Worker Quality, Wages and the Education Premium in the United States, 1980-2005[*,†]

Zhiqi Zhao

Abstract

The education premium in the U.S. has been increasing since the 1980s, but the rate of increase has slowed. One explanation for the slowdown is the decrease in the quality of workers who attended some college relative to the quality of workers who obtained a high school diploma. This paper develops a measure of worker quality to estimate the impact of college and high school graduates quality on wages and the education premium. The measure of worker quality uses a weighted average of an occupational skill index. I link occupational skill to the measurement of quality because the variance of college wages is increasing over time and is directly related to occupational choice. I find that a 1 percent increase in the quality of both high school and college graduates results in a 0.36 percent increase in the education premium, which is an economically meaningful change. One reason for the decline in the quality of college graduates comes from the increasing college enrollment over time. I find that a 1 percent increase in college enrollment leads to a decline in the quality of college graduates by 0.11 percent, and has nearly no effect on the quality of high school graduates.

[*]Department of Economics, Clemson University, 227 Sirrine Hall, Clemson, SC 29634. Email: zhiqiz@clemson.edu. Phone: 1-864-650-8819.

[†]I benefited from the comments of both labor and public workshops participants at Clemson University and 2015 Summer School on Socioeconomic Inequality in UChicago. Special thanks to Curtis Simon, Chungsang Lam, Howard Bodenhorn and Bill Dougan for their guidance. I also thank Christopher Taber, Jeffrey Smith, Steven Durlauf, Lars Lefgren, Scott Barkowski, Jaqueline Oliveira, Dan Wood, Michael Makowsky, Peter Blair, Robert Fleck, Caitlin Myers, Art Carden, Adam Millsap, Kelsey Roberts, Mallika Garg and Tim Bacon for helpful comments. All mistakes are my own.

JEL Codes: I24, J24, J31

Keywords: wage, worker quality, occupation, skill percentile, education premium

1 Introduction

The education premium, measured as the relative wages of college graduates to high school graduates, has increased dramatically since the 1980s. Acemoglu (2002) argues that skill-biased technology has caused the growth in wage inequality in the U.S during the past two decades. Autor, Katz, and Kierney (2008) show that education premium growth was especially large in the early 1980s, and continued at a slower pace through the mid-1990s.[1] An important question is the extent to which the slower growth is explained by the relative change of quality of college graduates?

This paper explores how the quality of college and high school graduates affects wages and the education premium. Figure 1 shows trends of the mean wage ratio and variance wage ratio for college graduates and high school graduates for full time-full year white male workers aged 26-60 in U.S.[2] The relative variance increases more than the relative mean wage over time. Thus, attending at least some college is a risky decision because some people after obtaining a college degree may be on the left tail of the returns to college and receive a very low wage. The distribution of the returns to college is related to the graduate's occupational choice. Hence, I link my new measurement of the quality of college graduates and quality of high school graduates to an occupation skill index based on a worker's place in the wage-skill distribution.

This new measurement of quality is a weighted occupational skill index in the state of birth. The reason why I want to link this new measurement to occupations is that the variance for the return to college is increasing over time, and the return to college is directly related to the occupation choice. The result shows that worker quality influences the wage for college graduates more than the wage for high school graduates.

[1]Pierce (2001) argues that inequality growth in broader measures of compensation (fringe benefits) slightly exceeds wage inequality growth over the 1981 to 1997 period. Since currently available data often lack information on benefit costs and multiple benefits categories, he used Employment Cost Index (ECI) data. Unfortunately, no worker demographic or human capital information is available in the ECI, which is the focus of my paper. Therefore, I do not consider the inequality or the education premium in nonwage compensation, which includes health and life insurance, several forms of leave, pension and savings plans, bonuses and unemployment insurance.

[2]The data comes from 1980, 1990, 2000 Census and 2005 American Community Survey.

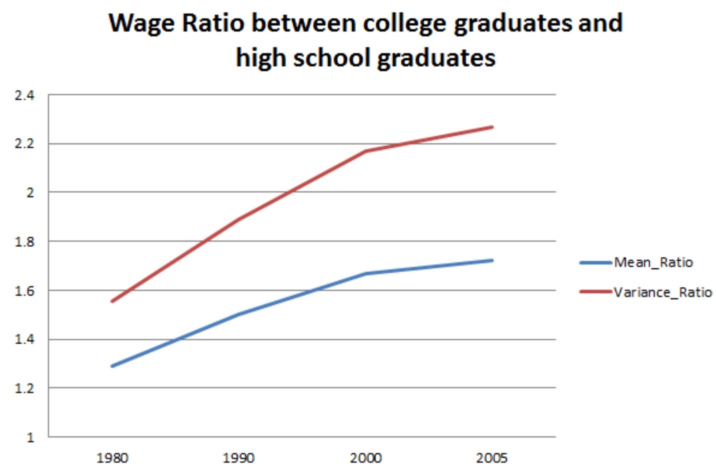

Figure 1

There are some classic measurements of human capital accumulation in a birth state in the literature. For example, the average SAT scores reflects the human capital accumulation before college in a birth state. The average ranking of colleges and universities shows the quality of colleges as well as the human capital accumulation in colleges. The new measurement in this paper includes all human capital accumulation before people start to work. Although the human capital accumulation in the new measurement has a longer period than the former two human capital accumulations, there may be some correlation among them. Hence, I will give evidence about the validity of the new measurement.

Carneiro and Lee (2011) argue that increased college enrollment led to a decline in the average quality of college graduates between 1960 and 2000, resulting in a decrease of 6 percentage points in the college premium. It is the first paper to decompose the college premium into price (quantity) effects and composition (quality) effects. They argued that college enrollment is an important factor influencing the quality of college graduates. They also regressed the wage for the high school group on college enrollment. However, college enrollment is not a good measurement for the average quality of high school graduates. Furthermore, college enrollment only affects the quality of college education, and it does not fully capture the quality of workers with a college degree, which is the quality of college graduates. Because quality of college education measures only human capital accumulation within college, but the quality of college graduates measures all human capital accumulation until individuals get a college degree. Therefore, I use a new measure of quality of college graduates and test how it influences the wage.

The structure of this paper is as follows: Section 2 describes the empirical strategy I use to measure worker quality and the validity of the new measurement. The empirical model is presented in section 3. The data and the empirical results are provided in sections 4 and 5, respectively. Section 6 shows sensitivity analysis. Section 7 concludes.

2 Methodology

I measure the quality of college and high school graduates by the average skill level in their state of birth using two steps. First, I revise the method of Autor and Dorn (2013) to calculate the skill percentile for each occupation in the whole nation level in every year. Second, I calculate average skill level in each state of birth for each group.

2.1 Skill percentile for occupations

First, in order to get consistent categories of occupations, I refer to the system developed by Dorn (2009). There are a new occupation system with 330 “occ1990dd” codes that provides a balanced panel of occupations covering the 1980, 1990, and 2000 Censuses and the 2005 ACS. Next, in each year, according to the mean real weekly wage for each occupation [3], I rank the skill level for each occupation by using cumulative share of employment. The methodology for developing skill percentile for occupation j in year t is described in the following function:

[3] Skill index is likely to be contaminated by price level effects. However, Glaeser and Resseger (2009) argue that bigger cities can attract more skilled workers and agglomeration makes workers more skilled too. In addition, more productive workers can live in cities because they can afford a higher cost of living. Thus, the contamination of skill index should not be an important issue. But in the section of robustness checks, I correct real wage by running wages on a vector of city dummies and use residual wages to construct the skill percentiles.

Table 1: 1980 Skill Percentile for Occupations

Occupation Category	Mean Real Weekly Wage	Employment Share	Cumulative Employment Share	Skill Percentile
Food preparation and service workers	89.10	0.00064	0.00064	0.064
Shoemaking machine operators	94.00	0.00024	0.00088	0.088
Hotel clerks	96.06	0.00015	0.00103	0.103
Child care workers	102.40	0.00025	0.00128	0.128
Clothing pressing machine operators	102.44	0.00015	0.00143	0.143
...	...	...	...	...
...	...	...	...	...
Optometrists	302.37	0.00024	0.99227	99.227
Health and therapy occupations	304.02	0.00014	0.99241	99.241
Podiatrists	359.22	0.00006	0.99247	99.247
Dentists	390.19	0.00124	0.99371	99.371
Physicians	395.38	0.00629	1.00000	100.00

Notes: Mean weekly wage is adjusted by CPI1999. The skill percentile is equal to the cumulative employment share multiplied by 100.

$$SkillPercent_{jt} = \sum_{k=1}^{j} EmployShare_{kt} * 100, \tag{1}$$

where occupation index j ranges from 1 to 330 and mean wage $\overline{w_k} < \overline{w_j}$ for $\forall$ k<j in year t.

Table 1 gives an example to illustrate how to calculate the skill percentile for occupations in 1980. The bottom 5 and top 5 occupations are showed according to their mean weekly wage. The skill percentile is equal to the cumulative employment share multiplied by 100. Similarly, the skill percentile for occupations is calculated in the same way for the year 1990, 2000 and 2005. So each year has its own occupation skill in the national level.

2.2 Average skill level in states of birth

The birthplace of workers is predetermined, though they can choose to work in any occupation in any state. Consistent with the previous literature, I assume workers are born and educated in the same state.[4]

[4]Card and Krueger (1992) show the evidence that there are relatively low mobility rates of preschool and school-age children. They assumed that an individual attends public elementary and secondary schools in his state of birth. However, the assumption that people attend college in the state of birth needs to be tested. Unfortunately, there is no cross section individual data available for the place of education in the state level. Following Winters (2012), one noisy and not ideal way to infer the college location of recent graduates (i.e. those who are ages 23-27 in the census and likely in school five years prior) is to use five-year migration data in the decennial censuses in 2000 and prior. There is one main five-year migration data in the state level in IPUMS: state or country of residence 5 years ago (MIGPLAC5). By using MIGPLAC5 and the variable of state of birth, the result shows that there is nearly 66 percent of college

Then I calculate average skill level in each state of birth for particular age group and education group in particular year. This index is a new measurement for the quality of workers which is directly related to where they were born. The quality here does not only capture the education resources in particular state of birth, it also reflect the average abilities and average family background of workers. These factors influence the wage received in their occupational state (not necessarily the same as the state of birth).

Since I want to focus on the education or skill premium[5], I separately calculate the average skill level for the high school graduates and college graduates at each state of birth in each year. Also, this quality index is different across the age groups. The following equation describes the calculation of the average skill level for the age group a, education group k, in state of birth b in year t:

$$AvegSkillPercent_{kabt} = \sum_{i=1}^{N_{kabt}} SkillPercent_{ikabt}/N_{kabt}, \tag{2}$$

where individual i born in state of birth b in age group a, education group k and work anywhere in a particular occupation corresponding with the skill percentile in year t. N_{kabt} is the total number of people in the sample born in state b belonging to the age group a, education group k in year t.

For example, how to calculate the quality of college graduates in the age group 25-30 born in New York State in year 1980? I average skill percentile for all people who are born and get college degree in New York State in the age group 25-30 in 1980. Each person in that group has occupation in a state (*not necessarily New York State*) and the occupation has skill level.

Table 2 illustrates how to calculate the quality index for college graduates in the age group 25-30 born in New York State in year 1980:

$$AverageSkillLevel = \frac{0.448 + 0.412 + 3.2 + 0.286 + 6.666 + ...595.362 + 297.723 + 595.582 + 4471.695 + 25700}{7 + 4 + 25 + 2 + 33 + ... + 6 + 3 + 6 + 45 + 257} = 58.744. \tag{3}$$

graduates studying in the same state of birth in U.S in the year 2000. Therefore, the assumption that people attend college in the state of birth is reasonable based on Census and ACS data. Besides this strategy, I also refer to a small sample from National Center for Education Statistics and it shows that more than 90 percent of first-time first-year students go to college in the same state of residence in 2000.

[5] In this paper, I do not distinguish the education premium and skill premium.

Table 2: In 1980, the average quality of college graduates, age 25-30, who were born in New York State

Occupation Category	Skill Percentile (1)	Number of observations (2)	Product(1)*(2)
Food preparation and service workers (444)	0.064	7	0.448
Hotel clerks (317)	0.103	4	0.412
Child care workers (468)	0.128	25	3.2
Clothing pressing machine operators (747)	0.143	2	0.286
Waiters and waitresses (435)	0.202	33	6.666
...	...	...	...
...	...	...	...
Optometrists (87)	99.227	6	595.362
Health and therapy occupations (89)	99.241	3	297.723
Podiatrists (88)	99.247	6	595.582
Dentists (85)	99.371	45	4471.695
Physicians (84)	100.00	257	25700

Notes: The skill percentile is calculated in the first step by the nation level. The number of observations means how many observations are for the occupations i in this group.

Table 3: The average mobility rate across state of birth and state of residence for FTFY white male workers,aged 25-60

Year	State Level	State−High School Graduates	State−College Graduates
1980	0.373	0.304	0.454
1990	0.394	0.296	0.459
2000	0.398	0.309	0.465
2005	0.399	0.312	0.460

Notes: Column 1 includes all the samples. Column 2 only includes workers who attended at least some colleges. Column 3 only includes workers who got at most high school diploma.

The key point of this methodology is that the average skill level captures the quality of the supply side. However, it is possible that mobility is so low across state of birth and state of work, and at the same time, the demand for some occupations is very high within a state of birth[6]. Thus, the demand shock for some occupations within the state may contaminate the measurement of quality. However, Table 3 gives evidence that average mobility rate across states is quite high, especially for the workers who attended at least some college−college graduates group. Therefore, the contamination from the demand shock should not be a big concern for the new measurement of quality.

[6]Thanks for this comment from Steven Durlauf and Christopher Taber in 2015 Summer School on Socioeconomic Inequality in UChicago.

2.3 The validity of the new measurement

This new measurement of quality captures the human capital accumulation in the state where workers are born and educated. How is it related to some traditional measurements of human capital accumulation in a birth state, like average SAT score across states? I will run the regression on the new measurement to SAT score for the college group to see whether they have any correlation.

The average SAT scores across states come from National Center for Education Statistics. Available data for SAT scores is discontinuous across years. Since the SAT test was re-centered in 1995, I converted the recentered scale in 1995 to the original scale before 1995 in order to compare the SAT scores across years [7].

Assume people take their SAT test around 16 years old. I will only focus on workers who are 26 years old and attended at least some college in order to match the years in my data with the years of the SAT data. In addition, 26 years old workers' SAT scores are more related to my measurement of quality because they have just worked for a few years rather than older workers whose wage may no longer reflect their SAT scores. Thus, I match 26 years old worker's quality in 1990 to SAT score in 1980; worker's quality in 2000 to SAT score in 1990; worker's quality in 2005 to SAT score in 1995.

Table 4 column 1 shows the regression result, and they have a negative correlation. However, there is one important explanation why this result is strange: perhaps people in different states have a different preference for SAT and ACT. It is possible that the average SAT score is very high in Wisconsin because only top students want to take the SAT in order to apply to universities in the northeast. However, the average quality of college graduates there may be very low. Thus, in column 2, I add one important independent variable: percent of graduates taking the SAT in 1993 [8].

Table 4 column 2 displays that there is nearly no correlation between average SAT scores

[7] http://research.collegeboard.org/programs/sat/data/equivalence/sat-mean

[8] This is the earliest data I can get for the percent of high school graduates taking the SAT across states. Thus, I assume it will not change much from 1980 to 1993, and I will use it for all years 1980, 1990, 1995.

Table 4: Correlation of average SAT scores and the new measurement of quality across states

	AvgSkill Level(1)	AvgSkill Level(2)
SAT Verbal	-0.025**	0.023
	(0.011)	(0.023)
Percent of graduates taking SAT		0.070**
		(0.029)
Observations	150	150
SAT Math	-0.022**	0.013
	(0.010)	(0.018)
Percent of graduates taking SAT		0.060**
		(0.026)
Observations	150	150

Notes: I exclude District of Columbia since there is no average SAT data in 1980.

and the new proposed measurement of quality. This was not expected. At least one important reason can explain why there is no correlation: SAT scores are actually not good to measure the quality of college graduates because the new measurement of quality includes all the human capital accumulation until people graduate from college, and SAT just measures human capital accumulation before people get into college. That means human capital accumulated in college or university is much more important than human capital accumulated before going to college. Thus, there is nearly no correlation between SAT score and my new measurement. I will further explore this point of view.

3 The Empirical Model

3.1 The wage structure

I utilize the framework of wage structure created by Carneiro and Lee (2011). Suppose that the wage of each individual i, of age a, at time t, who is born and goes to school in state b and works in state r (which may or may not be equal to b) can be written as:

$$W^k_{iatrb} = \Pi^k_{atr} * U^k_{i,t-a,b}, \quad (4)$$

where W^k_{iatrb} is the wage, Π^k_{atr} is the price of k-type skill for those with schooling level k [9], in age group a in year t working in state r, and $U^k_{i,t-a,b}$ is the individual specific endowment of k-type skill for those in cohort t-a and in state b. After averaging across individuals in each group, it follows from (4) that

$$W^k_{atrb} = \Pi^k_{atr} * U^k_{atrb}, \tag{5}$$

Thus, the average wage of workers in each group (k, a, t, r, b) is the sum of the price effect (or quantity) part and composition effect (or quality) part. The price effect, Π^k_{atr} is the interaction of the demand for skill with the quantity of skill supplied in the local labor market. The composition effect, U^k_{atrb} is determined by the quality of college and high school graduates across state of births b in a particular year.

Taking logs from equation (5):

$$w^k_{atrb} = \pi^k_{atr} + v^k_{atrb}. \tag{6}$$

This equation is the basis of the empirical models.

3.2 The composition/quality part

The composition part v^k_{atrb} varies over groups because of changes in the average quality of college and high school graduates. If a worker is born in the state with higher average quality, then he will get higher wage in the state where he works after controlling the price or quantity effect. Thus, the main variable of interest in the composition part is average quality of workers for the age group a in particular state of birth b in particular year t: P_{tba}.

Migration is one concern in order to specify P_{tba} in the composition part v^k_{atrb}. Since migration is not random (see, e.g., Heckman, Layne-Farrar, and Todd 1996, and Dahl 2002), it will lead to

[9] k equals two: high school graduates group and college graduates group.

self selection.[10] Thus, I control for selective migration by using time-varying interactions (state of birth)×(state of residence): γ_{ktrb}.

In addition, Carneiro and Lee (2001) hypothesize that the proportion of college-goers in cohort t-a born in region b is one of the important factor determining differences in the average quality of college graduates across regions of birth. They calculate an average proportion of individuals who attend at least some college for each cohort and region of birth (common across years), by averaging this number across all years. However, in my new measure of average quality of both college and high school graduates, I use the average skill level for each group (t,a,b). Moreover, comparing skill levels changes over time is difficult because they are a function of mean wages in each year. The skill percentile for particular occupation, for example CEO, is higher in 1990 than in 1980. This could be the result of higher quality of CEOs, higher demand for their work, or lower supply in 1990, which all influence the mean wage of CEOs. Thus, in the composition/quality part, I add a year fixed effect, γ_t.

School resources may very across states. Well endowed states could have higher quality schools, which simultaneously lead to higher worker quality and affect worker's wage in future. I account for this by using state of birth fixed effects, interacted with age: γ_{kba}

Therefore, the composition variable v^k_{atrb} is formulated as follows:

$$v^k_{atrb} = \gamma_{ktrb} + \gamma_{kba} + \gamma_t + \phi_k(P_{tba}), \tag{7}$$

where γ_{ktrb} is a state-of-birth by state-of-residence fixed effect that is interacted with year dummies (capturing the selective migration). The variables γ_t and γ_{kba} are year fixed effect and state of birth fixed effect. The form $\phi_k(P_{tba})$ is a function of average skill level for the college and high school graduates in state of birth b for each particular age group in particular year (capturing the quality

[10] The intuition is based on the cost-benefit analysis. Persons migrating long distances will tend to have higher earnings in all destination regions than those persons making short moves or no mores at all because their migration costs will be higher too.

effect).

3.3 The reduced form model

After inserting equation (7) into (6), I get the first main reduced form model :

$$w^k_{atrb} = \gamma_{katr} + \gamma_{ktrb} + \gamma_{kba} + \gamma_t + \beta_k * \phi_k(P_{tba}) + \epsilon_{atrb}, \quad (8)$$

where γ_{katr} is full interactions of age-time-state fixed effects (capturing the price effects). The first function form I use for $\phi_k(P_{tba})$ is P^k_{tba}.

Once I control for the price effect, γ_{katr}, on the wage, I can determine the composition effects β_k. The price effect is the interaction of the demand for skill with the quantity of skill supplied in the local labor market. The composition is determined by the quality of college and high school graduates in the state of birth b for particular age group in particular year.

3.4 Identification Problem

In the empirical model, I use state as the geographical unit, but Carneiro and Lee (2011) argue that the reason why they do not use the state as the geographical unit is that the resulting cell sizes would be too small for estimates to be reliable. It is reasonable because there are 2858 groups (51 states*7 age groups*2 education groups*4 years), which leads to no variation within some groups. Thus, the empirical model is hard to identify if I use state level instead of region level. However, it is not likely that quality of college graduates and quality of high school graduates is the same within the census region. Within a state, the quality is more likely to be the same. The identification problem and the quality similarity are the trade off in this scenario. I will omit some interaction of fixed effects in order to identify the model in the state level compared to the model in Carneiro and Lee (2011). Thus, the reduced model I want to estimate is

$$w^k_{atrb} = \gamma_{ra} + \gamma_b + \gamma_t + \gamma_m + \phi_k(P_{tba}) + \epsilon^k_{atrb}, \tag{9}$$

where γ_m is a dummy indicates whether on average there is migration from state of birth b to state of residence r in the group (a,t,k).

3.5 Instrumental Variable

The skill percentile is calculated for the information of mean wages and employment shares of occupations, so the average skill percentile P_{tba} may include some component of the mean wage as the dependent variable. To correct this endogeneity problem, I use an instrumental variable approach. The IV for P_{tba}, $P_{tb(a-1)(a+1)}$, is the weighted average skill level of the adjacent age groups, in the same year, same state of birth, and same education group. For example, for the quality of college graduates in the age group 31-35 in 1980, who were born in New York State, I use weighted average quality of college graduates of age group 25-30 and age group 36-40 in 1980, who were born in New York State, as the IV.

There are two main reasons for creation of this instrumental variable. First, the weighted average quality of college graduates of adjacent age groups is similar as the average quality of college graduates in this particular age group. Second, I can avoid the situation that the dependent variable and independent variable is in the same wage cell because now the IV's age group and dependent variable's age group are different.

However, empirically weak instruments can produce biased IV estimators. Stock and Yogo (2005) tabulate critical values that enable using F-statistic form of the Cragg-Donald (1993) statistic to test whether given instruments are weak. The Cragg-Donald Wald F statistic passes the 10 percent critical value. Thus, rejection of their null hypothesis represents the absence of a weak-instruments problem here.

4 Data

4.1 Data Description

I use cross-sectional data from the 1980, 1990, 2000 US Censuses and 2005 ACS. Compared to Careiro and Lee (2011), the data come from different census years, but the construction of the data are the same. I focus on white males aged from 25 to 60, and aggregate them into seven age groups: 25-30, 31-35, 36-40, 41-45, 46-50, 51-55, 56-60. I consider 51 states of birth and 51 states of residence eliminating from the sample those individuals who are foreign-born. For the education groups, the individuals are divided into two categories: high school graduate and college graduate. Thus, individuals are grouped into cells defined by five variables: year, state of residence, state of birth, schooling (high school or college) and age group.

For each cell I compute the relevant log of average weekly wages for full-time/full-year workers. Weekly wages for high school graduates are obtained by taking only males with exactly 12 years of schooling and dividing annual income by annual weeks worked. Weekly wages for college graduates are obtained in an analogous way for individuals with exactly 16 years of schooling.[11]

4.2 Summary Statistics

Descriptive statistics are summarized in Table 5. In Panel A, each column reports the sample mean and standard deviation of the dependent variable: real weekly wage for cell (a,r,b) in different years for college groups. In both panels, the sample mean and the standard deviation of the wage increases over time. In addition, Table 5 also reports the ratio of the mean wage for college graduates to the mean wage for high school graduates. It also increases over time, which suggests the education premium has increased. However, it increases at slower pace over time.

[11]Carneiro and Lee (2011) argues that in order to get measures of wages for clearly defined and relatively homogeneous groups of individuals, they ignore high school dropouts, those with some college, those with post-graduate studies.

Table 5: Summary Statistics

	1980	1990	2000	2005
Panel A. College				
Dependent Variable: Weekly Wage	233.19	684.23	1310.57	1800.20
	(91.25)	(346.60)	(765.08)	(1158.23)
Observations (groups)	11140	12731	13664	9000
Panel B. High School				
Dependent Variable: Weekly Wage	169.51	422.37	765.04	1018.23
	(54.77)	(170.92)	(368.47)	(552.25)
Observations (groups)	12717	12758	14258	9072
Ratio W^C/W^H	1.38	1.62	1.71	1.77

Notes:(1) The dependent variable is real weekly wage (adjusted to 1999 dollars). Weekly wages for high school graduates are obtained by taking only males with exactly 12 years of schooling. Weekly wages for college graduates are obtained in an analogous way for individuals with exactly 16 years of schooling. (2) Ideally, the number of observations should be 18207 (51 birth states*51 residence state* 7 age groups) in each education category each year. Since there are so many groups, some groups may be quite small, and there will be no observations. For example, there is no observation in age group 36-40, born in Delaware, working in Washington in 1980 for the FTFY white male workers.

5 Empirical Results

Table 6 reports estimation results that are obtained by implementing the econometric framework described in Section 3. Robust standard errors are reported in the table, clustered by year and state of birth.

In column 1 of Table 6, I only use the year fixed effect in the regression. Since we cannot compare the skill percentile for occupation across years as mentioned early, the year fixed effect is a very important control variable. Both estimates are significant for the two education groups.

Column 2 of Table 6 shows the estimation results for the reduced model (9). If the quantity part is fixed, both college graduates and high school wages respond substantially to changes in the quality of workers. If the quality of high school graduates increases by one percentile, on average their wages will increase by 0.67 percent. Further, if the quality of college graduates increases by one percentile, on average their wages will increase by 0.98 percent.

In column 3 of Table 6, I use the IV for the average skill level for the particular age group in each state of birth in each year. Column 3 only includes the year fixed effect. The magnitude of

Table 6: Regression of Wages on Quality of College and High School Gradues

	Reduced-form model (1)	Reduced-form model (2)	2SLS (3)	2SLS (4)
Panel A. College				
Quality of college graduates	4.075***	0.984***	3.792***	1.165***
	(0.131)	(0.119)	(0.194)	(0.254)
R^2	0.742	0.786	0.742	0.786
Observations	46535	46535	46535	46535
Panel B. High School				
Quality of high school graduates	3.023 ***	0.670***	2.867***	0.806**
	(0.114)	(0.111)	(0.157)	(0.387)
R^2	0.776	0.806	0.776	0.806
Observations	48805	48805	48805	48805
Included Year Fixed Effect	Yes	Yes	Yes	Yes
Included Other Fixed Effects	No	Yes	No	Yes

Notes: The dependent variable is log real weekly wage in each cell. The independent variable “Quality of college graduates” is measured by the average skill percentile in each age group, each region of birth and each year, which ranges from 0 to 1.

Significant codes: 0.01 '***' 0.05 '**' 0.1 '*'.

estimates is smaller than those in column 1 of Table 6. However, in column 4 of Table 6, when I use the IV and estimate the reduced form model (9) with all fixed effects included, the estimates are both statistically and economically significant and they are larger in magnitude than the estimates in column 2 of Table 6. If the quality of high school graduates increases by one percentile, on average the high school graduates wages will increase by 0.81 percent.

Since the college premium at time t is defined as $E({w_t}^C) - E({w_t}^H)$, this requires subtracting the college and high school wage equations and averaging across all ages, state of birth and state of residence. Given the wage equation is linear in all variables, I compute the effect of quality on the college premium as the difference between the coefficients on P_{tab} in college and high school equations. In column 4 of Table 6, the difference is 0.36, implying that the college premium increases by 0.36 percent if the quality of both high school and college graduates increases by one percentile.

6 Robustness Checks

In section 2, I argued that the skill index may be contaminated by price level effects. In order to test this conjecture, I first regress the weekly real wage on the city dummies, then use the residual wage to construct the skill percentile. The average skill level is calculated by the adjusted skill percentile according to section 2. The estimation results for the reduced form model (9) are showed in columns (1)-(4) of Table 7 . In column 4 of Table 7, the difference of the two estimates is 0.36, implying that the college premium increases by 0.36 percent if the quality of both high school and college graduates increases by one percentile. The magnitude of this result is similar as the magnitude calculated Table 6, which indicates the result is robust[12].

As an alternative specification, I use the functional form

$$\phi_k(P_{tba}) = P_{tab}/(1 - P_{tab})$$

for $\phi_k(P_{tba})$. As Carneiro and Lee (2011) argue, this function is a strictly increasing function of P_{tab} and can take any nonnegative value. The estimation results are shown in columns (5)-(6) of Table 7. Column 5 shows that for a quality increase of high school graduates from the 50th percentile to the 60th percentile (i.e. $\phi_k(P_{tba})$ increases from 1 to 1.5), the average high school graduates' wage increases by 0.10 percent. Column 6 shows that the education premium will decrease by 0.02 percent if the quality of both high school graduates and college graduates increases from the 50th percentile to the 60th percentile.

[12]Besides this method to check the price level effects, the skill percentile can be calculated by using real weekly wage. Here this real weekly wage is adjusted by cities' CPI index, not national CPI. It is considered for future research.

Table 7: Regression of Wages on Quality of College and High School Gradues

	Model (1)	Model (2)	2SLS (3)	2SLS (4)	Model (5)	2SLS (6)
Panel A. College						
Quality of college graduates	4.127***	0.999***	3.962***	1.356***	0.125***	0.131***
	(0.117)	(0.121)	(0.179)	(0.283)	(0.017)	(0.327)
R^2	0.744	0.786	0.744	0.786	0.786	0.786
Observations	46535	46535	46535	2268	46535	46535
Panel B. High School						
Quality of high school graduates	3.367***	0.703***	3.403***	0.998**	0.207***	0.173
	(0.119)	(0.122)	(0.166)	(0.480)	(0.041)	(0.129)
R^2	0.777	0.806	0.777	0.806	0.806	0.806
Observations	48805	48805	48805	48805	48805	48805
Included Year Fixed Effect	Yes	Yes	Yes	Yes	Yes	Yes
Included Other Fixed Effects	No	Yes	No	Yes	Yes	Yes
Using residual wage to construct skill percentile	Yes	Yes	Yes	Yes	No	No
New Function Form	No	No	No	No	Yes	Yes

Notes: The dependent variable is log real weekly wage in each cell. The independent variable "Quality of college graduates"' is measure by the average skill level in each age group, each region of birth and each year, which ranges from 0 to 1.

Significant codes: 0.01 '***' 0.05 '**' 0.1 '*'.

7 One Dimension of Quality of Workers: Evidence from College Enrollment

There are several important factors influencing the quality of workers, and one of them is the college enrollment. Figure 2 shows that the college enrollment is increasing over time, and it varies across states of birth in my sample. Two mechanisms in the literature explain why the increase of college enrollment will decrease the quality of college graduates. Since the college enrollment increases, the marginal student, whose quality is lower than the quality of average students in college, is drawn into college thus it results in a decline in average quality. The other mechanism focuses on the averages. The rise in college attendance means that resources at each college have to be spread more thinly across students, resulting in a lower quality education for each enrolled student. As Carneiro and Lee (2011) argues, it is hard to distinguish those two mechanism empirically.

It is certainly a challenge and very interesting question. However, I will not try to distinguish those two mechanism here because in the scenario of this paper, quality is actually a weighted occupation index. It is interesting to ask how the increase of college attendance affects the occu-

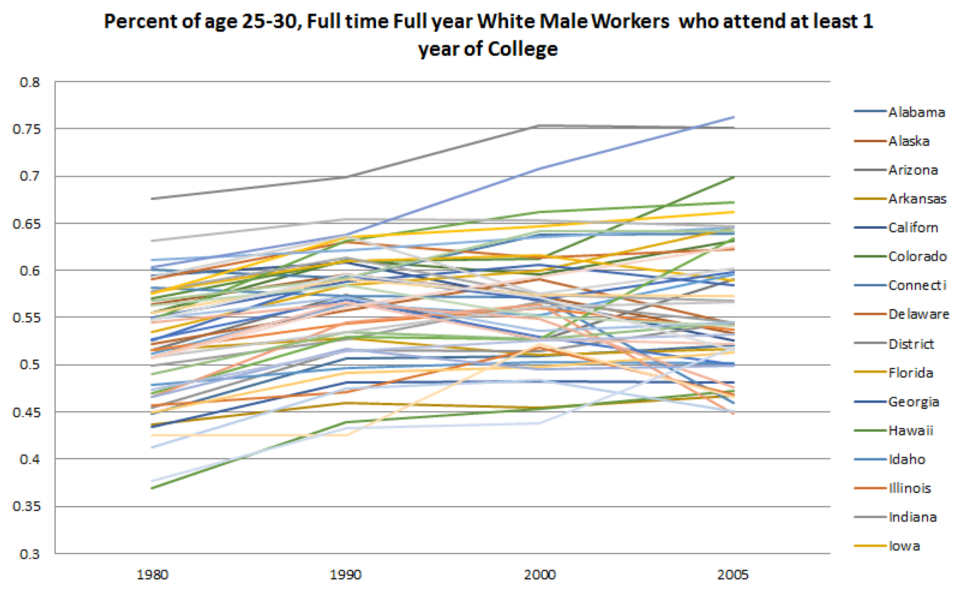

Figure 2

pation distribution of college graduates. The mechanism is that higher college enrollment pushes college graduates entering less skilled occupations, which indicates their quality decreases on average. Thus, college enrollment is causally effecting the quality of college graduates in the state of birth.

However, it is unlikely that college enrollment is the only important factor determining differences in the average quality of workers across the states of birth. For example, school resources may vary across states. Well endowed states (e.g. New York State) could have higher quality schools, which simultaneously lead to higher worker quality and higher college enrollment. Thus, in order to catch the causal effect, I need to account the state of birth fixed effects. I interact the state of birth with age: γ_{ba}, which indicates state of birth has different effect on different cohort. The empirical model is as following:

$$P^k_{tba} = \gamma_{ba} + \gamma_t + \beta_k E_{tba} + \epsilon_{atrb}, \quad (11)$$

where P^k_{tba} is the average quality of workers in age group a, state of birth b, education group k in

Table 8: Regression of Quality of College and High School Graduate on College Enrollment

	Reduced-form model
Panel A. College	
Proportion in college	-0.109***
	(0.012)
Observations	1428
Panel B. High School	
Proportion in college	-0.014
	(0.011)
Observations	1428

Notes: The dependent variable "Quality of college graduates" is measured by the average skill level in each age group, each state of birth and each year, which ranges from 0 to 1. The independent variable is the the proportion of individuals who attend at least some college in the sample, which ranges from 0 to 1.

Significant codes: 0.01 '***' 0.05 '**' 0.1 '*'.

year t; E_{tba} is the college enrollment for the same group[13]; γ_t is year fixed effect.

Table 8 reports estimation results. Robust standard errors are reported, clustered by state of birth and year. The results show that one percent increase in college enrollment leads to 0.11 percent decrease in average quality of college graduates. But one percent increase in college enrollment only causes 0.01 percent decrease in average quality of high school graduates. Therefore, the college enrollment has larger impact on the quality of college graduates than the quality of high school graduates.

It is clear that increase of college enrollment will decease the quality of college graduates. But why the estimate is negative even for the high school group, though it is not significant? The intuition is that the original marginal "good" high school graduates can now go to college because the college enrollment increases. This further pushes down the average quality of high school graduates.

[13]The college enrollment is defined as the proportion of individuals who attend at least some college in the sample.

8 Conclusion

In this paper, I develop a new measurement of quality of high school and college graduates, the average skill level in each state of birth. This measurement links the worker's quality to skill level of occupations. I use that measure to estimate the effect of relevant worker quality on the observed education premium, finding that increasing both the quality of college and high school graduates by one percent will cause 0.36 percent of increase in the education premium. In addition, one of the key factor influencing quality is the college enrollment in each state of birth. I find that increases in college enrollment will lead to the decline of quality of college graduates more than the quality of high school graduates by nearly 0.1 percent.

It is still puzzling why there is nearly no correlation between average SAT scores and the new proposed measurement of quality. Though those two measure different time periods of human capital accumulation, the details remain obscure. Thus, I will further explore that correlation.

References

[1] Acemoglu, Daron. 2000. Technical Change, Inequality and the Labor Market. *Journal of Economic Literature*, 40: 7-72.

[2] Autor, David H., Lawrence F. Katz, and Melissa S. Kearney. 2008. Trends in U.S. Wage Inequality: Revising the Revisionists. *Review of Economics and Statistics*, 90(2): 300-323.

[3] Autor, David H and Dorn, David. 2013. The Growth of Low-Skill Service Jobs and the Polarization of the U.S. Labor Market. *American Economic Review*, 103(5): 15531597.

[4] Brooks Pierce. 2001. Compensation Inequality. *The Quarterly Journal of Economics*, 116(4): 1493-1525.

[5] Card, David, and Alan Kruege. 1992. Does School Quality Matter: Returns to Education and the Characteristics of Public Schools in the United States, *Journal of Political Economy*, 100(1): 1-40.

[6] Coen Teulings. 1995. The Wage Distribution in a Model of the Assignment of skills to Jobs. *Journal of Political Economy*, 103(2): 280-315.

[7] Dorn David. 2009. Essays on Ineqaulity, Spatial Interaction, and the Demand for Skills. *Dissertation University of St. Gallen* no. 3613, September.

[8] Edward L. Glaeser, Matthew G. Resseger. 2010. The Complementarity between Cities and Skills. *Journal of Regional Science*, 50(1): 221-244 .

[9] Heckman, James, Anne Layne-Farrar, and Petra Todd. 1996. Human Capital Pricing Equations with an Application to Estimating the Effect of Schooling Quality on Earnings. *Review of Economics and Statistics*, 78(4): 562-610.

[10] Pedro Carneiro and Sokbae Lee. 2011. Trends in Quality Adjusted Skill Premia in the United States. *American Economic Review*, 101(6): 2309-2349.

[11] Stock, J. H., and M. Yogo. 2005. Testing for weak instruments in linear IV regression. In Identification and Inference for Econometric Models: Essays in Honor of Thomas Rothenberg, ed. D. W. K. Andrews and J. H. Stock, 80-108. *Cambridge: Cambridge University Press*.

[12] John, Winters. 2012. Differences in employment outcomes for college town stayers and leavers. *IZA Journal of Migration*: 1-11.