Title

Estimating the Labor Market Returns of the General Educational Development (GED)

Abstract

This paper is the first study to use the 2002 Educational Longitudinal Study (ELS:2002) dataset to investigate the economic returns to earning a General Educational Diploma (GED). The ELS:2002 data included a sample of about 15,000 10th grade students in 2002, and information on earnings and hours worked in 2011 are available. After controlling for pre-existing ability, dropouts and GED holders were indistinguishable in terms of subsequent income and hours worked in 2011. Traditional high school diplomas were associated with an increase in earnings of about 25%. Additionally, an interaction between GED and test scores was modeled to allow for the possibility of differential impacts of the GED based on cognitive ability; no significant interaction was found.

Purpose and Background

The General Educational Development (GED) has grown rapidly in recent years, to the point where approximately 12 percent of all high school diplomas are now issued via the GED test (Heckman, Humphries, and Kautz, 2014). The GED consists of a series of examinations in math, reading, and writing that collectively signal an academic skillset that is on par with the average high school graduate, and is often billed as a high school equivalency exam in both the job market and post-secondary education. As the GED Testing Service states, the GED gives dropouts "a fighting chance at 4 million jobs" (GED Testing Services, 2014, p. 1).

This paper evaluates this claim of equivalency in the labor market using the 2002 cohort of 10th graders from the Educational Longitudinal Study (ELS) dataset. A hierarchical linear model (HLM) approach is used to account for the clustered random sampling strategy used by ELS:2002. While high school diplomas were associated with a substantial improvement in later life earnings, the GED did not lead to an improvement in economic circumstances for its recipients, both in terms of income and hours worked.

Theoretical Framework

The GED could improve labor market outcomes through either human capital accumulation or market signaling (Becker, 1964; Spence, 1973). In a landmark study, Cameron and Heckman (1993) show that once pre-existing ability and background are controlled for, male GED certifiers appear much closer to uncredentialled dropouts than high school graduates in terms of key labor market outcomes. Any differences in wages between GED holders and high school graduates were almost entirely accounted for by their years of schooling.

Failing to find any direct effects of the GED, Cameron and Heckman (1993) also weigh the possibility of indirect effects that occur through granting access to post-secondary education. Here again, students with a GED were not equivalents to high school graduates: completion rates at four-year colleges was only 5% for GED recipients, as compared to 75% for high school graduates. The numbers were not much better for GED students at two-year colleges, where only 2% manage to earn a degree following enrollment. GED recipients that managed to graduate from college did almost as well as high school graduates, leading Cameron and Heckman (1993) to argue that the GED may be an effective route of poverty – on the condition that its recipients take advantage of further educational opportunities.

By employing a longitudinal design, Murnane, Willett, and Boudett (1995) find wage growth rates accelerated by 2.4% per year of actual work experience following GED acquisition – which represents an indirect effect of the GED which was modeled through an interaction term. Even so, the direct effects of GED on income were not significantly different from zero, leading Murnarne et al. (1995) to conclude that the modest increase in wage growth was not enough to help GED holders escape poverty. Later, Murnane, Willett, and Tyler (2000) argue that earlier models in the literature all made the mistake of assuming that the value of the GED does not depend on the skill levels of dropouts. By including an interaction between GED status and cognitive ability, Murnane et al. (2000) find that the main effect of the GED increased substantially and that lower-skilled individuals had more to gain from GED acquisition.

In a later study, Tyler, Murnane, and Willet (2000) argue that the GED is an effective market signal that boosts income, finding evidence that while dropouts with a GED increased their earnings by about 10-19%. Yet, they also found that the GED had no impact on the earnings potential of minorities. While Murnane and colleagues argued that the GED may boost subsequent income by acting as a market signal, Heckman, Hsse, and Rubinstein (2000) suggest the opposite is likely true: GED holders lack drive, discipline, and persistence, making the GED a "mixed" signal in the market.

Graduating high school is more than just a cognitive task – it requires persistence in the face of adversity, regular attendance, and maintaining positive social interactions with peers and faculty at the school. Heckman et al. (2000) argue that successfully graduating measures some latent trait(s) that a standardized test simply cannot: "the same traits that cause GED recipients to drop out of high school cause them to drop out of college as well as the military, marriage, and jobs" (Heckman et al., 2014, p. 177). In sum, although passing the GED may signal higher cognitive ability, the GED tends to "select" students with the weakest character skills from the dropout population. This reality makes the GED an inherently "mixed signal," and Heckman et al. (2000) back this up empirically by finding that GED holders earned less than dropouts in their study's sample. Heckman et al. (2014) later show that GED students are more likely to engage in risky behaviors as compared to high school graduates – being more likely to take drugs, cut class, or be involved in criminal activities as compared to their counterparts. While the GED may signal higher cognitive ability as claimed by Murnane and colleagues, it also signals deficits in various non-cognitive skills. This paper will contribute to the literature by using the ELS:2002 data to consider how GED acquisition alters one's economic future.

Methods and Sample

Data:

This study uses the ELS:2002 data to conduct a secondary analysis investigating the economic returns of the GED. ELS:2002 was collected by the National Center for Education Statistics (NCES) and is a nationally representative sample 10th graders in 2002 (Ingles et al. 2004, 2007). The original datafile included about 15,400 participating students from 750 schools. Data from the baseline year of 2002 and the third follow up (F3) in 2012 are used in this study. The dependent variables, income and hours worked in 2011, were collected nine years after the baseline data was collected. Individuals who had an income of $0 or greater than $200,000 in 2011 were dropped, as were any students that we enrolled in school in 2011. The final sample used in this paper

included 5,925 students in 739 schools. Descriptive statistics for the study variables are listed in Appendix A.

Analysis:
A hierarchical linear model (HLM) was used to account for the cluster random sampling design of the data. The *nlme* package in R was used in conjunction with the *clubSandwhich* package to estimate a two-level hierarchical linear model (HLM) with robust standard errors (O'Dwyer and Parker, 2014; Snijders and Bosker, 2012; R Core Team, 2017; Pinheiro, Bates, DebRoy, Sarkar and R Core Team, 2007; Pustejovsky, 2018). Following past literature, the natural logarithm transformation was used for both dependent variables to correct positive skew. The analytical model is included in Appendix C.

The interclass correlation coefficient (ICC) was estimated from an empty model. Since the ICC was only estimated at 4.17% and 1.21% for 2011 earnings and hours worked respectively, level 2 predictors were not included in the models.

Results

Table 1 presents unadjusted means and standard deviations for the study variables, disaggregated by high school graduation, GED, and dropout status. As can be seen, dropouts and GED holders are nearly indistinguishable from one another in terms of 2011 earnings and hours worked, while high school graduates command a slight advantage. Also note that in terms of cognitive ability, which is captured by the composite test score variable in this study, GED holders tend to fall between high school graduates and permanent dropouts.

Table 2 in Appendix C shows the regression results for the various specifications included in this analysis. Model 1 demonstrates that once cognitive ability is controlled for, dropouts are again statistically indistinguishable from GED holders. While high school graduates are estimated to earn wages that are 25% greater in 2011, there is no significant change in income associated with GED receipt. In fact, the negative coefficient lends support to Heckman et al.'s (2000) labeling of the GED as a "mixed signal." This negative but null finding has been found in several other studies that use dropouts as the reference group (Cameron and Heckman, 1993; Heckman and LaFontaine, 2006; Lofstrom and Tyler, 2004; Heckman et al., 2000). The second model attempts to measure the direct effects of GED acquisition by including two additional control variables, years of postsecondary education and a dummy for having ever received job-training. If the GED boosts later income indirectly through these mechanisms, then including them in the model should decrease the difference between GEDs and dropouts. While the difference does decrease, it is quite trivial (estimated at 1% of earnings), and the coefficient overall remains insignificant. This small change in the differential between GEDs and dropouts is likely because so few GEDs take advantage of further post-secondary education. Finally, model 3 included an interaction term with GED and cognitive ability to consider whether there is evidence for differential impacts of the GED according to the skill level of its recipients. While the main effect makes a large jump as in Murnane et al.'s (2000) study, it remains insignificant (possibly due to low power). The interaction term was not significant, contradicting Murnane et al.'s findings (2000).

Finally, model 4 shows that GED recipients do not work more hours than uncredentialled dropouts after controlling for background variables and cognitive ability. This null finding for labor supply is a replication of past research (Cameron and Heckman, 1993; Cao, Stromsdorfer, Weeks, 1996).

Significance and Conclusions

This paper replicated the conclusions reached by decades of research into the labor market effects of getting a GED: there are none in terms of later income or labor supply. The negative gap between GEDs and dropouts suggests that GED acquisition may even be associated with slightly depressed economic outcomes compared to dropouts, supporting Heckman et al.'s (2000) argument that the GED is a mixed signal in the labor market. These findings do not mean that getting a GED is not worthwhile, but instead that the GED should not be pursued as an end goal by itself. For GED holders who can take advantage of the enhanced opportunity to post-secondary education, the GED may very well be a lifeline.

References

Becker, G. S. (1964). Human capital theory. Columbia, New York, 1964.

Cameron, Stephen, and James Heckman. (1993). "The Nonequivalence of High School Equivalents." *Journal of Labor Economics* 11.1:1-47.

Cao, J., Stromsdorfer, E. W., & Weeks, G. (1996). The human capital effect of general education development certificates on low income women. Journal of Human Resources, 206-228.

GED Testing Services. (2014). "2014 Test for Educators." *GED Testing Services*. American Council on Education. Retrieved from https://www.gedtestingservice.com/educators/2014test

Heckman, James, Jingjing Hsse, and Yone Rubinstein. (2000). "The GED Is a 'Mixed' Signal: The Effect of Cognitive and Non-Cognitive Skills on Human Capital and Labor Market Outcomes." Paper presented at the American Economic Association Meeting in Boston, MA, January 2000.

Heckman, James, John Humphries, and Tim Kautz. (2014). *The Myth of Achievement Tests: The GED and the Role of Character in American Life*. Chicago: University of Chicago.

Ingles, Steven J., Daniel J. Pratt, James E. Rogers, Peter H. Siegel, Ellen S. Stutts, and Jeffrey A. Owings. (2004). Education longitudinal study of 2002: Base year data file user's manual. Washington, D.C.: National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education.

Ingles, Steven J., Daniel J. Pratt, David Wilson, Laura J. Burns, Douglas Currivan, James E. Rogers, and Sherry Hubbard-Bednasz. (2007). Education longitudinal study of 2002: Base-year to second followup data file documentation. Washington, DC: National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education.

Murnane, Richard, John Willett, and Kathryn Boudett. (1995). "Do High School Dropouts Benefit from Obtaining a GED?" *Educational Evaluation and Policy Analysis* 17.2: 133-47.

Murnane, R. J., Willett, J. B., & Tyler, J. H. (2000). Who benefits from obtaining a GED? Evidence from high school and beyond. Review of Economics and Statistics, 82(1), 23-37.

O'Dwyer, L. M., & Parker, C. E. (2014). A primer for analyzing nested data: Multilevel modeling in SPSS using an example from a REL study. Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Northeast & Islands. Retrieved from http://ies.ed.gov/ncee/edlabs/regions/northeast/pdf/REL_2015046.pdf.

Pinheiro, J., Bates, D., DebRoy, S., Sarkar, D., and R Core Team. (2007). nlme: Linear and Nonlinear Mixed Effects Models. R Package version 3.1-137.

Pustejovsky, James. (2018). clubSandwich: Cluster-Robust (Sandwich) Variance Estimators with Small-Sample Corrections. R package version 0.3.1. https://CRAN.R-project.org/package=clubSandwich

R Core Team (2017). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Snijders, Tom and Bosker, Roel. (2012). *An Introduction to Basic and Advanced Multilevel Modeling.* London: Sage Publishers.

Spence, Michael. 1973. "Job Market Signaling." Quarterly Journal of Economics, Vol. 87, No. 3 (August), pp. 355-74

Tyler, John, Richard Murnane, and John Willett. (2000). "Estimating the Labor Market Signaling Value of the GED." *The Quarterly Journal of Economics* 115.2: 431-68.

Appendix A: Descriptive Statistics for Study Sample

Table 1: Descriptive statistics for study variables, disaggregated by HS graduate, GED and permanent dropout status

	Conventional HS graduate	GED recipients	Permanent dropouts	All	
Number of observations	5,461	223	464	5,925 students, 739 schools	ELS:2002 Variable
Income in 2011	32,610 (21,896)	21,963 (16,887)	22,281 (17,668)	31,801 (21,772)	F3ERN2011, 2011 employment income
Log income 2011	10.11 (0.92)	9.61 (1.10)	9.60 (1.13)	10.07 (0.95)	Log of F3ERN2011
Hours worked in 2011	1,916 (754)	1,730 (848)	1,745 (802)	1,903 (760)	Product of F3C01 and F3C02, weeks worked and weekly hours worked in 2011
Log of hours worked in 2011	7.41 (0.69)	7.23 (0.91)	7.23 (0.82)	7.40 (0.71)	Log of product of F3C01 and F3C02
Female	0.52 (0.50)	0.37 (0.48)	0.39 (0.49)	0.51 (0.50)	BYSEX = 2, coded 0 or 1
Black or African American	0.08 (0.28)	0.14 (0.35)	0.16 (0.37)	0.09 (0.29)	BYRACE = 3, coded 0 or 1
Hispanic	0.11 (0.31)	0.19 (0.40)	0.25 (0.43)	0.12 (0.33)	BYRACE = 4 or 5, coded 0 or 1
Asian	0.07 (0.26)	0.02 (0.13)	0.05 (0.20)	0.07 (0.26)	BYRACE = 2, coded 0 or 1
White	0.69 (0.46)	0.55 (0.50)	0.47 (0.50)	0.67 (0.47)	BYRACE = 7, coded 0 or 1
Other	0.05 (0.21)	0.09 (0.29)	0.08 (0.28)	0.05 (0.22)	BYRACE = 1 or 6, coded 0 or 1
Socioeconomic status	0.21 (0.71)	-0.30 (0.66)	-0.41 (0.65)	0.17 (0.73)	BYSES1, socio-economic status composite
Composite test score	53.77 (9.09)	46.55 (7.81)	44.48 (8.70)	53.04 (9.40)	BYTXCSTD, Standardized test composite score-math/reading
Absence of biological parent	0.25 (0.43)	0.47 (0.50)	0.46 (0.50)	0.26 (0.44)	BYP61 = 1, coded 0 or 1
Years of post-secondary education	2.62 (2.21)	0.40 (1.03)	0.36 (0.97)	2.44 (2.23)	F3ATTAINMENT = 4, 5, 6, 7, 8, 10, coded with discrete values 0-8
On-the-job-training	0.47 (0.50)	0.38 (0.49)	0.33 (0.47)	0.46 (0.50)	F3B35 = 1, coded 0 or 1
GED*Composite test score	0.13 (2.45)	46.55 (7.81)	20.88 (23.77)	1.75 (9.00)	Product of F3EVRGED and BYTXCSTD
Maternal education	--	--	--	--	BYMOTHED, factor with 8 levels
Region of the US	--	--	--	--	F3REGION, factor with 4 levels

Note: Standard deviations are in parenthesis.

Appendix B: Analytical Model

Model 1 can be expressed as:

Level 1: $\ln(Income_{ij}) = \beta_{0j} + \beta_{10}GED_{ij} + \beta_{20}High\ school\ grad_{ij} + \beta_{30}Female_{ij} + \beta_{40}Black_{ij} + \beta_{50}Hispanic_{ij} + \beta_{60}Asian, Hawaiian, Pac.\ Isalnder_{ij} + \beta_{70}Other_{ij} + \beta_{80}Socioeconomic\ status_{ij} + \beta_{90}Abscence\ of\ biological\ parent_{ij} + \beta_{100}Composite\ test\ score_{ij} + \varepsilon_{ij}$

Level 2: $\beta_{0j} = \gamma_{00} + u_{0j}$

$\beta_{p0} = \gamma_{p0}$

Where:

Y_{ij} = Income in 2011 for student i in school j.

X_{pij} = value of covariate p for student i in school j.

γ_{00} = the grand mean of 2011 income for schools in the U.S.

γ_{p0} = the effect of a 1-unit increase in the p^{th} covariate on 2011 income.

ϵ_{ij} = the error for student i in school j.

u_{0j} = the random effect for school j.

Appendix C: Regressions for predicting 2011 earnings and hours worked

Table 2: Regression estimates for log 2011 earnings and log hours worked[1]

Variables	2011 Earnings			2011 Hours Worked
	Model 1	Model 2	Model 3	Model 4
GED	-0.12	-0.11	0.32	-0.12
	(0.08)	(0.08)	(0.38)	(0.06)
High school graduate	0.25***	0.18**	0.25***	0.05
	(0.06)	(0.06)	(0.06)	(0.05)
Female	-0.23***	-0.26***	-0.23***	-0.14***
	(0.02)	(0.02)	(0.02)	(0.02)
Black or African American	-0.18***	-0.20***	-0.18***	-0.08*
	(0.04)	(0.04)	(0.04)	(0.03)
Hispanic	-0.06	-0.07	-0.06	-0.02
	(0.04)	(0.04)	(0.04)	(0.03)
Asian	0.04	0.01	0.04	-0.07
	(0.05)	(0.05)	(0.05)	(0.04)
Other	-0.10	-0.10	-0.10	-0.07
	(0.06)	(0.06)	(0.06)	(0.04)
Socioeconomic status	0.09***	0.06*	0.09***	0.01
	(0.03)	(0.03)	(0.03)	(0.02)
Absence of biological parent	-0.09*	-0.06*	-0.09**	-0.04
	(0.03)	(0.03)	(0.03)	(0.02)
Composite test score	0.15***	0.11***	0.15***	0.04***
	(0.01)	(0.02)	(0.02)	(0.01)
On-the-job training		0.24***		
		(0.02)		
Years of post-secondary education		0.05***		
		(0.01)		
GED*Test score			-0.09	
			(0.08)	
Intercept	9.21***	9.30***	9.20***	7.23***
	(0.11)	(0.11)	(0.11)	(0.08)
Deviance	15,778	15,612	15,780	12,660
% of Variance Explained at				
School-Level	.656	.762	.659	.453
Student-Level	.053	.078	.053	.015

Note * = $p < .05$; ** = $p < .01$, *** = $p < .001$, standard errors in parenthesis.

[1] Note that these models also control for maternal education and regions of the U.S.