost of living is higher in cities where the share of college educated is higher, then wages may be higher to compensate workers, not because of productivity differences. Similarly, cities where the productivity of skilled workers is particularly high because of unobserved differences in industrial mix, technology or natural resources, pay higher wages and therefore attract more skilled workers. The computer industry boom, for example, has raised wages of skilled workers and therefore attracted a very skilled labor force to Silicon Valley. In this case, high wages cause the number of college graduates in the city to rise, not vice-versa. To abstract from city-level unobserved heterogeneity, I use instrumental variable estimates based on the 1970, 1980, and 1990 Censuses.

Second, there may be unobserved individual characteristics that are correlated with college share across cities. It is plausible that workers in cities with a well educated labor force have higher levels of unobserved ability. To control for this potential source of unobserved heterogeneity, I use data from the National Longitudinal Survey of Youths (NLSY), 1979-1994. Although much smaller than the Census, the NLSY follows a fixed set of individuals over time, and allows for richer models of non-random selection across cities.

The empirical section of the paper begins by presenting estimates based on 1980 and 1990 Census data. Cities differ widely in geographical location, industrial structure, weather and amenities. By pooling data from the two Censuses and examining changes in wages and college share, I abstract from any permanent city–specific characteristics that might bias a simpler cross-sectional analysis. First–differenced models may still be biased by the presence of time–varying factors that are correlated with changes in college share and wages across cities—for example, transitory productivity shocks that attract highly educated workers and raise wages. I attempt to estimate these shocks directly with a measure of demand shifts proposed by Katz & Murphy (1992).

To increase the robustness of these estimates, what is needed is an instrumental variable that is correlated with college share in a city and uncorrelated with factors that may affect wages directly. I consider three alternatives. The first is the presence of a land–grant college in the city. The Morrill Act of 1862, which established land–grant colleges and universities, was the first major

[4]The regression–adjusted average wage is obtained by conditioning on individual education and other individuals' characteristics.

federal program to support higher education in the U.S.; because the program was federal and took place more than one hundred years ago, the presence of a land–grant institution is unlikely to be correlated with local labor market conditions in the 1980s. The second instrumental variable is the change in cost of tuition in the state public universities in the period from 1975 to 1985. I show that increases in cost of tuition reduce the percentage of college graduates five years later. Even though it is possible that contemporaneous changes in the cost of tuition are correlated with the state business cycle, it is unlikely that they are correlated with the state business cycle five years ago. A third instrument is based on differences in the age structure of cities. The U.S. labor force is characterized by a long-run trend of increasing education since younger cohorts tend to be better–educated than older ones. To lessen any fear that age structure is endogenous, I use 1970 age structure to predict changes in college share between 1980 and 1990. Although the three instrumental variables are very different from each other, the resulting IV estimates are similar to each other and very close to the corresponding OLS estimates.

Having established that omitted *city characteristics* do not affect estimates, I investigate the hypothesis that the correlation between college share and wages is due to omitted *individual characteristics*, such as ability. I turn to the National Longitudinal Survey of Youth (NLSY), to test the hypothesis that individuals observed in cities with a large college share are better workers than individuals with the same observable characteristics who live in cities with a small college share. This type of sorting could take place if higher college share in a city is associated with higher return to unobserved ability, causing higher quality workers to move to cities with higher college share. In this case, the correlation between high wages and high college share might simply reflect higher unobserved ability of workers rather than higher productivity.

I account for unobserved ability by exploiting the longitudinal structure of the data. One source of identification comes from individuals who change city of residence.[5] A second source of identification comes from variation in college share within a city over time. By observing the same individual in the same city over time, I can control for unobserved factors that make an individual–city match particularly productive. The resulting econometric model is more general than standard individual fixed effects model.

The results from the NLSY sample are remarkably consistent with those based on Census data. In particular, I find that a 1 percentage point increase in college share in a city raises average wages by 1.2%–1.4% (after controlling for the private return to education). Of course finding that *average* wages are affected by the percentage of college graduates in the labor force does not necessarily indicate an externality effect: rather this finding may indicate complementarity between high and low education workers. Only by looking at the effect of an increase in the supply of educated workers on their own wage can one directly test for the presence of externalities. Estimates

[5] One of the advantages of the NLSY is that youths have higher mobility rate than adults, and this provides more identification.

of the effect of changes in the fraction of highly educated workers on different education groups show that a 1 percentage point increase in the labor force share of college graduates increases the wages of high-school drop-outs and high-school graduates by 2.3% and 1.4%, respectively. Surprisingly, it also increases wages of college graduates by 1.2%. 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, but still generates a rise in wages for the best–educated group.

The remainder of this paper is organized as follows. The next section provides a simple theoretical model. Section 3 reports first–differenced OLS estimates from the Census. Section 4 describes the three instrumental variables and the first–differenced IV estimates. Section 5 reports results from a panel data model estimated with NLSY data. In section 6, a more general model is estimated, where the effect of college share varies by education group. Section 7 estimates a more structural specification. Section 8 concludes.

2 Theoretical Framework

In this section I present a simple theoretical model that includes both standard demand and supply factors and externalities from education. A first goal of this framework is to identify the effect of an increase in the relative supply of college educated workers in a city on wages. A second goal is to identify potential sources of bias of the OLS estimator and suggests ways to account for them. The existing literature has generally ignored endogeneity issues by assuming that the overall level of education in cities is historically predetermined (Rauch 1993). 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 labor demand dominates heterogeneity in labor supply, the OLS coefficient is overestimated. If heterogeneity in supply dominates heterogeneity in demand, the OLS coefficient is underestimated.

I focus on local labor markets and identify the effect of human capital externalities by comparing the wages of otherwise similar individuals living in metropolitan areas with different shares of highly educated workers in the labor force. Alternative specifications are possible; externalities could arise at the national level[6] or along ethnic lines.[7] Nevertheless, metropolitan areas, being less arbitrary economic units than countries and having a higher degree of economic and social integration, represent a natural starting point.

[6]In the macroeconomics literature, changes in education levels across years have not been found to affect growth in a cross-section of countries (Benhabib & Spiegel 1994). Using the same data, Krueger & Lindahl (1998) show that this result is largely due to measurement error in aggregate years of schooling. After accounting for measurement error, Krueger & Lindahl (1998) find that the effect of changes in educational attainment on income growth is significantly larger than microeconomic estimates of the rate of return to years of schooling. Topel (1999) finds that an additional year of schooling raises average productivity by nearly 9%.

[7]Borjas (1995, 1997) reports that the human capital production function depends not only on parental inputs, but also on the average human capital of the relevant ethnic group.

To understand the effect of changes in the share of college educated workers in a city on wages, it is convenient to treat each city as a competitive economy that produces a single output good y traded on the national market. 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. Labor demand and supply differ across cities and education groups. Differences in labor demand reflect differences in productivity due to unobserved factors such as industrial mix, technology, natural resources, or institutions. Differences in labor supply reflect differences in the quality of life and amenities that appeal to different education groups. The model allows for fairly unrestricted specifications of heterogeneity in the labor demand and labor supply functions of different skill groups.

The production function in city c at time t is $y_{ct} = f(T_{ct}, L_{ct})$, where T_{ct} is a vector of non-labor inputs (land, capital, etc.), and L_{ct} is a CES–type aggregate of the quantities of labor N_{jct} in various skill categories j=1,..,J:

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

where θ_{jct} is a group–, city– and time–specific productivity shifter and $\sigma > 1$. If the selling price of output is normalized to 1, the condition that the marginal product of each type of labor is equal to its wage can be written as

$$\ln w_{jct} = \mu_{ct} - \frac{1}{\sigma} \ln N_{jct} + \frac{\sigma - 1}{\sigma} \ln \theta_{jct} \tag{2}$$

where $\mu_{ct} = \ln(f'_L(T_{ct}, L_{ct}) L_{ct}^{1/\sigma})$ is a city–time specific component shared by all skill groups.[8]

In the model, I introduce externalities by assuming that workers' productivity is a function of their own human capital and the human capital of other workers in the same city. This idea is old in economics. Marshall (1890) is often quoted as arguing that social interactions among workers in the same industry and location create learning opportunities that enhance productivity. More recent literature has build on Marshall's insight by assuming that human capital externalities arise because workers learn from each other, and they learn more from more skilled individuals.[9]

Consistent with the theoretical literature, I assume that workers who are in cities where the labor force is better educated are more productive than similar workers in cities where the labor force is less educated. In particular, the productivity of group j in city c at time t depends on the

[8] For simplicity I assume that the spillover takes place outside the firm, so that firm take the θ's as given. Qualitative results do not change if the spillover is internalized within the firm.

[9] An influential paper by Lucas (1988) suggests that human capital externalities may help explaining differences in long run economic performance of countries. The sharing of knowledge and skills through formal and informal interaction is viewed as the mechanism that generates positive externalities across workers. More recent models build on this idea by assuming that individuals augment their human capital through pairwise meetings with more skilled neighbors at which they exchange ideas (Glaeser 1997, Jovanovic & Rob 1989). Other authors focus on the importance of basic research in fostering technological innovation and productivity, the public good nature of the research and the resulting positive externalities in the form of knowledge spillovers (see, for example, Arrow (1962), Griliches (1979)) .

group's own human capital, captured by a group–specific effect θ_j, the share of college educated workers in the city, P_{ct}, and an error term:

$$\ln \theta_{jct} = \theta_j + \gamma P_{ct} + v_{jct} \tag{3}$$

where $\gamma \geq 0$. If $\gamma = 0$, the model reduces to the standard Mincerian model of wage determination without externalities. The term v_{jct} allows for unobserved geographic heterogeneity in labor demand, with an unrestrictive error component specification:

$$v_{jct} = u_j + u_c + u_t + u_{jc} + u_{jt} + u_{ct} + u_{jct} \tag{4}$$

Geographic heterogeneity in labor demand is due to permanent and transitory differences in technology, natural resource endowments, 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}.

Two goods are purchased in this economy: the consumption good y and housing. Workers maximize utility subject to a budget constraint by choosing quantities of the composite good y and housing, given the city's amenities. Because the composite good is traded, its price is the same everywhere. Variation in the cost of living depends only on the variation in the cost of housing. Workers and firms are perfectly mobile. In Appendix 1 I show that in the case of two cities and two education groups an equilibrium exists where both education groups are present in both cities.

Assume that the number of workers in each group in city c can be written as a simple function of the group's real wage in the city and the city's unobservable amenities, v'_{jct}:

$$\ln N_{jct} = \eta \ln \frac{w_{jct}}{r_{ct}} + v'_{jct} \tag{5}$$

where $\eta > 0$ and r_{ct} is the cost of living in the city. The error term v'_{jct} allows for unrestricted supply heterogeneity:

$$v'_{jct} = u'_j + u'_c + u'_t + u'_{jc} + u'_{jt} + u'_{ct} + u'_{jct} \tag{6}$$

Geographic heterogeneity in labor supply comes from differences in local amenities, such as weather conditions, coastal location, cultural services like museums, cinemas, concerts, etc. For example, if the weather in city c is particularly favorable, making the 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 $u'_c = 0$.[10]

The reduced form wage equation can be written as

$$\ln w_{jct} = d_j + d_c + d_t + d_{jc} + d_{jt} + \pi P_{ct} + \epsilon_{jct} \tag{7}$$

where the d's are dummies that capture any effect that is group–, city–, time–, group–city and group–time specific; and $\pi = \frac{(\sigma-1)\gamma}{\sigma+\eta}$ is the coefficient of interest. Equation 7 provides the theoretical basis for the empirical work in this paper. The equation relates equilibrium wages of group j in city c at time t to the share of educated workers in city c at time t. The goal of the empirical work is to consistently estimate the coefficient π. Note that P_{ct} is an equilibrium outcome itself. The share of college educated workers is a function of wages and therefore endogenous.

Unobserved city–time–specific heterogeneity may bias OLS estimates if it is correlated with college share in the city. Let $\mathbf{d} = d_j + d_c + d_t + d_{jc} + d_{jt}$. Then, the probability limit of the OLS estimate is given by:

$$plim\ \hat{\pi} = \pi + \frac{1}{\sigma+\eta}\left(\frac{cov(\epsilon^1_{jct}, P_{ct}|\mathbf{d})}{var(P_{ct}|\mathbf{d})} + \frac{cov(\epsilon^2_{jct}, P_{ct}|\mathbf{d})}{var(P_{ct}|\mathbf{d})} + \frac{cov(\epsilon^3_{jct}, P_{ct}|\mathbf{d})}{var(P_{ct}|\mathbf{d})}\right) \tag{8}$$

where $\epsilon^1_{jct} = (\sigma-1)(u_{ct} + u_{jct})$ represents heterogeneity in labor demand and $\epsilon^2_{jct} = -(u'_{ct} + u'_{jct})$, in labor supply. The last term, $\epsilon^3_{jct} = \sigma\mu_{ct} + \eta \ln r_{ct}$ includes the residual heterogeneity. The two main sources of endogeneity are unobserved demand heterogeneity and supply heterogeneity. Consider first unobserved *demand* heterogeneity. If demand for educated workers increases in city c at time t because of a productivity shock ($v_{jct} > 0$), then both the wage of educated workers and mean education in the city rise. Formally, this means that $cov(\epsilon^1_{jct}, P_{ct}|\mathbf{d}) > 0$, implying that OLS estimates are biased upward. 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 $cov(\epsilon^2_{jct}, P_{ct}|\mathbf{d}) < 0$, making OLS biased down. For example, Boston's amenities are considered so desirable that some educated workers are willing to accept a lower wage to live there.

In summary, the model shows that OLS is likely to be biased and the direction of the bias depends on the sources of unobserved heterogeneity. The true effect of college share on wages is larger or smaller than the OLS estimate depending on whether supply heterogeneity dominates

[10] Heterogeneity in labor supply may also come from differences in job characteristics. For example, being a teacher may be very different in Los Angeles than Ann Arbor. Parents' involvement, security and quality of students are important job characteristics that may be correlated with education in the city. 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.

demand heterogeneity or demand heterogeneity dominates supply heterogeneity. If variation in college share across cities is driven by unobserved *demand* shocks, OLS is biased upward. On the other hand, if variation in college share across cities is driven by unobserved *supply* shocks, OLS is biased downward.[11]

Human capital externalities may arise if human and physical capital are complements even in the absence of learning externalities (Acemoglu 1996).[12] The empirical specification adopted in this paper is consistent both with learning spillovers and spillovers arising from complementarity of human and physical capital (Acemoglu & Angrist 1999). The focus of the paper is on rigorously assessing whether externalities are present, and if so, their potential magnitude. Distinguishing between the two alternative sources of externalities is beyond the scope of this paper.

3 Empirical Specification and OLS Estimates Using Census Data

This section develops a simple estimation method that is used to identify externalities from education in local labor markets (equation 7), while recognizing the potential bias that affects a simple cross-sectional analysis. The model is estimated using data from the Public Use Micro Samples (PUMS) of the 1980 and 1990 Censuses of Population. The 5% public use samples provide relatively large sample sizes—1.69 million observations on working adults in 1980 and 1.98 working adults in 1990. I use the Metropolitan Statistical Area (MSA) as a local labor market. MSA's are defined to include local economic regions with populations of at least 100,000: most MSA's contain more than one county. A total of 282 MSA's can be identified and matched in 1980 and 1990 Census. The Data Appendix provides more detailed information on the procedures used to identify the MSAs in each year and match MSAs across them.

A two-stage econometric specification is adopted. In the first stage, the regression-adjusted mean wage in city c at time t, $\hat{\alpha}_{ct}$, is obtained from the following regression:

$$\ln w_{ict} = \alpha_{ct} + X_{it}\beta_t + v_{ict} \tag{9}$$

where w_{ict} is the hourly wage of individual i living in city c at time t; X_{ict} is a vector of individual

[11] Other, more subtle forms of endogeneity may arise in the form of general equilibrium effects of changes in the distribution of education. First, a change in the group composition of the city may shift the demand for housing. This may happen because of differences in tastes or income across groups. Hence, $cov(r_{ct}, P_{ct}|\mathbf{d}) \neq 0$. I address this possibility in section 4.4. In a more general model, where some of the output is locally consumed, a change in the group composition of the city may also alter the demand for city output and hence the demand functions for educated and uneducated labor (Altonji & Card 1991). Second, if human and physical capital are complements, increases in a city mean education may stimulate investment by firms. In this case, $cov(T_{ct}, P_{ct}|\mathbf{d}) > 0$.

[12] If firms and workers find each other via random matching and breaking the match is costly, equilibrium wages will increase with the average education of the work-force even without learning or technological externalities. The intuition is simple. 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 work-force. If a group of workers in a city increases its level of education, firms in that city, expecting to employ these workers, would invest more. 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 other cities.

characteristics including education, sex, race, Hispanic origin, U.S. citizenship and a quadratic term in potential experience; and α_{ct} is a set of city-time specific dummies that can be interpreted as a vector of adjusted city average wages. This equation is estimated separately for 1980 and 1990. In the second stage, $\hat{\alpha}_{ct}$ is regressed on college share in the city, controlling for city fixed–effects and time–varying city characteristics:

$$\hat{\alpha}_{ct} = d_c + d_t + \pi P_{ct} + \alpha Z_{ct} + \epsilon_{ct} \tag{10}$$

where P_{ct} represents the percentage of college educated workers in the labor force of city c in year t; Z_{ct} is a vector of observed city characteristics that includes the unemployment rate,[13] the proportion of blacks, Hispanics, females, US citizens; and d_c and d_t are city and year dummies, respectively. The interaction between a dummy for Southern states and time is also included, to account for unobservable factors that may affect changes in wages and education in the South. 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 number of observations per city. An alternative one-step estimation strategy could be obtained by substituting Equation 10 into 9. The one-step procedure gives rise to estimates that are identical to an appropriately weighted two-step procedure (Hanushek 1974).[14]

The coefficient of interest is π, which is the estimate of the effect of college share on average wages *after controlling for the private return to education.* It is well known that the private return to education has increased substantially during the 1980s. By letting β_t depend on the year, the specification of Equation 9 allows for changes over time in the private return to education. Later, I show that the results are very similar in a specification where the private return to education varies over time *and across cities* (β_{ct}).

In this paper, the externality is a function of the share of college educated workers in a city. This is not the only possible specification. For example, one could think of the externality as a function of average education in a city, instead of college share (Rauch 1993, Acemoglu & Angrist 1999). The specification based on average education is appropriate if the externality generated by an increase in years of schooling in the lower part of the education distribution has the same effect as one generated by an increase in the upper part of the distribution. This assumption has no theoretical justification, and it may not be very realistic in practice. For example, a one year increase in a city's average education obtained by a rise in the number of those who finish high–school may have a very different effect than the same increase in average education obtained by a rise in the number of those who go to college. The specification based on average education makes it difficult to understand what type of changes in the education distribution

[13]Unemployment rate is from the Bureau of Labor Statistics and is at the state level.

[14]The standard errors of the one-step estimator will be understated unless proper adjustments are made for the grouped structure of the data.

are responsible for the identification of the externality. On the contrary, the specification based on college share has the advantage of making explicit what part of the education distribution is identifying the externality. A disadvantage is that by focusing on college graduates, increases in a city's human capital that occur in the left tail of the distribution are missed. Results reported in this paper are based on the specification that includes college share. However, I re-estimated all the models presented here substituting college share with average education. Results are qualitatively the same. In almost all cases, a significant (insignificant) coefficient on college share in a regression is associated with a significant (insignificant) coefficient on average education in the same regression.

Table 1 shows mean wages for the 12 cities with the highest college share among the 282 cities in the sample as well as for the 12 cities with the lowest college share. The metropolitan area with the highest education is Stamford, CT; the one with the lowest education is Steubenville–Weirton, OH–WV. Figure 1 exemplifies the identification strategy used in this paper. The percentage of college graduates in 282 cities is plotted against the regression-adjusted average wage, $\hat{\alpha}_{ct}$. After controlling for the private return to education, wages are higher in cities where the labor force is better educated. Cross–sectional estimates of equation 10 obtained by ordinary least squares (OLS) confirm that the coefficient on percentage of college graduates is significantly larger than zero. Estimates of π in column 1 of Table 2 suggest that a one percentage point increase in the share of college graduates in a city would have raised wages by about 0.98% in 1990 and 0.78% in 1980. The coefficients on the remaining variables—not reported in the table—are consistent with findings in the literature of human capital earnings functions.[15]

Although cross-sectional estimates are purged of differences in cities' observable characteristics, Z_{ct}, the results may still be biased if there are unobserved differences in cities' characteristics that happen to be correlated with education. Row 3 of Table 2 reports estimates of the effect of college share when city–fixed effects are included. This specification compares changes in wages over time to changes in college share. By including city-fixed effects, all permanent unobserved heterogeneity that might be correlated with college share is absorbed. For example, New York is different than Brownsville in terms of the cost of living, industry composition, climate, cultural amenities, etc. When city fixed effects are included, the coefficient on college share is 1.5%. This coefficient is larger than the corresponding cross-sectional coefficients, suggesting that permanent unobserved heterogeneity across cities is negatively correlated with college share.

One explanation for the finding that average wages are higher in cities with a better educated labor force is that there are city-specific transitory productivity shocks that attract educated workers and at the same time raise wages. For example, the San Jose economy has experienced

[15]The private return to education is 0.085 (0.0001) in 1990 and 0.062 (0.0001) in 1980. The 1990 coefficients on the quadratic term in experience are 0.040 (0.0001) and -0.0006 (0.0000) for the linear and squared term, respectively. The corresponding coefficients for 1980 are 0.033 (0.0001) and -0.0004 (0.0000). The 1990 coefficients on black, Hispanic, female and US citizens are -0.119 (0.001), -0.076 (0.001), -0.313 (0.0008) and 0.132 (0.001), respectively. The corresponding coefficients for 1980 are -0.078 (0.001), -0.049 (0.002), -0.402 (0.0009), 0.093 (0.002).

an unprecedented expansion starting in the second half of the 1980s that was caused by the Silicon Valley computer industry boom. The same boom has attracted a highly educated labor force to San Jose. The simple model of section 2 suggests that this type of unobserved demand shock would bias the coefficient upward.

I account for transitory shocks in two ways. First, in the remainder of this section, I include direct estimates of the transitory shocks to local labor markets in the estimating equations. Second, in the next section, I use three different instrumental variables that predict variation in college share but are potentially uncorrelated with wages. To pursue the first of these approaches I adapt a measure of labor demand shifts proposed by Katz & Murphy (1992). This index, a generalization of a widely used measure of between-sector demand shifts, is based on nation-wide employment growth in industries, weighted by the city-specific employment share in those industries:

$$\text{shock}_c = \sum_{l=1}^{46} \left(\frac{E_{lc}}{E_c}\right)\left(\Delta \frac{E_l}{E}\right) \tag{11}$$

where l indexes two–digit industry; $\Delta \frac{E_l}{E}$ is the change in share of labor input in the U.S. economy in industry l between 1980 and 1990; and $\frac{E_{lc}}{E_c}$ is industry l's share of total employment in city c in the base year (1980). The index captures exogenous shifts in local labor demand that are predicted by the city industry mix. As noted by Bound & Holzer (1996), different cities specialize in the production of different goods, so that industry-specific demand shocks at the national level have a differential impact on cities. If employment in a given industry increases (decreases) nationally, cities where that industry employs a significant share of the labor force will experience a positive (negative) shock to their labor demand.[16]

When the Katz and Murphy index is included among regressors, the coefficient on college share does not change significantly. Column 2 in Table 2 reports the coefficient on college share conditional on the Katz and Murphy index and all other covariates. The coefficient is 1.4, lower than the corresponding fixed effects estimate in column 2, although the difference is statistically insignificant. The coefficient on the Katz and Murphy index is 0.052 (0.015), confirming that wages indeed rise in cities hit by positive demand shocks. The coefficient on the unemployment rate is -0.024 (0.002). The fact that the Katz and Murphy index does not significantly change the estimates lends support to the view that the bias introduced by demand shocks is not large.

4 Instrumental Variable Estimates Using Census Data

If the Katz and Murphy measure of demand shifts captures *some* transitory shocks to the local labor market, but not all, then OLS estimates in the previous section may still be inconsistent.

[16]Labor inputs is measured in two ways. First, I use hours worked. Second, following Katz & Murphy (1992), I measure labor input in "efficiency units", i.e. hours worked weighted by the value of each hour, which equals earnings. I report results using the "efficiency units" definition.

Instrumental variables (IV) can potentially yield estimates that are more robust, although less efficient. Appropriate instruments must be related to the share of college graduates in a city but orthogonal to unobserved shocks to the demand and supply of labor there. In this section I examine three instrumental variables: the presence of a land–grant college, the cost of tuition at colleges and universities, and the age structure in 1970. The IV estimates are very stable across instruments, and in no case are they significantly different from the corresponding OLS estimates.

4.1 Land–Grant Colleges

The presence of colleges or universities in a city tend to raise education there.[17] Using the presence of colleges and universities as an instrumental variable for college share is problematic, however, if their location is non–random. If colleges and universities tend to be located in wealthy areas, for example, then the exogeneity of the instrument is in question. A solution to this problem is to use as the instrument the presence of colleges and universities created in the nineteenth century following the "land-grant movement".[18]

In 1862, the U.S. Congress passed the Morrill Act, the first major federal program to support higher education in the United States. The act gave to every state that had remained in the Union a grant to establish colleges in engineering, agriculture and military science. A second act in 1890 extended the land–grant provisions to the sixteen southern states.[19] Altogether, 73 land-grant colleges and universities were founded, with each state having at least one. Although originally started as agricultural and technical schools, many of them have grown into large public universities that have educated almost one-fifth of all students seeking degrees in the United States. Even today, the presence of a land–grant college in a city remains a significant determinant of higher education there.[20] The top panel of Figure 2 plots the difference in the probability of schooling for cities with and without a land–grant college. In the figure, the probability of attending college is higher in cities with a land–grant college, while the probability of high–school graduation is lower. The figure suggests that some students who would otherwise stop studying after high-school, go to college if they live near a land–grant college. The difference in probability declines at lower grades. Figure 2 is consistent with the assumption that the

[17]First, having a local college lowers the cost of going to college. Second, a college graduate is more likely to stay and work in the city where she studied. Finally, high–tech and research firms that employ skilled workers are more likely to be located close to major research universities.

[18]I thank David Levine for suggesting this instrument.

[19]Since the act withheld funds from states that refused to admit nonwhite students unless those states provided "separate but equal" facilities, it encouraged the foundation of black colleges.

[20]For the complete list of land–grant colleges see Appendix in Nervis (1962). Only 39 of the original 73 land–grant colleges are located in a MSA. The following MSA have one or more land–grant colleges: Albany-Schenectady-Troy, NY; Athens, GA; Baton Rouge, LA; Boston, MA; Champaign-Urbana-Rantoul, IL; Columbia, MO; Columbia, SC; Columbus, OH; Des Moines, IA; Fargo-Moorhead, ND-MN; Fayetteville-Springdale, AR; Fort Collins-Loveland, CO; Gainesville, FL; Greensboro-Winston-Salem-High Pt., NC; Hartford, CT; Honolulu, HI; Knoxville, TN; Lafayette-West Lafayette, IN; Lansing-East Lansing, MI; Lexington-Fayette, KY; Lincoln, NE; Macon-Warner Robins, GA; Madison, WI; Minneapolis-St. Paul, MN-WI; Nashville, TN; Pine Bluff, AR; Portsmouth-Dover-Rochester, NH-ME; Providence, RI; Raleigh-Durham, NC; Reno, NV; Richmond-Petersburg, VA; Riverside-San Bernardino, CA; Sacramento, CA; San Francisco, CA; State College, PA; Tallahassee, FL; Tucson, AZ; Washington, DC-MD-VA; Wilmington, DE-MD.

presence of a land–grant college increases the probability of college graduation, and not vice versa. If the presence of a land–grant college captures unobservable characteristics of the area, such as tastes for education, then we might find that in cities with a land college not only the probability of going to college, but also the probability of graduating from high–school and attending some college is higher.

Table 3 quantifies the difference for 4 education groups. Row 1 and 2 report estimates of a first–stage regression of the percentage of college graduates on a dummy equal one if the city has a land–grant college for 1990 and 1980 respectively. Other wage equation covariates are also included. The presence of a land–grant college raises the percentage of college graduates in the city by 0.049 and 0.053 in 1990 and 1980 respectively.[21] This is a remarkable effect, given that average college share is 0.25.

Since the program that established land–grant colleges was *federal* and took place more than a century ago, the presence of a land–grant institution is likely to be uncorrelated with unobservable factors that affect wages in 1980 and 1990. Land–grant colleges were often established in rural areas, and their location was not dependent on natural resources or other factors that could make an area wealthier. In fact, judged from today's point of view, the geographical location of land–grant colleges seems close to random. (See Nervis (1962), Williams (1991) and Edmond (1978) for a history of the "land–grant movement"). It is true that many land–grant colleges are today close to other major educational institutions, but this is likely to be an effect of the original presence of the land–grant college in the area. Thus, land–grant colleges appear to be a legitimate instrument. An alternative instrument could be the presence of any college founded before, say, 1900. However, that instrument would be geographically unbalanced. Before the Morrill Act, only 6 out of 181 colleges and universities were located in Western states (See Map 1 in Tewksbury (1932)). In contrast, land–grant colleges are present in each state.

Column 4 of Table 2 shows IV estimates obtained by using the land–grant colleges as an instrument. The coefficients are 0.73 in 1990 and 0.58 in 1980, smaller than the corresponding OLS coefficients of column 2. A Hausman test suggests that IV and OLS estimates are not significantly different. This remains true in most of specifications presented in this paper. A potential concern is that in college towns such as Urbana (IL) or Gainsville (FL), the university may be a major employer. If teachers' and staff's wages are above the average, this could bias the results. To account for this possibility, I re-estimated the model excluding workers employed in colleges and universities, libraries and educational services. I also re-estimated the model including a college town dummy variable.[22] The results did not change in either case. This is not surprising, because cities in the sample are typically large metropolitan areas with many industries, and the university is rarely the only employer.

[21] The 1980 and 1990 coefficients are not statistically different.

[22] 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.

4.2 Cost of Tuition

The presence of a land-grant college in a city is good predictor of cross-sectional variation in college share. But a dummy for the presence of a land-grant college cannot be used as an instrument in the specification that includes city fixed effects, because it would be absorbed by the fixed effect. In this section and the following one, I propose two instruments that can be used when city fixed effects are included: the cost of tuition at state colleges and universities; and the demographic structure of cities. These instruments are correlated with changes in the share of college educated workers across cities between 1980 and 1990 but arguably uncorrelated with unobservable demand shocks.

The average cost of tuition at state public colleges and universities rose by 25% between 1976 and 1986, after adjusting for inflation. The increase was very different across states, ranging from -40% in Florida to 83% in California. Previous studies have used this variation in state public tuition levels as a natural experiment to identify the effect of tuition on college enrollment. Results suggest that the increase in cost of tuition significantly reduced college enrollment.[23] In a review of the literature, Leslie & Brinkman (1988) report a consensus estimate of a six to eight-percentage-point decline in enrollment rates for a $1000 change in net cost. To account for lags between college enrollment and labor force entry—and to abstract from possible endogeneity of tuition–setting—I use the change in average cost of tuition at state colleges and universities between 1976 and 1986 as an instrument for the change in the share of college graduates between 1980 and 1990.[24]

To assess the effect of increases in the cost of tuition on schooling, the bottom panel of Figure 2 plots the difference in the change in the probability of schooling for cities located in states with large and small increases in cost of tuition. I define an increase in tuition large if it is 50% or more.[25] The figure suggests that in cities where the increase in tuition costs was large, the share of workers with college degree did not grow as much as in cities where the increase in tuition costs was small. Table 3 (row 3) quantifies the difference for 4 education groups. The coefficients are obtained by regressing changes in the probability of belonging to one of the four groups on a dummy equal 1 if the increase in cost of tuition is large. Other wage equation covariates are also included. The results confirm that large increases in the cost of tuition at public 4-years colleges and universities lead to slower growth in the share of college educated workers in the labor force.[26]

One potential concern in using tuition costs as an instrument is that tuition increases are correlated with state–level economic trends. The large tuition increases of the 1980s were in part

[23]See Kane (1994 and 1995).

[24]Data on cost of tuition are the same used in Kane (1994) and are described there. I thank Philip Leslie for making them available.

[25]27 cities experience increases in cost of tuition larger than 50%.

[26]The percentage of high–school graduates seems also affected by the increase in cost of tuition, but the difference is small and barely significant.

caused by fiscal pressures on state budgets, which are not independent of the local business cycle. If tuition increases are counter-cyclical, rising when unemployment is high, then tuition changes may be correlated with other factors that influence wages, and yield inconsistent estimates when used as instruments. However, two features of the IV estimation procedure help to guard against potential endogeneity. First, the effect of local business cycle on wages is, at least partially, purged by the inclusion in the regression of the state unemployment rate and other observable time-varying city characteristics. Second, 5–years lags in tuition are unlikely to be correlated with current business cycle fluctuations.

Column 4 of Table 2 shows the IV estimate obtained by using changes in cost of tuition as instruments. The coefficient on college share is 1.6, not statistically different from the corresponding OLS estimates or from the other IV estimates based on demographic structure.

4.3 Age Structure

The third instrumental variable is based on exogenous differences in the demographic structure of cities. 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. For example, over the last 25 years, much of the increase in the number of college graduates can be attributed to the entry of the baby boom cohort. To the extent that the relative population shares of different cohorts vary across cities, this will lead to differential trends in college share across cities. Formally, the instrument for changes in mean years of education in city c is defined as a city–specific weighted average of national changes in educational achievement by age-gender group:

$$IV_{80} = \sum_m \omega_{mc}[P_{m90} - P_{m80}] \tag{12}$$

where m indicates age-gender groups (for example: men 58-60); P_{mt} is the mean in years of education for group m at time t at the national level and ω_{mc} is the share of group m in city c in 1980.[27]

To lessen any fear that age structure is itself endogenous, a similar instrument can be obtained from the 1970 age structure instead of the 1980 one. When 1970 data are used, ω_{mc} is the proportion of people living in city c in 1970 who, in 1980, would 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, ω_{mc} estimated with 1970 data would be on average equal to ω_{mc} estimated using 1980 data. In the presence of mobility, the two instruments 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.

[27] Weights ω_{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.

Rows 4 and 5 of Table 3 show results from the first stage regression of changes in college share on the instruments and all exogenous variables. Both instruments are good predictors of changes in average education, although IV_{80} is more precise, as expected. The instruments are not correlated with the share of workers with some college and high–school graduates, confirming that the secular trend in education increases mainly affect the upper part of the education distribution.

The instrument IV_{70} is exogenous if the 1970 demographic structure is orthogonal to unobservable labor demand shocks between 1980 and 1990. This condition does not require that wages at time t be uncorrelated to the age distribution at the same time. Instead, the condition requires that transitory shocks to labor demand experienced by a city between 1980 and 1990 not be associated with the age distribution in 1970.[28]

Some implications of this identification assumption can be tested. In particular, I can test whether demographic structure is correlated with geographical mobility, changes in labor force participation, and other labor market outcomes. 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. A second concern is the potential for correlation of demographic structure with labor force participation. Table 4 reports the correlation between the share of 3 age groups and the net inflow of domestic and international immigrants, the change in population size, labor force size and labor force participation rate.[29] The observation of a correlation between these labor market outcomes and age structure may cast some doubt on the exogeneity of the instruments. 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 4 columns are obtained similarly. All control variables that appear in the wage equations are included.

As expected, the demographic structure in 1970 is in general uncorrelated with changes between 1980 and 1990. The only exception is 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 peo-

[28] In terms of Equation 10, the instrument is exogenous if

$$cov[IV_{70}, (\epsilon_{c90} - \epsilon_{c80})|d_t, X, Z] = \sum_m (P_{m90} - P_{m80}) E[\omega_{mc}(\epsilon_{c90} - \epsilon_{c80})|d_t, X, Z] = 0 \qquad (13)$$

where expectation is taken over c. Equation 13 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: $E[\omega_{mc}(\epsilon_{c90} - \epsilon_{c80})|d_t, X, Z] = 0$, for each m. Since city fixed effects are included, the residual is purged of all *permanent* characteristics of the local labor market. Conditioning on X and Z eliminates any observable *time–varying* individual and city characteristics from the residual.

[29] Each individual in the 1990 Census was asked to identify the Public Use Micro Area (PUMA) 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.

ple in 1980. Changes in labor force participation rate are uncorrelated with both the 1970 and 1980 demographic structure of a city. From Table 4 I conclude that the demographic structure in 1970 is arguably a better instrument than the demographic structure in 1980. However, there is a trade-off between robustness and sample size. A total of 282 Metropolitan Statistical Areas are identified in the Census in 1980 and 1990, while only 115 MSAs are identified in 1970.

Instrumental variables estimates based on 1970 and 1980 age structure are in columns 5 and 6 of Table 2, respectively. These estimates imply that a one percentage point increase in college share would raise wages by 1.0% or 2.1% respectively, if 1970 or 1980 demographic structure is used. Since the 1970 estimate is arguably preferable, more confidence should be put on the lower estimate. A Hausman test suggests that neither estimate is statistically different from the corresponding OLS estimate.[30]

In summary, the most important feature of the results obtained so far is that, although the instruments are very different from each other, they all provide similar results. In column 7, 1970 age structure and cost of tuition are used as instruments. The coefficient is 1.3. Over-identifying restrictions are not rejected, confirming that the two first–differenced IV estimates are statistically equal (p-value is 0.46). I also test whether cross-sectional IV estimates are similar to first-differenced IV estimates and, for all possible pairs, I fail to reject the hypothesis that they are statistically equal.

A second interesting feature is that, regardless of the instrument used, IV and OLS estimates are close in magnitude, and the difference between them is never statistically significant. This suggests that any positive bias in the OLS estimate introduced by demand heterogeneity is offset by negative biases attributable to supply heterogeneity, once the time–varying city controls are included (Section 2).

The first–differenced instrumental variable estimates in column 7 of Table 2 imply that during the 1980s, a one percentage point increase in a city's college share was responsible for an increase in wages of about 1.3%. To help in assessing the magnitude of this effect, consider that the yearly increase in the proportion of college graduates in the labor force is quite small in most cities. The median increase in the percentage of college graduates from 1980 to 1990 was only 2%. Such an increase would imply an increase in wages by 2.6% over a ten year period. For a worker who earns $25,000, that effect amounts to $65 per year.

4.4 Accounting for Differences in Cost of Living

A potential concern is that cities where college share is higher may also have higher cost of living. To the extent that nominal wages adjust to higher cost of living, estimates could contain spurious

[30]The IV_{80} estimate obtained from the sample of 115 cities for which 1970 data are available is 1.792 (0.615), 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. I also estimated the same model using both IV_{70} and IV_{80}. An test of over-identification fails to reject the hypothesis of exogeneity, suggesting that the IV_{70} and IV_{80} estimates are not statistically different (p-value is 0.25).

correlation. This is not likely to be a serious problem for estimates in table 2. First, much of the variation in cost of living is permanent, as high–cost of living cities like New York are permanently more expensive than low cost of living cities, like Brownsville, TX. For example, the correlation coefficient between cost of housing in 1980 and 1990 in 282 cities is 0.78.[31] When city fixed effects are controlled for, most variation in in cost of living is absorbed. It is still possible that changes in cost of living between 1980 and 1990 are correlated with changes in college share and wages during the same period. But there is no reason to believe that IV estimates are biased, if the instruments are uncorrelated with cost of living.

To confirm that IV estimates in section 4 are not driven by unobserved heterogeneity in cost of living, I perform two types of tests. First, I re–estimate the model using only workers in manufacturing jobs. Second, I explicitly control for cost of housing.

The previous literature has pointed out that, if labor markets are perfectly competitive, *nominal* wage differences should reflect differences in the marginal product of labor in industries that produce tradable goods (Glaeser & Mare 1994, Rauch 1993). If this were not the case, firms would simply relocate to less expensive locations. For example, wages and cost of living in San Jose are among the highest in the nation. But if computer workers in Silicon Valley were not particularly productive there, computers firms could simply relocate to a cheaper area. This is not true for firms that produce non–tradable goods. By focusing on manufacturing wages, I have a direct test of whether estimates in the previous section are simply due to compensating differentials for higher cost of living or if they reflect true productivity differences.

Column 1 in Table 5 presents an estimate of the effect of changes in a city's college share on change in the wage of manufacturing workers. In this table, OLS estimates are shown, and each column represents a different specification. The model is identical to the first–differenced OLS specification in Table 2 (column 2, row 3). The coefficient for manufacturing workers is 1.66, almost identical to the corresponding coefficient for all workers. This result lends some support to the view that differences in nominal wages reflect productivity differences across cities.

The remaining columns of Table 5 show the influence of the percentage of college educated in non-manufacturing industries on the wage of manufacturing workers. In column 2, changes in a city's manufacturing wages are regressed on the share of college educated workers in manufacturing in that city. The coefficient, 0.83, is lower than the one in column 1. The specification in column 3 is similar to the one in column 2, but includes among the regressors college share in industries other than manufacturing. An interesting feature of column 3 is that, the effect of college share in manufacturing on wages in that industry is smaller than the similar effect in other industries. This evidence is consistent with the presence of spillovers across industries. The specification of column 4 is similar to the one in column 3, but college share in 'business

[31]As a measure of cost of housing I use the 'fair market rent', described below. The R^2 in a regression of 1990 housing cost on 1980 housing cost is 0.62.

services' industries replaces college share in all non-manufacturing industries.[32] Surprisingly, the effect of college share in 'business services' industries on wage in manufacturing is lower than the one of average education in all non-manufacturing industries. Columns 5 and 6 reproduce the specifications of columns 3 and 4, but include dummies for each region–period interaction to absorb any transitory regional shocks. Coefficients decrease significantly.

A second way to account for differences in cost of living across cities is to add cost of housing to the vector of control variables. Most of the variation in cost of living across cities is captured by variation in the cost of housing.[33] As a measure of the cost of housing, I use the logarithm of the 'fair market rent' calculated for each MSA by the Department of Housing and Urban Development.[34] Table 6 reports estimates obtained by augmenting the specifications of Table 2 with the log cost of housing. The positive coefficient on the cost of housing confirms that wages are indeed higher in more expensive cities.[35] However, the coefficients on average education are virtually unchanged from Table 2.[36]

4.5 Robustness Checks

In this section, I report estimates from several alternative specifications designed to probe the robustness of the main results of Table 2. At the top of Table 7 is the base case, the IV estimates in columns 4 and 5 of Table 2. I start by investigating the possibility that changes in the return to skills between 1980 and 1990 differ across cities. It is well known that the private return to education increased significantly during the 1980s. The base case specification is consistent with a uniform change across cities. In that specification, the coefficients on individual characteristics, β_t, vary by year (Equation 9). Row 2 of Table 7 reports results from a specification where the coefficients on individual characteristics vary by city and year.[37] This specification is correct if changes in the private return to education (or other individual characteristics) during the 1980s were different across cities. If this were the case, failure to let the coefficient on individual education vary by city and year could bias the coefficient on college share.[38] The coefficients in

[32] Business services include advertising, personnel supply services, business and consulting services, computer and data processing services, detective and protective services, other business services, finance, insurance, real estate.

[33] Non-housing cost of living are reported by the Bureau of Labor Statistics for only 28 cities and therefore are not used here. The non–housing cost of living indices for the 28 cities for which the indices are available vary by only plus or minus 4% around the national average (Beeson & Eberts 1989).

[34] The 'fair market rent' is used as a cost of living adjustment in some federal welfare programs. It corresponds to the 45^{th} percentile of rents for a 2–bedroom apartment in each MSA.

[35] One limitation of this specification is that cost of housing may be endogenous, if wage and rent are simultaneously determined.

[36] This result is not inconsistent with the Roback model of Appendix 1. The Roback model predicts that utility is the same across cities, not real wages. The effect of an increase in college share on real wages undetermined, as it depends on the form of preferences. An alternative explanation is that results of table 6 reflect imperfect adjustment toward equilibrium. Bartik (1991) presents evidence that adjustment takes longer than a decade to occur.

[37] A 2-stage econometric procedure similar to the one in section 3 is used. In the first stage, regression–adjusted mean wages are obtained by regressing individual wage on individual characteristics. This regression is done separately by city and year. The second stage is identical to the one described in section 3.

[38] In cities where the return to private education increased more, college share would be picking up part of that increase. The direction of the bias would depend on whether the private return to education increased more in cities with high or low share of college graduates .

row 2 increase, but the estimates are very imprecise.[39]

In row 3, I allow for the presence of inter-industry wage differentials. Including 29 industry dummies in the first-stage individual level regression takes any industry composition effect out of the adjusted mean wage. Estimates in row 3 suggest that industry composition does not drive the results. The occupation composition could also induce spurious correlation between wages and education. When four occupation dummies are included in the first-stage regression, results do not change appreciably (row 4).

College towns, where the population is younger and more educated, may pose a problem for the instrument based on the age structure, since demand for skills rose in the 1980s. To address this issue, a college town dummy variable is included (row 5). Not surprisingly, results do not change. Cities included in the sample are large metropolitan area whose economies rarely depend only on the presence of a university.

In this paper individuals are assigned a city on the basis of their residence. MSAs are large enough that for most people 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 6). As expected, results are not sensitive to the small number of workers who commute outside their MSA of residence.

In row 7, eight dummies for the interaction of region and year are included to capture region specific changes in wages between 1980 and 1990. Estimates becomes negative, but the standard errors are so large that estimates are not very informative. This is not surprising since the inclusion of a rich set of regional dummies takes away all the interregional variation needed for identification of the instruments, leaving only within-region variation. The latter represents only a small portion of the total signal and appears to be insufficient for identification. This is confirmed by a significant increase in standard errors and a positive OLS coefficient for the same specification: 1.09 (0.206).

Finally, the residuals are allowed to be spatially autocorrelated (row 8). Spatial autocorrelation may arise because cities that are in the same state are subject to similar unobservable shocks. Standard errors do not change significantly.

I conclude this section by examining whether there are differences in the spillover effect across racial groups. In section 2 it was suggested that externalities arise because social interactions among workers in the same location create learning opportunities that enhance productivity. If social interactions are more frequent among individuals of the same race, then we should find that human capital of white workers has a larger effect on wages of white workers than on wages of black workers, and that human capital of black workers has a larger effect on wages of black workers than on wages of white workers. Results in table 8 are consistent with this hypothesis. The table reports results obtained by estimating models for white and black workers separately

[39]The possibility that return to skills vary across cities in investigated in greater detail in section 5.

using 1990 data. For whites, the coefficient on the share of white college graduates is 0.89, while the coefficient on the share of blacks college graduates is only 0.30. For blacks, the opposite is true: the coefficient on the share of white college graduates is 0.75, while the coefficient on the share of blacks college graduates is only 0.84.[40]

4.6 Related Literature

Few other studies have attempted to estimate externalities from education. Rauch (1993) is the first to exploit differences in human capital across cities to identify externalities, but treats average schooling as historically predetermined.[41] On the contrary, Acemoglu & Angrist (1999) use state variation in child labor laws and compulsory attendance laws to instrument for average schooling. While their OLS estimates of the externality are qualitatively consistent with OLS estimates presented here, their IV estimates are smaller and in most cases not significantly different from zero.

There are two key differences between this study and Acemoglu & Angrist (1999). First, child labor and compulsory attendance laws affect educational attainment in the lower part of the distribution, mostly in middle school or high school. On the contrary, I identify externalities using variation in the number of college graduates, i.e. the upper part of the distribution. Both definitions refer to changes in the overall level of human capital in a city. But there is no theoretical reason to believe that the effect on wages has to be the same. It is possible that a one year increase in a city average education obtained by a rise in the number of those who finish high–school has a different effect than a similar increase in average education obtained by a rise in the number of those who go to college. The difference in the results presented here and in Acemoglu & Angrist (1999) suggests that this, in fact, may be the case in reality.

Another difference concerns the period under consideration. Most models in Acemoglu & Angrist (1999) are estimated for 1960–1980 data. When they add data from the 1990 Census, they find statistically significant positive estimates for the externality, when child labor laws are used as instruments. They suggest it may reflect a change in the social value of human capital, but is more likely due to increased measurement error in the schooling variable in 1990. In the 1980 Census, the survey asked respondents their *highest grade attended*, and whether they had completed that grade. Beginning with the 1990 Census, the survey has asked instead about individuals' *highest degree received*. In order to use a continous *years of schooling* variable in

[40]Measurement error in blacks college share is a problem. Not only there are fewer blacks than whites, but college graduates among blacks are more rare than among whites. For small cities, the percentage of blacks with college degree is very imprecisely estimated. Of the 282 cities, 2 do not have any black college graduate and are dropped from the sample. To alleviate attenuation bias, results in table 8 are obtaining by instrumenting blacks' college share in 1990 with blacks' college share in 1980. For this reason, fixed–effects estimates are not reported. Results in table 8 are obtained from four separate regressions. When both black and white college share are included in the same model, the coefficient on black college share is always zero, possibly because of measurement error.

[41]Rauch finds that a one year increase in average education raises wages by 3% to 5 % in 1980. The data used by Rauch are different from the data used here. Rauch uses the 1 in 1000 B PUMS, which identifies 69,910 individuals in 237 cities. I use the 5% version of the 1980 PUMS, that identifies 1.69 milion individuals in 282 cities.

regressions based on 1990 data, I impute years of schooling using tabulations from the 1992 CPS that contains responses to both educational attainement questions.[42] If there is measurement error in imputed years of schooling, the private return to education is underestimated and the external return is overestimated. Acemoglu & Angrist (1999) point out that the bias remains if average education is instrumented, but individual years of schooling are not.

This is unlikely to be a major problem for my results. The imputation procedure that I adopt is widely used, and has been shown to reproduce accurately estimates of the private return to education that would be obtained by using a continous years of schooling variable (Jaeger 1997).[43] If anything, some authors have suggested that the new measure has a 'decided advantage in accuracy and usefulness' over the old one (Kominsky & Siegel 1994).

I investigate the effect of measurement error by forcing the private return to education to be 10% or 15% larger than it would be in an unconstrained regression. If the external return to education is picking up some of the bias from the private return, then we should see that the coefficients on college share is significantly different in the constrained and unconstrained regression. Table 9 reports coefficients for the unconstrained and constrained regressions. Unconstrained estimates in row 1 are taken from table 2 and are reported only for convenience. When the private return to education is inflated by 10% (row 2), the estimated coefficient on college share decreases as expected, but the change is negligible. This remains true when private return to education is inflated by 15% (row 3). Row 4 refers to a specification where years of schooling is measured with error only in 1990, as Acemoglu & Angrist (1999) suggest. Inflating only the 1990 private return to education has little effect. From table 9 I conclude that attenuation bias in the estimate of private return to education, if exists, has a minor effect on estimates of the coefficient on college share.

5 Unobserved Ability, Comparative Advantage and Non-Random Selection

Results in the previous sections, based on repeated cross-sections from the Censuses, seem to indicate that increases in the percent of college educated in the labor force have a substantial effect on wages, even after controlling for the private return to education. First-differenced IV estimates suggest that this finding cannot be explained by unobserved, city–wide labor demand shocks that raise both wages and education in a city. Neither can cost of living differences explain those estimates.

[42] In particular, years of education are assigned to education codes used in 1990 Census following Table 1 in Kominsky & Siegel (1994).

[43] Jaeger (1997) finds that in those cases where the estimates obtained with the two variables are different, the estimated return to education obtained with the new variable is slightly *larger*. This finding suggests that estimates of the externality are unlikely to biased upward because of attenuation bias in the estimate of the private return to education.

Could there be an alternative explanation for this result? In this section I investigate the hypothesis that the correlation between college share and wages is due to omitted individual characteristics, such as ability. In particular, I test the hypothesis that individuals observed in cities with a high human capital are better workers than individuals with the same observable characteristics who live in cities with low human capital.

This type of sorting could take place if higher college share in a city is associated with higher return to unobserved ability, causing higher quality workers to move to cities with higher college share (Borjas, Bronars & Trejo 1992, Rauch 1993). Consider a conventional Roy model where different cities reward workers' skills—both observed and unobserved—differently, and mobility decisions are based on comparative advantage. In such a model, workers are not randomly assigned to cities, but choose the city where their skills are most valued and skill–price differentials determine the skill composition of migratory flows.[44] Cities that have an industrial structure that reward education more are also likely to pay a higher price for unobserved ability. The correlation between high wages and high college share documented in previous sections may simply reflect higher unobserved ability of workers rather than higher productivity.[45] In this case, OLS estimates (section 3) are clearly upward biased.

Suppose that the wage of individual i living in city c in period t has the following form

$$\ln w_{ict} = X_{it}\beta_c + \pi P_{ct} + \alpha Z_{ct} + \epsilon_{ict} \tag{14}$$

The only innovation relative to Equations 9 and 10 is that β_c now varies across cities. The residual follows a simple error component process:

$$\epsilon_{ict} = \mu_c\theta_i + \nu_c + \nu_t + \nu_{ct} + \nu_{ict} \tag{15}$$

where θ_i is a permanent unobservable component of human capital, such as ability or family background; μ_c is a factor loading which represents the return to unobserved skill in city c; ν_c represents permanent unobserved heterogeneity across cities, which may be correlated with P_{ct} (e.g., industrial structure, physical and cultural amenities); ν_t captures the national business cycle; ν_{ct} is a transitory shocks to labor demand and supply in city c in period t, which may be correlated with P_{ct}; ν_{ict} is the transitory component of log wages which is independently and identically distributed over individuals, cities and time.

Since β_c and μ_c are assumed to vary across cities but not over time, the model in equations 14 and 15 is not the most general model of self selection. Nevertheless, it allows for rich patterns

[44]Previous literature confirms that the return to education varies significantly across states and that this variation is a major determinant of internal migration flows (Borjas et al. 1992, Dahl 1997).

[45]For example, a high-school graduate working in a biotechnology firm in San Francisco is probably different along some unobservable dimension from a high-school graduate working in a shoes factory in Miami. Similarly, a lawyer working for a Wall–Street firm in New York is likely to differ from a lawyer in El Paso, TX.

of selection based on comparative advantage. For example, workers who appear identical to the econometrician may choose different cities based on different returns to unobserved ability. This is not the case in the standard panel data model with individual fixed effects, where unobserved skills are equally valued everywhere. If ability is equally rewarded everywhere, it does not affect mobility choices. While in the present model ability contributes to *comparative* advantage, in the standard model ability contributes to *absolute* advantage. The standard individual fixed effects estimator is thus not consistent in this model since it fails to eliminate θ_i for movers.

5.1 Fixed Effects Estimates Using NLSY Data

Cross-sectional estimates of Equation 14 are inconsistent because migrants are self-selected and individual earnings are only observed for the city in which they choose to live, based on their comparative advantage. It is possible to show that (Borjas et al. 1992)

$$E[\theta_i|\text{city } c \text{ is chosen}] > E[\theta_i|\text{city } g \text{ is chosen}] \quad \text{if and only if } \mu_c > \mu_g \tag{16}$$

In this section, the problem of non-random selection is addressed by exploiting the panel structure of the National Longitudinal Survey of Youths (NLSY). By observing the same individual over time and in different cities, I can explicitly control for the effect of unobserved skill differences across cities, and permanent differences in the return to these skills. In particular, I can use the longitudinal structure of earnings to condition on a rich set of city, time, and city×individual fixed effects to absorb ν_c, ν_t and $\mu_c\theta_i$, respectively. The key identifying assumption is that the return to unobserved ability varies across cities, but does not change over time or, if it does, the change is not systematically correlated with college share.[46]

The data come from a confidential version of the NLSY that identifies the metropolitan area (MSA) of residence. The sample has longitudinal earnings and city of residence information from 1979 to 1994. I drop youths under 23 because they may still be in school, so that only individuals who are 23 to 37 years old are in the sample. Using the confidential MSA codes, I match individual–level data with data on college share in the metropolitan areas of residence from the Census. I interpolate college share in the years for which Census figures are not available. The Data Appendix provides more detailed information on the NLSY sample and the matching.

Identification comes from two sources. First, individuals change city of residence. Every year, about 7.4% of individuals in the sample change metropolitan area. It is possible to identify π by comparing the wage before and after moving. One of the advantages of the NLSY is that youths have higher mobility rate than adults, and this provides more identification. The second source of identification comes from variation in college share within a city over time. Although there is

[46]The conditions under which the general form of a Roy model of self selection is identified are derived in Heckman & Honore' (1990). Heckman and Sedlacek (1985, 1990) present general empirical models of self selection with measured and unmeasured heterogeneous skills. Chamberlain (1993) shows that a Roy model with heterogenous sorting and dynamic feedbacks is not identified.

a secular national trend toward a more educated population, there is large variation in growth rates of per capita education across cities. In Sections 3 and 4, where Census data were used, only this second source of variation was available.

When the length of the panel is fixed, fixed effects estimators are consistent if the error term ν_{ct} has mean zero conditional on θ_i, ν_c, and all the leads and lags of X_{it}, P_{ct}, Z_{ct} and ν_t (X_{it}, P_{ct}, Z_{ct} and ν_t are strictly exogenous, conditional on θ_i and ν_c). The assumption of strict exogeneity states that, conditional on both observables and ν_c, ν_t and θ_i, people who change city do so randomly. Under strict exogeneity, workers considering moving to a given city calculate expected earnings on the basis of observable time-varying variables (X_{it}, P_{ct} and Z_{ct}) and *permanent* city-characteristics, such as the return to education in the city, β_c, the return to ability, μ_c, and other permanent city characteristics that affect wages, ν_c. This assumption rules out mobility decisions based on a transitory, city-specific shock to the local labor market, ν_{ct}.[47] In the next section, I use instrumental variables and a semiparametric estimator to relax the assumption of strict exogeneity. Results do not change appreciably.

Table 10 reports estimates of the effect of college share on wages obtained from the NLSY. Other covariates include the ones used in the Census regressions as well as year dummies to control for the national business cycle. The first row reports the coefficient on college share estimated ignoring the panel structure of the data. Observations for the same individual for different years are treated as if they referred to different individuals. The cross-sectional coefficient is 0.93, and is very close to the one estimated using the 1990 Census, 0.98 (Table 2, column 2). This indicates that the external effect for young people is similar to that of whole U.S.. Standard errors in all specifications of Table 10 are adjusted for the grouped nature of the data.

Since the college share is obtained by interpolation of the Census figures, measurement error is a concern. Measurement error is the difference between the true college share in a city and the estimated one. Using another data source may indicate the extent of measurement error. I use the March Current Population Survey (CPS) to calculate college share by city for each year from 1979 to 1994. Measurement error in college share obtained from the interpolated Census data is uncorrelated with measurement error obtained from CPS data, and information from the two data sources can be combined to assess the extent of the attenuation bias. Column 3 in Table 10 reports the reliability ratio, defined as the slope coefficient from a regression of the CPS measure of college share on the Census measure, conditional on all the covariates included in the wage equation.[48] The reliability of the cross-sectional estimate in row is 0.86 suggesting that the attenuation bias is about 14%.

[47] This assumption has been used in the literature on union wage differentials with regard to the choice between union and non-union sectors (Lemieux 1998). The assumption seems more realistic in the present context. Changing cities involves higher costs than moving from the union to the non-union sector. Given the large moving costs involved, it seems plausible to assume that *permanent* city characteristics, not *transitory* shocks, play an important role in the location decision.

[48] If one were to correct for measurement error by using Census estimates as an instrumental variable for CPS estimates, (1 - reliability ratio) would indicate the importance of the attenuation bias.

Cities differ widely in geographical location, industrial structure, weather and amenities. To control for permanent heterogeneity across metropolitan areas that may cause bias in the cross section, city fixed effects are included in row 2. The estimated external return conditional on city fixed effects is 1.1. There are two important points to be made here. First, the estimate from the same specification obtained using Census data is 1.5 (Table 2, column 2), which is remarkably close to the NLSY estimate. This is reassuring, because it implies that this result is not driven by some peculiar feature of one dataset. Second, the estimate obtained by conditioning on city effects is larger than the cross sectional one, suggesting that unobserved permanent heterogeneity across cities is negatively correlated with education.[49] Again, this result is similar to that found with Census data. The reliability ratio drops to 76%, suggesting that the attenuation bias is 24%.[50]

Row 3 reports results obtained including city and individual fixed effects. In this specification, any individual permanent characteristics, such as ability or family background, is controlled for by a set of 7,279 dummies, one for each individual. The fixed effects do not absorb all variation needed for identification. First, individuals may change cities. But even for stayers, the change in college share that occurs in a city over time provides useful variation. The estimated coefficient is 1.2. The precision of the estimator does not decrease when the large number of dummies is included. The reason is that variation in the variable of interest, college share, is at the city–time level, not at the individual level. Although there are 59,826 observations in the regression, the effective sample size for identification of the coefficient on college share is only 201 times 16 (number of cities times number of years).

If the return to individual ability varies across cities, a specification that includes individual fixed effects may be misspecified. In terms of Equation 15, individual fixed effects impose the restriction that the return to ability, μ_c, is equal to one everywhere. In a more general model where the return to ability differs across cities, an individual fixed effects specification fails to absorb the error component $\mu_c\theta_i$. To control for unobserved heterogeneity at the individual-city level, individual$\times$city fixed effects are included. In this specification, everything that is specific to an individual-city pair is absorbed by the fixed effect. This is arguably the most robust model presented. Robustness, however, comes at a cost. By keeping constant the individual-city match, one source of variation is lost. Variation that comes from movers is absorbed by the fixed effect. Identification is based on stayers and comes only from changes of college share in a city over time. Conditional on a city–individual match, the model estimates what happens to the individual's wage as college share around her increases. The coefficient in row 4 is 1.3.[51]

[49] In terms of Equations 14 and 15, this means that $cov(\nu_c, P_{ct}) < 0$.

[50] City fixed effects absorb more variation in the signal than the noise, so that the signal-to-noise ratio decreases. For the same reason, the standard error on college share doubles. This result is similar to that found by Krueger & Lindahl (1998).

[51] A limitation of this model is that the return to unobserved ability in a city, μ_c, is assumed to be fixed over time. In a more general model, where the return to ability varies by city and time (μ_{ct}), estimates conditional on city$\times$individual fixed effect would not be consistent.

5.2 Accounting for Transitory Shocks: Semiparametric Correction and Instrumental Variable Estimates Using NLSY Data

Consistency of the fixed effect estimator in section 5.1 depends on the assumption that transitory factors—other than those in Z_{ct}—are not an important determinant of individuals' mobility decisions. That assumption rules out some reasonable models of geographical mobility. It is possible that the unemployment rate and other observable time-varying city characteristics do not capture *all* transitory shocks that affect workers' mobility decisions.[52] In this section, I abandon the assumption of strict exogeneity. I retain the specification with city$\times$individuals effects—arguably the most robust specification presented so far—and I augment it in order to account for unobservable transitory shocks.

I start by including in the regression estimates of transitory shocks to local labor markets. Column 4 of Table 10 shows the coefficient obtained from a specification that includes estimates of transitory demand shocks based on the Katz and Murphy index described in section 3. The coefficient is 1.4, not statistically different from the corresponding coefficient in column 2. The coefficient on the Katz and Murphy index (not in the table) is 0.083 (0.041).

This results suggests that, conditional on city$\times$individuals effects, transitory shocks may not be large enough to dramatically alter estimates. But the Katz and Murphy measure of demand shock may not capture *all* transitory shocks that affect both wages and mobility decisions. To improve the robustness of the estimates, I consider two alternative approaches. First, I consider a semiparametric correction for sample selection based on the propensity score. Second, I use use an instrumental variable strategy. The propensity score estimator is more efficient than instrumental variables, because it absorbs only the part of variation in college share that is correlated with unobserved shocks, but is less robust, because it rests on assumptions about mobility. The instrumental variable estimate is robust, but inefficient. Inefficiency comes from the fact that all variation in the independent variable of interest that is not predicted by the instrument is assumed to be endogenous and ignored. Results are remarkably stable across estimators, and similar to estimates of section 5.1.

Consider the expected value of the wage of individual i who moves from city n to city c in period t. Conditional on city$\times$individual fixed effects, the expected value given that the individual is observed working in city c is $E(\ln w_{icnt}|\text{i is in city c}) = d_t + d_{ic} + X_{it}\beta_c + \gamma P_{ct} + \alpha Z_{ct} + E(\epsilon_{icnt}|\text{i is in city c})$ where d_{ic} is a vector of city$\times$individual fixed effects that absorbs all unobserved heterogeneity at the city–individual level. The term $E(\epsilon_{icnt}|\text{i is in city c})$ is the selectivity bias due to transitory shocks. Dahl (1997) and Angrist (1995) show that under a simple conditional independence assumption, the expected value of the selectivity bias is a function h of

[52] A second argument against strict exogeneity is the presence of dynamic feedbacks. If college share in a city depends on past unobservable shocks, fixed effect estimates are known to be inconsistent in short panels.

the probability that city c is chosen:

$$E(\epsilon_{icnt}|\text{i is in city c}) = h(\pi_{icnt}) \tag{17}$$

where π_{icnt} is the propensity score, defined as the probability that the individual moves from city n to city c in period t.

In practice, in order to estimate $h(\pi_{icnt})$ some restrictions are needed. I assume that individuals with the same level of education have the same propensity score π_{jcnt}, where j indexes education groups.[53] The function $h(\cdot)$ is usually approximated by a polynomial of degree three: $h(\hat{\pi}_{jcnt}) = h_1\hat{\pi}_{jcnt} + h_2\hat{\pi}^2_{jcnt} + h_3\hat{\pi}^3_{jcnt}$. The estimate conditional on the propensity score is shown in column 6 of Table 10. Identification relies on the restriction that wages depend only on the city of residence, while migration decisions are a function of both the city of residence at t-1 and potential city of residence at time t (See Dahl (1997) for discussion of this assumption). The estimated coefficients are close to the IV estimates, although the precision increases.[54] An F-test indicates that the coefficients on the polynomial in the propensity score are jointly significant (p-value is 0.014). The fact that the propensity score enters significantly but does not affect the coefficient on college share, suggests that unobserved transitory shocks play a role in mobility decisions, but they are mostly orthogonal to college share.

I now turn to instrumental variable estimation. Instrumental variable estimates are more robust than estimates obtained by conditioning on the Katz and Murphy index or the propensity score. A second advantage is that, unlike OLS estimates, IV estimates eliminate attenuation bias due to measurement error. I use the city demographic structure in 1970 as instrument.[55] The instrumental variable estimate in column 5 of Table 10 is 1.3. This estimate is similar to the corresponding OLS estimates but less precisely estimated, because of the small sample size. A Hausman test suggests that OLS and IV estimates are not statistically different.[56]

In summary, I have used a panel data model with comparative advantage and non–random selection to investigate the hypothesis that workers in cities with more college graduates have more unobserved ability. To implement the model, I exploited the longitudinal nature of the NLSY. By observing the same individual over time and in different cities, I was able to control for *permanent* factors that make an individual–city match particularly productive. The results

[53]Estimating the probability of moving from city to city is not feasible in the NLSY because of the very large number of possible combinations. I estimate the propensity score for each pair of cities, c and n, belonging to state s and p, respectively, as the fraction of workers in each education group who move from p to s in any given year. A different propensity score is estimated for each pair of states, education level and year.

[54]OLS standard errors are reported. Since the propensity score is estimated, the correct covariance matrix should be adjusted to account for the extra sampling variability of the propensity score.

[55]Since this instruments predicts changes in college share, not levels, I select all individuals with non–missing values in 1986, 1987, 1993 and 1994 (taking years further apart would have reduced the sample significantly). I estimate a wage equation with city×individual fixed effects, on this sample of 15,676 observations.

[56]If shocks are serially correlated, the exogeneity of the instrument may be at risk. The standard approach in the literature has been to take variables 2 or 3 years lagged as instrument, with the assumption that the shock has disappeared after such a lag. By using 1970 age structure as instrumental variable for the change in college share between 1986 and 1994, I am using a 16 years lagged instrument.

obtained from the NLSY panel are remarkably consistent with those based on repeated cross-sections from the Census, suggesting that unobserved individual ability does not play a major role in explaining the relationship between wages and college share.

6 Externality or Complementarity? Differences in the External Return Across Education Groups Using Census Data

Most of the analysis so far has focused on the issue of omitted variables. Findings, obtained by pooling all education groups together, show that increases in the share of college graduates in a city raise *average* wages. This result is in itself interesting, and is has several important policy implications.[57] Assuming that the correlation between average wages and college share is not explained by omitted variable, is it necessarily explained by externalities?

The answer is negative. In this section I first show that if educated and uneducated workers are complements (imperfect substitutes), finding that *average* wages are affected by the percentage of college graduates in the labor force does not necessarily indicate an externality effect: rather this finding may indicate complementarity between high and low education workers. I then present evidence that helps distinguish externalities from complementarity.

If high and low education workers are complements (imperfect substitutes)—as it is usually found in most studies (see, for example, Katz & Murphy (1992))— then low education workers benefit from the increases in the number of high education workers even in the absence of externalities. In the same way that increasing physical capital raises workers' productivity, increasing human capital increases low education workers' productivity. Consider a simplified version of the model of Section 2, where there are only two education groups, high education workers (N_1) and low education workers (N_2). Like before, the production function is $y = f(\theta_1 N_1, \theta_2 N_2, K)$, where N_j is the number of workers in group j and $\theta_j = b_j + \gamma(\frac{N_1}{N_1+N_2})$ is a productivity shifter that depends on a group specific effect, b_j, and the externality (see equation 3). The externality is a function of the share of high education workers. When $\gamma = 0$ the model becomes the standard model of wage determination without externalities. If we assume that wages are equal to the marginal product of each type of labor and that, for simplicity, the spillover is external to individual firms in the city but internal to the city as a whole, wages are

$$w_j = \theta_j \frac{\delta f}{\delta L_j} \tag{18}$$

Consider what happens to the wages of high education workers (w_1) and low education workers

[57]For example, from the point of view of a city, knowing the *average* effect of college share on wages may affect the type of residents the city wants to attract.

(w_2) when the number of high education workers increases in the city:

$$\frac{dw_1}{dN_1} = \theta_1 b_1 f''_{11} + \frac{\gamma}{(N_1+N_2)^2} q_1 \tag{19}$$

$$\frac{dw_2}{dN_1} = \theta_2 b_1 f''_{21} + \frac{\gamma}{(N_1+N_2)^2} q_2 \tag{20}$$

where $q_1 = N_2 f'_1 + \theta_1(N_2^2 f''_{12} + (N_1^2 + 2N_1N_2)f''_{11}) > 0$ and $q_2 = N_2 f'_2 + \theta_2(N_2^2 f''_{22} + (N_1^2 + 2N_1N_2)f''_{21}) > 0$. The wage of uneducated workers benefits for two reasons. First, an increase in the number of educated workers raises uneducated workers' productivity because of complementarity: $\theta_2 b_1 f''_{21} > 0$. Second, the externality further raises their productivity: $\frac{\gamma}{(N_1+N_2)^2} q_1 > 0$. The impact of an increase in the supply of educated workers on their own wage is determined by two competing forces: the first is the conventional supply effect which makes the economy move along a downward sloping demand curve: $\theta_1 b_1 f''_{11} < 0$. The second is the externality that raises productivity: $\frac{\gamma}{(N_1+N_2)^2} q_2 > 0$.[58] The important feature of equation 19 and 20 is that unskilled workers benefit from an increase in the share of educated workers in the city even if $\gamma = 0$, i.e. in the absence of any externalities. The effect on the wage of skilled workers, however, depends on the magnitude of the externality. If γ is large enough, the effect for skilled workers should be positive although lower than the one for unskilled workers. If $\gamma = 0$, the effect should be negative.

The effect on the average wage in a city is a weighted average of the effect on educated and uneducated workers' wages:

$$\frac{d\overline{w}}{dN_1} \approx \omega_1 \frac{dw_1}{dN_1} + \omega_2 \frac{dw_2}{dN_1} \tag{21}$$

where the weights are the fraction of educated (ω_1) and uneducated (ω_2) workers in the labor force. In the absence of externalities, $\frac{dw_2}{dN_1}$ is positive and $\frac{dw_1}{dN_1}$ is negative. It is possible that a rise in the number of educated workers in a city increases the average wage even in the absence of any externalities.

In this section, more direct evidence of the externality is provided by separately identifying the effect of college share on wages of different skills groups. There are four education groups: less than high-school, high-school, some college and college or more. The following equation is estimated for each education group:

$$\ln \hat{\alpha}_{jct} = d_{jt} + d_{jc} + \pi_j P_{ct} + \alpha_j Z_{ct} + \epsilon_{jct} \tag{22}$$

where $\hat{\alpha}_{jct}$ is the regression–adjusted mean wage of education group j in city c at time t (conditional on the individual's education and all the other individual characteristics X_{it}); and the d's are dummies that capture any effects that are group–time– or group–city–specific. Equation 22 is analogous to Equation 7 in the theoretical section.[59]

[58] Acemoglu (1998) and Ciccone, Peri & Almond (1999) make a similar point.

[59] The estimation procedure is similar to the one described in section 3. First, regression adjusted mean wage, α_{jct}, is

Equations 19 and 20 suggest that, irrespective of the magnitude of the externality, the coefficient on college share, π_j, should be unambiguously positive for unskilled workers. If the externality is strong enough, the coefficient for skilled workers should be positive although lower than the one for unskilled workers. If the standard supply effect is stronger than the externality, the coefficient for skilled workers should be negative. By looking only at the effect on wages of uneducated workers, one cannot separately identify complementarity from the externality. Only by looking at the effect of an increase in the supply of college graduates on their own wage, can one directly test for the presence of externalities.

The units of observation are city-education group cells. Since education-city cells are small in the NLSY, results in this section are based on Census data. Table 11 presents estimates of π_j, the effect of changes in college share on changes in each group regression-adjusted mean wage. Instrumental variable coefficients obtained using the 1970 demographic structure (column 2) confirm that the coefficient is larger for less educated groups, as predicted by a conventional demand and supply model. But even for college graduates, the effect of an increase in college share is positive, as predicted by a model that includes both conventional demand and supply factors and externalities. When cost of tuition (column 3) is used as an instrument, the inverse relationship between the coefficient on college share and the group education level is even more clear. Both instruments are used to obtain results in column 4. According to these estimates, a 1% increase in the share of college educated workers raises high-school drop-outs wages by 2.3%, high-school graduates wages by 1.4%, wages of workers with some college by 1.1% and those of college graduates by 1.2%. Over-identifying restrictions are not rejected, suggesting that the two instruments yield the same estimate (column 5).

To aid in the interpretation of the magnitude of the coefficients, consider an average worker earning \$25,000 a year, in a city like Bakersfield, CA, or Lancaster, PA, which experienced increases in the share of college graduates between 1980 and 1990 close to the median. The predicted increase in earnings caused by increases in college share is \$118 a year for an high–school drop-out, \$73 for an high–school graduate and \$60 for a college graduate (column 2 of table 12). Predicted earning increases for city with large and small increases in college share are shown in column 1 and 3, respectively.

This evidence is consistent with the prediction of the model presented in section 2, which includes both conventional demand and supply factors, and externalities. 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, which raises productivity. Even for college graduates, the external effect seems to be large enough

estimated for each group and city. This regression is done separately for 1980 and 1990. Second, each group's adjusted mean wage is regressed on college share, controlling for city characteristics, Z_{ct} and fixed effects (equation 22). This second stage estimation is performed separately for each education group. Second stage models are weighted by city sample size.

to generate a positive wage gain in better-educated cities. This finding implies that the existence of human capital externalities cannot be rejected. Standard demand and supply theory would predict that in the absence of externality, an increase in the supply of college educated workers would lower their wage.

A limitation of the framework adopted here is the implicit assumption that workers in the same education group are perfect substitutes. If educated workers are imperfect substitutes, then the effect of the changes in percentage of college graduates on wages of college graduates would not necessarily represent an externality, but could in part be explained by complementarity.

7 Structural Estimation Using Census Data

In this section I separately identify the effect of changes in college share and the effect of changes in the supply of different education groups. The effect of an increase in the proportion of college graduates on their own wage is plotted in figure 3. 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 level of human capital in the city has increased, the marginal product curve shifts from D1 to D2 as a result of the externality.

A test of the reliability of the model adopted in the paper may be obtained by separately identifying the effect of an increase in supply (1 to 2 in Figure 3) from the externality effect (2 to 3). The first one is expected to be negative and the second one to be positive. The following equation is estimated

$$\ln \hat{\alpha}_{jct} = d_j + d_c + d_t + d_{jt} + d_{jc} - \lambda f_{jct} + \delta P_{ct} + \alpha Z_{ct} + \epsilon_{jct} \tag{23}$$

where f_{jct} is the logarithm of the proportion of the labor force of city c in skill group j. Equation 23 allows separate identification of supply effect, λ, and the externality, δ. Equation 23 can be derived from Equation 2.[60]

The OLS estimates' of the coefficients λ and δ are shown in Table 13.[61] Separate estimates by education groups are not shown, as there is not enough variation to separately identify different coefficients for different groups. Robust standard errors that allow for correlated residuals among education groups in the same city are reported. Results are consistent with a model that includes standard demand and supply considerations, as well as externalities. Increases in the supply of

[60]Let N_{ct} be the number of workers in city c at time t. Adding and subtracting $(1/\sigma)\ln N_{ct}$ from Equation 2, marginal productivity of group j in city c at time t can be rewritten as

$$\ln w_{jct} = \mu'_{ct} - \frac{1}{\sigma}\ln\frac{N_{jct}}{N_{ct}} + \frac{\sigma-1}{\sigma}\ln\theta_{jct} \tag{24}$$

where $\mu'_{ct} = \mu_{ct} + (1/\sigma)\ln N_{ct}$ is a new city-time specific component. Equation 23 is the reduced form wage equation, where $\lambda = 1/\sigma$ and $\delta = ((\sigma-1)\gamma)/\sigma$.

[61]Instrumental variable estimates are not reported. IV estimates would require two separate instruments. The first instrument should predict changes in college share, the second one should predict changes in the share of other education groups. Although age structure and cost of tuition are available, the first stage regression is so weak that IV estimates are too noisy obe informative.

a given education group, keeping college share constant, decreases that group's wages. This is the standard supply effect: increases in a factor supply must reduce its price. Increases in college share, keeping the group's supply fixed, increase the group's wage, as predicted by a model that includes externalities.

8 Conclusion

This paper's main contribution is showing evidence that workers 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 for identifying and measuring the external return to education. Identifying the nature and causes of the externality, although important for policy implications, is beyond the scope of this paper.

The argument is divided in three parts. In the first part, the average effect across education groups of college share on wages is estimated using repeated cross-sections from the Census. I investigate the possibility that there are city–wide labor demand shocks that increase wages in a city and attract skilled workers there. I account for this possibility by both directly estimating these shocks with an index proposed by Katz and Murphy, and, indirectly, by instrumental variables techniques. Although the instruments are very different from each other, estimates are remarkably stable.

In the second part of the paper, I consider the possibility that cities with a more educated labor force also have higher levels of unobserved ability. I use a special version of the National Longitudinal Survey of Youth to control for a rich set of city×individual fixed effects, that absorb all permanent unobserved factors that make an individual–city match particularly productive. Conditional on the city–individual match, the wage of individuals rises as the share of college educated workers in the labor force increases. The coefficient on college share obtained from this specification using NLSY data is remarkably similar to the coefficient obtained from first–differenced IV using Census data. The most robust estimates are between 1% and 1.3%.

The last part of the paper estimates the effect of an increase in the percent of college graduates on the wages of four education groups. The coefficient on the percent of college graduates is larger for less educated groups, as predicted by a conventional demand and supply model. But even for college graduates, an increase in the percent of college graduates increases wages, as predicted by a model that includes both conventional demand and supply factors and externalities. A one percent increase in the proportion of college educated workers raises the wage of high-school drop-outs, high–school graduates, workers with some college, and college graduates by 2.3%, 1.4%, 1.1% and 1.2%, respectively.

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 eighth lowest level of average education in the US, experienced virtually no increase in the proportion of college graduates in the 1980's. The latter, with one of the highest levels of average education, witnessed a 5.1% increase in the proportion of college graduates. Findings in this paper suggest that in San Jose externalities from education may have accounted for wage increases among high-school graduates and college graduates of 6.1% and 5.1%, respectively. These increases occurred over the ten year period between 1980 and 1990. (Note however that the increase in the percent of college graduates experienced by San Jose is not typical at all, being more than twice as large as the median increase.) No wage increase was caused by human capital externalities in El Paso.[62]

The social return to education, which is the sum of private and external return, may be larger than previously thought. This finding is consistent with recent estimates by Krueger & Lindahl (1998). Using macro economic data on aggregate income and education for 68 countries between 1965 and 1985, and controlling for measurement error, Krueger & Lindahl (1998) find that a one year increase in average education increases GDP above the microeconomic estimates of the private return to education. They conclude that such large social return to education found in the cross-country models can be explained either by omitted variables problems or nationwide externalities, but that without a natural experiment similar to those that have been exploited in the literature on private returns to education they cannot separate the two alternative explanations.

While this paper does not identify a natural experiment, it uses two different individual–level datasets to allow for rich models of non–random selection, and thereby improves the existing estimates of the external return to education.

References

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

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

Acemoglu, D. & Angrist, J. (1999), 'How large are human capital externalities? evidence from compulsory schooling laws', *MIT – mimeo* .

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.

Angrist, J. (1995), 'Conditioning on the probability of selection to control selection bias', *NBER Technical Working Paper* .

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

[62]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): 'These companies say if you can give us a labor base with adequate skills, we would make you first choice to locate'. As a consequence, 'business leaders this year identified retraining workers as by far the region's top need'.

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

Arrow, K. (1962), Economic welfare and the allocation of resources for invention, *in* R. Nelson, ed., 'The Rate and Direction of Inventive Activity', Princeton University Press.

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

Bartik, T. (1991), *Who Benefits from State and Local Economic Development Policies*, W.E. Upjohn Institute.

Beeson, P. & Eberts, R. (1989), 'Identifying productivity and amenity effects in interurban wage differentials', *Review of Economics and Statistics* **71**, 443–452.

Benhabib, J. & Spiegel, M. (1994), 'The role of human capital in economic development: Evidence from aggregate cross-country data', *Journal of Monetary Economics* **34**, 143–174.

Blomquist, G. C., Berger, M. C. & Hoehn, J. P. (1988), 'New estimates of quality of life in urban areas', *American Economic Review* **78**(1), 89–106.

Borjas, G. J. (1995), 'Ethnicity, 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**.

Borjas, G. J., Bronars, S. G. & Trejo, S. (1992), 'Self-selection and internal migration in the united states', *Journal of Urban Economics* **32**, 159–185.

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

Chamberlain, G. (1993), 'Feedback in panel data models', *Manuscript* .

Ciccone, A., Peri, G. & Almond, D. (1999), 'Capital, wages and growth: Theory and evidence', *Manuscript* .

Dahl, G. B. (1997), 'Mobility and the return to education: Testing a roy model with multiple markets', *Mimeo - Princeton University* .

Edmond, J. B. (1978), *The Magnificent Charter: The Origin and Role of the Morrill Land–Grant Colleges and Universities*, Exposition Press.

Friedman, M. (1962), *Capitalism and Freedom*, University of Chicago Press.

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

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

Glaeser, E. L., Scheinkman, J. A. & Shleifer, A. (1995), 'Economic growth in a cross-section of cities', *Journal of Monetary Economics* **36**(1).

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

Griliches, Z. (1979), 'Issues in assessing the contribution of r&d to productivity growth', *Bell Journal of Economics* **10**, 92–116.

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

Heckman, J. J. (1999), 'Policies to foster human capital', *Manuscript – University of Chicago* .

Heckman, J. J. & Honore', B. E. (1990), 'The empirical content of the roy model', *Econometrica* **58**(5), 1121–1149.

Heckman, J. J. & Sedlacek, G. L. (1985), 'Heterogeneity, aggregation, and market wage functions: An empirical model of self-selection in the labor market', *Journal of Political Economy* **93**(6), 1077–1125.

Heckman, J. J. & Sedlacek, G. L. (1990), 'Self-selection and the distribution of hourly wages', *Journal of Labor Economics* **8**(1), S329–63.

Jaeger, D. A. (1997), 'Reconciling the old and new census bureau education questions: Recomendation for researchers', *Journal of Business and Economics Statistics* **15**(2), 300–308.

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.

Kane, T. J. (1994), 'College entry by blacks since 1970: the role of college costs, family background, and the returns to education', *Journal of Political Economy* **102**(5), 878–911.

Kane, T. J. (1995), 'Rising public college tuition and college entry: How well do public subsidies promote access to college?', *NBER Working Papers* **5163**.

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

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

Krueger, A. B. & Lindahl, M. (1998), 'Education for growth in sweden and the world', *Industrial Relation Section Working Paper, Princeton* .

Lemieux, T. (1998), 'Estimating the effects of unions on wage inequality in a panel data model with comparative advantage and nonrandom selection', *Journal of Labor Economics* **16**(2), 261–291.

Leslie, L. L. & Brinkman, P. T. (1988), *The Economic Value of Higher Education*, MacMillan.

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

MaCurdy, T., Mroz, T. & Gritz, R. M. (1998), 'An evaluation of the national longitudinal survey of youth', *The Journal of Human Resources* **33**(2), 345–436.

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

Nervis, A. (1962), *The State Universities and Democracy*, University of Illinois Press.

Rauch, J. E. (1993), 'Productivity gains from geographic concentration of human capital: Evidence from the cities', *Journal of Urban Economics* **34**, 380–400.

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

Roback, J. (1988), 'Wages, rents and amenities: Differences among workers and regions', *Economic Inquiry* **26**(1), 23–41.

Tewksbury, D. G. (1932), *The Founding of American Colleges and Universities Before the Civil War*, Columbia University.

Topel, R. (1999), Labor market and economic growth, *in* O. Ashenfelter & D. Card, eds, 'The Handbook of Labor Economics', North-Holland, Amsterdam.

Williams, R. L. (1991), *The Origins of Federal Support for Higher Education*, The Pennsylvania State University Press.

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

Appendix 1: Equilibrium with Externalities

In this appendix I use a standard general equilibrium framework proposed by Roback (1982, 1988) to show that an equilibrium exists when externalities are present. Roback's framework is often used to model worker and firm location decisions, with or without externalities (Rauch 1993, Beeson & Eberts 1989, Blomquist, Berger & Hoehn 1988). The model is a special case of the theoretical framework presented in section 2, when there are only two cities, two education groups, one period. As in section 2 there are two types of goods, a composite good y —nationally traded— and land h —locally traded. Each city is a competitive economy that produces y with a constant return to scale technology. Production requires inputs of the two types of labor—high education and low education workers—and land. For simplicity I ignore capital. An important assumption is that the two types of labor are imperfect substitutes (see, for example, Katz & Murphy (1992)). Productivity of workers in either education group depends on two factors: (i) their own human capital; (ii) the share of educated workers in the city labor force. If only the first factor is considered, the model reduces to a standard two–sector model of wage determination.

Workers maximize utility subject to a budget constraint by choosing quantities of the composite good and residential land, given the city amenity, v'_c. Workers and firms are perfectly mobile. Equilibrium is obtained when workers have equal utilities in all cities and firms have equal unit cost across cities. Because the composite good, y, is traded, its price is the same everywhere and set to 1. Variation in the cost of living depends only on variation in cost of land, p, which is the same for all workers in the same city, irrespective of the education group.

The equilibrium for the simple case of only two cities, A and B, is described in Figure A1. The upward sloping lines in each panel represent indifference curves for the two education groups. Indirect utility of workers belonging to group j, $V_j(w_j, p, v')$, is a function of the group's nominal wage, w_j, cost of land and the amenity. The indifference curves are upward sloping because workers prefer higher wages and lower rent. Since workers are free to migrate, utility of workers is equalized across locations: $V_1(w_1, p, v') = k_1$ and $V_0(w_0, p, v') = k_0$ for educated and uneducated workers, respectively. The downward sloping lines show combinations of wages and rents which hold constant firms' unit costs: $C_c(w_0, w_1, p) = 1$, where w_0 and w_1 are wages of uneducated and educated workers, respectively; and c indexes city.. A zero–profit condition for the firm assures that production must take place along the downward sloping curve. Thus the model has three equations (unit cost and indirect utility for each skill group) in three unknowns (w_0, w_1 and p).[63]

Point 1 in the left panel of Figure A1 represents the equilibrium combination of wage of educated workers and cost of land in city A. Point 1 in the right panel represents the same combination for uneducated workers. Rent is the same for both skill groups. If the two cities are identical, the equilibrium in city B is the same. Consider what happens to equilibrium wages

[63]Note that in section 2 I allow for heterogeneity in amenities (heterogeneity in labor supply) and productivity (heterogeneity in labor demand) across cities. Here, for simplicity, I allow heterogeneity in amenities, but I ignore heterogeneity in productivity. This assumption is easy to generalize.

when the relative supply of educated workers is higher in B than in A. Suppose for example that city B has an higher level of the local amenity than city A ($v'_B > v'_A$) and educated workers value the amenity, while uneducated workers don't. I interpret v' broadly, as any exogenous factor that increases the relative supply of educated workers. The indifference curve at level k_1 of educated workers in city B is to the left of the corresponding curve in A, while the indifference curve for uneducated workers does not change. If there are no externalities, the increase in the supply of educated workers in city B raises the wage of uneducated workers to w'_0 and lowers the wage of educated ones to w'_1 (point 2 in both panels of Figure A1). This is the standard result. Because of complementarity, uneducated workers are now more productive in city B. Because of the amenity, educated workers accept lower wages there.

If there are externalities, however, the combinations of wages and rents that hold firms' costs constant in city B is to the right of the corresponding combination in city A for both groups (point 3). For educated workers, the shift of the isocost curve is caused by the externality only; for uneducated workers the shift is caused by both complementarity (movement from 1 to 2) and the externality (movement from 2 to 3). In equilibrium, both skill groups are present in both cities.[64]

Data Appendix

Census: 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 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. 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 or SMSA 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 are assigned a MSA code on the basis of the County Group Equivalency files.[65] If over 50% of the PUMA population is attributable to a single MSA, I then assign all individuals in that PUMA to the majority MSA. The computer code for this assignment is available on request.

[64] Since workers are free to migrate from city A to city B, why are equilibrium wages—net of the compensating differential—not driven to equality? In this model, migration to high wage cities leads to higher rent, making workers indifferent between cities. Other models achieve the same result assuming that quality of life is declining in the size of the city (Glaeser et al. 1995). Whatever the equilibrating mechanism is, higher nominal wages in a city imply greater productivity. If workers weren't more productive, firms would leave high–wage cities and relocate to low wage cities.

[65] 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.

Since the MSA definition was changed after the 1980 Census, I redefine 1990 SMSAs to match the 1980 boundaries. The County Group Equivalency files are used to identify PUMAs that contain the affected counties in the 1990 Census. If the counties in question comprise more than half of the PUMA's population, all respondents are assigned to the pertinent MSA. If greater than 10% of a MSA's 1990 population is affected by the boundary changes and is unrecoverable from the County Equivalency files, I drop 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 are used to estimate the age structure. Only individuals in the labor force are used in the estimation of the 1980 and 1990 wage equations. Individual are coded as employed if they report positive earnings, including wage and salary and/or self-employment earnings, and positive weeks of work and positive usual hours per week. Workers employed in agriculture or in the military are excluded. Wages rates less than $1.00 per hour, or greater than $400 per hour, are 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 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 needed for one of the instrumental variables. 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 is matched to the 1980 MSA code using information in the Census Bureau publication Geographic Identification Code Scheme (1983, 11-17). I follow the same procedure used by Altonji & Card (1991). The computer code used for the matching is available on request. Of the 282 MSAs identified in 1980 and 1990, only 115 are identified in 1970.

National Longitudinal Survey of Youths: A special, confidential, version of the NLSY, where MSA of residence is identified, is used. I thank the Bureau of labor Statistics for making the version available. Using the MSA code, individual–level data are matched to city–level college share from the Censuses. I restrict the sample to the 9763 young men and women included in the 'cross-sectional sample' and the supplemental samples of Hispanics and Blacks. Following MaCurdy, Mroz & Gritz (1998) I exclude the economically disadvantaged Whites—discontinued in 1991—and the military sample —discontinued in 1983. I also exclude individuals who are 22 or younger. The maximum age is 37. The dataset follows individuals from 1979 to 1994. The panel is unbalanced. There are a total of 59826 observations with non–missing values for all the relevant variables. The wage definition used is 'hourly rate of pay in the current/most recent job'. Not all MSAs identified in the Census are in the NLSY sample. The main discrepancy is that, due to the fact that New England County Metropolitan Area are not comparable to MSA from the remainder of the country, most New England observations in the NLSY have their

MSA identifier missing, and they are not included in the analysis. Data on the share of college educated workers are available from the Censuses only for 1980 and 1990. I therefore interpolate the Census estimates for all years from 1979 to 1994. An alternative would have been to use estimates obtained yearly from the Current Population Survey. Given the smaller sample size of the CPS, results obtained by interpolating Census estimates turned out to be more precise than results obtained from CPS estimates.[66]

[66]The number of observations per city in the Census ranges from 935 to 99,371. The number of observations per city in the CPS ranges from 36 to 4950. A regression of college share in 201 U.S. cities in 1990 obtained from the CPS against the one obtained from the Census yields a slope equal 1.06 (0.06), not significantly different from one. The intercept is -1.48 (0.90), suggesting that there may be a systematic difference in measurement of education in the two surveys. The regression of Census estimates on CPS estimates yields an intercept equal 6.98 (0.39)and a slope equal 0.48 (0.03).

Table 1: Cities with the Highest and Lowest Percentage of College Graduates in 1990: Census Data

Rank		Percentage with College (1)	Regression–Adjusted Average Wage (2)
	Top		
1	STAMFORD, CT	0.497	1.10515
2	NORWALK, CT	0.446	1.08819
3	ANN ARBOR, MI	0.411	0.77543
4	WASHINGTON, DC-MD-VA	0.399	0.91655
5	BOSTON, MA	0.379	0.89903
6	RALEIGH-DURHAM, NC	0.364	0.71193
7	GAINESVILLE, FL	0.361	0.56829
8	SAN FRANCISCO, CA	0.357	0.90769
9	TALLAHASSEE, FL	0.353	0.61753
10	TRENTON, NJ	0.348	0.89106
11	SAN JOSE, CA	0.340	0.95537
12	HARTFORD, CT	0.325	0.92216
	Bottom		
270	VINELAND-MILLVILLE-BRIDGETON, NJ	0.126	0.77120
271	FALL RIVER, MA-RI	0.122	0.84395
272	ALTOONA, PA	0.122	0.56414
273	LIMA, OH	0.120	0.61863
274	FORT SMITH, AR-OK	0.119	0.48394
275	DANVILLE, VA	0.118	0.59524
276	ODESSA, TX	0.118	0.59739
277	JACKSONVILLE, NC	0.116	0.45634
278	HAGERSTOWN, MD	0.114	0.65215
279	GADSDEN, AL	0.113	0.55382
280	WILLIAMSPORT, PA	0.112	0.57051
281	HICKORY, NC	0.111	0.59252
282	STEUBENVILLE-WEIRTON, OH-WV	0.102	0.55030

NOTES: Regression adjusted average wage in column 2 is the average log hourly wage in each city after controlling for individual education, race, Hispanic origin, sex, work experience, and U.S. citizenship.

Table 2: The Effect of Changes in the Share of College Educated Workers on Wages: Census Data

	OLS		2SLS				
			Land Grant	Cost of tuition	Age Structure in 1970	Age Structure in 1980	2 IVs
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Cross–section - 1990							
(1) Coeff. of College Share	0.982		0.737				
	(0.070)		(0.230)				
Cross–section - 1980							
(2) Coeff. of College Share	0.785		0.582				
	(0.070)		(0.169)				
City fixed effects 1980–1990							
(3) Coeff. of College Share	1.547	1.453		1.666	1.079	2.167	1.391
	(0.137)	(0.150)		(0.446)	(0.614)	(0.500)	(0.537)
Hausman Test [p-value]			0.24	0.98	0.29	0.22	0.95
Overid. Test [p-value]							0.46
Katz-Murphy Index	no	yes	no	no	no	no	no
Cities	282	282	282	282	115	282	115

NOTES: Each entry is the coefficient of the percentage of college graduates in a separate regression. All models include individual characteristics (education, sex, race, Hispanic origin, US citizenship, and a quadratic in experience) and city characteristics (unemployment rate, percentage of Hispanics, females, US citizens). Eight regional dummies are also included in the cross-sectional specifications. The interaction South×year is included in the panel specification. In column 8, 1970 age structure and cost of tuition are used as instruments. The over-identifying restrictions are not rejected (test is 0.76; p-value 0.46). Hausman test refers to the hypothesis that IV estimates are different from the corresponding OLS estimates. The Hausman test in column 3 refers to 1990 cross-section. The corresponding p-value for 1980 is 0.16 . The coefficients obtained estimating specifications in columns 1 and 6 on the sample of 115 cities identified in the 1970 Census are 1.049 (0.108) (column 1, row 1); 0.938 (0.106) (column1, row 2); 1.664 (0.218) (column 1, row 3); 1.792 (0.615) (column 6). Standard errors corrected for clustering in parentheses.

Table 3: First–Stage Regressions. The Effect of the Instrumental Variables on the Distribution of Education in City: Census Data

	College + (1)	Some College (2)	High–School (3)	High–School Drop Out (4)
Cross–Section - 1990				
(1) Land–Grant	0.049	0.002	-0.021	-0.031
	(0.008)	(0.004)	(0.006)	(0.005)
Cross–section - 1980				
(2) Land–Grant	0.053	-0.001	-0.018	-0.033
	(0.007)	(0.004)	(0.006)	(0.006)
City Fixed Effects 1980–1990				
(3) Cost of Tuition	-0.019	0.001	-0.008	0.026
	(0.004)	(0.003)	(0.004)	(0.004)
(4) 1970 Age structure	0.287	-0.034	0.028	-0.281
	(0.070)	(0.049)	(0.064)	(0.075)
(5) 1980 Age Structure	0.228	-0.015	0.050	-0.263
	(0.046)	(0.033)	(0.044)	(0.043)

NOTES: Each entry is a separate regression. The dependent variable is the percentage of individuals in city belonging to one of the four education groups. Other wage equation regressors are included. R^2 in row 1: 0.37, 0.64, 0.62, 0.57. R^2 in row 2: 0.38, 0.70, 0.64, 0.49. R^2 in row 3: 0.19, 0.52, 0.28, 0.63. R^2 in row 4: 0.21, 0.66, 0.28, 0.63. R^2 in row 5: 0.18, 0.52, 0.28, 0.61. Standard errors corrected for clustering in parenthesis.

Table 4: Correlation Between 1980 and 1970 Demographic Structure and Changes in Population, Labor Force, Inflows of Domestic Workers, Inflow of Foreign Workers: Census Data

	Change Populat. (1)	Change Labor Force (2)	Change Participat. Rate (3)	Net Domestic Inflow (4)	Internat. Inflow (5)
1970 Age Structure					
Share of Young, 1970	0	0	0	0	-
Share of Middle-Aged, 1970	0	0	0	0	+
Share of Old, 1970	0	0	0	0	0
1980 Age Structure					
Share of Young, 1980	+	-	0	-	+
Share of Middle-Aged, 1980	0	-	0	0	0
Share of Old, 1980	0	+	0	+	-
City Effects	yes	yes	yes	yes	yes

NOTES: Each entry is a separate regression. 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 4 columns are obtained similarly. Other wage equation regressors are included. There are 115 cities in the lower panel and 282 cities in the upper panel. Young individuals are aged 16-28; middle-age ones 29-63 and old ones 64-70.

Table 5: The Effect of Changes in Share of College Educated Workers on Changes in Wages in Manufacturing: Census Data

	(1)	(2)	(3)	(4)	(5)	(6)
Coeff. of City College Share	1.661					
	(0.164)					
Coeff. of Manufacturing College Share		0.838	0.655	0.637	0.484	0.557
		(0.097)	(0.100)	(0.098)	(0.091)	(0.086)
Coeff. of Other Industries College Share			0.758		0.469	
			(0.151)		(0.139)	
Coeff. of Business Services College Share				0.214		0.088
				(0.100)		(0.086)
City Effects	yes	yes	yes	yes	yes	yes
Region×Year Effects	no	no	no	no	yes	yes
Cities	282	282	282	282	282	282

NOTES: Each column is a separate regression. All models include individual characteristics (education, sex, race, Hispanic origin, US citizenship, and a quadratic in experience) and city characteristics (unemployment rate, proportion of Hispanics, females, US citizens). The interaction South×year is also included. OLS coefficients are reported. Standard errors corrected for clustering in parentheses.

Table 6: The Effect of Changes in Share of College Graduates on Wages Controlling for Cost of Housing: Census Data

	OLS	2SLS		
		Age Structure in 1970	Cost of tuition	Both IVs
	(1)	(2)	(3)	(4)
Coeff. of college share	1.506	1.100	1.368	1.246
	(0.137)	(0.603)	(0.402)	(0.565)
Coeff. of log rent	0.052	0.047	0.055	0.043
	(0.022)	(0.037)	(0.025)	(0.036)
City Effects	yes	yes	yes	yes
Katz–Murphy Index	no	no	no	no
Cities	282	115	282	115

NOTES: Each column is a separate regression. All models include individual characteristics (education, sex, race, Hispanic origin, US citizenship, and a quadratic in experience) and city characteristics (unemployment rate, proportion of Hispanics, females, US citizens). The interaction South×year is also included. In column 4, the over-identifying restrictions are not rejected (test is 0.59; p-value 0.44). Standard errors corrected for clustering in parentheses.

Table 7: Robustness Checks: Census Data

	2SLS	
	1970 Age Structure (1)	Cost of Tuition (2)
(1) Basic specification	1.079	1.666
	(0.614)	(0.446)
(2) Coeff. on individ. charact. vary by city	2.467	2.921
	(1.788)	(1.541)
(3) Industry dummies	1.185	1.595
	(0.579)	(1.425)
(4) Occupation dummies	1.174	1.690
	(0.601)	(0.442)
(5) College town dummy	1.348	2.463
	(0.537)	(0.570)
(6) MSA of Work	1.136	1.613
	(0.628)	(0.484)
(7) 8 regional dummies	-0.545	-4.727
	(0.930)	(4.360)
(8) Spatial autocorrelation	1.079	1.666
	(0.732)	(0.554)
Cities	115	282

NOTES: Each entry is a separate regression. Entries are the coefficients of city college share.
(1) The base case is taken from column 5 and 6 of table 2;
(2) coefficients of individual characteristics vary by city;
(3) 29 industry dummies are included;
(4) 4 occupation dummies are included;
(5) a college town dummy is included;
(6) observations are assigned to MSA of work, not residence;
(7) region×period dummies are included ;
(8) residuals for cities in the same state are correlated.

Table 8: The Effect of Changes in Share of White and Black College Graduates on White and Black Wage: 1990 Census Data

	Dependent Variable: Wage of Whites	Dependent Variable: Wage of Blacks
Whites' College Share	0.898	0.754
	(0.069)	(0.064)
Blacks' college Share	0.307	0.849
	(0.137)	(0.142)

NOTES: All models include individual characteristics (education, sex, race, Hispanic origin, US citizenship, and a quadratic in experience), city characteristics (unemployment rate, percentage of Hispanics, females, US citizens) and eight regional dummies. Each entry is a separate regression. Standard errors corrected for clustering in parentheses.

Table 9: The Effect of Measurement Error in Individual Education on the Coefficient of College Share: Census Data

	OLS	2SLS		
		Age Structure	Cost of Tuition	2 IVs
	(1)	(2)	(3)	(4)
(1) Private Return to Education is Unconstrained				
[0.085 (0.0001) in 1990 and 0.062 (0.0001) in 1980]				
Coeff. of College Share (from table 2)	1.547	1.079	1.666	1.391
	(0.137)	(0.614)	(0.446)	(0.537)
(2) Private Return to Education +10%				
[0.093 in 1990 and 0.068 in 1980]				
Coeff. of College Share	1.498	1.045	1.586	1.337
	(0.137)	(0.612)	(0.445)	(0.536)
(3) Private Return to Education +15%				
[0.097 in 1990 and 0.071 in 1980]				
Coeff. of College Share	1.473	1.028	1.547	1.311
	(0.137)	(0.611)	(0.445)	(0.536)
(4) Private Return to Education +15% only in 1990				
[0.097 in 1990]				
Coeff. of College Share	1.445	1.089	1.502	1.339
	(0.137)	(0.608)	(0.445)	(0.536)
City Fixed Effects	yes	yes	yes	yes
Z_{ct}	yes	yes	yes	yes
Katz-Murphy Index	no	no	no	no
Cities	282	282	115	282

NOTES: All models include individual characteristics (education, sex, race, Hispanic origin, US citizenship, and a quadratic in experience) and city characteristics (unemployment rate, proportion of Hispanics, females, US citizens). The interaction South×year is also included. Standard errors corrected for clustering in parentheses.

Table 10: The Effect of Increases in Percentage of College Graduates on Wages: NLSY Data

	Census (from Table 2)	NLSY				
	OLS	OLS	Reliability of College Share	OLS Katz & Murphy Index	Propensity Score	2SLS
	(1)	(2)	(3)	(4)	(5)	(6)
(1) Cross Section	0.982	0.931	0.86			
	(0.070)	(0.145)	(0.05)			
(2) City Effects	1.547	1.162	0.76			
	(0.137)	(0.415)	(0.13)			
(3) City and Individual Effects		1.285	0.77			
		(0.361)	(0.14)			
(4) City×Individual Effects		1.362	0.77	1.426	1.243	1.365
		(0.380)	(0.15)	(0.380)	(0.388)	(1.293)
F–Test [p-value]					0.000	
Hausman Test [p-value]						0.99
Years	80,90	79-94	79-94	79-94	79-94	86,87,93,94

NOTES: Each entry is a separate regression. All models include individual characteristics (education, sex, race, Hispanic origin, US citizenship, and a quadratic in experience) and city characteristics (unemployment rate, proportion of Hispanics, females, US citizens). The interaction South×year is also included. The estimated reliability ratio is the slope coefficient from a regression of the measure of college share from the CPS on the one from the Census, conditional on all covariates included in the wage equation. One minus the reliability ratio is the attenuation bias. Hausman test refers to the hypothesis that IV estimates are different from the corresponding OLS estimates. F-test refers to the significance of the cubic term in the propensity score. Standard errors corrected for city–year clustering in parenthesis.

Table 11: The Effect of Changes in Share of College Graduates on Wage of Education Groups: Census Data

	OLS	2SLS			
		Age Structure in 1970	Tuition	Both IVs	Overid. [p-value]
	(1)	(2)	(3)	(4)	(5)
less than high-school	1.605	1.682	2.964	2.364	0.07
	(0.167)	(0.742)	(0.598)	(0.645)	
high-school	1.634	1.129	1.659	1.478	0.33
	(0.156)	(0.724)	(0.499)	(0.604)	
some college	1.580	0.857	1.366	1.103	0.45
	(0.139)	(0.657)	(0.448)	(0.547)	
college	1.270	1.225	1.219	1.231	0.98
	(0.122)	(0.555)	(0.392)	(0.477)	
City Effects	yes	yes	yes	yes	
Cities	282	115	282	115	

NOTES: Each entry is a separate regression. The dependent variable is the change in adjusted mean wage by city and education group. Entries are the coefficients on the percentage of college graduates. Standard errors corrected for clustering in parentheses.

Table 12: Predicted Effect on Earnings of an Increase in Percentage of College Graduates in City

	Large Increase in College Grads (+ 0.38 %)	Median Increase in College Grads (+ 0.2%)	Small Increase in College Grads (+0.06 %)
	(1)	(2)	(3)
less than high-school	224	118	35
high-school	139	73	22
some college	104	55	16
college	114	60	18

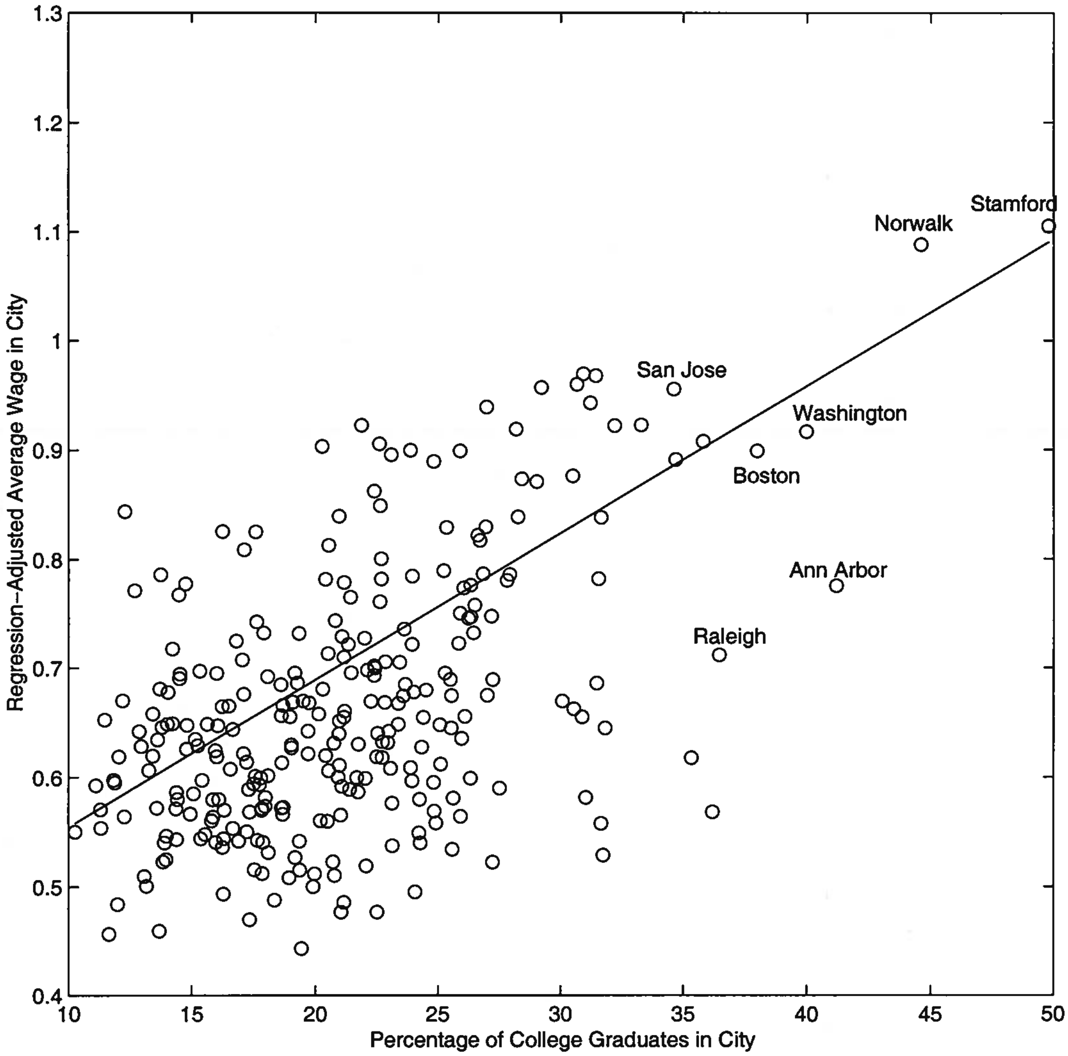
NOTES: Entries are predicted yearly earnings increases for a worker with initial earnings of $25,000.

Table 13: Effect of Changes in College Share and Share of Education Group on Education Group Wage: Census Data

	OLS (1)
Coeff. on Share of College Grads.	1.440
	(0.076)
Coeff. on ln Share of Education Group	-0.047
	(0.014)
City Effects	yes
Group Effects	yes
City–Educat. Groups	1128

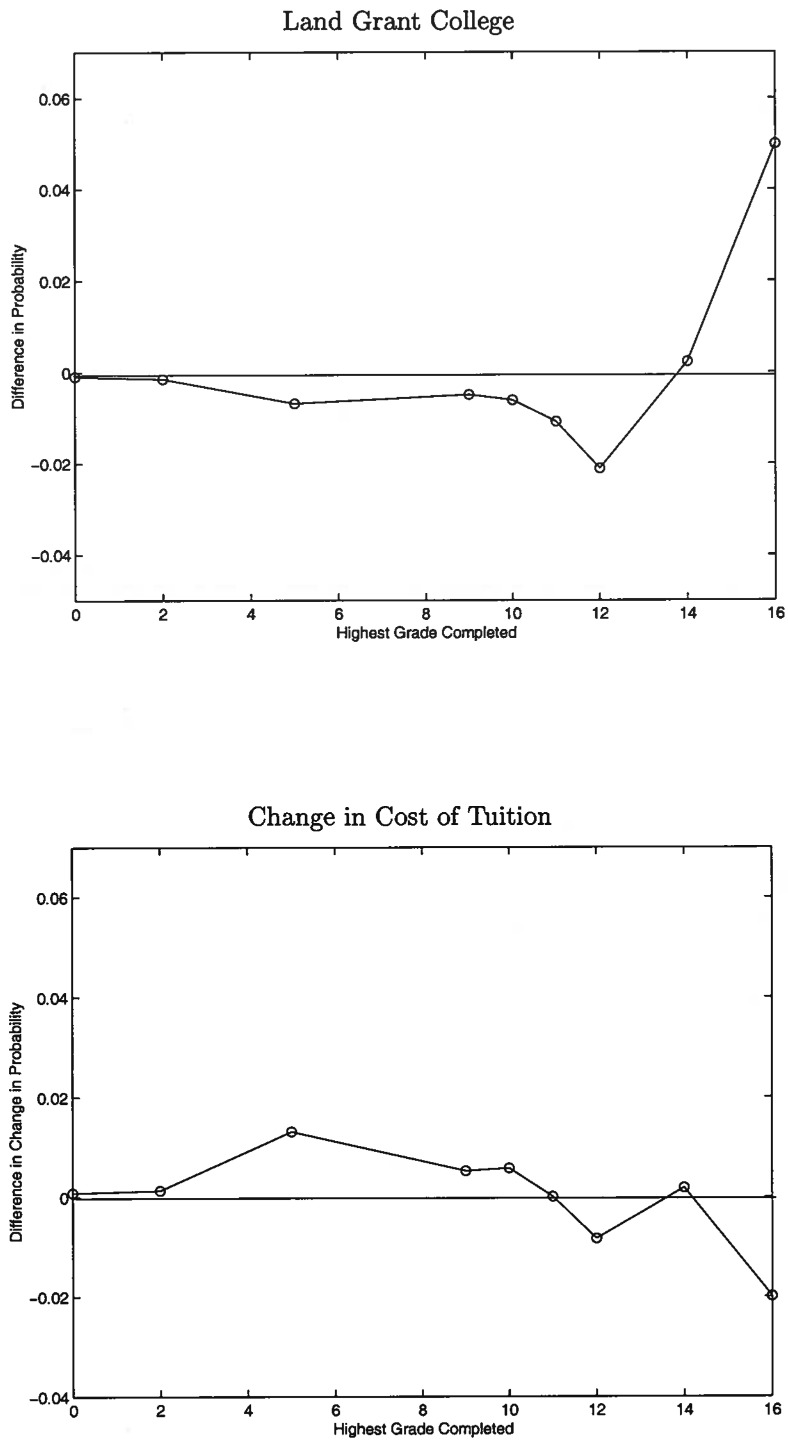
NOTES: The dependent variable is the change in regression–adjusted mean wage by city and education group. Standard corrected for group–city clustering in parentheses.

Figure 1: Correlation Between Regression-Adjusted Average Wage and Percentage of College Graduates in 282 Cities, in 1990.



NOTES: Regression–adjusted average wage is obtained by conditioning on individual education, gender, race, Hispanic origin, U.S. citizenship and work experience. Weighted OLS fit superimposed.

Figure 2: Difference in Distribution of Schooling, by Value of the Instrumental Variable.



Note: Panels show the difference in the probability of schooling at the grade level on the X-axis between cities with high and low values of the instrumental variables. The top panel shows the difference in the probability of schooling between cities with and without a land grant college. The bottom panel shows the difference in the change in the probability of schooling between cities with large and small increases in the cost of tuition.

Figure 3: The Effect of an Increase in Supply of College Graduates on College Graduates Wage.

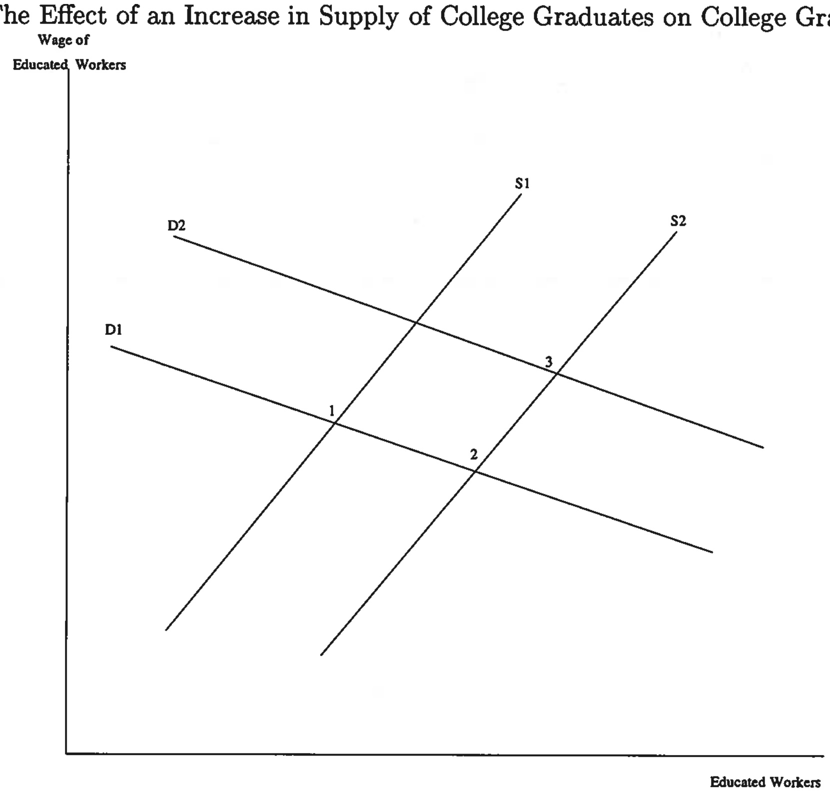
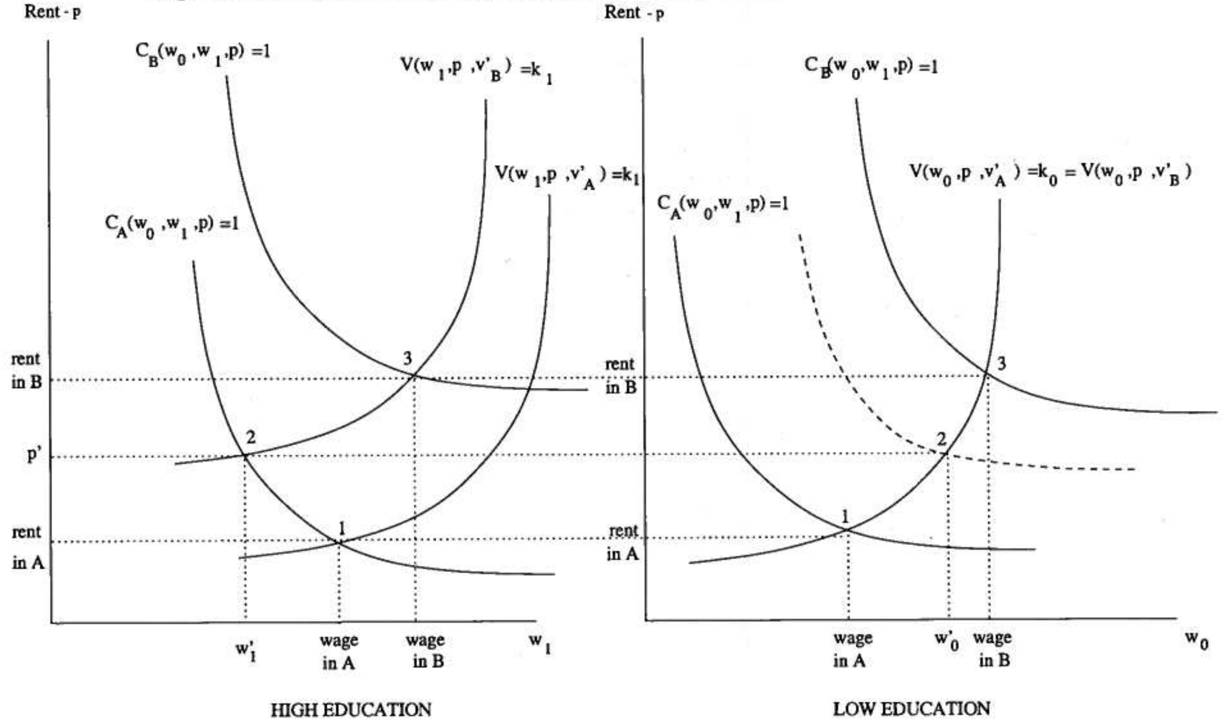


Figure A1: Equilibrium Wages of Educated and Uneducated Workers and Rent



NOTES: Point 1 is the equilibrium in city A. Point 2 is the equilibrium in city B without externality. Point 3 is the equilibrium in city B with externality. The dashed line in the right panel is the isocost curve in city B without externality. w_1 and w_0 are the nominal wage of educated and uneducated workers, respectively.