Appendix for “Evaluating Marginal Policy Changes and the Average Effect of Treatment for Individuals at the Margin”

Pedro Carneiro  
University College London  
IFS  
CeMMAP

James J. Heckman  
The University of Chicago  
Geary Institute,  
University College Dublin  
Cowles Foundation,  
Yale University

Edward Vytlacil  
Yale University  
American Bar Foundation

This Draft, July 10, 2009
A Description of the NLSY Data

We restrict the NLSY sample to white males (excluding the oversample of poor whites and military). We define high school graduates as individuals having a high school degree or below. We define participation in college as having attended some college or having completed more than 12 grades in school. The wage variable that is used is an average of deflated (to 1983) non-missing hourly wages reported in 1989, 1990, 1991, 1992 and 1993. We delete all wage observations that are below 1 or above 100. Experience is actual work experience in weeks (we divide it by 52 to express it as a fraction of a year) accumulated from 1979 to 1991 (annual weeks worked are imputed to be zero if they are missing in any given year). The remaining variables that we include in the $X$ and $Z$ vectors are mother’s years of schooling, schooling corrected AFQT, the presence of a four-year college in the SMSA of residence at age 14 (from Kling, 2001), average tuition in public four year colleges in the county of residence at age 17 (deflated to 1993), local average earnings in the SMSA of residence at 17 and local unemployment rate in the state of residence at age 17, local average earnings in the SMSA of residence in 1991 and local unemployment rate in the state of residence in 1991, and average (between 1979 and 2000) local earnings in the SMSA where the individual resided at age 17 and average (between 1979 and 2000) local unemployment rate in state where the individual resided at age 17. SMSA earnings correspond to the average wage per job in the SMSA constructed using data from the Bureau of Economic Analysis, deflated to 2000. The state unemployment rate data come from the BLS website. However, from the BLS website it is not possible to get state unemployment data for all states for all the 1970s. Data are available for all states from 1976 on, and for 29 states for 1973, 1974 and 1975. Therefore for some of the individuals we have to assign them the unemployment rate in the state of residence in 1976 (which will correspond to age 19 for those born in 1957 and age 18 for those born in 1958). Annual records on tuition, enrollment, and location of all public four year colleges in the United States were constructed from the Department of Education’s annual Higher Education General Information Survey and Integrated Postsecondary Education Data System “Institutional Characteristics” surveys. By matching location with county

2The distance variable we use is the one used in Kling (2001), available at the Journal of Business and Economics Statistics website.
of residence, we determined the presence of four-year colleges. Tuition measures are taken as enrollment weighted averages of all public four-year colleges in a person’s county of residence (if available) or at the state level if no college is available. SMSA and state of residence at 17 are not available for everyone in the NLSY, but only for the cohorts born in 1962, 1963 and 1964 (age 17 in 1979, 1980 and 1981). However, SMSA and state of residence at age 14 is available for most respondents. Therefore, we impute location at 17 to be equal to location at 14 for cohorts born between 1957 and 1962 unless location at 14 is missing, in which case we use location in 1979 for the imputation. Many individuals report having obtained a bachelors degree or more and, at the same time, having attended only 15 years of schooling (or less). We recode years of schooling for these individuals to be 16. This variable is only used to annualize the returns to schooling (divide returns to college by 4, which is the average difference in years of schooling between individuals in each schooling group. To remove the effect of schooling on AFQT we implement the procedure of Hansen, Heckman, and Mullen (2004).

References

