

Online Appendix: Priceless: The Nonpecuniary Benefits of Schooling

For Figures 1-5

The figures were created by first constructing a subsample of the combined 1972–2000 General Social Surveys. We restricted our data to nonimmigrants who were 14 years old since 1970 and at least 25 years old in the survey year. Dropping those born before 1956 cuts the sample substantially (from 34,173 to 6,811) and ends up limiting the upper age range to 45, but this allows us to focus the analysis on more recent cohorts. We did not find this to affect the overall results (if anything the results for schooling relationships were stronger for older cohorts). For some graphs, where the outcome variable was collected only over a subset of years, the sample is smaller.

The education variable is grouped by highest years of completed schooling: 0–11, 12, 13–15, and 16 or more, which we refer to generally as high school dropouts, high school graduates without college, some college, and college graduates, respectively. To graph the black bars in the figures, the outcome variable of interest is regressed on fixed effects for state of birth, age, year of birth, and survey year fixed effects (multiple survey years allow us to estimate these relative fixed effects simultaneously), plus dummy variables for gender and race, as well as father's occupational prestige score (interacted with a missing value indicator), father and mother's education attainment (and indicators if these values were missing), indicators for household composition at age 16 (living with both parents, mother only, etc.), whether mother worked (or if this value was missing), and whether family's income when 16 years old was far below average, below average, average, above average, far above average, or missing, as well as the education categories. The indicator for having 12 years of completed schooling was omitted. The remaining relative education coefficient estimates (the estimated outcome differences by school attainment relative to high school graduates) were added to the overall high

school graduate sample mean and presented in the graphs. The white bars show the same coefficients, but after adding fixed effects for family income bracket for each survey year. In earlier years, the number of bracket options were 13, but expanded to 24, with options ranging from “under \$1,000” to “\$110,000 or more,” after 1997.

For Table 1

The data used in Table 1 is derived from administrative registers and prepared for research by Statistics Norway. The cohorts include all 28–60 year-olds in 2005. The data set is a longitudinal population data set consisting of all people living in Norway born from the 1920s. Different registers for own family, parental family, educational outcomes, marriage, divorce, and other outcomes are merged using the same personal identifier. Persons are also matched to plants/firms as well as other labor market outcomes including unemployment, on welfare, retirement, and receiving disability pension. The month and year of birth is given in the data set and hence twins can be identified. Schooling is measured as the normalized number of completed years. An individual is classified as employed if he has a plant identification number at that time, and as unemployed if he does not and is registered with some months of unemployment during the year. Earnings or annual income is measured as annual income that provides pension points in the national security system. The included components are regular labor income, income as self-employed, and benefits received while on sick leave, being unemployed, and on parental leave. “Married” is registered as married in 2005. “Divorced” is registered as divorced in 2005. Teen fertility is whether the women had a child as a teenager. “Welfare” means the person received welfare in 2005. “Assortative mating” is measured in years of education for spouse in 2005. “Disability pension” identifies those disabled at the age of 58 or older from 1993–2005.

For Table 2

Table 2 uses an extract of native-born individuals aged 25 to 64 from the combined six decennial census microdata samples between 1950 and 2000 and the seven American Community Surveys between 2001 and 2007.¹ Hawaiian- and Alaskan-born respondents were excluded. We coded the schooling variable for individuals in the 1950 to 1980 data as highest grade completed. Average years of schooling were assigned to categorical values in the 1990 and 2000 Censuses and the American Community Surveys using an imputation method discussed in

¹ The specific Census files used were the 1950 General 1/330 sample (limited to those with long-form responses), the 1960 General 1 percent sample, the 1970 Form 2 State 1 percent sample, the 1980 Metro 1 percent sample, the 1990 1 percent unweighted sample, and the 2000 1 percent unweighted random sample. The data was downloaded from the IPUMS web site at (<http://usa.ipums.org/usa/index.shtml>).

Oreopoulos (2006b). The earnings variable—log weekly wage—was calculated by dividing annual wage and salary income by weeks worked, then taking logs. We used an indicator for welfare receipt from 1970 data and onwards. The 1950 to 1980 Censuses provided information on whether a respondent resided in a correctional or mental institution, which we used to create our outcome variables “in jail” and “in mental institution” respectively. Whether a mother gave birth as a teenager is calculated by determining whether the difference between a mother’s age and the age of her “eldest own child in household” is 19 years or less, for children 30 years old or less. We also used an occupational prestige score (the variable “PRENT” from the IPUMS web site). Among children in the household aged 8 to 16, an indicator for “ever repeated a grade” was calculated as whether a child was below the median grade level of other children the same age and quarter of birth in the same state and census year. The indicator was matched to the child’s mother and/or father. The voting indicator variable comes from the 1978 to 2000 November Current Population Surveys, the same data used by Milligan, Moretti, and Oreopoulos (2004). We restrict the sample to those of voting age (18 years old and older).

We restricted our data to individuals aged 16 between 1915 and 2000, aged 25 to 64 in the survey year. We also removed individuals with more than 12 years of completed schooling since these college educated were not substantially influenced by the minimum school leaving age (Oreopoulos, 2007). The results are generally the same, but less precise if we keep the larger sample. Similar to Oreopoulos (2006b), we measured each school-leaving age as the minimum between a state’s legislated dropout age and the minimum age required to obtain a working permit. We also record the dropout age under exceptions—for example if students could leave with permission from parents or if working full time. We grouped the small percent of the sample that faced school-leaving ages lower than 14 into one category (school-leaving age < 14). All others faced dropout ages of 14, 15, 16, 17, or 18. The dropout age is matched to individuals according to the year in which they were 16 years old and their state of birth (which is likely their state of residence when 16 as well).

The results reported in Table 2 are from regressing the outcome variables on fixed effects for birth cohort, state of birth, data year, gender, and race (for the full sample), a quartic polynomial for age, and years of completed schooling, with schooling instrumented by the dropout age under no exceptions and the dropout age under exceptions. The sample is weighted by the person weight variable from the IPUMS. Standard errors are clustered by state and year of birth.