Bias-Corrected Estimates of GED Returns

James J. Heckman, *University of Chicago and American Bar Foundation*

Paul A. LaFontaine, *Center for Social Program Evaluation, American Bar Foundation*

Using three sources of data, this article examines the direct economic return to General Educational Development (GED) certification for both native and immigrant high school dropouts. One data source—the Current Population Survey (CPS)—is plagued by nonresponse and allocation bias from the hot deck procedure that biases the estimated return to the GED upward. Correcting for allocation bias and ability bias, there is no direct economic return to GED certification. An apparent return to GED certification with age found in the raw CPS data is due to dropouts becoming more skilled over time. These results apply to both native-born and immigrant populations.

I. Introduction

There has been rapid growth in the fraction of persons who achieve high school certification by means of an equivalency exam rather than through the traditional route of classroom attendance and high school graduation.

This project was supported by the Mellon Foundation, the Joyce Foundation, the Pew Foundation, and NICHD R01-34598-03. We thank William Johnson and James Ziliak for helpful comments on the first draft. We discussed our findings with Barry Hirsch at Trinity University, San Antonio, March 2003. Additional material for this article, referred to in the text, is available at http://jenni.uchicago.edu/ged_imputation. Contact the corresponding author, James J. Heckman, at jjh@uchicago.edu.

[*Journal of Labor Economics,* 2006, vol. 24, no. 3]

0734-306X/2006/2403-0009$10.00

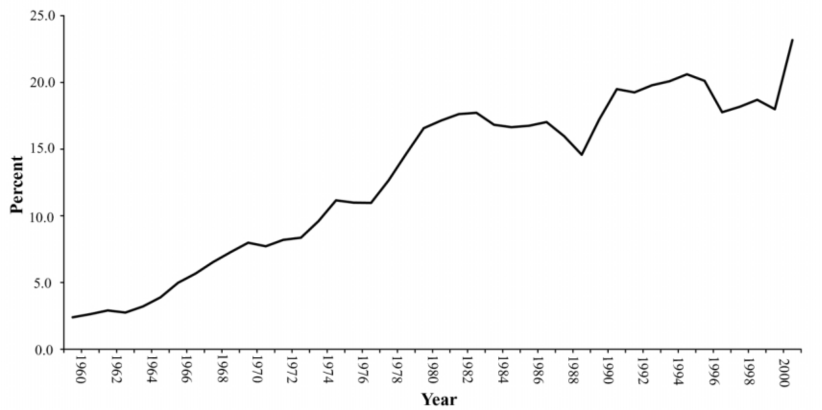

Fig. 1.—GED credentials issued as a percentage of public and private high school graduates, United States, 1960–2001.

The primary vehicle for high school equivalency certification is the General Educational Development (GED) program. In 1960, only 2% of all new high school certificates were awarded through equivalency exams in the United States. By 2001, over 20% of all new high school credentials were produced through GED certification (see fig. 1). This rapid growth in exam certification occurred despite apparently low direct economic returns to it. Using data from the National Longitudinal Study of Youth (NLSY), Cameron and Heckman (1993) find that, in terms of hourly wages, controlling for differences in ability, male exam-certified high school equivalents are statistically indistinguishable from high school dropouts who are uncertified. Any differences in wages among exam-certified equivalents and uncertified dropouts are completely accounted for by differences in ability. There is no causal effect of GED certification on wages.[1] Cameron and Heckman conclude that whatever economic return there is to GED certification must come through access to further postsecondary education and training that certification provides. However, GEDs are much less likely than ordinary high school graduates to complete 2- or 4-year colleges. A large body of subsequent work, summarized in Boesel, Alsalam, and Smith (1998), supports the position that GED recipients are more similar to dropouts than to high school graduates along many dimensions.

Advocates of the GED testing program raised some potentially valid criticisms of the Cameron and Heckman analysis following its publication (Murnane, Willett, and Boudett 1999; Boudett, Murnane, and Willett 2000; Jaeger and Clark 2006). First, Cameron and Heckman only considered

[1] Later work by Cameron (1994) found similar results for NLSY females.

labor market outcomes at ages 25 and 28. If GED certification opens up access to occupations that are closed to high school dropouts, then the effect of certification may not manifest itself until later in the life cycle. A second concern is the small sample sizes available in the NLSY data. Some argued that it would not be possible to assess the entire GED program based on a few hundred NLSY participants. Finally, there may be a disparate impact of the GED program across different race groups or other subpopulations. For instance, a GED may send a different signal for recent immigrants who acquire the credential than it does for native-born dropouts. This article addresses these questions.

In 1998, the monthly portion of the Current Population Survey (CPS) began distinguishing between the two types of high school completion. The large sample sizes for various racial and ethnic groups, as well as the wide range of available ages, appear to make the CPS ideal for addressing some of the limitations of the Cameron and Heckman analysis. However, four potentially serious problems and limitations plague the CPS data. First, the CPS contains no measure of ability. Cameron and Heckman found that the GED program is selective because it is the higher-ability dropouts who attain GED certification. Once differences in ability between GED recipients and uncertified dropouts are accounted for, wage differentials disappear. Second, as found by Hirsch and Schumacher (2004) in the context of estimating union-nonunion wage differentials, "match bias" can result from the CPS's method of imputing missing wages. We find that the estimated returns to GED certification are substantially upward biased because GED respondents who either refuse or fail to report their wage information are frequently assigned (matched to) the wages of traditional high school graduates. Third, CPS data show that a large fraction of workers have no reported earnings because they are unemployed or out of the labor force. Finally, bias may arise from low- and high-income earners refusing to report earnings.

This article addresses the first three of these problems. We show that, when estimation is performed carefully, the returns to GED certification and other educational estimates using CPS data are similar to those obtained from other, cleaner, data sources. We find that, after correcting for differences in ability, GED recipients who do not continue on to college earn the same wages as uncertified dropouts. This result applies to both males and females across the age spectrum. We find no evidence of postcertification life cycle wage growth attributable to the program. The apparent return to GED certification for older age groups in the raw data is due to a greater unobserved ability bias for older birth cohorts rather than from a causal effect of GED certification. After correcting for problems with the CPS data, the estimated GED-dropout difference in wages is the same in comparable NLSY and CPS cohorts. The positive wage returns to GED certification found in unadjusted CPS data arise from unobserved ability bias and improper allocation of GED missing wages. We also show that ability bias is greater when comparing foreign-born GED recipients and foreign-born dropouts. After adjusting for ability, no statistically significant effect of the GED on wages is discernible for both native and foreign-born males and females of all race and ethnic groups.

The plan of this article is as follows. In Section II, we discuss the CPS and compare evidence from it with evidence from the NLSY. In Section III, we present ability-bias-corrected returns to GED certification. In Section IV, we discuss the issues of age and cohort effects, using a variety of data sources. In Section V, we consider GED returns among immigrants. Section VI concludes.

II. The Importance of Wage Imputation and Nonresponse

A. CPS Data

We use the monthly outgoing rotation groups from the CPS for the period January 1998 to December 2003. Our sample consists of civilian males and females age 20–64 who are either in their fourth or their eighth month in the sample. We use a sample of dropouts, GED recipients, and high school graduates who have completed no college, along with a sample of 4-year degree holders for whom we cannot determine what type of high school certificate they hold.[2] For our wage analysis, we exclude those people who are enrolled in school, those who are self-employed, those who reported their ethnicity as Native American, Aleut, or Eskimo, and those who had their education status or years of schooling responses imputed. The self-employed are excluded because earnings are not available for these individuals. All regressions also exclude those who earn less than $.50 or more than $200 an hour (in 2000$). Data losses due to these exclusions are listed in table 1. The main exclusions are those who are not working or who are self-employed. For these groups, wage data are unavailable. Other sample restrictions only account for a small fraction of lost data.

B. CPS Problems and Limitations

Due to its large sample size, the long period over which it is collected, and its perceived quality, the CPS has become the primary data source for understanding a host of important economic issues, including the U.S. earnings structure, racial wage gaps, and returns to education. The growing nonresponse to income-related questions calls into question the quality

[2] Due to the structure of the CPS monthly questionnaire it is not possible to determine the GED status of those who continue on to college. For this reason, our estimates of GED returns using the CPS are limited to the direct effect of certification on outcomes. These estimates will be lower than an overall effect inclusive of the indirect effects of postsecondary training.

Table 1
Exclusion Restrictions by Data Source

	Native Males			Native Females			Foreign Males		Foreign Females	
	CPS	NLSY	NALS	CPS	NLSY	NALS	CPS	NALS	CPS	NALS
Potential observations	352,858	55,057	5,412	371,222	54,101	7,058	65,004	821	68,688	886
Not working	64,302	12,358	872	117,363	19,873	2,306	10,061	109	29,377	354
Working and enrolled	1,681	1,612	311	2,227	1,862	425	305	60	251	59
Self-employed*	40,311	3,334	0	21,064	2,107	0	5,772	0	2,921	0
Other race	3,065	0	30	2,761	0	38	124	. . .	128	. . .
Zero years of education	385	17	0	280	34	0	886	0	482	0
Imputed education	988	0	0	780	0	0	298	0	166	0
Earnings outliers	286	130	137	298	81	380	61	26	36	48
Total observations	239,400	37,961	4,106	225,517	30,621	3,952	47,295	629	35,174	429
% Not working	.182	.224	.161	.316	.367	.327	.155	.133	.428	.400
% Working and excluded	.170	.111	.096	.112	.105	.168	.139	.117	.105	.194

NOTE.—The total excluded observations is not the sum of the column since many individuals fall into multiple categories. Calculations are based on a sample of employed dropouts and GED recipients and high school graduates with no college plus 4-year college graduates. The sample ages are 20–64 for the CPS, 20–39 for the NLSY, and 20–64 for the NALS.

* It is not possible to determine years of schooling or self-employment in the NALS data.

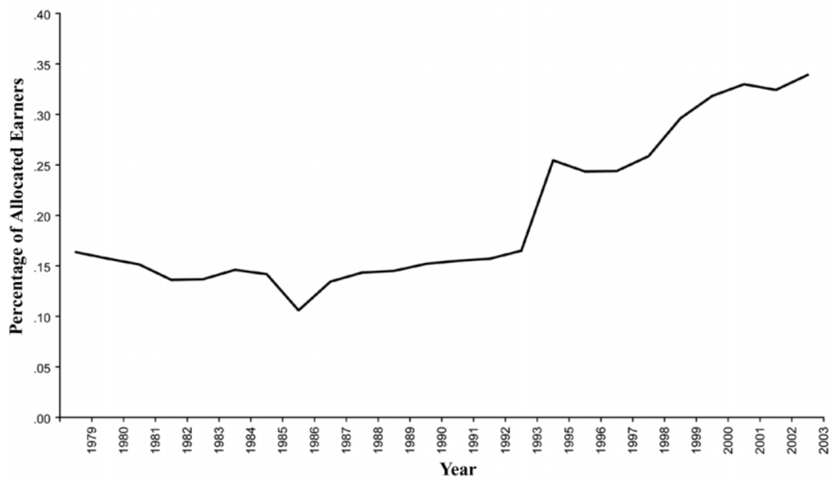

Fig. 2.—CPS monthly outgoing rotation groups, percentage of allocated earners, 1979–2003. Calculations are based on CPS Monthly Outgoing Rotation Groups from 1979 to 2003. The sample is restricted to individuals between the ages of 16 and 65 who are members of the civilian labor force and are earnings eligible. Allocation flags are unavailable from 1994 to August of 1995. Allocated earners from 1989 to 1993 are those people who have missing values for unedited weekly earnings since Census-provided allocation flags are unreliable.

of the data and its comparability across time. Figure 2 shows that prior to 1994 the percentage of those who chose not to report earnings was relatively stable at around 15%. After 1994, earnings nonresponse rose from a low of 24% in 1995 to nearly 34% in 2003.[3] Increasing rates of nonresponse, greater numbers of workers selectively withdrawing from the labor force, and the CPS practice of not collecting wage information from the self-employed have resulted in substantial fractions of respondents with missing wage data among certain race, sex, and age groups. Table 2 reveals that only about 50% of white and Hispanic males in each outgoing rotation group report earnings due to the combination of these factors. Wage data for black males are only available for around 38% of the sample due to higher rates of income nonresponse among the employed and higher incidence of unemployment among this population. The situation is worse for women due to their lower labor force participation rates. Unlike the NLSY, which surveys each person individually,

[3] The dramatic increase in allocation after 1994 is primarily due to the implementation of the newly redesigned CPS questionnaire. The new questionnaire asks a longer, more complex series of questions in order to determine weekly wages, and the new data processing procedures set weekly wages to missing if even one of these questions is met with either a refusal or a "don't know" response.

Table 2
Sources and Extent of CPS Missing Wage Data by Race for the Full Sample

	White Males				Black Males				Hispanic Males				Foreign Males
	20–29	30–39	40–49	50–59	20–29	30–39	40–49	50–59	20–29	30–39	40–49	50–59	20–59
Potential wage observations	95,928	122,760	140,283	109,744	12,956	14,819	15,032	10,654	21,697	20,714	14,283	8,049	87,019
Unemployed	4,580	3,529	3,748	2,648	1,311	846	763	392	1,286	884	561	323	3,523
Out of the labor force	11,003	6,734	10,341	16,999	2,851	1,891	2,663	2,947	2,109	1,366	1,375	1,509	8,696
Self-employed	4,132	14,577	22,713	19,102	265	749	834	719	611	1,331	1,288	714	8,493
Military	1,595	2,006	854	221	239	330	114	17	238	154	65	6	493
Nonresponse	21,753	27,912	33,282	24,692	3,384	4,433	4,790	3,087	5,034	4,781	3,367	1,875	21,714
Wage observations	52,865	68,002	69,345	46,082	4,906	6,570	5,868	3,492	12,419	12,198	7,627	3,622	44,100
% Reporting wages	.551	.554	.494	.420	.379	.443	.390	.328	.572	.589	.534	.450	.507
Proxy responses	3,0731	37,080	37,766	24,655	1,780	1,794	1,918	1,547	8,046	7,133	4,445	2,110	25,158
% Self-reporting wages	.231	.252	.225	.195	.241	.322	.263	.183	.202	.245	.223	.188	.218
	White Females				**Black Females**				**Hispanic Females**				**Foreign Females**
	20–29	30–39	40–49	50–59	20–29	30–39	40–49	50–59	20–29	30–39	40–49	50–59	20–59
Potential wage observations	99,672	128,051	145,846	113,599	17,926	20,086	19,987	13,893	20,726	20,741	15,244	8,826	89,359
Unemployed	3,658	3,210	3,139	2,006	1,650	1,173	855	341	1,115	966	600	240	3,236
Out of the labor force	21,677	28,971	28,539	32,172	4,620	3,932	4,389	4,687	7,513	6,786	4,422	3,684	30,779
Self-employed	2,727	8,900	11,819	9,710	216	515	564	407	276	634	686	441	4,540
Military	188	116	86	11	59	53	31	0	23	10	9	0	49
Nonresponse	18,930	23,214	30,358	22,277	4,023	5,682	6,047	3,741	3,111	3,295	2,887	1,411	16,784
Wage observations	52,492	63,640	71,905	47,423	7,358	8,731	8,101	4,717	8,688	9,050	6,640	3,050	33,971
% Reporting wages	.527	.497	.493	.417	.410	.435	.405	.340	.419	.436	.436	.346	.380
Proxy responses	23,274	19,702	21,934	14,447	2,337	2,114	2,101	1,145	3,981	3,281	2,553	1,259	14,983
% Self-reporting wages	.293	.343	.343	.290	.280	.329	.300	.257	.227	.278	.268	.203	.212

NOTE.—Based on CPS 1998–2003 monthly outgoing rotation groups. Potential wage observations are those people in their fourth or eighth month in samples who are in the civilian labor force. These are the individuals for whom wage and job information questions are asked.

the CPS survey is administered by telephone to one person who responds for his or her entire household. Potentially exacerbating the nonresponse problem is that the accuracy of the available wage information may also be questionable. For males, over 60% of the wage and labor force information is given by a proxy respondent, and these respondents may not be privy to all income-related information. The percent of available self-reported wages is extremely low—around 25% for males and 30% for females. The propensity to report earnings also varies across race groups. In particular, black males and females are 10% more likely not to report earnings than either their white or Hispanic counterparts.

Unfortunately, the CPS does not provide enough information to determine the nature of this response bias. We present some evidence on the severity of this potential bias using NLSY data. Nonresponse bias may not be large, since our estimates obtained from CPS data closely track those estimates from cleaner data sources where we can control for this potential bias.

C. CPS Imputation Strategy

To avoid computing national statistics based on a sample with a large proportion of missing data and in an attempt to correct for possible nonresponse bias caused by missing wage data, the CPS allocates missing earnings using a "hot deck" imputation method. A hot deck assigns the wages of respondents to nonrespondents based on a limited set of demographic, education, and occupational characteristics.[4] A common practice among researchers is to treat allocated values as observed when using CPS survey data. In a widely cited paper, Angrist and Krueger (1999) claim that CPS wage allocation is empirically unimportant. This article shows that CPS allocation methods and the resulting match bias are of first-order economic importance in estimating returns to GED certification.

"Match bias," a phrase due to Hirsch and Schumacher (2004), arises from the limited number of categories used to impute nonrespondent wages. Of particular interest to this article, the matching of wage nonrespondents to wage respondents is based on only three levels of educational attainment: high school dropout, high school graduate with up to but not including a bachelor's degree, and bachelor's degree and above. Given these education categories, it is clear that estimated returns for those who graduate high school and do not attend college will exhibit an upward bias since nonrespondents will frequently be matched to those who complete some college. On the other hand, estimated returns for those who complete above a bachelor's degree will be biased downward as a result of nonrespondents

[4] Currently, the CPS matches nonrespondents to respondents in the monthly data based on the following categories: gender (2), race (2), age (6), occupation (13), hours worked (8), education (3), and tips and overtime receipt (2).

being assigned the wages of those with only a bachelor's degree. Clearly, all CPS educational estimates will be affected by this type of educational mismatching within allocation cells. Hirsch and Schumacher (2004) and Bollinger and Hirsch (2006, in this issue) present a more detailed discussion of the CPS hot deck procedure and the resulting bias in estimates for various educational categories.

GED-allocated wages exhibit a particularly severe form of this type of misallocation bias since nonrespondents who hold GED credentials are frequently assigned the wages of high school graduates who may have postsecondary education up to but not including a bachelor's degree. If a wage differential exists between GED recipients and high school graduates, then this differential will approach zero as the proportion of GED nonrespondents increases. As nonresponse has grown from less than 15% to over 30% in recent years, the upward bias in estimated returns to the GED has increased proportionally.

Table 3 shows that, for native males, the estimated return to GED certification is overstated by over 35% when CPS-allocated wages are included in the sample. After dropping the allocated wages, the estimated return to GED certification drops from .14 log points to .09. For females, as shown in table 4, the bias tends to be generally smaller in magnitude but is still over 25%. The estimated return decreases from .15 log points to just under .11 for the full sample of females. As predicted, excluding allocated earners also decreases estimated returns to high school graduation and college completion. However, this decrease is not of the same magnitude as is found for GED recipients. The resulting reduction for the full sample of males is just over 5% for college graduates and just under 12% for high school graduates who did not attend college. The observed effects of CPS allocation for the female sample are similar. Overall, imputation tends to increase the estimated college dropout and high school dropout wage differentials and leaves the college–high school differential largely unaffected. The most serious bias is observed in the GED category.

Tables 3 and 4 show that the returns are different across racial, sex, and ethnic groups, although not dramatically so. Returns to certification are always higher for females compared to males, and minorities have higher returns than whites. Both Hispanic males and females show the highest returns to GED acquisition among all racial groups. However, the differences across groups are not dramatic. The largest estimated difference between pooled and separated race estimates is only .04 log points.

In order to assess how sensitive these estimates are to nonresponse and match bias, we implement a hot deck imputation procedure that differs from the CPS hot deck only in that it matches using more precisely defined educational groups. This is done both to show that it is the exclusion of

Table 3
CPS OLS Log Hourly Wage Regressions for Males by Race

	Model 1				Model 2				Model 3			
	Including Allocated Earners				Excluding Allocated Earners				Reallocating Missing Wages			
	All	Whites	Blacks	Hispanics	All	Whites	Blacks	Hispanics	All	Whites	Blacks	Hispanics
GED, no college	.137	.135	.146	.163	.088	.083	.105	.117	.086	.080	.092	.109
	(.005)	(.006)	(.016)	(.016)	(.006)	(.007)	(.020)	(.018)	(.007)	(.008)	(.021)	(.019)
High school, no college	.209	.209	.207	.209	.184	.180	.195	.197	.183	.181	.191	.203
	(.003)	(.004)	(.009)	(.010)	(.004)	(.005)	(.012)	(.012)	(.005)	(.005)	(.013)	(.012)
College graduate	.571	.570	.584	.591	.540	.534	.590	.573	.546	.540	.585	.584
	(.004)	(.004)	(.012)	(.015)	(.004)	(.005)	(.015)	(.017)	(.005)	(.005)	(.017)	(.016)
High school–Dropout	.209	.209	.207	.209	.184	.180	.195	.197	.183	.181	.191	.203
College-Dropout	.571	.570	.584	.591	.540	.534	.590	.573	.546	.540	.585	.584
College–High school	.362	.360	.377	.382	.357	.354	.395	.375	.362	.359	.394	.381
Adjusted R^2	.287	.272	.221	.282	.321	.306	.278	.313	.314	.312	.248	.299
Observations	236,666	203,012	21,182	11,824	158,314	137,892	11,868	8,100	236,666	203,012	21,182	11,824
F-test ($\Pr > F$):												
GED = Dropout	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000
GED = High school	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000

Note.—All dummy variables are defined exclusively. Dropouts are the excluded category. Persons enrolled in school at each age are deleted as are those people who have wages less than $.50 or more than $200 an hour (in 2000$), those who are self-employed, those who were not born in the United States, those who are younger than 20 years of age or older than 64, those who did not complete at least 1 year of schooling, those who are Aleut and Eskimo or Native American, or those had their completed schooling or GED status imputed by the CPS. Controls for central city status, married with spouse present, year of survey, region of residence, as well as a quadratic in age and race dummies (where appropriate), are included in each regression but not shown. Reported standard errors (in parentheses) are corrected for heteroscedasticity and clustering with the Huber-White sandwich estimator except when reimputing wages. Standard errors after reimputation are calculated using the method outlined in Shao and Sitter (1996).

Table 4
CPS OLS Log Hourly Wage Regressions for Females by Race

	Model 1 Including Allocated Earners				Model 2 Excluding Allocated Earners				Model 3 Reallocating Missing Wages			
	All	Whites	Blacks	Hispanics	All	Whites	Blacks	Hispanics	All	Whites	Blacks	Hispanics
GED, no college	.150	.140	.174	.196	.110	.102	.121	.157	.108	.099	.117	.144
	(.005)	(.006)	(.013)	(.016)	(.006)	(.007)	(.016)	(.018)	(.007)	(.007)	(.015)	(.017)
High school, no college	.237	.236	.217	.257	.215	.216	.191	.234	.210	.205	.199	.226
	(.003)	(.004)	(.007)	(.010)	(.004)	(.005)	(.009)	(.012)	(.005)	(.005)	(.009)	(.013)
College graduate	.673	.666	.712	.708	.647	.639	.698	.700	.639	.629	.689	.683
	(.004)	(.004)	(.009)	(.014)	(.004)	(.005)	(.011)	(.016)	(.004)	(.005)	(.010)	(.017)
High school–Dropout	.237	.236	.217	.257	.215	.216	.191	.234	.210	.205	.199	.226
College-Dropout	.673	.666	.712	.708	.647	.639	.698	.700	.639	.629	.689	.683
College–High school	.437	.430	.494	.450	.432	.423	.508	.466	.429	.424	.490	.457
Adjusted R^2	.277	.263	.308	.313	.307	.291	.355	.351	.305	.287	.342	.347
Observations	223,046	185,465	26,160	10,866	154,742	130,817	15,716	7,815	223,046	185,465	26,160	10,866
F-test ($\Pr > F$):												
GED = Dropout	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000
GED = High school	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000

Note.—See the note to table 3 for sample definitions and regression controls. Standard errors are in parentheses.

GED status as a match criterion in the CPS hot deck that causes the match bias and to correct for possible nonresponse bias in our final estimates.

We impute wages using the CPS hot deck with an added GED educational category. In order to account for the uncertainty associated with the imputed wage estimates of nonrespondents, we use the bootstrapping algorithm of Shao and Sitter (1996). This procedure produces unbiased estimates of standard errors by reimputing missing wages for the bootstrap replicates.

The last columns of tables 3 and 4 show that the estimates obtained from either reallocating wages or dropping those who do not report earnings are nearly identical. This is entirely consistent with the findings of Hirsch and Schumacher (2004), who show similar results comparing the wages of union and nonunion workers. This evidence does not prove the absence of nonresponse bias, but it is suggestive that for our present analysis the nonresponse bias is minimal.

Bollinger and Hirsch (2006, in this issue) present a detailed description of the CPS allocation procedure and an analysis of its shortcomings. They also present an analysis of the implications of census allocations on other outcomes besides wages. The primary focus of our article is on estimating the direct effects of GED certification on the wages of dropouts. CPS imputation bias is only a part of our story, but it is the main thrust of the Bollinger-Hirsch analysis.

We focus on estimating the return to the GED using a variety of data sets and methods to adjust for selection effects and ability bias to show that estimated direct returns to GED certification are very low. For the remainder of this article, we use the most expedient method of dealing with allocated values—and the one advocated by Bollinger and Hirsch (2006, in this issue)—by dropping employed workers who do not report earnings rather than imputing missing wages. Due to the richer set of conditioning variables available in the NLSY compared to the CPS, we are able to correct NLSY-based estimates for sample selection bias due to employment status using both parametric and semiparametric selection correction models; this is described in more detail in the next section.

III. Ability Bias

Even though the exclusion of allocated earners dramatically reduces the size of the estimated return to a GED credential, the resulting wage relative to the wage of dropouts is both positive and statistically significant for both males and females across all race groups. Cameron and Heckman (1993) found that positive returns to GED certification could be attributed entirely to ability bias. Those who choose to take the GED examination are a select group from the dropout pool. The distributions of measured ability of the people who choose to take the GED and those who do not are very different. The CPS data do not include any measures of ability. Unobserved ability may be driving the observed wage differences between education categories. Accordingly, we turn to other strategies to control for ability bias and to richer data sets.

A. NLSY Data

This section uses the National Longitudinal Survey of Youth (NLSY79) to control for ability bias. The NLSY is a representative sample of young Americans who were between the ages of 14 and 21 at the time of the first interview in 1979. The NLSY is made up of three subsamples: (1) a random sample of 6,111 noninstitutionalized civilian youths, (2) a supplemental sample of 5,295 youths designed to oversample civilian Hispanics, blacks, and economically disadvantaged whites, and (3) a sample of 1,280 youths who were ages 17–21 as of January 1, 1979, and who were enlisted in the military as of September 30, 1978. The NLSY collects information on parental background, schooling decisions, labor market experiences, and cognitive test scores. Our sample includes only the random sample and the black and Hispanic oversamples of the 1979–2000 waves. Our wage analysis is carried out separately for males and females and excludes those who are enrolled in school, those who have wages less than $.50 or greater than $200 per hour, and those who are self-employed.

In 1980, the Armed Services Vocational Aptitude Battery (ASVAB) was administered to all NLSY respondents, with a completion rate of about 94% for the sample. We use the AFQT test score as our measure of ability.[5] Figure 3 presents the distributions of AFQT scores by education and race for the NLSY. The differences in ability between GED recipients and dropouts for both males and females of all races are large and statistically significant.[6] In fact, GED recipients have nearly the same measured ability as high school graduates who do not continue on to college across all races.

B. Estimation

In order to determine the importance of ability bias in generating the estimated returns to GED certification using CPS data, we compare CPS estimates to those obtained in the NLSY, both including and excluding the

[5] The ASVAB consists of a battery of 10 tests: general science, arithmetic reasoning, word knowledge, paragraph comprehension, numerical operations, coding speed, auto and shop information, mathematics knowledge, mechanical comprehension, and electronics information. The Armed Forces Qualification Test (AFQT) is the sum of word knowledge, arithmetic reasoning, paragraph comprehension, and numeric operations components of the ASVAB and is a general measure of trainability used by the military for enlistment screening and job assignment.

[6] Wilcoxon rank sum tests of stochastic dominance show strong differences.

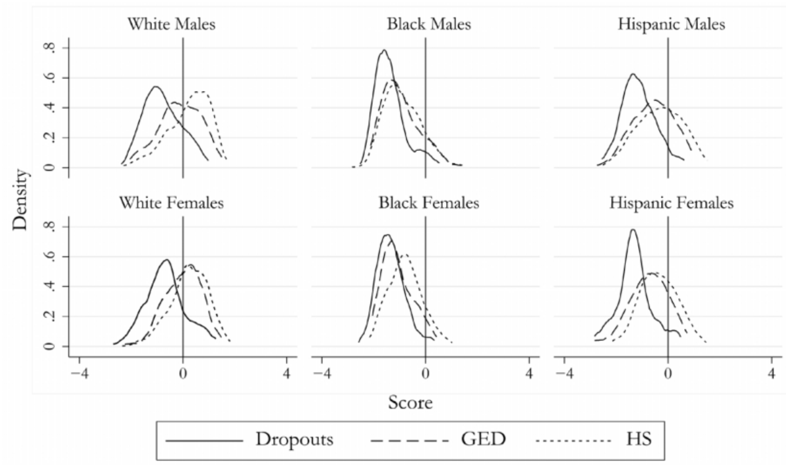

Fig. 3.—Density of NLSY AFQT scores by race and gender

AFQT score. Tables 5 and 6 show that the estimated returns to certification across race groups using NLSY data for respondents who are between 20 and 39 years of age are similar to those obtained from the CPS. The exception is for black males. For this group, the CPS estimate is higher. Returns to GED certification are also positive and, in all but one case, they are statistically significant across all race groups using standard significance levels. However, when the AFQT score is controlled for, the estimated GED effect is essentially zero for males. The estimated effect for females is still slightly positive across all race groups, but it is always statistically insignificant. All wage differentials between GED recipients and dropouts can be eliminated by accounting for ability. The positive GED effects obtained in the CPS arise from an unobserved ability bias that results from high-ability dropouts self-selecting into the GED program.

To test the robustness of the NLSY estimates to sample selection bias problems that may arise from excluding workers on the basis of their labor force status, we estimate a parametric selection correction model due to Heckman (1979).[7] As shown in the last columns of tables 5 and 6, accounting for selective participation in the work force does not overturn the conclusion that GEDs are paid the wages of high school dropouts at the same ability level.

One method for controlling for unobserved ability is to use fixed effects

[7] At our Web site (http://www.jenni.uchicago.edu/ged_imputation), we report estimates based on a semiparametric factor model structure (based on Carneiro, Hansen, and Heckman [2003]). The parametric and semiparametric estimates agree (see tables A1–A4, which are at the Web site or available from the authors).

Table 5
NLSY OLS and Parametric Selection Corrected Hourly Wage Regressions for Males by Race

	Model 1				Model 2				Model 3			
	No Selection on AFQT				Including AFQT				Controlling for AFQT and Selection*			
	All	Whites	Blacks	Hispanics	All	Whites	Blacks	Hispanics	All	Whites	Blacks	Hispanics
GED, no college	.065	.068	.049	.092	−.004	−.004	.003	.006	−.008	.004	.000	−.020
	(.020)	(.035)	(.030)	(.040)	(.021)	(.034)	(.031)	(.042)	(.021)	(.034)	(.031)	(.044)
High school, no college	.131	.165	.080	.140	.044	.071	.026	.032	.035	.065	.029	.001
	(.014)	(.021)	(.024)	(.031)	(.015)	(.023)	(.025)	(.032)	(.015)	(.022)	(.024)	(.034)
College graduate	.477	.472	.500	.523	.274	.276	.312	.253	.257	.261	.307	.207
	(.018)	(.024)	(.037)	(.055)	(.022)	(.031)	(.043)	(.057)	(.022)	(.031)	(.043)	(.057)
AFQT score	. . .	. . .	. . .	. . .	.113	.109	.113	.125	.110	.104	.111	.123
	. . .	. . .	. . .	. . .	(.008)	(.012)	(.015)	(.016)	(.008)	(.012)	(.015)	(.018)
High school–Dropout	.131	.165	.080	.140	.044	.071	.026	.032	.035	.065	.029	.001
College-Dropout	.477	.472	.500	.523	.274	.276	.312	.253	.257	.261	.307	.207
College–High school	.346	.307	.420	.383	.230	.205	.286	.221	.221	.197	.279	.206
Adjusted R^2	.303	.299	.261	.212	.331	.324	.296	.250	. . .	. . .	. . .	. . .
Observations	33,573	18,199	9,009	6,365	32,054	17,351	8,735	5,968	36,706	19,126	11,168	6,412
F-test ($\Pr > F$):												
GED = Dropout	.001	.055	.107	.022	.842	.899	.925	.882	.701	.909	.993	.650
GED = High school	.000	.004	.257	.219	.010	.018	.390	.516	.021	.053	.284	.605

NOTE.—All dummy variables are defined exclusively. Dropouts are the excluded category. Persons enrolled in school at each age are deleted as are those people who have wages less than $.50 or more than $200 an hour (in 2000$), are younger than 20 years of age or older than 39, or are self-employed. Controls for central city status, married with spouse present, year of survey, and region of residence, as well as a quadratic in age and race dummies (where appropriate), are included in each regression but not shown. Reported standard errors (in parentheses) are corrected for heteroscedasticity and clustering with the Huber-White sandwich estimator.

* We use a parametric model selection correction due to Heckman (1979). For both males and females, the participation equation includes race dummies, family income in 1979, mother's and father's education, broken home status at 14, urban status at 14, south at 14, number of siblings, local unemployment rate, age, and age squared. For the female model, spouse's income, number of children in the household, and dummies for the presence of a baby or toddler in household are also included.

Table 6
NLSY OLS and Parametric Selection Corrected Hourly Wage Regressions for Females by Race

	Model 1				Model 2				Model 3			
	No Selection on AFQT				Including AFQT				Controlling for AFQT and Selection			
	All	Whites	Blacks	Hispanics	All	Whites	Blacks	Hispanics	All	Whites	Blacks	Hispanics
GED, no college	.113	.093	.122	.111	.027	.012	.033	.027	.017	.000	.015	.032
	(.021)	(.029)	(.039)	(.041)	(.021)	(.030)	(.035)	(.043)	(.021)	(.031)	(.034)	(.045)
High school, no college	.225	.199	.248	.247	.130	.123	.141	.123	.101	.096	.107	.116
	(.016)	(.023)	(.032)	(.030)	(.016)	(.024)	(.029)	(.034)	(.016)	(.027)	(.028)	(.036)
College graduate	.651	.607	.667	.769	.429	.413	.415	.507	.376	.372	.345	.475
	(.019)	(.026)	(.037)	(.041)	(.023)	(.032)	(.041)	(.052)	(.023)	(.038)	(.039)	(.054)
AFQT score	. . .	. . .	. . .	. . .	.131	.118	.151	.146	.126	.123	.135	.131
	. . .	. . .	. . .	. . .	(.009)	(.118)	(.016)	(.019)	(.009)	(.012)	(.016)	(.021)
High school–Dropout	.225	.199	.248	.247	.130	.123	.141	.123	.101	.096	.107	.116
College-Dropout	.651	.607	.667	.769	.429	.413	.415	.507	.376	.372	.345	.475
College–High school	.426	.408	.419	.522	.299	.290	.274	.384	.276	.276	.238	.359
Adjusted R^2	.309	.298	.312	.307	.339	.323	.349	.349	. . .	. . .	. . .	. . .
Observations	28,489	16,225	7,341	4,923	27,567	15,645	7,195	4,727	42,707	22,186	12,923	7,598
F-test (Pr > F):												
GED = Dropout	.000	.002	.002	.007	.187	.689	.268	.534	.428	.999	.673	.394
GED = High school	.000	.000	.000	.001	.000	.000	.001	.013	.000	.001	.002	.049

Note.—See the notes to table 5 for discussion of the selection model and exclusions. Standard errors are in parentheses.

models. Although the CPS was not originally intended as a longitudinal data set, many researchers construct 2-year panels from the fourth and eighth survey months. We now exploit this longitudinal structure in an attempt to correct for ability bias using the CPS sample.

A number of important caveats need to be given before presenting the estimates based on a fixed effect analysis. First, the CPS survey follows households and not individuals from one survey to the next. A person who moves out of a household will not appear in the next survey. This is of particular importance for our estimation because the subpopulation we are interested in—those who attain a GED between survey rounds—tends to be younger in age and significantly more likely to move between survey rounds compared to older individuals. This biases longitudinal samples toward those who are more stable, that is, toward those who do not move between surveys. Second, changes in GED status could be due to the mismatching of individuals or errors in reporting education from proxy responses. While every effort is made to eliminate error due to the first consideration by matching individuals on a number of demographic characteristics, the second source of error is less easily dealt with. Because the CPS surveys one member of a household, and he or she responds for the entire household, changes in the educational status of an individual, particularly their GED status, occur quite frequently when different members of a household respond. This type of misreporting may be particularly severe for GED recipients given that a proxy respondent may be unaware that someone has a GED and because a GED is often assumed to be the same degree as a regular high school diploma and therefore is frequently reported as such. Finally, if a person does not report wages or is not working in either the fourth or eighth survey months, then we cannot use them in the estimation. Using only households with wages reported in both interviews leads to a small sample of individuals for whom we can estimate a fixed effect model, and the bias inherent in the sample from this exclusion is unknown.

We present estimates from two longitudinal models using CPS data that attempt to control for the ability bias that plagues OLS estimates. The first model is a standard fixed effects regression that differences out individual specific effects. The second model identifies those who obtain a GED any time during the sample period and then enters a dummy variable into the wage equation indicating whether an individual is in a pre–GED attainment state or a post-GED state. Comparing the pre-GED and post-GED coefficients helps to determine the causal effect of GED certification on wages. No difference in pre- and post-earnings indicates that the GED effect is zero and that cross-section estimates seriously overstate the value of a GED. A positive difference in pre- and post-GED earnings is evidence that supports the claim that the GED has a direct effect on earnings. In addition, if pre-GED individuals are already earning significantly more than dropouts before certification, then this is evidence that preexisting productive factors, such as unmeasured ability, are driving the higher wage returns of GED recipients, not any true direct effect of the GED.

Excluding allocated earners is of particular importance when estimating longitudinal models in the CPS. Tables 7 and 8 show that dramatically different conclusions are reached depending on the treatment of allocated earners. Including allocated earners results in large differences in earnings pre- and postcertification for both males and females. After dropping allocated earners, we find no evidence of a positive treatment effect of the GED on earnings. GED recipients earn the same in both the pre- and post-GED states, and they earn more before certification than other dropouts. Fixed effects models strengthen the conclusion that positive GED returns from cross-section estimates are not causal. The inclusion of allocated earners once again generates an apparently large and statistically significant positive effect of certification for both males and females. Dropping allocated observations results in a zero estimated direct effect of certification after controlling for unobserved individual effects. Estimates from the NLSY sample confirm the conclusions drawn from the CPS.

IV. Cohort versus Age Effects and Further Evidence on Ability Bias

Proponents of the GED program argue that a GED title may confer little initial benefit but that, after time, GED holders will experience higher wage growth than dropouts who do not certify. This claim is based on an analogy with the returns to college. In the early years after completing schooling, college graduate earnings do not exceed those of high school graduates of a comparable age. In later years, their earnings far exceed the returns to high school graduates as returns to investment are harvested. If the GED is an investment with long-term yields, we would expect to see higher wage differentials between GED recipients and high school dropouts at older ages. Tables 9 and 10 shed light on this question by estimating the return to GED certification in the CPS by age groups for white males and females, respectively. We focus on whites because the minority samples in NALS and the NLSY are too small.[8] We consider only GED recipients who get no further education in order to estimate direct effects of certification on wages. For white males, we find evidence that apparently supports the notion that the GED is an investment. GED recipients in each successive age category have higher estimated returns

[8] The estimates for minorities are consistent with those for whites, but the cells are small and the standard errors are large. See tables A5 and A6 in the table appendix (on the Web site at http://www.jenni.uchicago.edu/ged_imputation or available from the authors) for these results.

Table 7
OLS Pre- versus Post-GED and Fixed Effects Estimates for CPS and NLSY White Males

	CPS Pre-GED vs. Post-GED OLS Wage Regressions*		CPS Fixed Effects Wage Regressions†		NLSY Pre-GED vs. Post-GED OLS Wage Regressions‡	NLSY Fixed Effects Wage Regressions§
	With Allocations	Without Allocations	With Allocations	Without Allocations		
Pre-GED	.049	.073	. . .	. . .	.098	. . .
	(.037)	(.043)	. . .	. . .	(.037)	. . .
Post-GED	.124	.078	.116	.033	.082	−.036
	(.009)	(.011)	(.057)	(.056)	(.036)	(.050)
High school, no college	.196	.167	. . .	. . .	.188	. . .
	(.006)	(.007)	. . .	. . .	(.021)	. . .
College graduate	.556	.523	. . .	. . .	.524	. . .
	(.006)	(.008)	. . .	. . .	(.025)	. . .
Adjusted R^2	.318	.355	.012	.009	.320	.086
Observations	96,711	67,232	14,451	9,981	18,956	4,041
			F-test: (Pr > *F*)			
Pre-GED = Dropout	.185	.094	. . .	. . .	.009	. . .
Pre-GED = Post-GED	.042	.091	. . .	. . .	.709	. . .
Post-GED = Dropout	.000	.000	.041	.550	.026	.475

* See the notes to table 3 for sample definitions and regression controls. The exception is this sample is matched white males between the ages of 18 and 49.
† High school and college graduates are omitted in fixed effects regressions as well as any time invariant controls listed under table 3.
‡ See the notes to table 5 for sample definitions and regression controls. The only exception is that this sample is between the ages of 18 and 46.
§ High school and college graduates are omitted in fixed effects regressions as well as any time invariant controls listed under table 5.

Table 8
OLS Pre- versus Post-GED and Fixed Effects Estimates for CPS and NLSY White Females

	CPS Pre-GED vs. Post-GED OLS Wage Regressions*		CPS Fixed Effects Wage Regressions†		NLSY Pre-GED vs. Post-GED OLS Wage Regressions‡	NLSY Fixed Effects Wage Regressions§
	With Allocations	Without Allocations	With Allocations	Without Allocations		
Pre-GED	.029	.065	. . .	. . .	.056	. . .
	(.035)	(.037)	. . .	. . .	(.049)	. . .
Post-GED	.127	.096	.011	−.074	.086	−.050
	(.009)	(.011)	(.058)	(.056)	(.032)	(.054)
High school, no college	.213	.203	. . .	. . .	.207	. . .
	(.007)	(.008)	. . .	. . .	(.025)	. . .
College graduate	.662	.644	. . .	. . .	.616	. . .
	(.007)	(.008)	. . .	. . .	(.028)	. . .
Adjusted R_2	.297	.327	.004	.013	.298	.055
Observations	87,139	62,945	9,603	6,933	16,989	2,671
			F-test: (Pr > *F*)			
Pre-GED = Dropout	.409	.081	. . .	. . .	.254	
Pre-GED = Post GED	.005	.390	. . .	. . .	.520	
Post-GED = Dropout	.000	.000	.857	.186	.009	.352

Note.—Standard errors are in parentheses.
* See the note to table 4 for sample definitions and regression controls. The exception is this sample is matched white females between the ages of 18 and 49.
† High school and college graduates are omitted in fixed effects regressions as well as any time invariant controls listed under table 4.
‡ See the notes to table 6 for sample definitions and regression controls. The only exception is that this sample is between the ages of 18 and 46.
§ High school and college graduates are omitted in fixed effects regressions as well as any time invariant controls listed under table 6.

Table 9
CPS-NLSY Comparison, OLS and Selection Corrected Hourly Wage Regressions for White Males

	CPS				CPS NLSY Cohort*	NLSY				NLSY AFQT and Selection	
						Excluding AFQT		Including AFQT			
	20–29	30–39	40–49	50–59	34–46	20–29	30–39	20–29	30–39	20–29	30–39
GED, no college	.031	.082	.104	.130	.076	.052	.067	-.031	−.040	−.024	−.031
	(.011)	(.013)	(.014)	(.017)	(.014)	(.035)	(.043)	(.035)	(.042)	(.035)	(.042)
High school, no college	.112	.173	.234	.220	.195	.152	.206	.057	.062	.047	.057
	(.008)	(.008)	(.010)	(.012)	(.009)	(.020)	(.027)	(.022)	(.029)	(.022)	(.029)
College graduate	.363	.544	.615	.589	.598	.387	.584	.198	.318	.175	.305
	(.009)	(.009)	(.010)	(.012)	(.010)	(.024)	(.030)	(.031)	(.038)	(.031)	(.037)
AFQT score	. . .	. . .	. . .	. . .	. . .	. . .	. . .	.111	.153	.104	.149
	. . .	. . .	. . .	. . .	. . .	. . .	. . .	(.012)	(.015)	(.012)	(.015)
High school–Dropout	.112	.173	.234	.220	.195	.152	.206	.057	.062	.047	.057
College-Dropout	.363	.544	.615	.589	.598	.387	.584	.198	.318	.175	.305
College–High school	.250	.371	.381	.369	.403	.235	.377	.142	.256	.128	.248
Adjusted R^2	.246	.283	.267	.228	.294	.214	.278	.244	.317	. . .	. . .
Observations	29,120	40,190	38,916	24,418	34,184	10,625	8,284	10,180	7,930	11,795	8,501
F-test ($\Pr > F$):											
GED = Dropout	.001	.000	.000	.000	.000	.134	.113	.367	.344	.487	.459
GED = High school	.000	.000	.000	.000	.000	.003	.001	.006	.009	.026	.025

Note.—See the notes to tables 3 and 5 for sample definitions and controls. Parametric selection model estimates are shown. See the note to table 5 for details of the estimation procedure. Standard errors are in parentheses.

* This is a cohort of persons from the CPS in the years 1998–2003 who were born in the years 1957–64, the birth years of the NLSY cohort.

Table 10
CPS-NLSY Comparison, OLS and Selection Corrected Hourly Wage Regressions for White Females

	CPS				CPS NLSY Cohort*	NLSY				NLSY AFQT and Selection	
						Excluding AFQT		Including AFQT			
	20–29	30–39	40–49	50–59	34–46	20–29	30–39	20–29	30–39	20–29	30–39
GED, no college	.095	.102	.119	.105	.108	.084	.119	.011	.011	−.004	−.001
	(.013)	(.014)	(.014)	(.018)	(.015)	(.033)	(.044)	(.034)	(.046)	(.034)	(.046)
High school, no college	.164	.229	.251	.229	.243	.172	.222	.092	.118	.029	.108
	(.009)	(.010)	(.011)	(.011)	(.011)	(.026)	(.033)	(.028)	(.036)	(.027)	(.035)
College graduate	.527	.703	.683	.619	.704	.483	.732	.298	.510	.194	.501
	(.010)	(.010)	(.011)	(.012)	(.012)	(.029)	(.035)	(.036)	(.044)	(.036)	(.044)
AFQT score	. . .	. . .	. . .	. . .	. . .	. . .	. . .	.126	.142	.120	.151
	. . .	. . .	. . .	. . .	. . .	. . .	. . .	(.013)	(.019)	(.013)	(.018)
High school–Dropout	.164	.229	.251	.229	.243	.172	.222	.092	.118	.029	.108
College-Dropout	.527	.703	.683	.619	.704	.483	.732	.298	.510	.194	.501
College–High school	.363	.474	.432	.391	.461	.310	.510	.207	.393	.165	.393
Adjusted R^2	.323	.321	.261	.230	.290	.217	.300	.244	.327	. . .	. . .
Observations	26,307	35,136	38,342	25,211	31,642	9,442	6,914	9,110	6,671	13,182	9,307
F-test ($\Pr > F$):											
GED = Dropout	.000	.000	.000	.000	.000	.010	.007	.738	.811	.899	.991
GED = High school	.000	.000	.000	.000	.000	.001	.006	.002	.005	.213	.004

Note.—See the notes to tables 3 and 5 for sample definitions and controls. Parametric selection model estimates are shown. See the note to table 5 for details of the estimation procedure. Standard errors are in parentheses.

* This is a cohort of persons from the CPS in the years 1998–2003 who were born in the years 1957–64, the birth years of the NLSY cohort.

to certification. For white females, the pattern of returns is quite different, being nearly constant across age groups.

It is not clear whether the higher returns to GED certification for males at older ages are due to age or cohort effects. It is not possible to answer the age versus cohort question using cross-sectional data such as the CPS (see Heckman and Robb 1985). It may be that the acquisition of the GED title causes the wage differential to increase between male GED recipients and dropouts at older ages, or it may be that older birth cohorts exhibit higher returns due to unobservable differences in quality between GED recipients and dropouts that are not present in more recent birth cohorts. Comparing CPS to NLSY data and data from the National Adult Literacy Survey (NALS) that will be discussed further in Section IV.A, we find that higher estimated returns for older groups are due to cohort differences and not increased wage growth resulting from GED acquisition.

By comparing GED estimates for a cohort comparable to the NLSY cohort in the CPS to estimates reported by Cameron and Heckman (1993) at younger ages, Jaeger and Clark (2006) claim to find evidence of strong GED life cycle wage growth. They report that estimated returns to GED certification in the monthly CPS data for the NLSY cohort—those born between 1957 and 1964—far exceed the estimates reported at age 25 and 28 in Cameron and Heckman's analysis. They conclude that by the time GED recipients are in their late thirties to early forties, the GED title has helped them "catch up" to high school graduates and to far exceed the wage growth exhibited by high school dropouts who do not exam certify.

Tables 9 and 10 show that this conclusion arises as an artifact of inclusion of allocated earners in the Jaeger and Clark samples.[9] We construct an NLSY birth cohort in the CPS. It is the sample in the CPS survey years 1998–2003 that was born in the period 1957–1964, the same years in which the NLSY cohort is born. In 1998, these people are ages 34–41. In 2003 they are ages 40–46. After excluding those who do not report their earnings, the estimated GED returns for the NLSY-comparable cohort constructed from the CPS data are nearly identical to the estimates obtained from the NLSY when the sample is in their twenties and again in their thirties.

Both data sources show that GED recipient wage growth is not greater than that exhibited by high school dropouts. Furthermore, the positive wage differences between GED recipients and uncertified dropouts is completely accounted for by the inclusion of the ability measure for males and females of all ages. However, the returns to college remain. This is

[9] The log hourly wage regressions in the NLSY and CPS comparisons include similar covariates and are based on the same sample restrictions to make the estimates comparable.

clear evidence of investment occurring in college. However, there is no investment occurring in GED certification.

Tables 11 and 12 strengthen this conclusion by comparing male and female estimates of the CPS-NLSY cohort with cross-sectional estimates obtained from the NLSY sample at ages 25, 28, 30, 35, and 38. We again see that the estimated returns to GED certification and high school graduation for this cohort are remarkably similar between the two data sources and across ages. The estimated GED-dropout difference at ages 35 and 38 is no different than those previously found by Cameron and Heckman (1993) at ages 25 and 28. According to official published statistics from the GED testing service, over 75% of GED recipients acquire the degree before the age of 25. Therefore, the majority of the wage sample at ages 35 and 38 have had their diplomas for over 10 years, which is ample time for any positive net benefits to accrue. If GED recipients have not shown positive wage growth within 10 years of obtaining the title, it is highly unlikely that they will do so later. Both the NLSY and CPS data strongly reject the hypothesis of postcertification life cycle wage growth posited by Jaeger and Clark (2006), as well as Murnane et al. (1999) and Boudett et al. (2000), once match bias is accounted for and estimation is performed on comparable cohorts.

Controlling for ability differences in the NLSY data produces no statistically significant differences in wages between GED recipients and dropouts who do not certify for both males and females at all ages. It is possible that the differences in wages between GED recipients, high school graduates, and dropouts observed in the CPS can be completely accounted for by unobserved ability differences as well. Given that the NLSY cohort shows little life cycle wage growth, it is also plausible that the higher returns to GED certification seen for older birth cohorts in CPS data are due to a growing difference in this ability bias between GED recipients and dropouts. Two—not necessarily mutually exclusive—possibilities may explain the data. The first is that, as the GED program has expanded rapidly over the last 30 years, the quality of GED recipients may have declined. Second, the quality of dropouts may have improved. Figure 4 shows that the quality of dropouts, as measured by their years of completed schooling, has improved across cohorts, while GED quality has remained roughly constant. Male and female dropouts of all races have obtained greater levels of schooling, while the completed secondary schooling levels of GEDs are nearly constant across all birth cohorts. The greater schooling attainment of dropouts may indicate that the skill gap between GED recipients and dropouts is closing across cohorts, or it may be the consequence of social promotion. Both factors may be at work. We now turn to the National Adult Literacy Survey (NALS) data to explore this issue further. It provides data on literacy skills of successive cohorts.

Table 11
CPS-NLSY Comparison, OLS Log Hourly Wage Regressions for Males by Age

	NLSY Excluding AFQT Score					CPS NLSY Cohort*	NLSY Including AFQT Score				
	25	28	30	35	38	34–46	25	28	30	35	38
GED, no college	.059	.043	.015	.050	.079	.085	−.034	−.037	−.065	−.043	−.084
	(.038)	(.035)	(.034)	(.041)	(.061)	(.015)	(.039)	(.038)	(.036)	(.041)	(.062)
High school, no college	.170	.141	.161	.157	.194	.208	.059	.044	.056	.029	−.015
	(.024)	(.023)	(.024)	(.028)	(.041)	(.010)	(.026)	(.024)	(.025)	(.030)	(.041)
College graduate	.373	.443	.496	.650	.714	.611	.134	.235	.269	.382	.333
	(.033)	(.029)	(.029)	(.034)	(.048)	(.011)	(.040)	(.035)	(.037)	(.040)	(.053)
AFQT score	. . .	. . .	. . .	. . .	. . .	. . .	.134	.124	.133	.149	.192
	. . .	. . .	. . .	. . .	. . .	. . .	(.013)	(.013)	(.013)	(.014)	(.019)
High school–Dropout	.170	.141	.161	.157	.194	.208	.059	.044	.056	.029	−.015
College- Dropout	.373	.443	.496	.650	.714	.611	.134	.235	.269	.382	.333
College–High school	.203	.302	.335	.493	.520	.403	.074	.191	.214	.353	.348
Adjusted R^2	.168	.229	.258	.314	.358	.308	.207	.269	.294	.352	.415
Observations	2,247	2,367	2,400	2,287	1,088	30,549	2,165	2,254	2,298	2,196	1,039
F-test ($\Pr > F$):											
GED = Dropout	.119	.217	.650	.219	.193	.000	.390	.330	.067	.287	.173
GED = High school	.001	.003	.000	.003	.032	.000	.006	.018	.000	.038	.189

Note.—See the notes to tables 3 and 5 for sample definitions and controls. Standard errors are in parentheses.
* This is a cohort of persons from the CPS in the years 1998–2003 who were born in the years 1957–64, the birth years of the NLSY cohort.

Table 12
CPS-NLSY Comparison, OLS Log Hourly Wage Regressions for Females by Age

	NLSY Excluding AFQT Score					CPS NLSY Cohort*	NLSY Including AFQT Score				
	25	28	30	35	38	34–46	25	28	30	35	38
GED, no college	.096	.117	.109	.114	.149	.107	.014	.007	−.014	.028	.022
	(.047)	(.048)	(.048)	(.046)	(.062)	(.015)	(.048)	(.049)	(.049)	(.047)	(.065)
High school, no college	.210	.234	.275	.272	.315	.237	.113	.123	.125	.161	.160
	(.033)	(.035)	(.034)	(.035)	(.051)	(.011)	(.034)	(.037)	(.036)	(.038)	(.055)
College graduate	.489	.640	.728	.799	.858	.700	.277	.417	.432	.573	.570
	(.036)	(.039)	(.037)	(.038)	(.062)	(.011)	(.042)	(.048)	(.045)	(.048)	(.073)
AFQT score	. . .	. . .	. . .	. . .	. . .	. . .	.137	.142	.180	.142	.146
	. . .	. . .	. . .	. . .	. . .	. . .	(.014)	(.016)	(.017)	(.018)	(.023)
High school–Dropout	.210	.234	.275	.272	.315	.237	.113	.123	.125	.161	.160
College-Dropout	.489	.640	.728	.799	.858	.700	.277	.417	.432	.573	.570
College–High school	.279	.405	.452	.527	.543	.463	.164	.294	.307	.412	.411
Adjusted R^2	.176	.261	.318	.319	.311	.297	.213	.295	.361	.350	.342
Observations	1,855	1,832	1,873	1,857	913	29,452	1,803	1,782	1,812	1,800	885
F-test ($\Pr > F$):											
GED = Dropout	.041	.014	.025	.013	.017	.000	.765	.879	.783	.554	.733
GED = High school	.003	.002	.000	.000	.001	.000	.011	.003	.001	.001	.004

Note.—See the notes to tables 3 and 5 for sample definitions and controls. Standard errors are in parentheses.
* This is a cohort of persons from the CPS in the years 1998–2003 who were born in the years 1957–64, the birth years of the NLSY cohort.

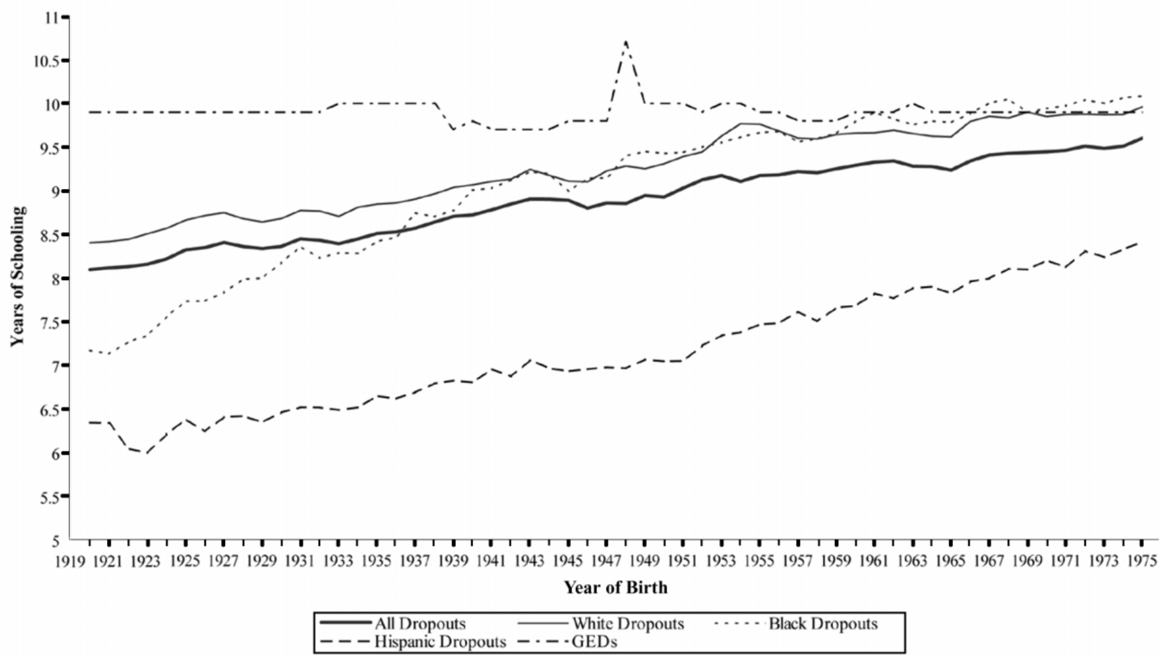

FIG. 4.—Average years of secondary schooling for dropouts and GEDs by year of birth

A. NALS Data

The National Assessment of Literacy (NALS) is a decennial survey administered by the NCES to a random sample of the U.S. adult population to determine their literacy skills. The 1992 sample used in this section consists of a random sample of 13,600 adults ages 16 and over and a state supplement of 11,344 adults. The NALS testing battery consists of three separate tests designed to measure three types of skills: prose, document, and quantitative skills. Unlike the CPS, the NALS sample does not ask respondents to report their hours of work. Therefore, all comparisons between CPS and NALS data are based on weekly wage regressions. These regressions exclude those individuals who have weekly wages less than $100 or more than $4,000 (in 2000$), those who are younger than 20 years of age or older than 64, and those who are Aleut, Eskimo, or Native American. Controls for central city status, married with spouse present, year of survey, and region of residence, as well as a quadratic in age and race dummies (where appropriate), are included in each regression.[10]

B. NALS Test Scores

As measured by the NALS test scores, people who choose to take the GED test are as capable in their basic cognitive skills as high school graduates and are more capable than high school dropouts who choose not to certify. Figure 5 shows the distributions of total NALS test scores derived from the average over all three components of the NALS battery by race and education status for the native born. The distributions of NALS scores for high school graduates and GED recipients are nearly identical across all races, while dropouts have lower scores. In terms of basic literacy skills, the GED exam effectively sorts between those who pass the exam and those who do not.

Since the gap in years of schooling completed between dropouts and GED recipients is narrowing across birth cohorts, we might expect to find the cognitive skill gaps between the groups to be narrowing as well. Figure 6, which presents NALS score distributions across different birth cohorts, shows that this is indeed the case. The distributions of scores for GEDs have remained nearly identical to those of high school graduates across all birth cohorts for males and females. As dropouts have obtained more years of schooling, their test score distributions are becoming more similar to GEDs across birth cohorts, but they are still statistically significantly different, even in the most recent cohort. This pattern of test scores could produce the cross-section finding of greater return to the

[10] The amount of data lost due to these exclusion restrictions for the NALS sample is comparable to data loss generated from similar restrictions on the CPS sample.

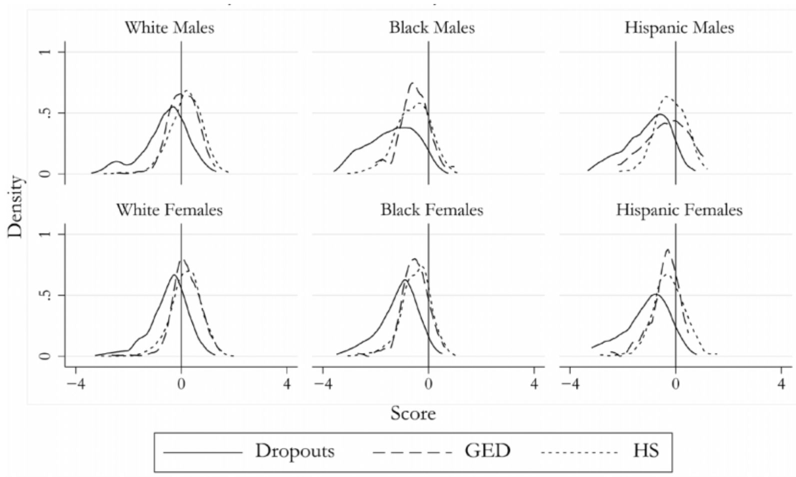

FIG. 5.—Density of NALS test scores by race for the native born

GED by age solely as a consequence of diminished ability bias for more recent cohorts. In addition, the rise in GED certification may be due in part to diminishing participation costs of preparing for the exam by uncertified dropouts. Whereas in 1950, when passing the GED would have required a substantial investment and skill acquisition for a sixth-grade droput, the average dropout from today's public school system with 10 years of education requires only minor preparation to pass the exam.[11]

C. Estimation

The returns to GED certification found in the NALS92 sample for males and females ages 20–64 closely match those found in the CPS 1998–2003. Tables 13 and 14 show that male GED recipients have 6.6% higher weekly wages than dropouts before controlling for ability. Female GED recipients earn 9.4% more than dropouts. However, these positive returns to certification are completely eliminated once we control for the NALS test score. As with the male NLSY sample, GED recipients earn less than dropouts at the same level of ability. Once again, this effect is not statistically significant. Female GED recipients show a small, but statistically insignificant, positive return to certification, adjusting for ability, much as we saw in the NLSY data. It is evident that not controlling for ability in CPS data leads to an overestimate of the wage returns to GED certification. All positive returns to certification can be completely accounted for by selection into the GED program based on ability.

[11] A 1980 study found that the median study time of GED examinees was only 20 hours. By 1989, they were preparing 30 hours for the test (Boesel et al. 1998).

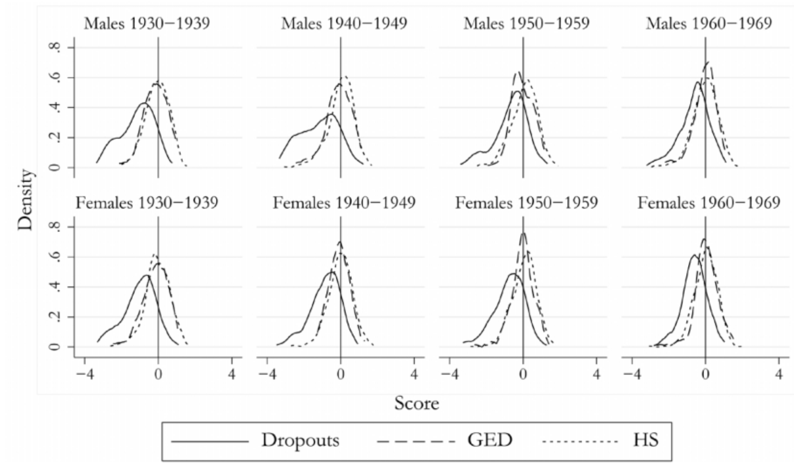

Fig. 6.—Density of NALS test scores by birth cohort for the native born

The NALS distribution of test scores across birth cohorts shows that the ability differential between GED recipients and dropouts is diminishing. In a cross section, this results in the pattern of wage returns to the GED across ages that is observed for males in the CPS. Older age groups show a higher return to certification. This is a spurious age pattern due solely to a greater ability gap between GEDs and dropouts in earlier cohorts.

Table 15 makes this point clearly by comparing estimated weekly wage returns in both the CPS and NALS for two birth cohorts. The first is the pre-NLSY cohort (those born before 1957 in CPS and NALS), and the second includes the NLSY cohort (those born 1957–64 in the CPS and NALS) and those born afterward. Once again, we see the pattern of higher returns for the older cohort in both the NALS and CPS data. However, controlling for the NALS test score, across all birth cohorts, there is no statistically distinguishable wage benefit for both male and female GED recipients. The available evidence suggests that the GED program has always selected the most able from the dropout pool and that the direct wage benefits across all certification cohorts range from small to nonexistent once we account for this selection on ability.

V. GED Returns among Immigrants

Jaeger and Clark (2006) argue that the GED has an even greater signaling effect for immigrants than for the native born. However, their study does not control for the ability differences between education groups. It is possible that the GED program is even more selective in the immigrant population than it is for natives, so that only the most able immigrants with

Table 13
NALS-CPS Comparison, OLS Log Weekly Wage Regressions for Males by Race

					NALS							
	CPS				Excluding Test Score				Including Test Score			
	All	Whites	Blacks	Hispanics	All	Whites	Blacks	Hispanics	All	Whites	Blacks	Hispanics
GED, no college	.085	.079	.107	.115	.066	.079	.057	.008	−.022	−.003	-.028	−.079
	(.007)	(.007)	(.023)	(.021)	(.043)	(.050)	(.108)	(.138)	(.043)	(.050)	(.110)	(.140)
High school, no college	.193	.190	.200	.212	.221	.241	.163	.191	.126	.147	.092	.092
	(.004)	(.005)	(.013)	(.013)	(.024)	(.030)	(.046)	(.079)	(.025)	(.031)	(.049)	(.080)
College graduate	.577	.571	.616	.619	.658	.664	.688	.639	.441	.443	.514	.407
	(.005)	(.005)	(.016)	(.019)	(.026)	(.031)	(.063)	(.091)	(.032)	(.038)	(.076)	(.111)
NALS score	. . .	. . .	. . .	. . .	. . .	. . .	. . .	. . .	.148	.156	.115	.129
	. . .	. . .	. . .	. . .	. . .	. . .	. . .	. . .	(.013)	(.016)	(.031)	(.047)
High school–Dropout	.193	.190	.200	.212	.221	.241	.163	.191	.126	.147	.092	.092
College–Dropout	.577	.571	.616	.619	.658	.664	.688	.639	.441	.443	.514	.407
College–High school	.384	.381	.415	.407	.437	.423	.525	.447	.315	.297	.422	.316
Adjusted R^2	.316	.301	.261	.303	.389	.352	.354	.316	.407	.371	.371	.337
Observations	158,603	136,796	11,704	8,026	4,077	3,236	589	245	4,077	3,236	589	245
F-test ($\Pr > F$):												
GED = Dropout	.000	.000	.000	.000	.122	.115	.596	.955	.603	.948	.799	.574
GED = High school	.000	.000	.000	.000	.000	.000	.314	.182	.000	.001	.264	.215

Note.—All dummy variables are defined exclusively. Dropouts are the excluded category. Persons enrolled in school at each age are deleted, as are those people who have weekly wages less than $100 or more than $4,000 (in 2000$); are not born in the United States; are younger than 20 years of age or older than 64; or are Aleut, Eskimo, or Native American. Controls for central city status, married with spouse present, year of survey, region of residence, and a quadratic in age and race dummies, where appropriate, are included in each regression but are not shown. Robust standard errors are shown in parentheses.

Table 14
NALS-CPS Comparison, OLS Log Weekly Wage Regressions for Females by Race

	CPS				NALS							
					Excluding Test Score				Including Test Score			
	All	Whites	Blacks	Hispanics	All	Whites	Blacks	Hispanics	All	Whites	Blacks	Hispanics
GED, no college	.127	.110	.142	.158	.094	.088	.083	.087	.023	.019	.008	.054
	(.007)	(.009)	(.020)	(.021)	(.037)	(.047)	(.085)	(.083)	(.037)	(.047)	(.084)	(.086)
High school, no college	.241	.234	.235	.266	.229	.215	.252	.233	.158	.149	.179	.192
	(.005)	(.006)	(.011)	(.014)	(.023)	(.031)	(.046)	(.067)	(.024)	(.032)	(.046)	(.070)
College graduate	.704	.686	.783	.766	.737	.706	.860	.731	.561	.530	.678	.637
	(.005)	(.006)	(.013)	(.019)	(.026)	(.033)	(.056)	(.099)	(.032)	(.039)	(.065)	(.119)
NALS score	. . .	. . .	. . .	. . .	. . .	. . .	. . .	. . .	.145	.154	.135	.064
	. . .	. . .	. . .	. . .	. . .	. . .	. . .	. . .	(.015)	(.018)	(.029)	(.051)
High school–Dropout	.241	.234	.235	.266	.229	.215	.252	.233	.158	.149	.179	.192
College-Dropout	.704	.686	.783	.766	.737	.706	.860	.731	.561	.530	.678	.637
College–High school	.463	.453	.548	.499	.508	.492	.607	.497	.403	.382	.499	.445
Adjusted R^2	.252	.235	.336	.309	.304	.279	.371	.379	.320	.295	.387	.384
Observations	150,841	126,097	15,272	7,577	3,952	2,950	750	238	3,952	2,950	750	238
F-test ($\Pr > F$):												
GED = Dropout	.000	.000	.000	.000	.011	.059	.330	.295	.528	.671	.927	.533
GED = High school	.000	.000	.000	.000	.000	.002	.041	.094	.000	.002	.034	.109

Note.—See the notes to tables 3 and 13 for sample definitions and regression controls. Standard errors are in parentheses.

Table 15
NALS-CPS Comparison, OLS Log Weekly Wage Regressions by Cohort of Birth

	NALS Males						NALS Females					
	Males, 1940–56			Males, 1957–69			Females, 1940–56			Females, 1957–69		
	CPS	NALS	NALS	CPS	NALS	NALS	CPS	NALS	NALS	CPS	NALS	NALS
GED, no college	.126	.108	.003	.084	.067	.013	.136	.106	.048	.128	.073	−.009
	(.013)	(.069)	(.071)	(.011)	(.062)	(.062)	(.013)	(.057)	(.057)	(.012)	(.061)	(.062)
High school, no college	.226	.254	.139	.209	.196	.136	.253	.244	.180	.259	.243	.171
	(.008)	(.040)	(.043)	(.007)	(.038)	(.038)	(.009)	(.039)	(.041)	(.009)	(.040)	(.040)
College graduate	.616	.757	.515	.636	.523	.353	.691	.833	.675	.751	.662	.476
	(.009)	(.042)	(.052)	(.008)	(.041)	(.049)	(.010)	(.041)	(.051)	(.009)	(.044)	(.050)
NALS score	. . .	. . .	.154	. . .	. . .	.132	. . .	. . .	.121	. . .	. . .	.165
	. . .	. . .	(.020)	. . .	. . .	(.023)	. . .	. . .	(.024)	. . .	. . .	(.025)
High school–Dropout	.226	.254	.139	.209	.196	.136	.253	.244	.180	.259	.243	.171
College-Dropout	.616	.757	.515	.636	.523	.353	.691	.833	.675	.751	.662	.476
College-High school	.390	.503	.376	.427	.327	.217	.438	.589	.495	.491	.419	.305
Adjusted R^2	.248	.360	.380	.310	.319	.336	.206	.304	.314	.248	.281	.302
Observations	51,798	1,730	1,730	61,594	1,530	1,530	53,104	1,754	1,754	55,810	1,432	1,432
F-test ($\Pr > F$):												
GED = Dropout	.000	.119	.964	.000	.283	.838	.000	.063	.405	.000	.234	.882
GED = High school	.000	.019	.030	.000	.022	.027	.000	.007	.009	.000	.002	.001

NOTE.—See the notes to tables 3 and 13 for sample definitions and regression controls. Standard errors are in parentheses.

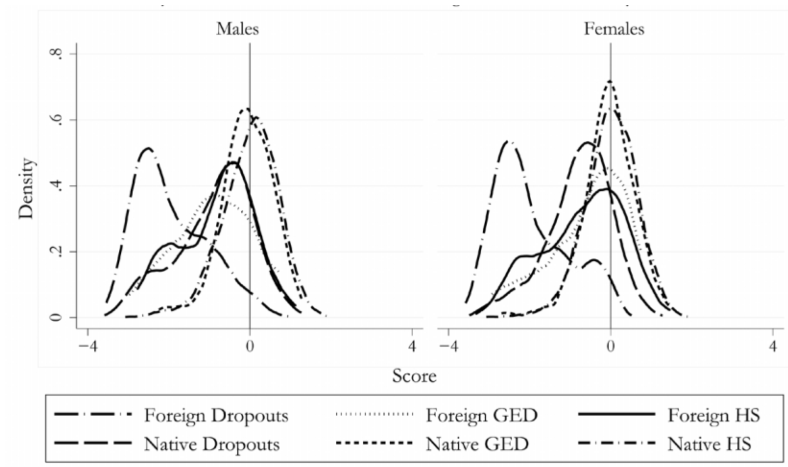

FIG. 7.—Density of NALS test scores for the foreign and native born by education

higher skills GED certify. Failure to control for these factors would cause an even wider disparity between the GED and dropout literacy and cognitive distributions than is found in native born populations, which would result in a higher perceived return for this subpopulation if these differences were not accounted for in estimation. Figure 7 reveals that the distributions of literacy levels for foreign-born dropouts, GED recipients, and high school graduates are dramatically different. While GED recipients and high school graduates are nearly identical in terms of literacy, immigrant dropouts have extremely low literacy and quantitative skills. These vast differences in basic skills among foreign-born educational groups call into question the comparability of wage returns between them, since the types of jobs available to them will be very different as well. This evidence suggests that it is even more important to adjust for literacy and cognitive skill differences among the foreign born than it is for native-born populations in order to accurately determine the value of a GED credential for immigrants.

Immigrants who take the GED also come into the country with higher levels of completed schooling in their home countries than immigrants who do not take the GED. Table 16 shows that foreign-born GED recipients and high school graduates are far more likely than foreign-born dropouts to have attended secondary schooling in their native country. The majority of immigrant dropouts only complete elementary school or less. Both high school graduates and GED recipients are also more likely to have been schooled solely in the United States, as evidenced by the percentage who did not attend school before arriving in the United States. GED recipients also have the highest probability of entering the country having completed a postsecondary vocational training program. All of these factors point

Table 16
NALS Foreign Years of Schooling Completed before Entering the United States

	Males			Females		
	Drop-outs	GED	High School	Drop-outs	GED	High School
Did not attend school	.104	.154	.145	.084	.158	.126
Primary (grades K–3)	.151	.039	.039	.158	.000	.049
Elementary (grades 4–8)	.494	.115	.089	.524	.263	.113
Secondary (grades 9–12)	.223	.577	.648	.197	.474	.635
Vocational training	.002	.077	.011	.009	.053	.014
College	.007	.000	.017	.006	.000	.005
Other	.000	.000	.006	.004	.000	.005
NA	.019	.039	.045	.018	.053	.054
Observations	431	26	179	513	38	162

toward the possibility that the GED program is even more selective for immigrants than it is for natives and that large wage differences exist between foreign GED recipients and foreign dropouts before they certify.

We now present CPS and NALS estimates of the returns to GED certification among the foreign born. We estimate the same regression model as was used to analyze the native-born population, except that we also add controls for country of birth, citizenship status, and cohort of entry into the United States. Table 17 shows that the CPS match bias that results from matching foreign-born nonrespondent GED recipients and high school graduates to native wage donors by the hot deck overstates the value of both degrees by about .05 log points for males and .06 log points for females. In contrast to the results for the native born, if we drop the unallocated workers, we cannot reject the hypothesis that GED certification is equivalent to high school graduation for both males and females, using a 10% level of statistical significance as the criterion. The data reject the null hypothesis that there are no direct wage benefits of obtaining a GED compared to staying in the dropout state, so that there appears to be a positive effect of GED certification over the dropout state. The positive estimated returns to GED certification among the foreign born in the CPS appear to be driven by unobserved ability bias. Figure 7 shows that, in the NALS data, GED recipients and those dropouts who choose not to certify have very different skill distributions. Table 17 shows that unobserved skill differences account for all differences between GED recipients and uncertified dropouts and that the positive wage returns to certification estimated in CPS data are spurious due to selection on ability.

Another interesting comparison that can be made in the NALS and CPS data is one between native and foreign-born educational groups shown in table 18. Not adjusting for ability, the ordering in the returns to education between the groups is as expected, except for the ordering

Table 17
NALS-CPS Comparison, OLS Log Weekly Wage Regressions for the Foreign Born

	CPS				NALS			
	Including Allocated Values		Excluding Allocated Values		Excluding Test Score		Including Test Score	
	Males	Females	Males	Females	Males	Females	Males	Females
GED, no college	.186	.157	.134	.090	.109	.086	.012	-.045
	(.016)	(.018)	(.019)	(.020)	(.113)	(.112)	(.110)	(.111)
High school, no college	.159	.189	.100	.138	.093	.095	−.024	−.049
	(.006)	(.007)	(.007)	(.008)	(.057)	(.067)	(.058)	(.070)
College graduate	.603	.641	.574	.591	.614	.659	.319	.397
	(.009)	(.009)	(.011)	(.012)	(.064)	(.071)	(.076)	(.084)
NALS score	. . .	. . .	. . .	. . .	. . .	. . .	.155	.153
	. . .	. . .	. . .	. . .	. . .	. . .	(.024)	(.029)
High school–Dropout	.159	.189	.100	.138	.093	.095	−.024	−.049
College-Dropout	.603	.641	.574	.591	.614	.659	.319	.397
College–High school	.445	.452	.474	.453	.521	.564	.343	.446
Adjusted R^2	.337	.309	.376	.325	.508	.350	.540	.391
Observations	46,912	33,996	31,498	22,747	629	429	629	429
F-test ($\Pr > F$):								
GED = Dropout	.000	.000	.000	.000	.333	.446	.914	.685
GED = High school	.084	.086	.063	.027	.885	.935	.751	.973

Note.—In addition to the regression controls listed in the note to table 13, all regressions include additional controls for cohort of entry, world region of birth, and whether or not the person is a citizen of the United States. Standard errors are in parentheses.

Table 18
NALS-CPS Comparison, Native versus Foreign Born

	CPS		NALS: Excluding Test Score		NALS: Including Test Score	
	Males	Females	Males	Females	Males	Females
Native, dropout	.151	.103	.108	.070	−.055	−.104
	(.007)	(.008)	(.043)	(.050)	(.043)	(.052)
Foreign, GED	.215	.192	.231	.148	.069	−.010
	(.020)	(.020)	(.118)	(.122)	(.116)	(.121)
Native, GED	.235	.229	.171	.170	−.089	−.080
	(.008)	(.009)	(.051)	(.057)	(.054)	(.061)
Foreign, high school	.152	.217	.209	.187	.061	.004
	(.007)	(.008)	(.057)	(.067)	(.056)	(.068)
Native, high school	.344	.341	.323	.301	.059	.051
	(.006)	(.007)	(.041)	(.048)	(.045)	(.052)
Foreign, college gradute	.609	.679	.723	.729	.393	.425
	(.012)	(.013)	(.056)	(.067)	(.060)	(.071)
Native, college graduate	.730	.803	.766	.812	.367	.451
	(.007)	(.007)	(.042)	(.049)	(.050)	(.058)
NALS score	. . .	. . .	. . .	. . .	.161	.154
	. . .	. . .	. . .	. . .	(.011)	(.013)
Adjusted R^2	.351	.262	.406	.306	.430	.326
Observations	183,759	167,142	4,735	4,412	4,735	4,412
F-test ($\Pr > F$):						
Native dropout = Foreign dropout	.000	.000	.011	.165	.209	.045
Native GED = Foreign GED	.316	.062	.612	.859	.154	.566
Native high school = Foreign high school	.000	.000	.026	.049	.967	.410
Foreign GED = Native dropout	.000	.000	.290	.514	.205	.428
Foreign GED = Native high school	.000	.000	.428	.199	.672	.599
Native GED = Foreign dropout	.000	.000	.001	.003	.098	.187
Native GED = Native dropout	.000	.000	.115	.015	.382	.560

Note.—Regression controls are as listed in the note to table 13. Foreign dropouts are the excluded category. Allocated earners are excluded. Standard errors are in parentheses.

for GED recipients. Despite the lower cognitive ability of foreign-born GED recipients, as shown in figure 7, they earn the same on average as native GED recipients for both males and females. After adjusting for ability in the NALS data, an interesting result emerges. Both male and female native dropouts and GED recipients earn less than their foreign counterparts, although this difference is not always statistically significant. This finding would not be predicted by a one-ability model of earnings. We conjecture that the foreign born have compensating favorable non-cognitive traits such as motivation and industriousness that offset their lower cognitive ability levels. A recent paper by Heckman, Stixrud, and Urzua (2006, in this issue) finds that both native GED recipients and dropouts have low noncognitive skills that account for their relatively poor economic and social outcomes. Our evidence suggests that foreign-born GED recipients may differ from native-born recipients in these important traits. These issues are explored more fully in a forthcoming book (see Heckman and LaFontaine 2007).

Given the small immigrant sample available in the NALS data, we must be cautious in drawing any firm conclusions about the value of GED certification among the foreign born. However, the evidence suggests that those immigrants who choose to GED certify are very different from those who do not and that any study of the value of GED certification among this population needs to be able to account for this selection.[12]

VI. Conclusion

This article shows the importance of accounting for the CPS hot deck procedure in order to obtain unbiased estimates of the return to education using CPS data. Misallocation of nonrespondent GED recipients to high school status results in a sizable overestimate of the value of GED certification. This bias does not arise from nonresponse and is more sizable among certain subpopulations such as the foreign born. Correcting for match bias is important in order to have conceptually comparable estimates of the returns to the GED across different data sources. Researchers should pay closer attention to how missing wages are allocated. Alternative allocation procedures may dramatically affect their conclusions. The importance of this warning is highlighted by our finding of a low direct wage return and zero life cycle wage growth for GED certification, in contrast to the evidence presented by Jaeger and Clark (2006), who used a biased sample.

Our evidence suggests that direct returns to GED certification are low. Selection into the GED program on the basis of cognitive ability can account for all wage differentials between those dropouts who do not certify and those who choose to do so. The gap in cognitive skills appears to be greater for older birth cohorts, and it is this greater ability bias that produces the apparent growth in the return to the GED with age that is found in the CPS data. No empirically significant life cycle wage growth can be attributed to the GED title itself. Cognitive ability differences also account for the positive effects found for GED certification among immigrants in the CPS. This evidence highlights the importance of using data with a rich set of family background and cognitive variables in order to evaluate the true impact of social programs. When we control for ability and other person-specific invariant components using longitudinal models

[12] Our estimate of the effect of the GED on immigrants may be understated because the test score we use to control for ability may be raised by preparation time qualifying for the GED. Our NALS measure is post-GED.

in the CPS, we find no causal effect of the GED. While CPS data provide a foundation from which to begin an analysis of the GED program, it cannot be considered a definitive data source. For this reason, we are currently engaged in a more refined analysis of NLSY data and other data sources to determine the treatment effects of GED certification among different groups and to expand on the analysis of differences in GED certification across cohorts reported here.

The available evidence suggests that GED certification for those who do not obtain postsecondary schooling has little or no direct causal effect on wages among men, women, older and more recent cohorts, and the foreign born. All measured differences between GED recipients and dropouts who do not certify can be accounted for by cognitive skill differences, and these are highly correlated with schooling.

While the direct benefits of GED certification appear low, there may still be an economic value to GED certification in opening postsecondary schooling and training opportunities. We discuss this issue elsewhere (see Heckman and LaFontaine 2007). As previously noted, from the CPS, we do not know the GED status of those who go on to attend institutions of higher learning. Thus, we cannot use these data to compute option values from attaining the GED. From the NLSY data, we know that about 40% of the GEDs go on to college. However, only a small percentage finish 2-year or 4-year schools. The GED opens doors to opportunities that are not realized. Overall, 3% of GEDs complete 4-year college and 5% complete an associate's degree at a 2-year college. Those who obtain vocational skills certificates do so at the same rates as high school dropouts.

What is true today was true 60 years ago when the GED program was first started: there are no cheap substitutes for classroom instruction and training.

References

Angrist, Joshua D., and Alan B. Krueger. 1999. Empirical strategies in labor economics. In *Handbook of labor economics*, vol. 3A, ed. Orley Ashenfelter and David Card, 1277–1366. New York: North-Holland.

Boesel, David, Nabeel Alsalam, and Thomas M. Smith. 1998. *Educational and labor market performance of GED recipients*. Washington, DC: Office of Educational Research and Improvement, National Library of Education, U.S. Department of Education.

Bollinger, Christopher R., and Barry T. Hirsch. 2006. Match bias due to earning imputation: The case of imperfect matching. *Journal of Labor Economics* 24, no. 3:483–519.

Boudett, Katherine, Richard J. Murnane, and John B. Willett. 2000. "Second chance" strategies for female school dropouts. *Monthly Labor Review* 123, no. 12:19–32.

Cameron, Stephen V. 1994. Assessing high school certification for women who drop out. Unpublished manuscript, Department of Economics, University of Chicago.

Cameron, Stephen V., and James J. Heckman. 1993. The nonequivalence of high school equivalents. *Journal of Labor Economics* 11, no. 1, pt. 1: 1–47.

Carneiro, Pedro, Karsten Hansen, and James J. Heckman. 2003. Estimating distributions of treatment effects with an application to the returns to schooling and measurement of the effects of uncertainty on college choice (2001 Lawrence R. Klein Lecture). *International Economic Review* 44, no. 2:361–422.

Heckman, James J. 1979. Sample selection bias as a specification error. *Econometrica* 47, no. 1:153–62.

Heckman, James J., and Paul A. LaFontaine. 2007. *America's dropout problem: The GED and the importance of social and emotional skills*. Chicago: University of Chicago Press (forthcoming).

Heckman, James J. and Richard Robb. 1985. Using longitudinal data to estimate age, period, and cohort effects in earnings equations. In *Cohort analysis in social research: Beyond the identification problem*, ed. William M. Mason and Stephen E. Fienberg. New York: Springer-Verlag.

Heckman, James J., Jora Stixrud, and Sergio Urzua. 2006. The effects of cognitive and noncognitive abilities on labor market outcomes and social behavior. *Journal of Labor Economics* 24, no. 3:411–82.

Hirsch, Barry T., and Edward J. Schumacher. 2004. Match bias in wage gap estimates due to earnings. *Journal of Labor Economics* 22, no. 3: 689–722.

Jaeger, David A., and Melissa A. Clark. 2006. Natives, the foreign-born, and high school equivalents: New evidence on the returns to the GED. *Journal of Population Economics* (forthcoming).

Murnane, Richard J., John B. Willett, and Kathryn Parker Boudett. 1999. Do male dropouts benefit from obtaining a GED, postsecondary education, and training? *Evaluation Review* 22, no. 5:475–502.

Shao, Jun, and Randy R. Sitter. 1996. Bootstrap for imputed survey data. *Journal of the American Statistical Association* 91, no. 435:1278–88.