Online Appendix to "Trends in Quality-Adjusted Skill Premia in the United States, 1960-2000"

By Pedro Carneiro and Sokbae Lee*

This online appendix provides a more detailed description of data and some additional empirical results.

Appendix

Data

We use data from the 1960, 1970, 1980, 1990 and 2000 US Censuses (1% sample). Our data was extracted from `http://www.ipums.umn.edu/` (see Ruggles et al., 2004). We focus on white males, between the ages 25 and 60. We exclude from the analysis all those who are foreign born. We consider 5 year age groups (with the exception of the first, which has 6 years): 25-30, 31-35, 36-40, 41-45, 46-50, 51-55.

For 1960, 1970 and 1980, the education variable we use is "highest grade of schooling", while for 1990 and 2000 we use "educational attainment recode". We group individuals into four categories: high school dropout, high school graduate, some college, college graduate or more. For 1960, 1970 and 1980, we consider high school dropouts those who have completed less than 12 years of schooling, high school graduates are those with exactly 12 years of completed schooling, college graduates have completed 4 or more years of college, and some college is the category for those with more than 12 but less than 16 years of schooling. For 1990 and 2000, dropouts are those with up to 11 years of schooling, high school graduates have exactly 12 years of schooling, those with some college have 1 to 3 years of college, and college graduates have four or more years of college. Our final classification has two groups only, one comprising of high school graduates and high school dropouts, and the other comprising of those with some college, a college degree, or above.

We compute weekly wages by dividing annual wage and salary income by annual weeks worked. We deflate all wages to 1990 values using the CPI-U from the Economic Report of the President. In order to compute average log wages for each year-age group-region of residence-region of birth cell (the main outcome variable in our analysis) we drop all observations for whom real wages are below 50 dollars per week.

For each year, region of birth and five year cohort we also estimate the proportion of individuals who attend at least some college. However, even among adults, educational attainment increases over time. Therefore we 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, using as weights the number of individuals in each year, cohort and region of birth cell.

We consider 9 regions of birth and 9 regions of residence, and we drop from the sample those individuals who are foreign born. In particular, we use the regions defined by the

* Carneiro: Department of Economics, University College London, Gower Street, London WC1E 6BT, United Kingdom, and the Institute for Fiscal Studies (e-mail: p.carneiro@ucl.ac.uk); Lee: Department of Economics, Seoul National University, 599 Gwanak-ro, Gwanak-gu, Seoul, 151-742, Republic of Korea, and the Institute for Fiscal Studies (e-mail: sokbae@gmail.com).

Census: New England (Connecticut, Maine, Massachusetts, New Hampshire, Rhode Island, Vermont), Middle Atlantic (New Jersey, New York, Pennsylvania), East North Central (Illinois, Indiana, Michigan, Ohio, Wisconsin), West North Central (Iowa, Kansas, Minnesota, Missouri, Nebraska, North Dakota, South Dakota), South Atlantic (Delaware, District of Columbia, Florida, Georgia, Maryland, North Carolina, South Carolina, Virginia, West Virginia), East South Central (Alabama, Kentucky, Mississippi, Tennessee), West South Central (Arkansas, Louisiana, Oklahoma, Texas), Mountain (Arizona, Colorado, Idaho, Montana, Nevada, New Mexico, Utah, Wyoming), Pacific (Alaska, California, Hawaii, Oregon, Washington).

We group individuals into cells defined by five variables: schooling (high school or college), year, age group, region of residence, and region of birth. The reason we do not use the state as the regional unit is that the resulting cell sizes would be too small for our estimates to be reliable.

For each cell we compute the relevant average weekly log wages for full-time/full-year workers, log total weeks worked (a measure of labor supply) and the proportion of individuals in college (a measure of composition). Full-time/full-year workers are individuals who work at least 35 hours a week and 40 weeks a year. In order to construct hours worked per week we use usual hours worked in a week for the year of 1980, 1990 and 2000, and hours worked the week before the interview for 1960 and 1970. We exclude farm workers from the group of full-time/full year workers. The construction of wages and weeks worked described in this section is based on Card and Lemieux (2001). Weekly wages for high school graduates are obtained by taking only males with exactly 12 years of schooling and dividing annual income from wages by annual weeks worked. Weekly wages for college graduates are obtained in an analogous way, but considering only individuals with exactly 16 years of schooling. Unfortunately, for the 1960 and 1970 US Censuses weeks worked are only available in intervals: 1 to 13, 24 to 26, 27 to 39, 40 to 47, 48 to 49, 50 to 52. For these two years we take the midpoint of each interval as our estimate of weeks worked.

Log weeks worked by high school graduates (or high school equivalents) are a weighted sum of weeks worked by white males in each region of residence, who can be high school dropouts, high school graduates, and even individuals with some college. Log weeks worked by college graduates are a weighted sum of weeks worked by white males with some college, a college degree, or post-graduate studies.

The weights for these sums are defined as follows. Each high school dropout week is only a fraction of a regular high school graduate week. This fraction corresponds to the relative wage of high school dropouts and high school graduates. Similarly, each week worked by an individual with a post-graduate degree has a larger weight than a week worked by a college graduate, and the weight is given by the relative wage of post-graduates and college graduates. Finally, in order to construct the weights for some college weeks, we first look at the difference between high school graduate and college graduate wages. If the difference between some college wages and high school graduate wages is say, one third of the difference between college and high school wages, then we assign one third of some college weeks to high school, and two thirds to college. We allow these weights to vary across age groups, but not across year or region.

Average and Marginal Students

Table A-1 in the appendix simulates the effects of declining quality of marginal college graduates on their average log wages for different specifications of the model, and at different levels of P. We consider a simple economy that consists of only two schooling groups: high school graduates and college graduates. Let w_1 be the average log wage of

Table A-1—Changes in the log wages of marginal college graduates

(1) Baseline College Enrollment	(2) Changes in College Enrollment	(3) Proportion of Marginal College Graduates	(4) Changes in Log Wages of Marginal College Graduates	(5)
Model Specification			Linear in $P/(1-P)$	Linear in P
Panel A - Absolute Increase of 10% in College Enrollment				
0.10	0.10	0.50	-0.02	-0.10
0.20	0.10	0.33	-0.05	-0.16
0.30	0.10	0.25	-0.08	-0.21
0.40	0.10	0.20	-0.14	-0.26
0.50	0.10	0.17	-0.25	-0.30
0.60	0.10	0.14	-0.51	-0.37
Panel B - Proportional Increase of 10% in College Enrollment				
0.10	0.01	0.09	-0.01	-0.06
0.20	0.02	0.09	-0.03	-0.11
0.30	0.03	0.09	-0.06	-0.17
0.40	0.04	0.09	-0.11	-0.23
0.50	0.05	0.09	-0.21	-0.29
0.60	0.06	0.09	-0.42	-0.34

Notes: This table shows the effects of declining quality of marginal college graduates on their average log wages. Consider an economy that consists of only two schooling groups: high school graduates and college graduates. Let w_1 be the average log wage of the baseline college graduates and w_2 be the average log wage of the marginal college graduates. Note that the average log wage after college expansion, say w_*, has the form:

$$w_* = (1 - \text{column}(3)) \times w_1 + \text{column}(3) \times w_2.$$

If the model is linear in the odds of college enrollment ($P/(1-P)$), then $w_* = w_1 - 0.086 \times$ (changes in $P/(1-P)$). Thus, solving for w_2 gives $w_2 = w_1 + \text{column}(4)$, where $\text{column}(4) = -0.086 \times$ (changes in $P/(1-P)$)/column(3). If the model is linear in the level of college enrollment (P), then $w_* = w_1 - 0.517 \times$ (changes in P). Thus, $w_2 = w_1 + \text{column}(5)$, where $\text{column}(5) = -0.517 \times$ (changes in P)/column(3). Panel A of the table shows the effects of absolute increases of 10% in college enrollment and Panel B shows those of proportional increases of 10% in college enrollment.

the baseline college graduates and w_2 be the average log wage of the marginal college graduates. Note that the average log wage after college expansion, say w_*, has the form:

$$w_* = (1-\mu) \times w_1 + \mu \times w_2, \quad \text{(A-1)}$$

where μ is the proportion of marginal college graduates after college expansion. Since our model is linear in the odds of college enrollment ($P_{t-a,b}/(1-P_{t-a,b})$), in view of the estimation result in column (1) of Table 1 in the main text, we have

$$w_* = w_1 - 0.086 \times (\text{changes in } P_{t-a,b}/(1-P_{t-a,b})). \quad \text{(A-2)}$$

Thus, combining (A-1) and (A-2) and solving for w_2 gives

$$w_2 = w_1 - 0.086 \times \mu^{-1} \times (\text{changes in } P_{t-a,b}/(1-P_{t-a,b})).$$

The column (4) of Panel A of Table A-1 shows the effects of absolute increases of 10% in college enrollment across different baseline college enrollment rates and the same column of Panel B shows those of proportional increases of 10% in college enrollment. For example, if the baseline college enrollment rate is 40%, then an absolute increases of 10% in college enrollment leads to a decrease of 14% in marginal college wages, thereby implying that the quality of the average marginal college graduate is 14% lower than that of the average baseline college graduate. It can be seen that marginal college graduates are always of lower quality than baseline college graduates, but the magnitude of the decrease in quality is larger when there is a higher proportion of baseline college graduates. In other words, as college expands, it would induce lower and lower quality college graduates. Our main qualitative results do not change when an alternative model specification is used. column (5) of Table A-1 reports that the main results are the same when the model is linear in the level of college enrollment ($P_{t-a,b}$).[1]

Notice that changes in worker quality resulting from increases in college participation do not necessarily have to lead to decreases in the college premium. Theoretically, the adjustment in the college premium could go either way, depending on how individuals sort into different levels of schooling, and on how important heterogeneity is in high school and college (see Carneiro and Lee, 2009). Our results indicate that: i) skill heterogeneity and self selection into schooling are important phenomena, so that if college enrollment went from 40% to 50%, the average marginal student's quality would be 14% lower than that of the average student in college; ii) those individuals with the highest college skills select into college; iii) there is no clear relationship between the type of skills used in high school occupations and selection into college.

Additional Specification Checks

In this section we present additional specification checks. In Table A-2 we examine how results change when we drop one of the nine regions at a time from the sample, to check whether results are driven by one region alone as opposed to being a national phenomenon. The table has 10 columns. The first one corresponds to our original specification, where all regions are included, and in the remaining ones we drop from the sample one region

[1] In this case, the corresponding coefficient for $P_{t-a,b}$ is -0.517, implying that

$$w_2 = w_1 - 0.517 \times \mu^{-1} \times (\text{changes in } P).$$

Table A-2—Sensitivity Analysis: Dropping One Region at a Time

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	All	Dropping the Following Region:								
	Regions	New England	Middle Atlantic	East North Central	West North Central	South Atlantic	East South Central	West South Central	Mountain	Pacific
Panel A - College										
Odds of Proportion in College	-0.086	-0.085	-0.065	-0.093	-0.092	-0.091	-0.083	-0.089	-0.080	-0.102
	[0.036]**	[0.035]**	[0.031]**	[0.039]**	[0.038]**	[0.044]**	[0.037]**	[0.037]**	[0.034]**	[0.038]**
Observations	2598	2361	2294	2287	2325	2289	2337	2300	2305	2286
Panel B - High School										
Odds of Proportion in College	-0.032	-0.034	-0.045	-0.034	-0.034	-0.025	-0.034	-0.032	-0.031	-0.039
	[0.022]	[0.022]	[0.020]**	[0.022]	[0.022]	[0.021]	[0.023]	[0.023]	[0.024]	[0.026]
Observations	2692	2441	2391	2379	2388	2381	2416	2384	2379	2377

Notes: The dependent variable is log weekly wage in each cell. All regressions are weighted by the inverse of the sampling variance of average log wages in each cell. Robust standard errors in brackets, clustered on the region of residence-schooling-year cell. * significant at 10%; ** significant at 5%; *** significant at 1%.

Table A-3—Sensitivity Analysis: Dropping One Year at a Time

	(1) All Years	(2) 1960	(3) 1970	(4) 1980	(5) 1990	(6) 2000
		Dropping the Following Year:				
Panel A - College						
Odds of Proportion in College	-0.086	-0.087	-0.082	-0.051	-0.101	-0.116
	[0.036]**	[0.037]**	[0.043]*	[0.028]*	[0.053]*	[0.051]**
Observations	2598	2148	2099	2068	2043	2034
Panel B - High School						
Odds of Proportion in College	-0.032	-0.026	-0.031	-0.037	-0.029	-0.037
	[0.022]	[0.022]	[0.029]	[0.020]*	[0.020]	[0.034]
Observations	2692	2203	2155	2140	2137	2133

Notes: The dependent variable is log weekly wage in each cell. All regressions are weighted by the inverse of the sampling variance of average log wages in each cell. Robust standard errors in brackets, clustered on the region of residence-schooling-year cell. * significant at 10%; ** significant at 5%; *** significant at 1%.

at a time. Across columns, the coefficient of interest in the college equation is always negative and statistically significant, while in the high school equation the coefficient is small and insignificant.

TABLE A-4—REGRESSION OF WAGES WITH MEASURES OF WITHIN GROUP COMPOSITION

	(1) Basic	(2) Add Within Group Composition	(3)	(4)
Panel A - College				
Odds of Proportion in College	-0.086	-0.086	-0.087	-0.087
	[0.036]**	[0.037]**	[0.035]**	[0.035]**
Within Odds of Proportion in Some College		0.016		0.008
		[0.115]		[0.106]
Within Odds of Proportion in Post-College Enrollment			-0.078	-0.074
			[0.221]	[0.196]
Observations	2598	2598	2598	2598
Panel B - High School				
Odds of Proportion in College	-0.032	-0.021		
	[0.022]	[0.022]		
Within Odds of Proportion in Dropout		-0.023		
		[0.011]*		
Observations	2692	2692		

Notes: The dependent variable is log weekly wage in each cell. The regression includes basic controls as in column (1) of Table 1 in the main text. Robust standard errors in brackets, clustered on the region of residence-schooling-year cell. * significant at 10%; ** significant at 5%; *** significant at 1%.

In Table A-3 we examine the sensitivity of our results to dropping one year at a time.[2] In the first column of the table we present the original specification, while in the remaining five columns we drop from the sample one year at a time. Across columns there are relatively small changes in the coefficient of interest, both in the college and in the high school equation.

Second, since we aggregate individuals into two levels of schooling only, one may worry that changes in college attainment also lead to changes in composition within each of these aggregates (between dropouts and high school graduates, those with some college and college graduates, or those between college graduates and post-college-educated workers). This is likely to be the case, but it does not seem to affect our results. Table A-4 shows that the main results are basically unchanged once we include measures of within group composition (the within odds of proportion in dropout, the within odds of proportion in some college, and the within odds of proportion in post college enrollment).

Table A-5 presents estimation results for the high school wage equation with alternative specifications and samples. Basically, there is no composition effect for high school graduates across different specifications.

Figure A-1 shows estimated functional forms of ϕ_k with alternative specifications in the reduced-form model for $k = C, H$. The top panels of the figure show alternative specifications of ϕ_C for log college wages in the reduced-form model. In particular, the alternative forms of ϕ_C are as follows: linear in $P/(1-P)$, P, $\log P/(1-P)$, or

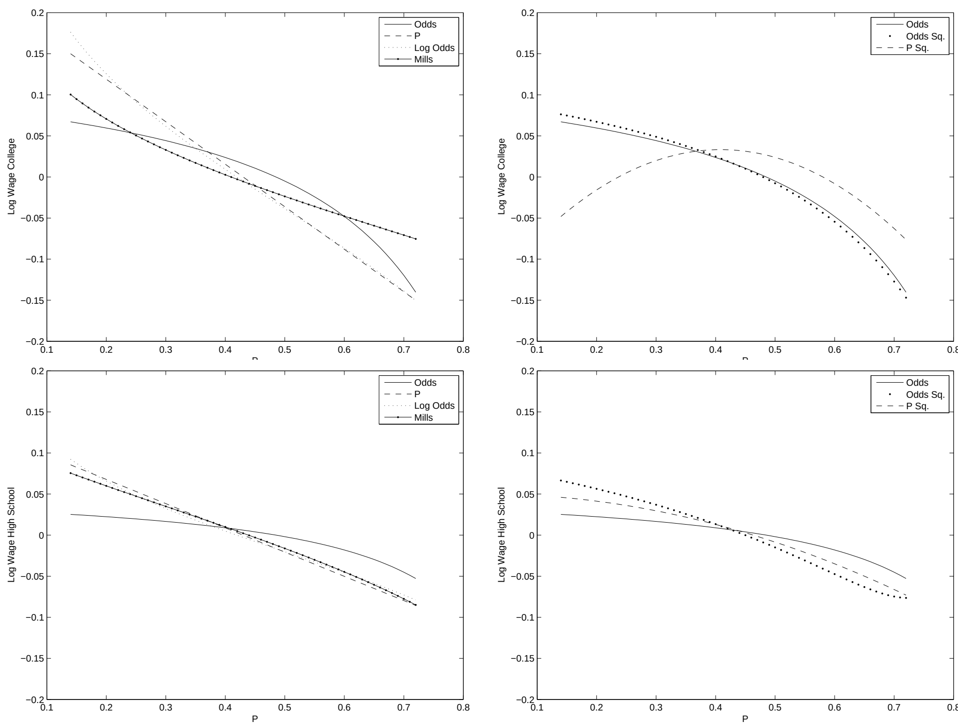
[2]If we estimate the model decade by decade the coefficient on variable of interest is always negative but not statistically significant.

Table A-5—Sensitivity Analysis: Alternative Specifications and Samples

	(1)	(2)	(3)	(4)
The Dependent Variable: Log High School Wages				
White Males (Original Sample)				
Odds of Proportion in College	-0.032 [0.022]	-0.125 [0.067]*		
(Odds of Proportion in College)^2		0.024 [0.016]		
Proportion in College			-0.295 [0.186]	0.005 [0.542]
(Proportion in College)^2				-0.245 [0.377]
Observations	2692	2692	2692	2692
P-value	0.14	0.14	0.12	0.16
White Females				
Odds of Proportion in College	-0.011 [0.023]	-0.009 [0.061]		
(Odds of Proportion in College)^2		-0.001 [0.014]		
Proportion in College			-0.134 [0.167]	-0.209 [0.332]
(Proportion in College)^2				0.081 [0.259]
Observations	2538	2538	2538	2538
P-value	0.63	0.88	0.43	0.73
Both White Males and Females				
Odds of Proportion in College	-0.027 [0.016]	-0.107 [0.048]**		
(Odds of Proportion in College)^2		0.023 [0.013]*		
Proportion in College			-0.204 [0.126]	0.107 [0.392]
(Proportion in College)^2				-0.286 [0.298]
Observations	2754	2754	2754	2754
P-value	0.11	0.07	0.11	0.06

Notes: Robust standard errors in brackets, clustered on the region of residence-schooling-year cell. * significant at 10%; ** significant at 5%; *** significant at 1%.

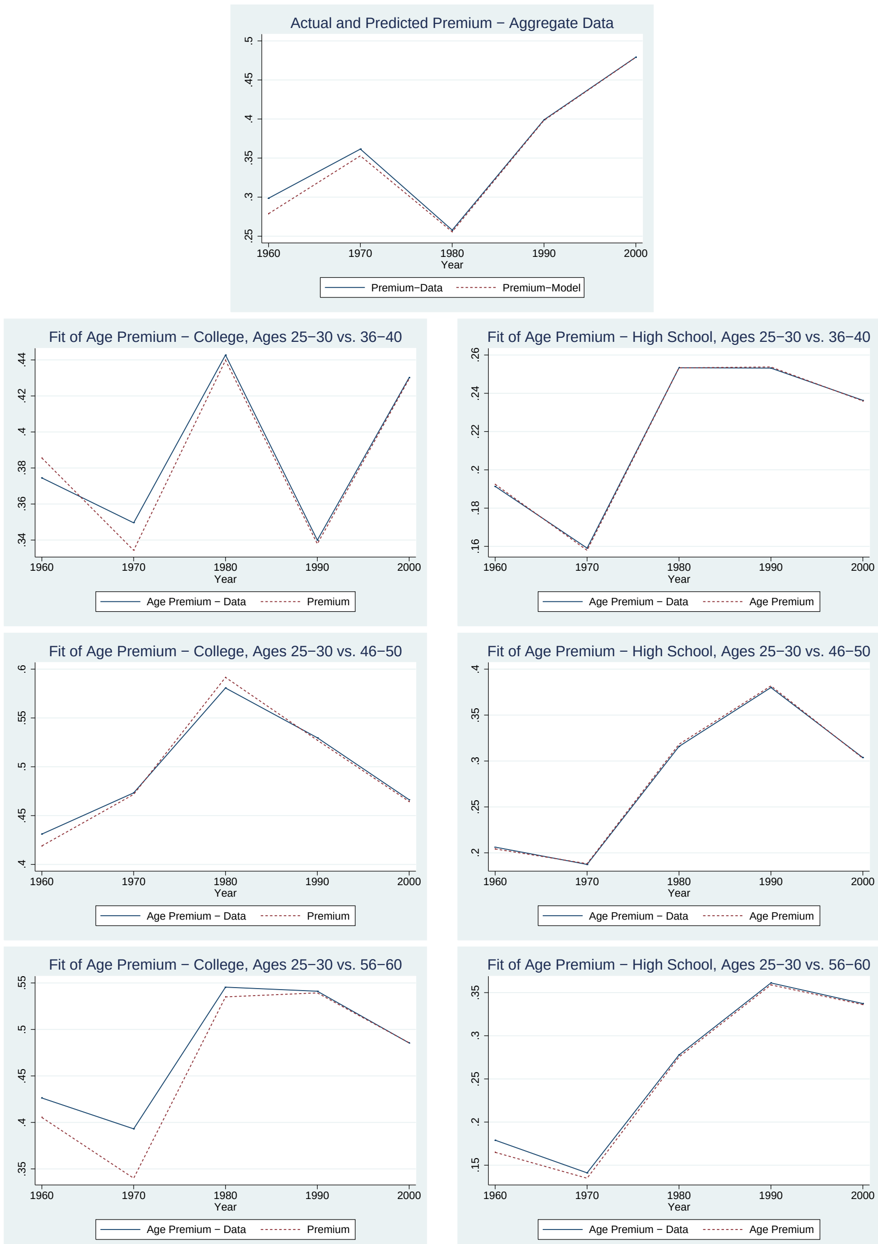
FIGURE A-1. ALTERNATIVE SPECIFICATIONS OF ϕ_k IN THE REDUCED-FORM MODEL



*

Notes: The top panels of the figure show alternative specifications of ϕ_C for log college wages in the reduced-form model. In particular, the alternative forms of ϕ_C are as follows: linear in $P/(1-P)$, P, $\log P/(1-P)$, or $\phi(\Phi^{-1}(P))/P$ (the inverse Mills ratio based on the normality assumption) [top-left panel] and also linear and quadratic in $P/(1-P)$ and P [top-right panel], where P is the birth-cohort/region-of-birth specific college enrollment rates. The bottom panels show alternative specifications of ϕ_H for log high-school wages in the reduced-form model. The alternative forms of ϕ_H are the same as the college equation except that the inverse Mills ratio for high school is now $\phi(\Phi^{-1}(P))/(1-P)$. In the college equation, all the p-values for the hypothesis that ϕ_C is significantly different from zero are less than or equal to 0.06, except for the cases of an inverse Mills ratio (p-value: 0.39) and the log of odds (p-value: 0.12). In the high school equation, the smallest and largest p-values across all specifications are 0.10 and 0.26.

Figure A-2. The Fit of the Reduced Form Model



*

Notes: This figure compares data with the fit of model based on column (1) of Table 1 in the main text. The top panel shows the fit of the college premium and the three left-side panels and the three right-hand side panels show the fit of the college age premia and the high school age premia, respectively.

$\phi(\Phi^{-1}(P))/P$ (the inverse Mills ratio based on the normality assumption) [top-left panel] and also linear and quadratic in $P/(1-P)$ and P [top-right panel], where P is the birth-cohort/region-of-birth specific college enrollment rates. The bottom panels show alternative specifications of ϕ_H for log high-school wages in the reduced-form model. The alternative forms of ϕ_H are the same as the college equation except that the inverse Mills ratio for high school is now $\phi(\Phi^{-1}(P))/(1-P)$. In the college equation, all the p-values for the hypothesis that ϕ_C is significantly different from zero are less than or equal to 0.06, except for the cases of an inverse Mills ratio (p-value: 0.39) and the log of odds (p-value: 0.12). In the high school equation, the smallest and largest p-values across all specifications are 0.10 and 0.26.

The fit of the reduced form model is presented in Figure A-2. The figure compares data with the fit of model based on column (1) of Table 1 in the main text. The top panel shows the fit of the college premium and the three left-side panels and the three right-hand side panels show the fit of the college age premia and the high school age premia, respectively. Overall, the fit of the model is very good, especially for 1980-2000.

Additional Direct Evidence on Declining Quality of College Workers

In this section we provide additional direct evidence on declining quality of college workers by analyzing data from the original cohorts of the National Longitudinal Survey of Young Males (NLS66), the National Longitudinal Survey of Youth of 1979 (NLSY79), and the National Longitudinal Survey of Youth of 1997 (NLSY97). For simplicity, we present results comparing only North and South, but we have available results for finer regional partitions which show similar patterns. We take white males, and we divide them in two groups: those growing up in the southern regions of the US, and those growing up elsewhere. For each individual in each region we compute his percentile in the distribution of test scores within region (the AFQT in the NLSY79 and NLSY97, and IQ in NLS66).[3] Finally, we divide the sample into those with some college or more, and those with less than college, and calculate the average percentile in the test score distribution for each education group and region. The question we ask is: in regions where college participation is higher is the average college student in a lower percentile of the within-region test score distribution?

Results are shown in table A-6. The number in each cell in columns (1) and (2) corresponds to the average percentile in the within-region test score distribution for each dataset, region and schooling group. Columns (3) and (4) show the proportion of individuals in each sample with at least some college, and the corresponding odds. For example, in the NLSY79 the college participation rate is 56% in the North and 48% in the South, and the average percentile of college student in the within-region test score distribution is 64% in the North and 69% in the South. Therefore, in the North, where the levels of college attendance are higher, the average college student has lower quality (relatively to the other residents in the region) than in the South, where levels of college attendance are lower. This is also true in NLS66 and NLSY97. Notice that percentiles of the test score distribution are taken within region, not across regions, because there could be systematic differences in test scores across regions which we want to abstract from. The question we ask is whether, in regions with high college participation, college students have lower ability *relatively to other residents in the region* than in regions with low college participation.

One could ask whether the magnitude of these test score differences is enough to explain the wage differences we observe in the Census. If we regress log hourly wages in

[3]We use schooling corrected AFQT in the NLSY79, as in Carneiro, Heckman and Vytlacil (2009).

Table A-6—Average Ability of College Attendees and Proportion of Going to College

	(1)	(2)	(3)	(4)
Data: NLS66				
	Average Percentile (High School)	Average Percentile (College)	Proportion in College (P)	Odds in P
North	0.35	0.60	0.59	1.45
South	0.33	0.64	0.55	1.22
Difference	0.02	-0.04	0.04	0.22
Data: NLSY79				
	Average Percentile (High School)	Average Percentile (College)	Proportion in College (P)	Odds in P
North	0.32	0.64	0.56	1.27
South	0.33	0.69	0.48	0.92
Difference	-0.01	-0.05	0.08	0.35
Data: NLSY97				
	Average Percentile (High School)	Average Percentile (College)	Proportion in College (P)	Odds in P
North	0.36	0.64	0.51	1.04
South	0.37	0.66	0.46	0.86
Difference	-0.01	-0.02	0.05	0.17

Notes: Data are from the original cohorts of the National Longitudinal Survey of Young Males (NLS66), the National Longitudinal Survey of Youth of 1979 (NLSY79), and the National Longitudinal Survey of Youth of 1997 (NLSY97). In each dataset, white males are divided into two groups: those growing up in the southern regions of the US, and those growing up elsewhere. For each individual in each region, his percentile is computed in the distribution of test scores within region (the AFQT in the NLSY79 and NLSY97, and IQ in NLS66). Then, the sample is divided into those with some college or more, and those with less than college, and the average percentile in the test score distribution is computed for each education group and region. Each number in columns (1) and (2) corresponds to the average percentile in the within region test score distribution for each dataset, region and schooling group. For each dataset and region columns, (3) and (4) show the proportion of individuals in the sample with at least some college and its corresponding odds.

1994 on within region AFQT percentile for white males in the NLSY79 with some college and residing in the North we get a coefficient of about 0.5 (and essentially the same coefficient, 0.6, if we use those in the South instead). This magnitude of the estimated coefficient translates to an increase of about 14.4% (= $0.5 \times 28.8\%$, where 28.8% is the standard deviation of within region percentile test scores) in hourly wages with respect to the one standard deviation increase in the AFQT percentile. This estimate is within the range of previous estimates for the return to the test scores. For example, Bishop (1989) finds that a one standard deviation increase in test score is associated with a 19% increase in earnings for male household heads using data from the 1971 Panel Study of Income Dynamics (see equation (3) of Bishop, 1989); Neal and Johnson (1996) report an estimate of about 17% increase (see Column 3 of Table 1 of Neal and Johnson, 1996); Murnane, Willett and Levy (1995) obtain an estimate of about 8% increase with respect to the one standard deviation increase in the mathematics score using data from High School and Beyond (see tables 2 and 3 Murnane, Willett and Levy, 1995). Our estimation result from NLSY79 means that the difference of 5 percentile points between those college participants residing in the South and those residing elsewhere, corresponds to a difference in wages of 2.5 percentage points (= $0.5 \times 5\%$) if we use the estimate from the North (3 percentage points, respectively, if we use the estimate from the South). Given that the college participation rates are 48% in the South and 56% elsewhere, if we were to use our estimates of the reduced-form equation, we would expect college wages to be about 3 percentile points higher in the south than elsewhere ($0.086 \times$ a difference of 0.35 in the odds of P). These magnitudes are reassuringly similar to each other. This provides suggestive and direct evidence that increases in college attainment lead to declines in the quality of college attendees.

REFERENCES

Bishop, John H. 1989. "Is the Test Score Decline Responsible for the Productivity Growth Decline?" *The American Economic Review*, 79(1): 178–197.

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

Carneiro, Pedro, and Sokbae Lee. 2009. "Estimating distributions of potential outcomes using local instrumental variables with an application to changes in college enrollment and wage inequality." *Journal of Econometrics*, 149(2): 191–208.

Carneiro, Pedro, James J. Heckman, and Edward Vytlacil. 2009. "Estimating Marginal Returns to Education." unpublished manuscript, Department of Economics, University College London.

Murnane, Richard J., John B. Willett, and Frank Levy. 1995. "The Growing Importance of Cognitive Skills in Wage Determination." *The Review of Economics and Statistics*, 77(2): 251–266.

Neal, Derek A., and William R. Johnson. 1996. "The Role of Premarket Factors in Black-White Wage Differences." *The Journal of Political Economy*, 104(5): 869–895.

Ruggles, S., M Sobek, T Alexander, C. Fitch, R. Goeken, P. Hall, M. King, and C. Ronnander. 2004. "Integrated Public Use Microdata Series: Version 3.0 [Machine-readable database]." Minneapolis, MN: Minnesota Population Center [producer and distributor].