1995

Essays in Empirical Labor Economics: Dissertation Summary

David Jaeger
University of Michigan

Essays in Empirical Labor Economics

David A. Jaeger

This dissertation contains three essays which are loosely organized around the role that educational attainment and skill plays in the labor market. The first essay examines the effects of the large changes in the supply of immigrant labor to the U.S. labor market during the 1980s. The second essay estimates sheepskin effects in the returns to education using a data set that contains information on both degrees received and the highest grade that individuals completed. This essay was written jointly with Marianne Page. The third essay proposes a reconciliation scheme for the old and new educational attainment questions in the Current Population Survey and the decennial U.S. Census.

I. Skill Differences and the Effect of Immigrants on the Wages of Natives

In this paper I examine the effects of the large changes in the supply of immigrant labor during the 1980s on the wage structure of natives. I disaggregate *both* the immigrant and native populations by skill level (as proxied by educational attainment). I first estimate the elasticity of substitution between immigrants and natives within 8 sex $\times$ skill cells. Even after accounting for various possible biases, as well as adjusting for changes in the quality of immigrant and native workers, I estimate that immigrants and natives possessing similar skills are nearly perfect substitutes in production. I then use this finding, along with a simple aggregate production function model, to examine the effects of the large relative changes in the supplies of immigrants on the wage structure of natives. I find that, in the aggregate, changes in the supply of immigrant labor accounted for a relatively small percentage of the increase in the native college–high school wage gap during the 1980s. The magnitude of the effect was not homogenous across regions, however, and was strongest on the Pacific coast..

To estimate the elasticity of substitution between immigrants and natives with sex and skill groups, I assume that the aggregate production technology in the economy can be characterized by a nested function of the form

$$Q = f(g_{md}(X_{nmd}, X_{imd}), g_{mh}(X_{nmh}, X_{imh}), g_{ms}(X_{nms}, X_{ims}), g_{mc}(X_{nmc}, X_{imc}), g_{fd}(X_{nfd}, X_{ifd}), g_{fh}(X_{nfh}, X_{ifh}), g_{fs}(X_{nfs}, X_{ifs}), g_{fc}(X_{nfc}, X_{ifc}), K) \quad (1)$$

where X_{ijk} is the quantity of labor supplied by nativity $\times$ sex $\times$ skill group ijk (i={native, immigrant}; j={male, female}; k={high school dropout, high school graduate, some college, college graduate}); g_{jk} is a function that aggregates native and immigrant labor in sex $\times$ skill group jk; and K is the capital stock. This form of the production function assumes that the 8 sex × skill groups are weakly separable from one another and from capital, e.g. dropout skill makes affect the wages of college graduate females only through the relationship between g_{md} and g_{fc}. This specification allows me to examine the relationship between natives and immigrants within sex × skill groups independently of the other groups and capital.

Under the additional assumption that the g_{jk} are linear homogeneous the elasticity of substitution, σ_{jk}, between natives and immigrants in group *jk* is

$$\sigma_{jk} = \frac{d \log (X_{ijk} / X_{njk})}{d \log (w_{njk} / w_{ijk})}. \tag{2}$$

A simple estimate of the elasticity of substitution between immigrants and natives can be obtained by using the changes in relative wages and relative labor supply during the 1980s. This estimate assumes that *all* of the change in relative wage between the groups is due to changes in the relative supplies.

The elasticity of substitution can also be estimated in a regression context, using the variation across metropolitan areas in the change in supply of immigrant labor. In the two-factor linear homogeneous model described above the within-skill × sex group elasticity of complementarity between immigrants and natives, c_{jk}, is merely the inverse of σ_{jk}: $c_{jk}=1/\sigma_{jk}$. Estimating c_{jk} therefore gives an estimate of σ_{jk}. If any measurement errors in

H_{ijk} and H_{njk} are orthogonal to the error in the regression equation (an assumption that is relaxed later in the paper), an unbiased estimate of c_{jk} can be obtained with

$$e = \alpha + \beta h + \epsilon, \tag{3}$$

where I have suppressed the sex $\times$ skill subscript, jk, and the metropolitan area subscript, c, on all variables for notational simplicity, $e = d\log(E_i / E_n)$ is the change in log relative immigrant earnings and $h = d\log(H_i / H_n)$ is the change in log relative immigrant labor supply hours, and ϵ is an error term. With this specification, $\hat{c}_{jk} = 1\text{-}\hat{\beta}$ and therefore $\hat{\sigma}_{jk} = 1/(1\text{-}\hat{\beta})$.

Table 1 presents a summary of the paper's results on the elasticity of substitution between immigrants and natives within sex $\times$ skill groups. The "simple" and OLS results are computed as described above. IV results are also presented because e and h may be measured with error, with common error components in each variable. I instrument the change in relative hours with the change in the population ratio of immigrants to natives (taken from a different data source) in each of the 50 metropolitan areas used in the regressions. Under the assumption that the population ratios are uncorrelated with any measurement error in the change in relative hours to relative earnings, IV will give consistent estimates of c_{jk}. The quality-adjusted results control for within-

Table 1
Estimated Elasticity of Substitution Between Immigrants and Natives Within Sex and Skill Groups

	Unadjusted					Quality-Adjusted				
	Simple	OLS		IV		Simple	OLS		IV	
Skill Level	$\hat{\sigma}$	$\hat{c}$	$\hat{\sigma}$	$\hat{c}$	$\hat{\sigma}$	$\hat{\sigma}$	$\hat{c}$	$\hat{\sigma}$	$\hat{c}$	$\hat{\sigma}$
Men										
Dropouts	12	.104	9.6	.105	9.5	27	.198	5.0	.117	8.6
		(.055)		(.069)			(.066)		(.088)	
High School	12	.116	8.6	.162	6.2	-12	.145	6.9	.181	5.5
		(.053)		(.062)			(.056)		(.065)	
Some College	9	.024	42.5	.063	16.0	-10	.085	11.8	.093	10.7
		(.068)		(.102)			(.082)		(.119)	
College	8	.084	11.9	.217	4.6	40	.148	6.7	.054	18.7
		(.095)		(.133)			(.101)		(.061)	
Women										
Dropouts	14	-.009	-106.5	.024	41.9	20	.021	48.2	.054	18.7
		(.050)		(.061)			(.050)		(.061)	
High School	13	.170	5.9	.207	4.8	425	.180	5.6	.223	4.5
		(.058)		(.078)			(.058)		(.077)	
Some College	12	-.077	-13.0	.041	24.3	2221	-.078	-12.9	.026	38.5
		(.054)		(.096)			(.059)		(.107)	
College	14	.194	5.2	-.036	-28.2	16	.203	4.9	-.097	-1.3
		(.091)		(.171)			(.099)		(.196)	

SOURCE: Calculations from the five percent Public Use Microsamples of the 1980 and 1990 U.S. Census.
NOTE: Standard errors are in parentheses. Dependent variable is $d\log(E_i/E_n)$ and independent variable is $d\log(H_i/H_n)$ (unadjusted) or $d\log(H_i/H_n) + (\ell_i - \ell_n)$ (quality-adjusted). Instrument for all IV models is the population $d\log(I/N)$.

nativity × sex × skill group changes in demographic composition. The characteristics controlled for include age and race/ethnicity for all groups and time in the U.S., language ability, country of origin for immigrants.

The results in Table 1 clearly indicate that immigrants and natives are close substitutes for one another with sex × skill cells. Few of the estimated c_{jk} are statistically significantly different from zero (implying estimated σ_{jk} that are not different from infinity), while the estimated c_{jk} that *are* significantly different from zero are small enough to imply large σ_{jk}. The paper also examines several possible biases of the estimates. I directly estimate the effect of errors-in-variables bias and adjust the OLS results to give consistent estimates and also consider omitted-variables bias due to demand factors. I find that neither factor is large enough to alter the conclusion that immigrants and natives are essentially perfect substitutes in production.

Having established the approximate perfect substitutability of immigrants and natives, the paper then turns to estimating the effect of immigration on the growth in the college high-school wage gap during the 1980s. I now assume that the aggregate production function

$$y = F(G(L, H), K) \tag{4}$$

where F and G are linear homogeneous functions, L is the effective supply of low-skill labor in the economy, H is the effective supply of high-skill labor in the economy, and K is the capital stock. To simplify the analysis, I continue to assume that K is weakly separable from $G(L, H)$. Under the additional assumption of a constant elasticity of substitution, ρ between the high- and low-skill labor aggregates, the effect off changes in the supply of immigrants on change in the native log wage differential between time t and time τ is

$$\delta^{\tau} - \delta^{t} = -\frac{1}{\rho}\left[\log\left(\frac{H^{\tau}/L^{\tau}}{H_n^{\tau}/L_n^{\tau}}\right) - \log\left(\frac{H^{t}/L^{t}}{H_n^{t}/L_n^{t}}\right)\right]. \tag{5}$$

I assume that immigrants and natives are perfect substitutes and also assume that men and women are perfect substitutes. After adjusting for changes in the efficiency of each of the 16 nativity × sex × skill groups, I then aggregate these into high (i.e. college equivalent) and low (i.e. high school equivalent) groups and examine how immigrants have changes the skill composition of the labor force, and therefore, how they have affected the relative wages of high- and low-skill natives. I find that immigrants increased the relative supply of low-skill workers by 1.2 percentage points in 1980 and 2.4 percentage points in 1990. Table 2 presents the effects of immigrants on the native

Table 2
Estimated Effect of Immigrants on the Native College/High School Wage Differential

	Levels		Change
	1980	1990	
Native College / High School Wage Diff.	.343	.474	.131
Change in wage differential from change in relative supplies due to immigrants			
$\rho = 1.00$	.012	.024	.012
$\rho = 1.41$	.009	.017	.008
$\rho = 2.00$	.006	.012	.006
Percent of change in wage differential explained by immigration			
$\rho = 1.00$			9
$\rho = 1.41$			6
$\rho = 2.00$			4

SOURCE: Five percent Public Use Microsamples of the 1980 and 1990 U.S. Census.

college/high school wage gap in 1980 and 1990 as well as for the change over the 1980s under different plausible values for ρ I find that immigrants can account for approximately .8 percentage points (or six percent) of a 13.1 percentage point change in the native college-high school wage gap over the 1980s.

These effects were not uniform throughout the country however. Table 3 presents similar results for the largest metropolitan areas within each of the 9 Census divisions. These results make clear that the largest effects of immigration on growth of the native college-high school wage differential were on the Pacific, where immigrants contributed 2.9 percentage points (or 23 percent) to an increase of 12.9 percentage points. The

Table 3
Effect of Immigrants on the College / High School Wage Differential Within Census Divisions: 1980-1990

	Change in native log (*H/L*) Wage Diff.	Effect of Imm. on Change in log(*H/L*) Supplies	Effect of Immigrants on log(*H/L*) Wage Diff., Assuming ρ =			% of (1) Explaine d by (4): [(4)/(1)]
			1.00	1.41	2.00	
Metropolitan Area	(1)	(2)	(3)	(4)	(5)	(6)
New England	.047	-.001	.001	.001	.001	2
Middle Atlantic	.111	-.008	.008	.005	.004	5
South Atlantic	.117	-.011	.011	.008	.006	7
East North Central	.154	-.003	.003	.002	.001	1
West North Central	.120	-.001	.001	.001	.000	0
East South Central	.138	.004	-.004	-.003	-.002	-2
West South Central	.140	-.025	.025	.018	.012	12
Mountain	.141	-.004	.004	.003	.002	2
Pacific	.129	-.041	.041	.029	.021	23

SOURCE: Five percent Public Use Microsamples of the 1980 and 1990 U.S. Census.

paper also examines these effects within metropolitan areas and finds the largest effects in Miami (5.1 percentage points of a 13.1 percentage point change) and Los Angeles (4.6 percentage points of a 11.6 percentage point change). Surprisingly in half of the 50 metropolitan areas examined, changes in the supply of immigrants actually *increased* the relative supply of college-equivalent workers.

II. Degrees Matter: New Evidence on Sheepskin Effects in the Returns to Education*

The second essay, written jointly with Marianne Page, examines "sheepskin effects" (the return to receiving a degree above and beyond that embodied by the human capital imparted by years of schooling) in the returns to education. We use a data set created from the 1991 and 1992 March Current Population Surveys that includes information on both degrees received and completed years of education for the same individuals. This allows us to improve on previous estimates which relied on spline functions with breaks at the "usual" degree years to estimate approximate sheepskin effects. Having additional information on

*This essay is forthcoming in *The Review of Economics and Statistics*.

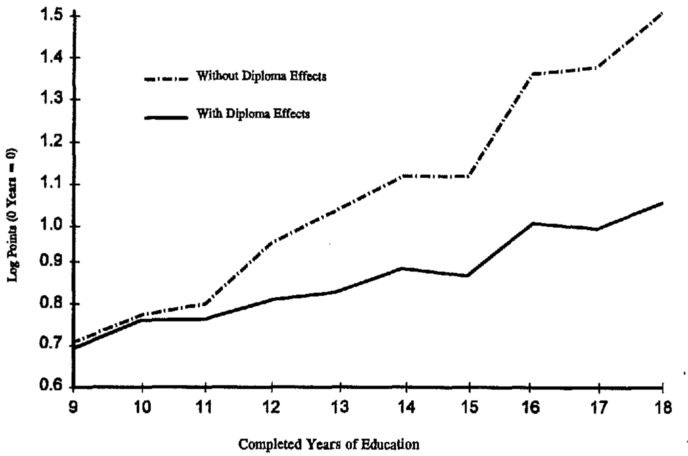
degrees received allows us to estimate directly the return to possessing degree, while controlling for time spent in school.

We found that the sheepskin effects estimated using the true information on degree receipt were more than twice as large as those using spline function specifications with only completed years of education. For example the credentialling effect for receiving a high school diploma was 3.4 percent using a spline function model and 10.6 percent using true information on degrees received. Similarly, the estimated sheepskin effect for completing 16 years of education was 11.4 percent while that for a Bachelor's degree was 27.3 percent. Thus, the wage premium associated with high school or college graduation would appear to be substantially higher than previous estimates would suggest. Diploma effects also comprise a substantial portion of the total return to education. Figure 1 plots the total return to education using model with only dummy variables for each completed year of education along with the same model with the addition of indicators for various degrees received. Diploma effects explain approximately one quarter of the total return to completing 16 years of education and more than half of the return to completing 16 years relative to 12 years.

The paper also examines differences in sheepskin effects across demographic groups and between different degrees. The

Figure 1

Total Returns to Completed Years of Education for White Men



estimated diploma effects for white and black men and women are presented in Table 4. The model estimates the effects for all four groups in one regression by interacting demographic group dummy variables with each of the years of education and diploma dummy variables.

We find relatively little support for the hypothesis that high school and college diploma effects should vary by race and sex and cannot reject the hypothesis the high school and college diploma effects are the same for all four demographic groups.

Table 4

Estimated Diploma Effects for Four Demographic Groups

Coefficient	Men		Women	
	White	Black	White	Black
Completed Years of Education				
9	-0.097 (0.060)	-0.065 (0.152)	-0.344 (0.067)	-0.196 (0.151)
10	-0.032 (0.054)	0.003 (0.140)	-0.229 (0.058)	-0.267 (0.127)
11	-0.039 (0.051)	-0.280 (0.128)	-0.173 (0.055)	-0.157 (0.113)
12	ref.	ref.	ref.	ref.
13	0.018 (0.033)	0.043 (0.110)	0.016 (0.033)	0.134 (0.095)
14	0.071 (0.031)	0.001 (0.106)	0.084 (0.033)	0.232 (0.093)
15	0.045 (0.044)	0.122 (0.156)	0.154 (0.044)	0.243 (0.123)
16	0.167 (0.044)	0.046 (0.137)	0.271 (0.046)	0.159 (0.126)
17	0.150 (0.057)	-0.201* (0.307)	0.250 (0.060)	0.365 (0.200)
18 or more	0.215 (0.054)	-0.280 (0.245)	0.335 (0.059)	0.208 (0.176)
Diploma Effects				
High School	0.125 (0.041)	0.119 (0.108)	0.062 (0.044)	0.105 (0.097)

Table 4
(Continued)

Coefficient	Men		Women	
	White	Black	White	Black
Marginal Effect Over High School				
Some College, No Degree	0.083 (0.027)	0.133 (0.092)	0.101 (0.027)	0.079 (0.076)
Occupational Associate's	0.075 (0.043)	0.162 (0.151)	0.305 (0.041)	-0.008 (0.133)
Academic Associate's	0.188 (0.046)	0.145 (0.166)	0.241 (0.044)	-0.103 (0.140)
Bachelor's	0.245 (0.045)	0.305 (0.134)	0.223 (0.046)	0.394 (0.130)
Marginal Effect Over Bachelor's				
Master's	0.055 (0.041)	0.658 (0.223)	0.155 (0.043)	0.125 (0.155)
Professional	0.286 (0.059)	0.541* (0.331)	0.488 (0.091)	0.570* (0.378)
Doctoral	0.083 (0.066)	0.792* (0.382)	0.103 (0.097)	0.352* (0.367)
Sample Size	8,957	711	8,122	909

NOTE: * denotes cells with 10 or fewer observations. Dependent variable is log hourly wages. Estimated using ordinary least squares. Standard errors are in parentheses ( ). Calculated from a matched sample of individuals 25 to 64 years old from the 1991 and 1992 March Current Population Survey. Model also includes separate intercepts for each group, dummy variables for zero through eight years of education for each group, Potential Experience, and Potential Experience Squared as covariates. The R^2 for the model is 0.229; the mean squared error is 0.367.

We do find significant differences for occupational Associate's degrees, where white women appear to earn more than the other demographic groups. This may be due to difference in occupational selectivity, in particular the extent to which these groups do or do not select into nursing. Contrary to some prior research, our results also suggest that the labor market values Associate's and Bachelor's degrees differently, and that these wage differences do not result solely from differences in years of education.

Our paper also contributes to the literature on the returns to post-graduate degrees, where we find that white men and women with professional degrees earn 33 percent and 63 percent more, respectively than those who earn only a Bachelor's degree but complete the same number of years of school. For Doctoral degrees these differences are 9 and 11 percent while for Master's degrees the differences are 6 and 17 percent, respectively. The effects for blacks are generally larger, although most of these estimates are based on very small samples.

III. Reconciling the Old and New Census Bureau Education Questions: A Recommendation for Researchers

The third essay of the dissertation proposed a recoding scheme to reconcile the "old" and "new" Census Bureau educational attainment question employed in the 1990 Census and in the Current Population Survey since 1992. The new question focuses on degrees received while the old question focused on highest grade completed. Researchers who formerly relied on the continuous variable "Highest Grade Completed" to capture differences in educational attainment must now find new ways to represent this variation. I use the data from Chapter II, which contained information from a matched sample of the Current Population Survey before and after the change in the educational attainment question—giving answers to both questions.

The proposed recoding scheme is presented in Table 5. Most notable the way in which I reconcile the two variables is including individuals who completed 16 or 17 years of education as college graduates. This does not correspond to the usual practice of imputing a college diploma from 16 years of completed education. Nevertheless, the match is improved by including those who completed 17 years with those who completed 16 years.

The effect of the recoding scheme on the distribution of educational attainment using both the old and new questions is presented in Table 6. The distribution is very similar using either the new or the old variable and the match rate between them is 85.3 percent.

Table 5
Recoding Scheme for Old and New Education Questions

	Old Question Highest Grade Attended		New Question
Recoded Category	Not Completed	Completed	Code
	Current Population Survey		
0 Years	0,1		31
1 - 8 years	2,3,4,5,6,7,8,9	1,2,3,4,5,6,7,8	32,33,34
9 - 11 years	10,11,12	9,10,11	35,36,37
12 years		12	38,39
< 4 years college	13,14,15,16	13,14,15	40,41,42
4 - 5 years college	17,18	16,17	43
> 5 years college		18	44,45,46

As a case study, I estimate the returns to education using both the old and new recoded variables. I find that the results are generally quite comparable, with the returns for categories greater than 12 years being somewhat higher using the new variables. This may be due to the explicit inclusion of GED recipients with those with traditional high school diplomas.

I also find that the new question has somewhat more explanatory power in wage equations than the old question. This

Table 6
Distribution of Recoded Old and New Education Variables

	Old Variable			New Variable		
Category	Freq.	Percent	Cum. Dist.	Freq.	Percent	Cum. Dist.
0 years	95	0.004	0.004	91	0.003	0.003
1 - 8 years	1,445	0.054	0.058	1,370	0.051	0.055
9 - 11 years	2,352	0.088	0.145	2,114	0.079	0.134
12 years	10,241	0.383	0.528	10,404	0.389	0.523
< 4 years college	6,174	0.231	0.759	6,403	0.239	0.762
4 - 5 years college	4,421	0.165	0.924	4,174	0.156	0.918
> 5 years college	2,023	0.076	1.000	2,195	0.082	1.000

NOTE: Tabulated from a matched sample of individuals aged 25 to 64 years from the 1991 and 1992 March Current Population Survey. Sample size is 26,751.

is true even if the samples are limited to those with educational attainment between a high school degree and Bachelor's degree. Because the old question contained more detail about the educational attainment for individuals with less than a high school degree and the new question contains more information about individuals with more than a college degree, this effect is likely due to importance of sheepskin effects, as discussed in Chapter II.