# 1 DATA

The data are from the 1979 youth cohort of the National Longitudinal Survey of Labor Market Experience (NLSY79). The NLSY79 consists of 12,686 individuals, approximately half of them men, who were 14-21 years old as of January 1, 1979. The sample consists of a core random sample and an oversample of blacks, Hispanics, poor whites, and the military. One unique characteristic of the NLSY79 is that from 1985, it has comprehensive asset information for each respondent. In any intertemporal labor supply model, the shadow price of assets, or marginal utility of wealth, plays an important role as linking period-by-period decisions intertemporally. In the past, the Panel Study of Income Dynamics (PSID) was frequently used to estimate such models, and researchers either first differenced away the shadow price of assets, as in MaCurdy (1981), or used the marginal utility of food consumption as a proxy for the shadow price of assets, as in Altonji (1986) or Shaw (1989). It was necessary for researchers analyzing the PSID to use food consumption data because that is the only consumption data it contains. Here we use the asset data directly to either measure the shadow price of assets, or, using the intertemporal budget constraint, back out total consumption.

We use the white male sample of the NLSY79 data. We only use males who are at least 20 years old and have completed schooling. In our analysis, we treat schooling as exogenous. Since people can either accumulate human capital by on the job experience or schooling, omission of the schooling decision can be an important source of bias. By only using data beginning from the year after the respondent last attended school, we hope to minimize the potential bias.[9] Also, we censor anybody who served in the military from the sample. Appendix A.3 describes in more detail how we constructed the data.

Since the NLSY79 only has asset data beginning in 1985, and the asset data in

---

[9]The failure to treat school attendance as a choice variable potentially creates two types of biases. Suppose that once people leave school they rarely return. If people decide when to leave school based on the wage draws they receive (i.e., the shock to human capital $\epsilon_{1,t+1}$ in Equation (??)), then people will tend to have relatively high levels of wages (human capital) in the first period after leaving school. If the wage process exhibits any subsequent mean reversion, this may lead to understatement of the gradient of wages with respect to experience in the early postschool years. This would, in turn, cause us to underestimate the return to human capital investment in the early postschool years, which would cause us to underestimate the i.e.s. Suppose on the other hand, that people often return to school in periods when they receive very low wage draws. Failure to account for this is analogous to ignoring corner solutions in labor supply, which we have argued would be likely to bias estimates of the i.e.s. toward zero. In any case, omitting the schooling choice, as well as omitting the choice of working zero hours, may result in bias in our estimated parameters. We left those choices out of the model because (1) it imposes even more computational burden in the estimation routine, and (2) most of the labor supply literature, such as MaCurdy (1982), Altonji (1986), and Shaw (1989) focus exclusively on workers with positive hours and omit schooling choices from their models as well. But estimating intertemporal labor supply models with corner solutions and schooling choices would be a promising future line of research. Keane and Wolpin (2001) estimate such a model, but they discretize the hours choices.

Table 1:
MEAN AGE PROFILES

| Age | Hourly Wage | Hours | Total Wealth |
| --- | --- | --- | --- |
| 20 | 5.785 (2763) | 1531.7 (2837) | 6334.8(207) |
| 21 | 5.998 (3220) | 1578.6 (3348) | 6178.5(568) |
| 22 | 6.697 (3396) | 1725.7 (3536) | 8933(956) |
| 23 | 7.12 (3420) | 1866.7 (3557) | 9163(1361) |
| 24 | 7.342 (3308) | 1989.4 (3480) | 2889.9(1751) |
| 25 | 7.99 (3187) | 2042 (3360) | -9276.6(1885) |
| 26 | 9.294 (2987) | 2101.6 (3160) | 1205.6(2189) |
| 27 | 9.098 (2871) | 2134.7 (3048) | 4651.3(2504) |
| 28 | 10.1 (2785) | 2182.1 (2976) | -29888(2538) |
| 29 | 9.426 (2499) | 2200.7 (2670) | -37101(2235) |
| 30 | 12.36 (2028) | 2224.3 (2188) | -25726(1903) |
| 31 | 15.24 (1649) | 2243.9 (1785) | -12168(1614) |
| 32 | 13.6 (1206) | 2238.1 (1307) | 19201(1313) |
| 33 | 22.98 (851) | 2253.1 (922) | 27379(1098) |
| 34 | 11.39 (554) | 2246.7 (608) | 55264(649) |
| 35 | 11.57 (291) | 2294.2 (325) | 84001(299) |
| 36 | 10.01 (65) | 2283.5 (71) | 58172(67) |

NOTE: Sample sizes are in parentheses

1991 is missing, we recover the missing assets using the intertemporal budget constraint as discussed in the previous section. Table 1 gives the sample means of wages, hours of labor supply, and total wealth of individuals. Also, Table 2 gives the quantiles of the wage and labor supply distribution. Notice that the sample mean of the wages far exceeds the median. This indicates that there are some very high wage values. In order to remove the effect of outliers, we removed the top and bottom 2.5% of the wage and hours distributions. Also, following Keane and Wolpin (2001), who also used the NLSY79 data, we only used assets that satisfy the following fomula:

$$-2500 \times (t - 10) \leq A(t) \leq 10,000 \times (t - 15)$$

where $t$ is the age of the individual. This was necessary because there were some assets whose values were either extremely high or low. After censoring the data, the sample means are closer to the medians (see Table 3). In the estimation, we treat outliers as missing values.

In the data, the percentage of the individuals with zero hours supplied is 17% at age 20, falls below 10% at age 25, and is 9.6% overall. Hence, for the estimation of this article, we adopt the conventional approach of the prime age male labor supply literature and assume only interior solutions. We start

2

Table 2:
QUANTILE AGE PROFILES

| | Wage Quantiles | | | Hours Quantiles | | |
|---|---|---|---|---|---|---|
| Age | 25% | 50% | 75% | 25% | 50% | 75% |
| 20 | 3.26 | 4.752 | 6.641 | 800 | 1640 | 2080 |
| 21 | 3.509 | 5.138 | 7.188 | 892 | 1694 | 2095 |
| 22 | 3.912 | 5.645 | 7.98 | 1127 | 1948 | 2185 |
| 23 | 4.14 | 6.096 | 8.623 | 1402 | 2080 | 2294 |
| 24 | 4.527 | 6.533 | 9.203 | 1640 | 2080 | 2357 |
| 25 | 4.739 | 7 | 9.833 | 1760 | 2080 | 2395 |
| 26 | 5.079 | 7.404 | 10.38 | 1880 | 2080 | 2438 |
| 27 | 5.389 | 7.754 | 10.84 | 1925 | 2080 | 2486 |
| 28 | 5.633 | 8.142 | 11.35 | 2000 | 2120 | 2531 |
| 29 | 5.673 | 8.318 | 11.65 | 2016 | 2145 | 2560 |
| 30 | 5.851 | 8.611 | 12.02 | 2060 | 2162 | 2580 |
| 31 | 5.989 | 8.76 | 12.43 | 2051 | 2170 | 2600 |
| 32 | 6.077 | 8.983 | 12.98 | 2068 | 2134 | 2600 |
| 33 | 6.166 | 8.917 | 12.66 | 2080 | 2165 | 2600 |
| 34 | 6.737 | 9.378 | 13.24 | 2080 | 2161 | 2600 |
| 35 | 6.581 | 9.316 | 13.72 | 2080 | 2200 | 2619 |
| 36 | 6.388 | 10.18 | 12.37 | 2080 | 2151 | 2667 |

estimating the model on individuals after they finish schooling and when they start working positive hours, and we exclude individuals who have a year of zero hours of work after that. The estimation of the intertemporal labor supply model with corner solutions using the dynamic programming ML approach is left for future research. Because of the computational burden of the dynamic programming and estimation routine, we restrict the heterogeneity to be only on the observed educational outcome. That is, we divide individuals into four educational types: type 1 individuals are those who did not graduate from high school (final grade less than 12 years), type 2 are those who graduated from high school (final grade equal to 12 years), type 3 are those who attended some college (final grade less than 16 years), type 4 are those who graduated from college (final grade more than or equal to 16 years).[10]

The total sample of white males who are at least 20, have completed schooling, are not in the military, and whose schooling record is known is 4418 individuals. Among them, 718 individuals are high school dropouts, 1980 individuals are high school graduates, 869 individuals attended some college, and 851 individuals are college graduates. Among them, we only use the individuals who

---

[10]While we do not include unobserved heterogeneity in the model, persistence in wages is captured by the persistent nature of the shocks to human capital.

MEAN AGE PROFILES AFTER CENSORING

| Age | Mean Wage | Sample Size | Mean Hours | Sample Size | Zero Hours | Mean Assets | Sample Size |
|---|---|---|---|---|---|---|---|
| 20 | 5.147 | 2624 | 1567.5 | 2769 | 17.14 | 3478.9 | 202 |
| 21 | 5.582 | 3059 | 1616.3 | 3266 | 14.94 | 4633.2 | 561 |
| 22 | 6.09 | 3227 | 1766.5 | 3451 | 13.63 | 5404.6 | 931 |
| 23 | 6.603 | 3249 | 1910.9 | 3470 | 11.74 | 6676.6 | 1337 |
| 24 | 7.07 | 3143 | 2034.3 | 3398 | 10.24 | 8501.6 | 1706 |
| 25 | 7.5 | 3028 | 2089.9 | 3277 | 8.37 | 10333 | 1831 |
| 26 | 7.965 | 2837 | 2146.1 | 3088 | 6.81 | 11195 | 2128 |
| 27 | 8.34 | 2728 | 2183 | 2973 | 5.95 | 12312 | 2434 |
| 28 | 8.748 | 2646 | 2231.2 | 2903 | 5.97 | 13818 | 2439 |
| 29 | 8.943 | 2374 | 2250.5 | 2605 | 5.42 | 15733 | 2143 |
| 30 | 9.349 | 1927 | 2272.1 | 2135 | 6.09 | 14695 | 1818 |
| 31 | 9.656 | 1566 | 2289.9 | 1744 | 5.1 | 14431 | 1530 |
| 32 | 10.04 | 1145 | 2284.6 | 1276 | 6.24 | 13029 | 1265 |
| 33 | 9.866 | 808 | 2300.3 | 900 | 6.49 | 14730 | 1060 |
| 34 | 10.51 | 527 | 2292.1 | 594 | 7.03 | 24411 | 604 |
| 35 | 10.81 | 276 | 2337 | 318 | 8.19 | 32658 | 252 |
| 36 | 9.768 | 62 | 2283.5 | 71 | 7.79 | 22175 | 59 |

completed schooling before age 25 and whose wage and hours data are available for at least 6 years. The total sample size of those left is 2143 individuals. Then, we remove those people that have zero hours at some point after the starting age. After that, the remaining sample is 1972 individuals.

Finally, we restricted the sample size for the estimation by randomly choosing 1000 people out of 1972 individuals. The total number of person year observations is 7465. The total number of wage observations is 7465, that of the hours observations is 7465, and that of the asset observations is 4323. Notice that people with missing data for several periods are still carried forward (as described in the Appendix).