

Appendix I. Data

This study exploits data from several sources. Firstly, I use the Current Population Survey (CPS) March Supplements 1968-2001 for wages and educational attainment of a representative sample of the U.S. population. Although the data set is available beginning in 1964, the first four surveys were excluded because questions on earnings have changed since 1968. Sample selection criteria follow those of the Bureau of Labor Statistics (BLS (1993)) as closely as possible. To be included in the sample, individuals should have a job when surveyed and report positive weeks and hours worked for the past year. Individuals working for the government or who are self-employed are excluded. Age restrictions are imposed so that individuals between the ages of 16 and 65 in the year they worked are included in the sample. Relaxing the age restriction appears to have little effect on the results.

All individuals in the sample reported their years of schooling, weeks worked in the past year, usual hours worked (or hours worked in the past week for surveys before 1976), wage or salary income in the past year, and other individual characteristics such as gender, race, marital status, census division of residence, and standard metropolitan statistical area (SMSA) status. Since weeks worked were separated into intervals through 1975, mean weeks worked by gender in each interval in the 1976 survey were substituted for previous surveys. Annual hours are obtained by the product of weeks worked and hours worked per week. Top-coded earnings before 1996 were multiplied by 1.5, following Katz and Murphy (1992). Hourly wages are obtained by dividing annual earnings by annual hours worked. Since earnings and hours in the survey data are for the previous calendar year, hourly wages cover the period from 1967 through 2000 even if the survey years are 1968 through 2001. These wages are deflated using the personal consumption expenditure (PCE) price index—using the consumer price index (CPI) instead changes the results little. The BLS started to code years of schooling in intervals in 1992. I assigned 0, 3, 6, 8, 9, 10, 11, 12, 14, 16, and 17 for the categories of none, 1-4, 5-6, 7-8, 9, 10 years, 11 years of schooling or 12 years with no diploma, 12 years, some college with no degree or an associate's degree, bachelor's degree,

and master's and Ph.D. degrees. Throughout the survey, the highest years of schooling were recoded as 17 years. The CPS March Supplements person weights are used for all descriptive statistics and regressions.

The time series of educational expenditures is taken from the National Center for Education Statistics (1993) and the National Center for Education Statistics (2004). Specifically, yearly total expenditure per pupil in public elementary and secondary schools is exploited for the analysis in this paper. Since the data were collected biennially in the mid-20th century, a cubic spline is used to interpolate the series.

In order to avoid overstating the rise in real school expenditure, school expenditures are deflated by the price index for PCE on education services. Since the deflator is not available before 1929 (the earliest cohort considered in this paper started school in 1908), I use the projection of the price index for PCE on education on the CPI by splicing it to actual data since 1929. For the years before 1913, during which the CPI was not available, the price index in Warren and Pearson (1935) is used.

Appendix II. Proof of Ability Sorting in Schooling Choice

Consider individuals 1 and 2 who were born in the same year with learning ability $\tilde{\gamma}_0$ and $\tilde{\gamma}_0'$, where $\tilde{\gamma}_0' > \tilde{\gamma}_0$, respectively. Suppose that individual 1 chooses \tilde{s}^* as his optimal years of schooling. Assuming an interior solution, \tilde{s}^* satisfies equation (2) for individual 1, that is, the marginal benefit (left-hand side (LHS) of equation (2)) of an additional year in school equals its marginal cost (right-hand side (RHS) of equation (2)) if individual 1 chooses \tilde{s}^* years of schooling at the optimum.

Consider the marginal benefit (LHS of equation (2)) of schooling associated with \tilde{s}^* years for individual 2 and compare it with that for individual 1. Note that both individuals 1 and 2 face the same skill prices and the same educational expenditure (recall that the educational expenditure is chosen by the median ability person.). Let $h(\tilde{s}^*; \tilde{\gamma}_0)$ and $h(\tilde{s}^*; \tilde{\gamma}_0')$ denote human capital stocks accumulated through \tilde{s}^* years of schooling by individuals 1 and 2,

respectively. The first term (for human capital increment) in the LHS of equation (2) of individual 2 relative to that of individual 1 is then written as $\frac{\widetilde{\gamma}_0' h(\widetilde{s}^*; \widetilde{\gamma}_0')^{\gamma_1}}{\widetilde{\gamma}_0 h(\widetilde{s}^*; \widetilde{\gamma}_0)^{\gamma_1}}$. The ratio of the third term in the LHS (for utility gain) can also be written as $\frac{\widetilde{\gamma}_0' h(\widetilde{s}^*; \widetilde{\gamma}_0')^{\gamma_1}}{\widetilde{\gamma}_0 h(\widetilde{s}^*; \widetilde{\gamma}_0)^{\gamma_1}}$, while the second term (for gain from skill premium) of individual 2 relative to that of individual 1 is given by $\frac{h(\widetilde{s}^*; \widetilde{\gamma}_0')}{h(\widetilde{s}^*; \widetilde{\gamma}_0)}$.

If one repeats the same comparison with each component of the marginal cost of schooling (RHS of equation (2)), then he will see that the first and the third terms of individual 2 relative to individual 1 are $\frac{h(\widetilde{s}^*; \widetilde{\gamma}_0')}{h(\widetilde{s}^*; \widetilde{\gamma}_0)}$ and both individuals have the same value for the second term. Using the property that $\frac{1 + \widetilde{\gamma}_0'(1 - \gamma_1) \int_0^{\widetilde{s}^*} d(a)^{\gamma_2} da}{1 + \widetilde{\gamma}_0(1 - \gamma_1) \int_0^{\widetilde{s}^*} d(a)^{\gamma_2} da} < \frac{\widetilde{\gamma}_0'}{\widetilde{\gamma}_0}$, it can be easily shown that $1 < \frac{h(\widetilde{s}^*; \widetilde{\gamma}_0')}{h(\widetilde{s}^*; \widetilde{\gamma}_0)} < \frac{\widetilde{\gamma}_0' h(\widetilde{s}^*; \widetilde{\gamma}_0')^{\gamma_1}}{\widetilde{\gamma}_0 h(\widetilde{s}^*; \widetilde{\gamma}_0)^{\gamma_1}}$. This implies that for individual 2, the marginal benefit associated with \widetilde{s}^* years of schooling is greater than its marginal cost. Thus, at the optimum, individual 2 will choose to stay in school longer than \widetilde{s}^* years.

Appendix III. Estimation

Obtaining Time Series of Skill Prices

Log wage y_{it} of an individual i at time t is observed with a measurement error:

$$y_{it} = \ln w_t(s_i) + \ln h(s_i) + \phi(x_{it}) + \varepsilon_{it}$$

where ε_{it} is a measurement error exogenous to any known individual characteristics at time t and x_{it} is potential experience at time t . Changes in the mean log wages of each cohort with s years of schooling between periods t and $t + 1$ include changes in the skill price for education level s , return to accumulated experience, and a difference in mean errors (because the mean human capital stocks accumulated from schooling cancel out) as follows:

$$\bar{y}_{t+1} - \bar{y}_t = \ln w_{t+1}(s) - \ln w_t(s) + \phi(x_t + 1) - \phi(x_t) + \bar{\varepsilon}_{t+1} - \bar{\varepsilon}_t$$

Here variables with an upper bar are mean values for the same cohort with s years of schooling. Assuming that the error term in the observed wages of an individual i at time

t is *i.i.d.*, a difference in mean errors approaches 0 for a large sample. Given parameters governing the experience function ϕ , taking the average of cohort mean log wage changes for education level s over all cohorts determines the changes in the corresponding skill price as residuals. Note that this procedure identifies growth rates of the skill prices, not their initial levels. Log skill price for each education level in the earliest year of the sample are obtained as residuals from the mean log hourly wages of different skill groups in the 1968 CPS survey after taking out the mean log human capital stocks for a given set of parameters. This procedure is repeated for four education groups: i) high school dropouts, ii) high school graduates, iii) workers with some college, and iv) college graduates. For high school dropouts and workers with some college, the linear projections of their average wages on the wages of high school and college graduates are used to extract the extent to which their wages move with the wages of high school and college graduates. The resulting wages are used for model-based Mincer regressions. Solving the model requires longer time series of skill prices than data permits. Thus, I assume that log skill prices grow linearly over the entire period with their growth rates determined by the average growth rates in the CPS data during the period of 1967 to 2000. The trend wages of the four education groups are allocated to workers with 0, 12, 14, and 17 years of schooling and wages for in-between education levels are linearly interpolated.

Computing Cohort-Level Educational Spending

Although individual data on educational spending at different grade levels are required to construct model moments, school expenditures per pupil are available only as an aggregate time series. As an alternative, I assume that educational expenditures in the first grade grow at a constant rate across cohorts. Note that the first-order condition for the quality margin of schooling (equation (1)) implies that school spending grows at the rate of $\frac{g_p - r}{\gamma_2 - 1}$, where g_p is the continuous growth rate of the relative price of educational goods, as one proceeds to a higher grade regardless of cohort. Thus, generating cohort-level school expenditures requires two parameters: first grade spending of the earliest cohort and the growth rate of the first grade spending across cohorts. These two parameters can be computed based on

the initial level and the growth rate of the time series of school expenditures.

Running a Mincer Regression with the Mean Human Capital Stocks

Running a Mincer regression based on the mean log human capital stocks yields the same coefficients as what I would obtain with individual human capital stocks. The estimated Mincer return to schooling based on individual human capital stocks is $\frac{Cov(\ln h(\gamma_0^i, T, s_i) + \ln w(s_i), s_i)}{Var(s_i)}$, while the Mincer coefficient from a regression with the mean human capital stocks for all cells is $\frac{Cov(E(\ln h(\gamma_0^i, T, s_i) + \ln w(s_i) | s_i, T), s_i)}{Var(s_i)}$. Since there is no variation in years of schooling within each schooling-potential experience cell, it can be easily shown that $Cov(\ln h(\gamma_0^i, T, s_i) + \ln w(s_i), s_i) = Cov(E(\ln h(\gamma_0^i, T, s_i) + \ln w(s_i) | s_i, T), s_i)$.

References

- BLS, 1993. Labor Composition and U.S. Productivity Growth, 1948-90. BLS Bulletin 2426.
- Katz, L. F., Murphy, K. M., Feb. 1992. Changes in relative wages, 1963-1987: Supply and demand factors. *Quarterly Journal of Economics* 107 (1), 35–78.
- National Center for Education Statistics, 1993. 120 Years of American Education: A Statistical Portrait. U.S. Department of Education, Washington D.C.
- National Center for Education Statistics, 2004. Digest of Education Statistics. U.S. Department of Education, Washington D.C.
- Warren, G. F., Pearson, F. A., 1935. *Gold and Prices*. John Wiley and Sons, New York.