Supplement to "Stepping stone and option value in a model of postsecondary education"

(*Quantitative Economics*, Vol. 6, No. 1, March 2015, 223–256)

Nicholas Trachter
Federal Reserve Bank of Richmond

S1. Accounting for low enrollment and graduation rates at four-year colleges when returns to graduation are high

Several studies claimed that there is underinvestment in education. Judd (2000) combines Capital Asset Pricing Model (CAPM) techniques with the indivisibility of human capital to compare the return to four-year college graduation with assets of similar risk to find an excess of return to the college investment option. Heckman, Lochner, and Todd (2008) evaluate the internal rate of return of the four-year college investment option relative to work to find that since 1960, internal rates of return have been around 10 percent or higher, depending on the cohort and different specifications of labor markets and taxes. Cunha, Heckman, and Navarro (2005), using data from NLSY/1979, extend the analysis to evaluate the internal rate of return for the marginal student, the agent with the lowest observable measures of ability who enrolls in four-year college, to find an unexplained wedge in returns.[1] Cunha, Heckman, and Navarro (2005) concluded that this wedge is explained by nonpecuniary costs of education, namely, tastes for school, risk aversion, and other.

The evidence presented in this paper points in a different direction: the wedge in returns is explained by the existence of academic two-year colleges (in fact, it is actually more general because vocational schools explain the wedge in returns for students who enroll in academic two-year colleges), as high school graduates sort across the different enrollment alternatives. In particular, high school graduates who enroll in four-year colleges have measures of observables that lie above those who enroll in academic two-year colleges. It follows that the high school graduate with the lowest measures of observables who enrolls in a four-year college is indifferent to enrollment at academic two-year colleges as opposed to being indifferent to joining the labor force. Using the parameterized version of the model, it is possible to quantify the return for this marginal student and how much of average returns for the population that enrolls in four-year colleges is explained by the return of the marginal student. The return for the student

Nicholas Trachter: nicholas.trachter@rich.frb.org
The views expressed in this article are mine and do not necessarily represent the views of the Federal Reserve Bank of Richmond or the Federal Reserve System.

[1] See also the *Handbook of Economics of Education* (Heckman, Lochner, and Todd (2006)).

DOI: 10.3982/QE326

at the threshold between academic two-year colleges and four-year colleges is 2.56 percent, which accounts for 63 percent of the measured average return to four-year college enrollment. To obtain a monetary value for the return for the marginal student, recall that $\Sigma_i(p)$ is the compensating variation of enrollment in institution i relative to joining the workforce and that it is measured in units of h^w, the wage for agents with no degree. It follows that the monetary value in 1984 dollars is simply $\Sigma_i(p) \times 16{,}454.41$. Then the monetary value of enrollment for the marginal student is \$96,600.

S2. Reasons to use NLS-72

The choice of NLS-72 over other data sets is not arbitrary. *High School and Beyond* (HS&B) follows a cohort from 1982 to 1990. *National Education Longitudinal Study of 1988* (NELS:88) follows a cohort from 1992 to 2000. Relative to NLS-72 and HS&B, NLS-72 presents longer horizon wage information (13 years vs. 8 years after high school graduation in the newer data sets). Also, the design of the questionnaire of NLS-72 includes questions regarding the type of two-year college education at any point in time (broadly speaking, two-year colleges are a combination of academic two-year colleges and vocational school). These questions were not available in the newer data sets. Last, NLS-72 has a more detailed analysis of the cost structure of postsecondary education. An alternative would be to use *National Longitudinal Survey of the Youth* (NLSY), which presents better life-cycle earnings information, but requires extensive data mining (in particular, there is no straightforward way to disentangle vocational school from academic two-year colleges). Furthermore, many community colleges have extended their scope to offer both types of programs, making it increasingly difficult to distinguish one from the other.

S3. Construction of data set from NLS-72

The *National Longitudinal Study of the High School Class of 1972* (NLS-72) is a panel that follows the educational histories of high school seniors in 1972. Participants in the study were selected when they were seniors in high school in the spring of 1972 and in a supplementary sample drawn in 1973. The records include the "base year" survey and followup surveys in 1973, 1974, 1976, 1979, and 1986.

The analysis in this paper is specialized to high school graduates; high school dropouts are not allowed to enroll in four-year colleges, so these agents were dropped from the sample. Also, this paper contains no theory as to why agents have discontinuous spells of education. Why do these people work first? Or why do some of them drop out, work for a while, and then enroll again? Tastes, learning, and credit constraints can be touted as possible explanations. Furthermore, there is no clear pattern in terms of observable measures of ability that explain discontinuities. As such, agents with discontinuous spells of education were excluded. Also, data limitations do not allow for an analysis of how graduate school fits into the model logi, so agents who pursued graduate education were also excluded.

This paper relies on observable measures of ability to link priors to data. The variables used here are race, gender, socioeconomic status of family, maximum educational

level of father, rank in high school class, and location of high school. Agents with missing values for any of these variables were discarded.

The cost of education present many missing values. Instead of excluding these observations, a cost regression was performed to predict these values. The regression runs cost of education on observable measures of ability and geographic location of students, and then the estimates are used to compute the expected cost of education for agents presenting missing values.

References

Cunha, F., J. Heckman, and S. Navarro (2005), "Separating uncertainty from heterogeneity in life cycle earnings." *Oxford Economic Papers*, 57 (2), 191–261. [1]

Heckman, J. J., L. J. Lochner, and P. Todd (2006), "Earnings functions, rates of return, and treatment effects: The mincer equation and beyond." In *Handbook of the Economics of Education*, Elsevier Science, Amsterdam. [1]

Heckman, J. J., L. J. Lochner, and P. E. Todd (2008), "Earnings functions and rates of return." *Journal of Human Capital*, 2 (1), 1–31. [1]

Judd, K. L. (2000), "Is education as good as gold? A portfolio analysis of human capital investment." Unpublished. [1]

Submitted November, 2012. Final version accepted March, 2014.