Spatial Changes in the College Wage Premium

Joanne Lindley* and Stephen Machin**

July 2012

*  Department of Economics, University of Surrey

**  Department of Economics, University College London and Centre for Economic Performance, London School of Economics

Abstract
We study spatial changes in the US college wage premium for states and cities using US Census and American Community Survey data between 1980 and 2010. We report evidence of significant spatial variations in the evolution of the college wage premium through time, for US states and MSAs. We report evidence that changes in school quality measures are significantly related to changes in state of birth college/high school wage differentials. Finally, we use estimates of spatial relative demand and supply models to calculate implied relative demand shifts for college graduates vis-à-vis high school graduates and show that relative demand has shifted faster in states and cities which have experienced faster increases in R&D intensity.

JEL Keywords: College wage premium; Relative demand; Spatial changes.
JEL Classifications: J31; R11.

Email of corresponding author: j.lindley@surrey.ac.uk.
1. Introduction

Over the last twenty years study of the evolution of the wage distribution over time has become a major preoccupation of empirical economists. A widening of the wage distribution showing rising wage inequality in a number of countries has been very clearly documented in this work.\(^1\) Empirical studies in this area have highlighted the temporal evolution of particular wage differentials linked to, for example, education or experience emphasising increase in the college wage premium or the return to experience that have gone hand-in-hand with rising wage inequality.

Despite there being a sizable urban economics literature studying the urban wage premium, work studying the spatial dimensions of rising wage inequality is more sparse. In part, this is because within/between type decompositions show that much of the increase in overall wage inequality, or in particular wage differentials, has been within, rather than between, spatial units of observation like regions, states, cities or local labour markets.

Nonetheless, at a given point in time there are very sizable spatial differences in wages and in wage differentials between different groups of workers. Whilst these spatial differences have traditionally been highly persistent through time, showing if anything evidence of regional convergence in income differences, it is interesting that in the period since wage inequality started to rise in the US (since the mid-to-late 1970s) this has stopped. Since then mean reversion in spatial wage differences is absent as wages have risen faster in places with higher initial wages. Moretti (2010), for example, shows plots of the wages of college graduates and high school graduates in 288 US Metropolitan Statistical Areas (MSAs) in 1980 and 2000 where wages grow faster in MSAs with higher wage levels in 1980 for both groups of workers. We find the same using data between

\(^1\) See Katz and Autor (1999) or Acemoglu and Autor (2010) for reviews of the large literature in labour economics and Hornstein et al. (2005) for a review of the work in macroeconomics.
1980 and 2010 for 219 MSAs, as demonstrated in Figure 1.\(^2\) But Figure 2 also shows the college/high school wage differential in 2010 to be higher in MSAs that had higher college/high school wage differentials in 1980.

We study changing patterns of spatial college wage premia in detail in this paper. We begin by documenting the nature of changes in the college wage premium across different spatial units, looking at the evolution over time at MSA and state level. We use US Census and American Community Survey (ACS) data in which we have information on individual’s MSA of residence and state of residence. We also know which state individuals were born in. We thus consider changes in spatial college wage premia for each of these spatial definitions.

Whether differences in state of birth wage premia by education can be linked to differences in school quality has been studied in 1980 Census data by Card and Krueger (1992). Given that we find changing college wage premia over time, a natural question to ask is whether these changes can also be empirically linked to changes in school quality over time. We therefore update the Card-Krueger approach to link changes in state of birth level college wage premia to measures of school quality from the state that individuals grew up in.

In the wage inequality literature, rising wage gaps between college educated and high school educated workers have been connected to shifts in the relative demand and supply of these groups of workers. Indeed, aggregate evidence shows that a key drivers of rising wage college premia has been an increased relative demand for college educated workers (see Katz and Murphy, 1992; Katz and Autor, 1999; Acemoglu and Angrist, 2010). The presence of rising spatial college wage premia at different rates also suggests

\(^2\) We use the 5 percent 1980, 1990 and 2000 Censuses and the 1 percent 2010 ACS which we collapse into 51 US states and 219 consistently defined MSAs. For 1980 and 2000 Census data Moretti (2010) reported slope coefficients (and associated standard errors) of 1.82 (0.89) for high school graduates and 3.54 (0.11) for college graduates.
there may be differential relative demand shifts occurring at the spatial level. We thus modify the commonly used relative demand and supply model so as to calculate the extent of spatial relative demand shifts and examine the variations in their evolution through time.

Previewing our key results, we report evidence of significant spatial variations in the evolution of the college wage premium through time, for US states and MSAs. We report evidence that school quality measures do significantly explain state of birth college/high school wage differentials. Finally, we use estimates of spatial relative demand and supply models to calculate implied relative demand shifts for college graduates vis-à-vis high school graduates and show that relative demand has shifted faster in states and cities that have experienced faster increases in R&D intensity.

The rest of the paper is structured as follows. Section 2 offers a descriptive analysis of changing educational wage differentials in the US at different levels of spatial aggregation. Section 3 connects the state of birth educational wage differentials to measures of school quality. Section 4 then considers the differentials in a relative supply-demand framework, showing spatial variations in the nature of relative demand shifts over time. Section 5 concludes.

2. Spatial Employment Shares and Wage Differentials - Descriptive Analysis

To investigate spatial changes in the college wage premium we use data from the US Census in 1980, 1990 and 2000, and the American Community Survey (ACS) of 2010. We use the 5 percent Census sample and the 1 percent ACS to study the evolution of wage differentials for the 48 contiguous US states (dropping Alaska, Hawaii and the District of Columbia) and for 219 consistently defined MSAs. We focus on US born individuals aged 26-55 throughout our analysis.
It has been widely documented that the employment shares of more educated workers have increased over time in the US (Acemoglu and Autor, 2010). Table 1 shows that, on average across the 48 states we look at, that employment shares of college graduates rose by 10.8 percentage points between 1980 and 2010, going from a share of 22 percent in 1980 to just over 33 percent in 2010. The comparable average increase is slightly larger at 11.3 percentage points across MSAs, as expected this is because metropolitan areas do not contain rural geographies and are likely therefore to contain more college graduates.

There are sizable spatial disparities in the change in the college share of employment. Whilst the mean rises by 0.108 across states and 0.113 across MSAs, the interquartile ranges are for states (MSAs) have increased from 0.061 (0.064) in 1980 to 0.076 (0.095) in 2010. Figure 3 plots the 1980 and 2010 shares for states of residence and MSAs showing them to be significantly correlated and that there is considerable spatial variation. The size of the circles reflect the population weighting and it is interesting to note the significantly larger college shares in the larger states and cities in the Figure. This clearly remains the case over time: high spearman rank correlations of 0.98 (p-value = 0.00) for states and 0.93 for MSAs (p-value = 0.00) shows that there is very strong spatial persistence in college employment shares over time.

At the same time as the shares of the college educated in employment has risen, so have their relative wage differentials. Table 2 presents mean composition adjusted wage differentials for college graduates relative to high school graduates across states and MSAs. One can see the well known average pattern of increasing wage payoffs to college

---

3 These wages are compositionally adjusted using log wage equations estimated separately for each year, year of birth cohort, as well as 48 states and 219 MSAs respectively. The equations include dummies for gender and race and state of residence (for the state of birth estimates) and state of birth (for the state of residence estimates). Three education dummies are included for college plus (16 or more years of education), some college (13 to 15 years of education) and high school dropouts (less than 12 years of education).
graduates as the college/high school log wage premium rises by 0.265 percentage points between 1980 and 2010 across states of residence and by a slightly smaller increase of 0.258 percentage points across MSAs.

We can also consider state of birth college wage premia using the Census/ACS data, for the state and MSA samples. These log wage differentials are also summarised in Table 2 and again show significant increases at the average, going up by 0.258 percentage points with states of birth fixed effects and by 0.260 percentage points with the MSA fixed effects.

Previous research by Black, Kolesnikova and Taylor (2009) has noted that, at a point in time (in their case the cross-sections from the 1980, 1990 and 2000 US Census), there are sizable spatial disparities in education related wage differentials. This is also the case for our analysis, both in terms of the yearly variations across spatial units (thus confirming the Black, Kolesnikova and Taylor findings) and in terms of the increase in the college premium. For example, the inter-quartile range for states (MSAs) has remained constant at 0.166 (0.131) in 1980 and 2010. Our study of their variation over time shows that these spatial disparities are highly persistent. Figure 4 plots the 1980 and 2010 composition adjusted wage premia for states and MSAs. It shows the 1980 composition adjusted college/high school wage differential to be positively and significantly correlated with that in 2010.

The final column of Table 2 corroborates the notion that spatial units have experienced different rates of increase in the college wage premium by showing that, for both states and cities, most of the growth in the college wage premium over time has occurred within geographies and not between. For state and MSA of residence the within

---

education) relative to the omitted group of high school graduates (12 years of schooling). The college/high school graduate log wage differential is the estimated coefficient on the college graduate variable.
component is 86 percent and 77 percent respectively. It is even higher for the state of birth estimates, at around 88 percent for both samples we consider.

Given these spatial disparities, in the remainder of the paper we consequently we look for explanations of within-state changes in graduate wage differentials. Firstly, in the next Section of the paper, we investigate the role of changing school quality and secondly, in Section 4, we estimate spatial demand and supply models so as to calculate spatial shifts in relative demand that we can relate to measures of technical change that vary by spatial geographies.

3. Changes in State Wage Differentials and School Quality

So far we have found that there are significant within-state and within-MSA increases in the college/high school wage premium over time. We now explore whether these are related to the school quality in the state where individuals grew up. We build upon the paper by Card and Kreuger (1992) which looked at years of education wage differentials by state of birth across birth cohorts in the 1980 Census and how they related to school quality measures. We extend this to a cross-Census comparison, using more up to date birth cohorts and retain our focus on the change over time in the college/high school wage differential.

Like Card and Krueger (1992), we adopt a two stage estimation procedure, estimating composition adjusted wage differentials by state of birth and birth cohort in the first stage and relating them to cohort specific measures of state school quality in the second stage. The school quality measures at state level we consider are the pupil-teacher ratio and the relative teacher wage (these are available for most years by state from the Biennial Survey of Education up to 1958 and in the annual Digest of Education Statistics
from 1960 onwards). We confine our analysis to the post-war period, therefore using state level school quality measures from 1950 onwards.

Table 3 describes the age structure of our data. Basically we have 6 birth cohorts across the whole data set, defined at 9 year intervals, for workers born between 1925 and 1984. The second column shows the years in which these birth cohorts attended school. We must assume therefore that these individuals went to school in the states where they were born. We then seek to explain state level earnings differentials based on the schooling quality in their state of birth at the time when they attended school. We can follow the subsequent earnings of these birth cohorts to form a pseudo panel across the 1980-2010 Census and ACS data. This allows us to track wage differentials over the lifecycle at different age points, although we only observe a full lifecycle (age 26-35, 36-45 and 46-55) for two of our birth cohorts, those born in 1945-54 and 1955-64, who went to school in the 1950s and 1960s. Because our school quality measures run from 1950 onwards, we can only study nine of the twelve cohorts as three (the 36-45 and 46-55 age groups in the 1980 Census and the 46-55 age group in the 1990 Census) were at school prior to 1950.

Table 4 shows log weekly state of birth level college/high school graduate wage differentials, estimated from our first stage individual level wage regressions, averaged across the nine birth cohorts and time.\(^4\) Firstly, comparing down the diagonals shows that over time the average college/high school wage differential has increased for all age groups. Secondly, looking within the columns we can see that in the 1980s and 1990s the average college/high school wage differential is linear in age, but from 2000 onwards it

---

\(^4\) These are state level differentials are derived from individual level log wage equations estimated separately by state (for 48 states) for each year and year of birth cohort. The equations include dummies for gender and race and state of residence. Three education dummy variables are included for college plus (16 or more years of education), some college (13 to 15 years of education) and high school dropouts (less than 12 years of education) relative to the omitted group of high school graduates (12 years of schooling). The college graduate/high school graduate log wage differential is the estimated coefficient on the college graduate dummy variable.
becomes quadratic in age peaking at age 36-44 falling thereafter. Thirdly, the third and fourth rows (for birth cohorts 1945-54 and 1955-64) provide the pseudo panel comparison. These indicate that the state average college/high school wage differential is linearly increasing over time for people who went to school in the 1950s and 1960s. We do not yet have the data to see if this is still the case for people who went to school in the 1980s and 1990s.

In order to try study whether differences in state level wage differentials are related to school quality measures we estimate a second stage regression, where we regress the composition adjusted college/high school wage differentials (estimated from the first stage) at the state of birth, birth cohort and year level on our school quality measures also measured at the state and birth cohort level. We estimate three models. The first model uses the pupil-teacher ratio as a measure of school quality, whilst the second model uses the relative teacher salary. The final model includes both of these measures to provide a horse race between them. Given the geographic differences in the cost of living and in the level of alternative wages available to teachers, we normalize teacher wages in each state by the level of average wages in the state. We divide the relative teacher wage in each state by the national average of this ratio in the same year. This eliminates any time-series variability in the average value of relative teacher salaries, while preserving the interstate variation in relative teacher wages at a point in time.

It is worth noting at this point that we only have up to three birth cohorts in each year, and nine overall, so with 48 states of birth, this gives us a total sample size of 432. We include state, birth cohort and year fixed effects in all of our regressions and these are weighted using the inverse sampling variance of the estimated returns.\(^5\) Table 5 reports the

\(^5\) Following Card and Krueger (1992) we weight our regressions to take account of the covariances between the estimated returns for different states. The reason for this is that the estimated returns by state of birth are not quite independent since there is some covariation between them that arises from the fact that the same regression parameters are used to adjust for other control variables in the first-stage regressions.
results for our three models. The first two columns show that individually both variables are correlated with college/high school wage differentials, with estimated parameters (standard errors) of -0.005 (0.002) and 0.126 (0.022) for the pupil-teacher ratio and relative teacher salaries respectively. However, as column (3) shows, if we include both of the school quality measures only the relative teacher salary is statistically significant. This is also what Card and Krueger (1992) find and is likely a consequence of multicollinearity among the quality variables (which is not so surprising, given their levels of variation). However, the magnitudes of the estimated school quality coefficients suggest a quantitatively important effect of school quality on the college/high school wage differentials.

4. Relative Supply and Demand Models.

Table 1 showed that the relative supply of college graduates vis-à-vis high school graduates has increased over time when averaged across states and MSAs. At the same time, Table 2 showed that mean college graduate wage differentials have also increased. Consequently we now draw upon the Katz and Murphy (1992) canonical model of relative supply and demand to see if there are differential relative demand shifts by state and MSA.

The starting point in this approach is a Constant Elasticity of Substitution production function where output in for state or MSA i in period t ($Y_{it}$) is produced by two education groups ($E_{1it}$ and $E_{2it}$) with associated technical efficiency parameters ($\theta_{1it}$ and $\theta_{2it}$) as follows:

$$Y_{it} = (\theta_{1it}E_{1it}^{\rho} + \theta_{2it}E_{2it}^{\rho})^{1/\rho}$$  \hspace{1cm} (1)

where $\rho = 1 - 1/\sigma_E$, where $\sigma_E$ is the elasticity of substitution between the two education groups.
Equating wages to marginal products for each education group, taking logs and expressing as a ratio leads to the relative wage equation

\[
\log\left(\frac{W_{1t}}{W_{2t}}\right) = \log\left(\frac{\theta_{1t}}{\theta_{2t}}\right) - \frac{1}{\sigma_E} \log\left(\frac{E_{1t}}{E_{2t}}\right)
\]

(2)

that can be transformed by parameterising the demand shifts term as

\[
\log\left(\frac{\theta_{1t}}{\theta_{2t}}\right) = \alpha_0 + \alpha_i f(t) + a_i + e_{it}
\]

(3)

where \(f(t)\) is a function of time (e.g. proxied by a time trend in the economy wide approaches of Katz and Murphy, 1992, Autor et al., 2008, and Card and Lemieux, 2001), \(a_i\) are state level fixed effects which are captured using state/MSA level dichotomous variables and \(e_{it}\) is an error term, to give

\[
\log\left(\frac{W_{1t}}{W_{2t}}\right) = \alpha_0 + \alpha_i f(t) + a_i + \gamma \log\left(\frac{E_{1t}}{E_{2t}}\right) + e_{it}
\]

(4)

where \(\gamma = -1/\sigma_E\).

We specify the \(f(t)\) function in its most general way, as a full set of time dummy variables, so that the estimating equation expresses the relative wage as a function of a time, state/MSA fixed effects and relative supply. The typical approach for estimating equations like (4) focuses on the college only/high school wage differential and model relative supply in terms of college equivalent and high school equivalent workers. To define equivalents within the college and high school groups, individuals with different education are assumed to be perfect substitutes, but are given different efficiency weights. In this regard we follow the same approach here (see Autor, Katz and Kearney, 2008, or Lindley and Machin, 2011, for more detail).

Our equation (4) differs from that usually estimated in the literature since we are estimating at the state and MSA level and including fixed effects. However, in this case the correlation between relative wages and relative supply may arise because of geographical
migration or because of potential endogeneity. Also, when estimating at the spatial level it is more likely that our variables may be measured with error which will bias downwards the $\gamma$ parameter in equation (4). Consequently, we adopt a Two Stage Least Squares (2SLS) approach.

To instrument relative labour supply, we draw upon the paper by Ciccone and Peri (2005) who use data from Acemoglu and Angrist (2000) to show that changes in state level compulsory attendance laws are correlated with the labour supply of high school graduates relative to high school drop outs. For our purposes it was necessary to update the data used in Acemoglu and Angrist (2000) to include state level compulsory attendance laws up to 2002.\textsuperscript{6} We then use these laws as instruments for changes in the relative supply of college graduates at the state and MSA level.\textsuperscript{7} The basic information is summarized in five dichotomous variables associated with each individual in the sample. The dummies CA8, CA9, CA10, CA11 and CA12 are equal to 1, and all other compulsory attendance law dummies are equal to 0, if the state where the individual is likely to have lived when aged 14 had compulsory attendance laws imposing a minimum of 8, 9, 10, 11 and 12 plus years of schooling. The five dummies are used to calculate the share of individuals for whom each of the five dummies is equal to 1 in each state and MSA of residence. We omit the share for CA8 (as the variables add up to 1) and these are used as instruments for the relative supply of college graduates. The data do not include precise information on where individuals lived when aged 14, we therefore assume (as in the study of school quality in Section 3 above) that at age 14 individuals still lived in the state where they were born.

\textsuperscript{6} The Acemoglu and Angrist (2000) data run from 1915 to 1978 and are available from the authors on request. We updated these data by obtaining compulsory schooling data up to 2002 from the digest of education statistics, supplemented by looking up the laws themselves in the state statutes. We cross checked our new data with that derived in Oreopoulos (2009) and found we had very similar measures.

\textsuperscript{7} Ciccone and Peri (2005) adopt a similar instrumentation approach when they estimate state level demand and supply models for high school graduates relative to high school drop outs using 1950-1990 Census data.
**Relative Demand and Supply Models**

Table 6 reports the first stage regression results for the state of birth instruments on the relative supply of college graduates. These include time and state/MSA fixed effects. These are negative and significant overall with F statistics of 4.46 for state and 7.58 for MSA level equations. Table 6 also reports correlations between the state of birth instruments and the composite components of relative supply to help us to understand why the compulsory school attendance (CA) shares are negatively correlated with our measure of relative supply. We can see that for the state of residence equations, the CA shares are more correlated with non-college equivalents than with college equivalents, but both are positive. This suggests that in states where people had longer compulsory schooling laws there is a greater supply of non-college equivalents than college equivalents explaining the negative correlation with relative labour supply. One might expect this is compulsory school attendance laws are local and impact less on the education levels of those who would have gone to college anyway.

For the MSA equations there is a negative correlation for college equivalents suggesting that there is lower labour supply of college workers in MSAs which are inhabited by more workers exposed to longer state of birth level compulsory schooling laws. This correlation is still positive for non-college equivalents (although only significant for CA11). This suggests that metropolitan areas with relatively more college graduates also have more workers raised in states with lower compulsory schooling laws. Perhaps this is because college graduates raised in states with longer compulsory schooling laws are more likely to live in rural areas than in MSAs.

Table 7 provides the OLS and 2SLS estimates for equation (4). Again, these include time and geographical fixed effects. The estimate of the labour supply parameter $\gamma$ is negative, as expected, in all cases. Moreover, this parameter becomes more negative
using 2SLS compared to OLS as one would expect with attenuation bias. For the state level regressions the 2SLS estimate is -0.235 which provides an estimated elasticity of substitution ($\sigma_E$) of 4.25. For the MSA level estimates, this is slightly smaller at 3.60. These are slightly larger substitution elasticities than in the existing literature: for example, Lindley and Machin (2011) derive an estimate of around 2.6 using aggregate CPS data from 1963 to 2010 which is in the same ballpark as Autor, Katz and Kearney's (2008) estimate of 2.4, who also use the CPS data from 1963 to 2005.

The estimated parameters on the time dummies in Table 7 tell us something about the relative demand shifts that have occurred on a decade by decade basis. Relative the 1980s the relative demand for college graduates has increased across all time periods although these incremental changes are getting smaller. This supports the idea of a quadratic relationship between relative labour demand and relative wages. Although the demand for college graduates vis-à-vis high school graduates is still increasing the relative demand shifts are slowing down.

*Implied Relative Demand Shifts*

One other reason for estimating spatial demand and supply models is to compare the scale of relative demand shifts across states, but also to relate these to changes to other potential drivers of rising wage inequality that can be directly measured at the state level. Such correlations are usually undertaken at the industry level by estimating cost share equations. But we can directly correlate our state level demand shifts to potential drivers of inequality. One example is state level changes in Research and Development (R&D) intensity, which we define as state R&D expenditure divided by state GDP. To do this, we refer back again to equation (2) which can be written as

$$\log \left( \frac{W_{1t}}{W_{2t}} \right) = D_\alpha - \frac{1}{\sigma_E} \log \left( \frac{E_{1t}}{E_{2t}} \right)$$

(5)
where \( D_{it} = \log \left( \frac{\theta_{1it}}{\theta_{2it}} \right) \). Hence we can estimate the \( \gamma = -1/\sigma_E \) parameter from equation (4) and derive implied demand \( D_{it} \) for state \( i \) at time \( t \) using

\[
D_{it} = \log \left( \frac{W_{1it}}{W_{2it}} \right) + \frac{1}{\sigma_E} \log \left( \frac{E_{1it}}{E_{2it}} \right)
\]

Figure 5 uses the 2SLS estimates from Table 7 to show such implied demand shifts between 1980 and 2010 by state. These are ranked in descending order of magnitude. Figure 6 then plots the state level implied demand shifts from Figure 5 against state level changes in R&D performance between 1977 and 2007, taken from the National Industrial Productivity Accounts (NIPA). There is a clear positive correlation between changes in state level R&D and college graduate relative demand shifts, with a regression slope of 2.89. We find a similar positive relationship between changes in state level computer use and our college relative demand shifts (slope of 0.220 and a standard error of .209) using the proportion of workers using a computer at work from the October 1984 and 2003 CPS.

5. Concluding Comments

The focus of this paper is the spatial dimension of changing wage inequality, with a specific focus being placed on spatial changes over time in the US college wage premium. It is well known that the wage premium of college graduates over high school graduates has risen very sharply over the last thirty years. What is less well understood, and much less studied, is whether there are important spatial dimensions to the temporal evolution of the college/high school wage gap (though see Black et al.’s, 2009, study of spatially varying educational wage differentials).

We study spatial changes in the US college wage premium for states and cities using US Census and American Community Survey data between 1980 and 2010. We report evidence of significant spatial variations in the evolution of the college wage
premium through time, for US states and MSAs. We report evidence that changes in school quality measures are significantly related to changes in state of birth college/high school wage differentials. Finally, we use estimates of spatial relative demand and supply models to calculate implied relative demand shifts for college graduates vis-à-vis high school graduates and show that relative demand has shifted faster in states and cities that have experienced faster increases in R&D intensity. Thus, spatial variations in the quality of schooling and in innovative activity have played a role in the widening of the college/high school wage gap within US states and cities over time.
References


Notes: Each panel plots the nominal wage in 1980 against the nominal wage in 2010 by metropolitan area. The top panel is for high school graduates and the bottom panel is for college graduates. These are weighted using the number of workers in the relevant metropolitan area and skill group in 1980. There are 219 metropolitan areas. The regression line is the predicted log wage in 2010 from a weighted OLS regression. The slope is 1.302 (0.228) for high school graduates and 4.701 (0.267) for college graduates. Data are from the Census of Population. The sample includes all full time US born workers age between 26 and 55 who worked at least 40 weeks in the previous year.
Figure 2:
Change Over Time in the Average College/High School Graduate Wage Differential by Metropolitan Area

Notes: Figure 2 plots the average college graduate wage relative to the average high school graduate wage in 1980 against the same in 2010 by metropolitan area. These are weighted using the number of workers in the relevant metropolitan area and the two skill groups in 1980. There are 219 metropolitan areas. The regression line is the predicted log wage in 2010 from a weighted OLS regression. The slope is 0.675 (0.074). Data are from the Census of Population. The sample includes all full time US born workers age between 26 and 55 who worked at least 40 weeks in the previous year.
Figure 3:
Change Over Time in the Employment Shares of College Graduates
by Geography of Residence

1. State of Residence

![State of Residence Graph]

Spearman Rank (P-Value) = 0.98 (0.000)
Slope = 0.966 (0.027)

2. MSA of Residence

![MSA of Residence Graph]

Spearman Rank (P-Value) = 0.932 (0.000)
Slope = 0.969 (0.014)

Notes: Weighted using state/MSA population shares. Standard errors in parentheses. The sample includes all US born workers age between 26 and 55.
Figure 4:
Change Over Time in the Composition Adjusted Wage Differentials by Geography of Residence

1. State of Residence

Spearman Rank (P-Value) = 0.498 (0.000)
Slope = 0.655 (0.133)

2. MSA of Residence

Spearman Rank (P-Value) = 0.337 (0.000)
Slope = 0.492 (0.079)

Notes: Weighted using state/MSA population shares. Standard errors in parentheses. The sample includes all US born full time full year workers age between 26 and 55.
Figure 5: State-Specific Relative Demand Shifts, 1980 - 2010 (Based on 2SLS Estimates)

Notes: These are 1980-2010 changes in implied demand estimated using the parameter estimates and data from Table 7.

Figure 6: State-Specific Relative Demand Shifts and Change in Industrial R&D Performance (Based on 2SLS Estimates)

Notes: The R&D data are taken from the October 1977 and 2007 NIPA data for Total funds for Industrial R&D Performance in millions of Dollars as a Proportion of Nominal GDP.
Table 1: Average State/MSA Employment Shares by Education Group

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>State of Residence</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N = 48</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>College Plus</td>
<td>0.222</td>
<td>0.261</td>
<td>0.295</td>
<td>0.331</td>
<td>0.108 (0.009)</td>
</tr>
<tr>
<td>Some College</td>
<td>0.219</td>
<td>0.301</td>
<td>0.316</td>
<td>0.253</td>
<td>0.034 (0.007)</td>
</tr>
<tr>
<td>High School Graduate</td>
<td>0.353</td>
<td>0.329</td>
<td>0.300</td>
<td>0.335</td>
<td>-0.018 (0.009)</td>
</tr>
<tr>
<td>High School Drop Out</td>
<td>0.205</td>
<td>0.109</td>
<td>0.088</td>
<td>0.081</td>
<td>-0.124 (0.008)</td>
</tr>
<tr>
<td><strong>MSA of Residence</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N = 219</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>College Plus</td>
<td>0.248</td>
<td>0.292</td>
<td>0.324</td>
<td>0.361</td>
<td>0.113 (0.006)</td>
</tr>
<tr>
<td>Some College</td>
<td>0.234</td>
<td>0.312</td>
<td>0.318</td>
<td>0.250</td>
<td>0.016 (0.004)</td>
</tr>
<tr>
<td>High School Graduate</td>
<td>0.334</td>
<td>0.299</td>
<td>0.274</td>
<td>0.309</td>
<td>-0.025 (0.005)</td>
</tr>
<tr>
<td>High School Drop Out</td>
<td>0.184</td>
<td>0.097</td>
<td>0.083</td>
<td>0.080</td>
<td>-0.104 (0.004)</td>
</tr>
</tbody>
</table>

Notes: Employment shares are for all US born workers age 26-55.
### Table 2: Composition Adjusted College Wage Premia

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>State of residence</td>
<td>0.283</td>
<td>0.517</td>
<td>0.587</td>
<td>0.549</td>
<td>0.265 (0.012)</td>
<td>0.86</td>
</tr>
<tr>
<td>MSA of residence</td>
<td>0.263</td>
<td>0.493</td>
<td>0.558</td>
<td>0.522</td>
<td>0.258 (0.009)</td>
<td>0.77</td>
</tr>
<tr>
<td>State of birth with state of residence fixed effects</td>
<td>0.332</td>
<td>0.465</td>
<td>0.541</td>
<td>0.593</td>
<td>0.260 (0.008)</td>
<td>0.88</td>
</tr>
<tr>
<td>State of birth with MSA fixed effects</td>
<td>0.305</td>
<td>0.427</td>
<td>0.492</td>
<td>0.541</td>
<td>0.235 (0.008)</td>
<td>0.88</td>
</tr>
</tbody>
</table>

Notes: The composition adjusted state of birth college plus/high school log wage differential are derived from estimated log wage equations estimated separately for each year, year of birth cohort, as well as 48 states and 219 MSAs respectively. The equations include dummies for gender and race and state of residence (for the state of birth estimates) and state of birth (for the state of residence estimates). Three education dummies are included for college plus (16 or more years of education), some college (13 to 15 years of education) and high school dropouts (less than 12 years of education) relative to the omitted group of high school graduates (12 years of schooling). The college graduate/high school graduate log wage differential is the estimated coefficient on the college graduate variable. For the change 2010-1980 standard errors are in parentheses.
### Table 3: Birth Cohort and Age Structure of the Census and ACS Data

<table>
<thead>
<tr>
<th>Birth cohort</th>
<th>Years when ages 5-15</th>
<th>Ages, 1980 Census</th>
<th>Ages, 1990 Census</th>
<th>Ages, 2000 Census</th>
<th>Ages, 2010 ACS</th>
</tr>
</thead>
<tbody>
<tr>
<td>1925-34</td>
<td>1930-49</td>
<td>46-55</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1935-44</td>
<td>1940-59</td>
<td>36-45</td>
<td>46-55</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1945-54</td>
<td>1950-69</td>
<td>26-35</td>
<td>36-45</td>
<td>46-55</td>
<td></td>
</tr>
<tr>
<td>1955-64</td>
<td>1960-79</td>
<td>26-35</td>
<td>36-45</td>
<td>46-55</td>
<td></td>
</tr>
<tr>
<td>1975-84</td>
<td>1980-99</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Table 4: Composition Adjusted State of Birth College/High School Graduate Log Wage Differentials by Cohort

<table>
<thead>
<tr>
<th>Birth cohort</th>
<th>Mean College/High School Graduate Log Wage Differential</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1980</td>
</tr>
<tr>
<td>1945-54</td>
<td>0.256</td>
</tr>
<tr>
<td>1955-64</td>
<td>0.419</td>
</tr>
<tr>
<td>1965-74</td>
<td></td>
</tr>
<tr>
<td>1975-84</td>
<td></td>
</tr>
</tbody>
</table>

Notes: Individual level log wage equations are estimated separately by state (48 states) for each year and year of birth cohort. Equations include dummies for gender and race and state of residence. Three education dummies are included for college plus (16 or more years of education), some college (13 to 15 years of education) and high school dropouts (less than 12 years of education) relative to the omitted group of high school graduates (12 years of schooling). The college graduate/high school graduate log wage differential is the estimated coefficient on the college graduate variable.
### Table 5:
State of Birth College /High School Graduate Log Wage Differentials and School Quality

<table>
<thead>
<tr>
<th>State of Birth College/High School Log Wage Differential</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pupil-teacher ratio</td>
</tr>
<tr>
<td>Relative teacher salary</td>
</tr>
<tr>
<td>Birth cohort fixed effects</td>
</tr>
<tr>
<td>State fixed effects</td>
</tr>
<tr>
<td>Year fixed effects</td>
</tr>
<tr>
<td>Sample size</td>
</tr>
<tr>
<td>R-squared</td>
</tr>
</tbody>
</table>

Notes: Dependent variable is the composition adjusted state of birth college/high school log wage differential. Weighted by the inverse sampling variances of the dependent variable. Standard errors in parentheses.
## Table 6: First Stage 2SLS Regressions (Relative Supply = College Equivalents/Non-College Equivalents).

<table>
<thead>
<tr>
<th>State of Residence</th>
<th>MSA of Residence</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Relative Supply</td>
</tr>
<tr>
<td>CA9</td>
<td>-0.118 (0.096)</td>
</tr>
<tr>
<td>CA10</td>
<td>-0.272 (0.102)</td>
</tr>
<tr>
<td>CA11</td>
<td>0.091 (0.187)</td>
</tr>
<tr>
<td>CA12</td>
<td>-0.363 (0.136)</td>
</tr>
<tr>
<td>Partial R Squared</td>
<td>0.11</td>
</tr>
<tr>
<td>F-Test</td>
<td>4.46</td>
</tr>
<tr>
<td>P-Value</td>
<td>0.002</td>
</tr>
<tr>
<td>Geographical Fixed Effects</td>
<td>Yes</td>
</tr>
<tr>
<td>Year Dummies</td>
<td>Yes</td>
</tr>
<tr>
<td>Sample Size</td>
<td>192</td>
</tr>
</tbody>
</table>

Notes: Where CA9 is the proportion of individuals residing in a geographical area with 9 years Compulsory Schooling Attendance (CSA) in the state in which they were born, when they were age 14. Similarly CA10 is the same but for CSA of ten years, CA11 for CSA of 11 years and CA12 for CSA of 12 years and over. College equivalents contain the hours of college graduates and half of the hours for some college. Non-college equivalents include the hours for high school drop outs, high school graduates and half of the hours for some college.
Table 7:
Estimates of Supply-Demand Models of College Plus/High School Wage Differentials
(Relative Supply = College Equivalents/Non-College Equivalents).

<table>
<thead>
<tr>
<th></th>
<th>State of Residence</th>
<th>MSA of Residence</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>OLS</td>
<td>2SLS</td>
</tr>
<tr>
<td>Log(Relative Supply)</td>
<td>-0.092 (0.032)</td>
<td>-0.235 (0.088)</td>
</tr>
<tr>
<td>Year = 1990</td>
<td>0.150 (0.011)</td>
<td>0.194 (0.027)</td>
</tr>
<tr>
<td>Year = 2000</td>
<td>0.241 (0.016)</td>
<td>0.307 (0.040)</td>
</tr>
<tr>
<td>Year = 2010</td>
<td>0.299 (0.017)</td>
<td>0.371 (0.043)</td>
</tr>
<tr>
<td>Geographical Fixed Effects</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Sample Size</td>
<td>192</td>
<td>192</td>
</tr>
<tr>
<td>R-Squared</td>
<td>0.96</td>
<td>0.95</td>
</tr>
</tbody>
</table>

Notes: The 2SLS method uses State Compulsory School Leaving Laws as instruments.