The Heterogeneity and Dynamics of Individual Labor Supply over the Life Cycle: Facts and Theory *

Andrès Erosa† IMDEA
Luisa Fuster‡ IMDEA
Gueorgui Kambourov§ University of Toronto

VERY PRELIMINARY AND INCOMPLETE

This version: November, 2009

Abstract

In this paper we first document various facts about the labor supply decisions of male workers in the US over their life-cycle and we then build a theory of individual life-cycle labor supply that builds on the neoclassical growth model. The theory features heterogeneous agents, incomplete markets, nonlinear wages, and variation in labor supply along the extensive and intensive margin. While the model economy is calibrated without explicitly targeting the facts on hours worked, the predictions of the theory on labor supply are quantitatively close to the observations documented in the data. We show that in the calibrated model economy there is a disconnect between the theoretical elasticity of labor supply obtained from the FOC of the individuals' problem, the empirical elasticity recovered from simulated micro data using standard econometric techniques, and the macro elasticity obtained by simulating the aggregate labor supply response to a one period unanticipated tax increase on labor earnings. While the empirical elasticity of labor supply is 0.15 (a value well within the range [0,0.5] in the empirical literature), the theoretical and the macro elasticities are higher by a factor of three and six (with values of .46 and .85, respectively). Moreover,

*We thank seminar participants at Alberta, the Chicago Fed, Concordia, the Minneapolis Fed, Pennsylvania, Queen's, Rochester, UAB, UBC, the 2008 Yale-Cowles Summer Conference, 2008 CEA Conference, 2008 SED Conference, III Dynamic General Equilibrium Macroeconomics workshop, 2008 WCU-Seoul Conference on Recent Development in Macroeconomics, 2009 NBER Summer Workshop, and 2009 Minnesota Macroeconomic Workshop, for useful comments and discussions.

†Madrid Institute for Advanced Studies, Calle Isaac Newton 2, Tres Cantos, 28760 Madrid, Spain. E-mail: andres.erosa@imdea.org

‡Madrid Institute for Advanced Studies, Calle Isaac Newton 2, Tres Cantos, 28760 Madrid. E-mail: luisa.fuster@imdea.org

§Department of Economics, University of Toronto, 150 St. George St., Toronto, ON, M5S 3G7 Canada. E-mail: g.kambourov@utoronto.ca.

we simulate the model for different preference parameters and find that the empirical elasticity virtually does not change, a finding that underscores that the empirical elasticity of labor supply is not a good target for identifying preference parameters in our model economy. The labor supply response along the extensive margin explains why the macro elasticity is much bigger than the theoretical elasticity. Time aggregation makes (measured) annual wages a noisy measure of the returns to work faced by individuals during the year and this effect is important in accounting for the low empirical elasticity of labor supply on the annual data.

JEL Classification: D9, E2, E13, E62, J22.

Keywords: Labor Supply, Heterogeneous Agents, Life-Cycle.

VERY PRELIMINARY AND INCOMPLETE

WE EXPECT TO POST MORE COMPLETE VERSIONS OF THIS PAPER AT:

http://www.chass.utoronto.ca/ gkambour/research/labor_supply/labor_supply.htm

1 Introduction

The aggregate response of labor supply to macroeconomic shocks is the subject of a heated debate among economists. The workhorse in macroeconomics analysis is the representative agent neoclassical growth model calibrated to fit aggregate time series data. Macroeconomists typically assume that preferences are such that the income and substitution effects of permanent wage changes cancel out, rendering a constant aggregate labor supply in an economy with sustained technological progress. This research has led to numerous papers in macroeconomics making the case of a large aggregate labor supply elasticity.[1] Empirical studies based on micro level data, however, find a much smaller labor supply elasticity (MaCurdy (1981) and Altonji (1986)). Hence, recently, macroeconomists modeling heterogeneous agents and individual labor supply assume away "balance growth preferences" and technical progress (see, Castaneda, Díaz-Giménez, and Ríos-Rull (2003)). Despite these recent contributions, the main objective of our paper is to show that the neoclassical growth model with "balanced growth preferences," as used in standard business cycles and growth studies, can be reconciled with the micro evidence on labor supply. We show that an heterogeneous agent model with preferences consistent with balance growth can go along way in capturing salient features of labor supply over the life-cycle for male workers. Moreover, using simulated data from the model economy, we find that the empirical elasticity of labor supply is not a good calibration target for pinning down preference parameters in our model economy.

In this paper we first document various facts about the labor supply decisions of male workers in the US over their life-cycle. For cohorts of college and non-college men in the PSID, we study the life-cycle profiles of average hours worked, the fraction of individuals with positive hours worked during the year, and the coefficient of variation of hours. In addition, we analyze the persistence in labor force participation and hours worked over the life-cycle. Then, we develop a theory of individual labor supply that builds on the neoclassical growth model. The theory models life-cycle behavior to better relate the model predictions to the data. We also model heterogeneous agents, incomplete markets, nonlinear wages, and variation in labor supply along the extensive and intensive margin. All of these features, as discussed in our literature review, can create a disconnection between the theoretical and the empirical elasticity of labor supply. Heterogeneity is introduced by assuming that

[1] Kydland and Prescott (1982) and Prescott (1986) find that aggregate labor supply is very responsive to business cycle shocks while Prescott (2004) finds a large response to a change in taxes.

individuals are subject to uninsured labor productivity risk. The period utility function is allows for balance growth and for variation in labor supply along the intensive and extensive margin. To model nonlinear wages, we follow Hornstein and Prescott (1993). To calibrate the model, we estimate an annual wage process on the PSID data and, given that we use a model period of a quarter, we find a *quarterly* stochastic process on labor productivity that makes the predictions of the model economy consistent with the annual wage process estimated in the data. While the model economy is calibrated without explicitly targeting the facts on hours worked, the predictions of the theory on labor supply are quantitatively close to the observations documented on the data.

We show that in our calibrated model that there is a disconnect between three alternative ways of computing the Frisch elasticity of labor supply. First, assuming an interior solution in the labor supply decision and using a linear approximation to the first order conditions, we obtain the "theoretical" elasticity of labor supply which is an explicit function of preference parameters in the utility function. Second, we simulate the model economy and use standard econometric techniques to estimate in the simulated data an "empirical" elasticity of labor supply. Third, we compute a "macro" labor supply elasticity by assessing the aggregate labor supply response to a one period (quarterly) unanticipated tax increase on labor earnings. The results from these experiments could not be more striking: While the empirical elasticity of labor supply is 0.15 (a value well within the range [0,0.5] in the empirical literature), the theoretical and the macro elasticities are higher by a factor of three and six (with values of .46 and .85, respectively). Moreover, we simulate the model for different preference parameters and find that the empirical elasticity virtually does not change, a finding that underscores that the empirical elasticity of labor supply is not a good target for identifying preference parameters in our model economy.

We find that the labor supply response along the extensive margin explains why the macro elasticity is much bigger than the theoretical elasticity, which should be intuitive as the theoretical elasticity was derived assuming an interior solution in the labor supply decision. Our findings point that the extensive margin accounts for about 70% of the aggregate labor supply response to the tax change and that non-linear wages play a crucial role in generating the large labor supply response along the extensive margin. We find that time aggregation, together with an operative extensive margin at the quarterly frequency, play an important role in generating the low estimates of the empirical elasticity. To understand this point, note that empirical studies that use household survey data (such as the PSID) use a wage rate that is obtained as the ratio of annual earnings over annual hours. We show that, even in the absence of measurement error in hours and earnings, this wage rate gives a noisy measure of the returns to work faced by individuals during the year. Intuitively, while temporary low wage shocks may induce individuals not to work in a given quarter, there are no traces of these low wage shocks in the annual wage data (since the low quarterly wage shocks are unobserved in the annual wage data when individuals do not work). To assess the role of time aggregation, we use quarterly simulated data to regress changes in log-leisure on changes in labor productivity. We focus on changes in leisure (rather than labor supply) to avoid the log-zero problem when individuals do not work. In particular, when individuals do not work their leisure time is set to 100% of the time endowment and their wage rate is set to the realized value of labor productivity in that quarter. Again, the results are striking: We obtain an elasticity of leisure of $-.50$ which is quite close to the macro elasticity of leisure of $-.54$. Thus, when using quarterly level data and labor productivity to measure the returns to work, the empirical elasticity now recovers the macro elasticity measured in the tax experiment. In a second experiment, we construct in our simulated data an annual measure of the returns to work that uses information on quarterly level productivity (where it is important that we *sum the logs* of quarterly productivity rather than use *log of the sum* of quarterly productivity) and we find that a regression on annual level data now recovers the macro elasticity.

Several important recent contributions have argued that the micro elasticity of labor supply need not be related to the aggregate labor supply response. In a theory of a representative agent with indivisible labor, Rogerson (1988) is the first one to show that individual and aggregate labor supply elasticities are effectively unrelated. By modeling heterogeneity, Chang and Kim (2006) go one step further and show that the slope of the aggregate labor supply schedule is determined by the distribution of reservation wages rather than by the willingness to substitute leisure intertemporally, establishing that when the extensive margin is operative aggregation plays a crucial role in determining aggregate labor supply responses. Moreover, borrowing constraints (Domeij and Flodén (2006)) and nonlinear wage functions (Rogerson and Wallenius (2006)) have been shown to generate a disconnect between the individual labor supply elasticity and the theoretical labor supply elasticity implied by preference parameters. Despite the insights provided by these recent contributions, each of them is usually consistent with only a small set of facts regarding labor supply at the micro level. This motivates us to build a theory of labor supply of an economy with heterogeneous agents which is consistent with micro level data and captures salient features of labor supply over the life cycle in several important dimensions. This is important for disciplining the aggregation from individual to aggregate labor supply responses, and for proceeding with some confidence to evaluate the aggregate labor supply response to macroeconomic shocks. Perhaps, the closest paper to ours is Imai and Keane (2004). These authors use a life-cycle model disciplined by micro level data to show that the empirical labor supply elasticity is unrelated to the preferences parameters when agents accumulate human capital. Relative to these authors, we contribute by focusing on the role of the extensive margin, time aggregation, and borrowing constraints in accounting for the disconnect. Moreover, we study an economy with preferences consistent with balance growth. Finally, Kimmel and Kniesner (1998) provide empirical evidence that the interaction of time aggregation and the extensive margin plays an important role in accounting for the low estimated empirical elasticities of leisure in the literature.

The paper proceeds as follows. Section 2 presents empirical facts on labor supply using data from the PSID. Section 3 develops a life-cycle theory of individual labor supply with heterogeneous agents. The calibration of the model economy is discussed in Section 4. Section 5 discusses the performance of the baseline economy in accounting for the facts documented on labor supply and the disconnect between the empirical and theoretical elasticities of labor supply.

2 Empirical Facts

2.1 The Data

The main dataset used in our analysis is the Michigan Panel Study of Income Dynamics (PSID) for the period 1968-1997.[2] The sample is restricted to males between the ages of 18 and 65. We do not place other restrictions on the sample. In particular, note that we do not restrict to heads of household − we use the information on annual hours worked provided by the PSID for those males who are listed as "wives" as well as the information on annual hours worked, whenever available in the individual files, on males who are dependents. This allows us to provide a more representative overview of the facts on labor supply as compared to the related literature which has mainly focused on male workers with strong labor market attachment.[3] Appendix I provides a detailed description of the construction of the dataset

[2]We have performed a similar empirical analysis using the Survey of Income and Program Participation (SIPP). The results obtained on the SIPP data are largely consistent with those obtained on the PSID data. These are available from the authors upon request.

[3]See for example Storesletten, Telmer, and Yaron (2001), Heathcote, Storesletten, and Violante (2004), Kaplan (2007), Badel and Huggett (2007).

and the variables used in the analysis.

The analysis is focused on the labor supply of men. A cohort is defined to consist of all individuals who turn 18 years old in a given year − for example, the 1967 cohort consists of all individuals who turn 18 years old in 1967. Since the PSID is a relatively small dataset, we grouped our sample into age and cohort groups. By age, individuals are grouped into 12 age groups each consisting of four ages − for example, age-18 group on the graphs include individuals between the age of 18 and 21, while age-22 group includes all individuals between the ages of 22 and 25. We have 17 cohort groups each consisting of three cohorts − for instance, the 1976 cohort group includes cohorts 1976, 1977, and 1978 while the 1985 cohort includes cohorts 1985, 1986, and 1987. We drop all cohorts smaller than 1940 and all cohorts greater than 1990.[4] We use PSID sample weights in the analysis.

As the data suggests, cohort effects are not significant in the case of men. Figures 1-4 show the following labor supply statistics over the life-cycle for various cohorts of men and women: mean annual hours worked, mean annual hours worked for those with positive hours, the fraction reporting positive annual hours, the variance of log annual hours, and the coefficient of variation of annual hours. While more recent cohorts of women have dramatically different labor supply behavior than older cohorts of women, that is not the case for men − while we do observe some differences across cohorts the labor supply behavior of more recent cohorts does not differ much from that of older cohorts. As a result, at this point we do not take out cohort effects for men.

Next we proceed with the empirical analysis and document a wealth of facts regarding the labor supply of men over the life-cycle. The patterns that we see in the data will be motivating the main features which will be introduced in the model. The most important patterns are as follows:

- We see a very pronounced life-cycle pattern in the labor supply behavior of men. We see the life-cycle trend in the mean annual hours worked, the participation rate, the dispersion of annual hours, the persistence in annual hours worked, and in the persistence of labor market participation.

- There is a substantial dispersion of annual hours worked at every point in the life-cycle.

- For most individuals, and for most ages during the life-cycle, annual hours are quite persistent.

[4]When we conduct the analysis by education groups, our last cohort is 1985 in order to be able to classify individuals as either high school or college.

- The labor supply behavior of high school and college graduates is different enough to warrant a separate analysis for each of these groups.[5]

2.2 Facts on the Life-Cycle Labor Supply of Men

Figures 5-11 and Tables 1 - 3 provide various facts on the life-cycle labor supply of men.

2.2.1 Average Annual Hours over the Life-Cycle

Figures 5 shows that mean annual hours worked clearly exhibit an inverted U-shape over the life-cycle − they increase early in life until the late 20s, stay constant after that until the late 40s, and decline monotonically after the age of 50. The second panel shows that college and non-college graduates have different life-cycle profiles − college graduates initially work less (while studying) while working more after the age of 26. In addition, the mean annual hours of high-school workers start declining earlier at the age of 50.

Figures 6 and 7 illustrate the intensive and extensive margins of the labor supply of men over the life-cycle. Between ages 30 to 46 working hours are quite constant and average annual hours are about 2,200 for non-college and 2,300 for college graduates. The extensive margin matters early in life until the age of 26, but is especially quantitatively important late in life after the age of 50. Furthermore, it is interesting to point out that the participation rate of those with high-school starts declining in the late 40s while the participation rate of those with college start declining significantly only in the late 50s.

2.2.2 Dispersion of Annual Hours over the Life-Cycle

Figure 8 displays the dispersion of annual hours over the life-cycle as measured by the coefficient of variation of annual hours. This figure illustrate three facts of particular importance. First, the dispersion in annual hours is U-shaped − it is high early in the life-cycle until the age of 26, then declines and is constant until the late 40s, and increases substantially after the age of 50. Second, the degree of dispersion is quite substantial as the variance of log hours is between 0.25 and 0.9. Indeed, inequality in hours worked is large relative to the inequality wages or consumption observe over the life cycle. While the variance of log wages is at most 0.3, the variance of log consumption is at most 0.15 during the lifecycle

[5] In this version of the paper, we consider an individual to be high school if he or she has at most 13 years of education while those with 14 years of education or more are considered to be college graduates. A sensitivity analysis with respect to the education cut-off separating high-school and college graduates indicates the current partition is a sensible one.

according to Kaplan (2007). Third, even though the dispersion of hours over the life cycle has the same shape for both College and non-college, the variance of log hours is higher in the case of non-college than college for all ages after 22. Other authors (Kaplan (2007) Heathcote, Storesletten, and Violante (2004)) have measured the inequality of hours work using the PSID data. They report a variance of log hours of about 0.08 which is much lower than what we find. The reason for the disparity is that we do not restrict the sample data as they do.[6]

2.2.3 Persistence in Annual Hours Worked

In this section, we investigate the extent to which annual hours worked are persistent over the individual's life. For that purpose each year we divide individuals into four groups: 1 − those with annual hours less than 100; 2 − those with annual hours between 100 and 1500; 3 − those with annual hours between 1500 and 2800; and 4 − those with annual hours greater than 2800.[7] We then construct transition matrices where cell ij indicates the fraction of all individuals in cell i in year t who moved to cell j in year $t+1$. We document the facts for all men as well as for high school graduates and college graduates.

Table 1 presents the transition matrix and the relative size of each group for men in three age groups: young workers between the ages of 18 and 29, middle-aged workers between the ages of 30 and 54, and old workers between the ages of 55 and 65. We found it useful to present graphically some of these results. In particular, Figure 9 graphs the relative size of each of the four groups as well as the fraction of workers who stay in each of these groups in two consecutive years (i.e. the diagonal elements from the transition matrix). Note that this graphical representation makes it easy to consider 12 age groups rather than the 3 age groups considered in Table 1.

Three important findings are worth pointing out. First, the group of full-time workers with annual hours between 1500 and 2800 is by far the largest, with the exception of the first and very last years of the life-cycle, and exhibiting very high persistence in annual hours worked − over 70% of men are in this labor supply group and more than 80% of those who are in this group in year t remain in it in year $t+1$. Table 1 further shows that, between

[6] Both Kaplan (2007) and Heathcote, Storesletten, and Violante (2004) restrict their sample to heads of households working more than 500 and less than 5086 hours annually and between ages 20-60.

[7] The cut-offs were chosen in order to capture four broad patterns of labor market behavior − no labor market participation (group 1), part-time labor supply (group 2), full-time labor supply (group 3), and very high labor supply (group 4). Slight changes in these cut-offs do not significantly change the main patterns documented here.

the ages of 30 and 54, most of those who move out of this group move temporarily into the group with large labor supply and work more than 2800 hours. That indicates that for the most part of the life-cycle, especially between the ages of 30 and 50, annual hours worked are quite persistent for most men. Second, the fraction of men who work less than 100 hours is quite small throughout the life-cycle, but starts increasing gradually after the age of 46. Furthermore, with age, this group becomes an absorbing state − after the age of 46, more than 80% of men who are in this group in year t will be there in year $t + 1$. Furtermore, as Table 1 shows, those who move out of it later in life, move temporarily into the part-time labor supply group. Third, the other two groups − those working between 100 and 1500 hours and those working more than 2800 hours − do exhibit a life-cycle pattern but are relatively small. In addition, each of these two groups seem to represent a temporary state in one's labor market history since the probability of remaining there is not very high.

These broad patterns are observed also for each of the two education groups − high-school and college men. As one can expect, after the age of 26, the group of full-time workers with annual hours between 1500 and 2800 (i) is bigger for the college men than the high-school men, (ii) starts declining gradually for high-school men at the age of 46 while for college men declines only after the age of 62, and (iii) is more persistent for college than for high-school men.

3 The Model

We develop a life-cycle theory of the labor supply of individuals. For simplicity, we abstract from the labor supply decisions of women and model males only. We consider a small open economy facing a fixed interest rate. We follow Hornstein and Prescott (1993), in modeling a production technology that give rise to a competitive equilibrium with non-linear wages.

3.1 Population, preferences, and endowments

We consider an economy populated by overlapping generations of individuals. Individuals face uncertain lifetimes and can live, at most, J periods. They differ in terms of their education (college versus non-college) and labor-productivity. The date-t utility function takes the form

$$u_t = u(c_t, l_t) = \ln c_t + \varphi \frac{l_t^{1-\sigma}}{1-\sigma},$$

where c_t is consumption and l_t denotes leisure. The utility function is consistent with balance growth. This assumption is motivated by the observation that there are no important cohort effects in the labor supply of men. It also allows the theory to be consistent with the fact that there are large permanent differences in labor productivity across individuals (heterogeneity in fixed effects) but not in their lifetime labor supply. The Frisch elasticity of leisure is given by $\frac{-1}{\sigma}$. Note that by modeling utility of leisure (rather than disutility of labor), the theory allows for an active extensive margin. In particular, the specification $u(c_t, h_t) = \ln c_t - \varphi \frac{h_t^{1-\sigma}}{1-\sigma}$ does not deliver an active extensive margin and, moreover, it often implies that individuals work 100% of their available time.

Individuals maximize lifetime expected utility

$$E_t \sum_{j=t}^{J} \beta^{t-j} u(c_j, l_j),$$

where E_t denotes expectations at date-t. Individuals face mortality shocks each period and uncertainty regarding their labor productivity z up to age 65 when labor productivity is zero (mandatory retirement). An individual's time endowment in each period is one. The amount of time that can be allocated to work is $h_j = 1 - l_j$. The college decision is exogenous. The education type of an individual determines the stochastic processes driving the mortality and labor productivity shocks.

3.2 Technology

There are a large number of plants and each plant is a collection of jobs. We assume that plants can operate jobs at zero costs. The production function of a job is given by

$$f(K, h, A\ z) = h^{\varepsilon} K^{1-\theta} (A\ z)^{\theta}, \quad \text{with } \theta \leq \varepsilon \leq 1$$

where h denotes the workweek, K is the amount of capital for the job, and Az is effective labor in the job. Effective labor in the job is given by the product of the worker productivity z and the level of technology A. Note that, for a fixed workweek, the job technology exhibits constant returns to scale in capital and effective labor. Moreover, as discussed in Osuna and Ríos-Rull (2003), when $\varepsilon = \theta$ the job technology reduces to the standard Cobb-Douglas technology where total hours of effective labor is what matters. When $\varepsilon > \theta$ the hours and effective labor are imperfect substitutes and the composition between these two inputs matters. When $\varepsilon = 1$ the technology is linear in hours and corresponds to the case where workers are not subject to fatigue.

3.3 The Plant's Problem

The plant takes as given the wage schedule $\tilde{w}(h, N)$ and the interest rate r. For each job, the plant manager chooses hours of work h, capital K, and effective labor N. In equilibrium, the wage schedule is a non-linear function of the workweek h and a linear function of effective labor N. To show this point, consider a job hiring a worker for h hours and with N units of effective labor. The optimal amount of capital K solves

$$\pi = \max_K h^{\varepsilon} K^{1-\theta} N^{\theta} - K(r+\delta) - \tilde{w}(h, N).$$

The solution to this problem implies

$$\frac{K}{N} = k^*(h, r) = \left[\frac{(1-\theta)h^{\varepsilon}}{r+\delta}\right]^{1/\theta}.$$

Next, notice that a job is open only if profits are non-negative. Free entry, and the fact that jobs can be opened at zero costs, imply that in equilibrium plants will make zero profits (will not extract economic rents from workers). Hence, competition for workers implies that the wage bill $\tilde{w}(h, N)$ is determined from

$$\pi = h^{\varepsilon} \left[N k^*(h, r)\right]^{1-\theta} N^{\theta} - N\ k^*(h, r)(r+\delta) - \tilde{w}(h, N) = 0,$$

which gives

$$\begin{aligned} \tilde{w}(h, N) &= w(h)\ N, \text{ where} \\ w(h) &\equiv (r+\delta)\frac{\theta}{1-\theta}\left[\frac{(1-\theta)h^{\varepsilon}}{r+\delta}\right]^{1/\theta}. \end{aligned}$$

It follows that the wage schedule $\tilde{w}(h, N)$ is linear in effective labor N and non-linear in hours of work h. When $\varepsilon = \theta$ earnings are also linear in h. When $\varepsilon > \theta$ earnings increase with h. In this case, households would be better off by selling employment lotteries to firms (Hornstein and Prescott (1993)). However, we rule out this possibility by assuming that households cannot commit to work when the realization of the employment lottery imply that they should work.

3.4 Government, Annuity, and Credit market

The government taxes consumption, capital income, and labor income. The tax revenue is used to finance government expenditures. Individuals can insure mortality risk in fair annuity markets. Denoting by R the gross interest rate net of capital income taxes τ_k, the gross interest rate faced by an individual j years old with education e is given by

$$R_j^e = 1 + \left(\frac{1+r}{\pi_j^e} - 1\right)(1-\tau_k),$$

where π_j^e is the conditional probability that an age $j-1$ individual with education e survives to age j. We assume that individuals can't borrow.

Social Security. The government also administers a pay-as-you-go social security system. To finance pensions for retired individuals, the government uses a payroll tax τ_{ss}. Individuals retire at age 65. Social security benefits depend on the average earnings made by individuals over the 35 highest years of earnings. Denoting this average earnings by $\overline{w}$, social security benefits can be expressed as $b_s(\overline{w}, j)$, where j denote the age of individuals.

Social security benefits are a function of the Average Indexed Monthly Earnings (AIME) over the 35 highest earnings years. Given that the model period is a quarter, for computational simplicity we compute average quarterly earnings over the 35×4 highest earnings quarters as follows

$$\overline{w}_{j+1} = \overline{w}_j + zw(h_j)/(35 \times 4) \text{ for } j \leq 35 \times 4, \quad (1)$$

$$\overline{w}_{j+1} = \overline{w}_j + \max\{0, (\min\{zw(h_j), \widehat{y}\} - \overline{w}_j)/(35 \times 4)\} \text{ for } j > 35 \times 4, \quad (2)$$

where $\widehat{y}$ is the maximum taxable earnings by the social security administration, which is set at 2.47 the average earnings in the economy). We express (1)-(2) in a compat way setting:

$$\overline{w}' = \Gamma_{ss}(\overline{w}, zw(h)) \quad (3)$$

At retirement, the Social Security Administration computes the Primary Insurance Amount (PIA) which is the sum of three portions of the Average Index Monthly Earnings (AIME). The bend points in the PIA formula are 0.2 and 1.24 of the average earnings in the economy when individuals file for social security($\overline{W}$)[8]. The social security benefit is given by

$$b(\overline{w}) = \begin{cases} 0.90 \times \overline{w} \text{ for } \overline{w} < .2\overline{W}, \\ 0.90 \times 0.2\overline{W} + 0.33 \times (\overline{w} - 0.2\overline{W}) \text{for } \overline{w} \in (.2\overline{W}, 1.24\overline{W}], \\ 0.90 \times 0.2\overline{W} + 0.33 \times (1.24\overline{W} - 0.2\overline{W}) + 0.33 \times (\overline{w} - 1.24\overline{W}) \text{ for } \overline{w} > 1.24\overline{W}, \end{cases} \tag{4}$$

3.5 The Individual's Problem

We use the recursive language to describe the problem of an individual. To simplify the notation, we abstract from the fact that the education type of an individual determines his earnings and mortality processes. The state of an individual is given by his age j, assets a, average lifetime earnings $\overline{w}$, and earnings shock z.

Since individuals live at most J periods, we set $V_{J+1}(x) = 0$. When a person is retired (has applied for social security benefits) his value is given by

$$\begin{aligned} V_j(a, b_s, z) &= Max\{u(c,1) + \beta\pi^e_{j+1}E[V_{j+1}(a, b_s', z')]\} \\ a_{j+1} &= b_s + R_j a - c(1+\tau_c), \\ a_{j+1} &\geq 0. \end{aligned}$$

The value of a person that has not retired is

$$\begin{aligned} V_j(a, \overline{w}, z) &= Max_{\{c,h,a'\}}\{u(c,l) + \beta\pi_{j+1}E[V_{j+1}(a, \overline{w}', z)]\} \\ & \quad \textit{subject to} \\ a_{j+1} &= (1 - \tau_{ss} - \tau_h + 0.5\tau_{ss}\tau_h)\ \min\{zw(h), \widehat{y}\} + (1-\tau_h)\max\{zw(h) - \widehat{y}, 0\} \\ & \quad ... + R_j a - c\ (1+\tau^c), \\ a_{j+1} &\geq 0, \\ \overline{w}' &= \Gamma_{ss}(\overline{w}', zw(h)),\ \text{and } l + h = 1. \end{aligned}$$

Using that the wage function $w(h, N)$ is an affine function of effective labor, the earnings of an individual working h hours and supplying z units of labor can be expressed as $z\ w(h)$. The individual takes as given the wage schedule $w(h)$, the function Γ_{ss} in (3) determining the evolution of average lifetime earnings. Note that the Social Security Administration does not tax earnings above $\widehat{y}$. Half of the social security taxes are payed by the employer and are not subject to personal income tax τ_h.

[8] $\overline{W}$ is the average earnings in the economy in the year when the individual becomes 62 years old.

4 Calibration

The calibration for most of the parameters in the Baseline Economy is quite standard so that we fix these parameters using available estimates in the literature. The crucial task in our calibration is the parameterization of the stochastic process on labor productivity. To model variation in employment within a year, the calibration sets the model period to 1 quarter. In calibrating a *quarterly* stochastic process on labor productivity, one difficulty is that the PSID only reports earnings and hours of work at an annual frequency. Moreover, in using wage data to calibrate a stochastic process on labor productivity we need to take a stand on how hours of work affect labor productivity and we need to consider that the data only report wages for individuals that work. To deal with these problems, we proceed as follows:

1. Estimate an annual wage process for college and non-college workers from the PSID data.

2. Use estimates from Aaronson and French (2004) on nonlinear wages to pin down the parameter ε determining how hours of work affect labor productivity in the model economy.

3. Feed a quarterly labor productivity process into the model economy.

4. Simulate the model economy to obtain quarterly data on employment, hours of work, and earnings.

5. Aggregate the quarterly data to an annual period.

6. Estimate an annual wage process for college and non-college workers in the model generated data.

7. Feed a new quarterly labor productivity process (go back to step 3), until the ”same” annual wage process is obtained in the model and in the data.

Below we describe the calibration in detail. We first discuss the calibration of the “macro” parameters. We then discuss the calibration of the labor productivity process and how we deal with the possibility of measurement error in hours and earnings in the PSID data.

4.1 Calibration of preferences, technology, and macro parameters.

The model period is set at one quarter. The model economy is solved in partial equilibrium for a fixed interest rate. The quarterly interest rate is chosen so that the implied annual rate of return on capital (net of depreciatoin) is 4%. [9]

Preference parameters, time endowments, and mortality rates. Following Kaplan and Violante (2008), the discount rate β is chosen to match an asset to income ratio of 2.5. This is the wealth to income ratio when the top 5% of households in the wealth distribution are excluded from the Survey of Consumer Finances. The reason for excluding the richest households in computing an aggregate wealth to income ratio is that the PSID undersamples the top of the wealth distribution. Following Prescott (2004) and Osuna and Ríos-Rull (2003), the time endowment is set at 5200 hours a year (100 hours per week). The preference parameter φ determining taste for leisure is chosen so that prime age individuals work about 42% of their available time. The curvature parameter on leisure (σ) is set at 3 implying a Frish-elasticity of leisure of .33 and a Frish-elasticity of labor supply of .5. In the sensitivity analysis we also calibrate economies with σ equal to 2 and 4. The mortality risk for college and non-college individuals is taken from Bhattacharya and Lakdawalla (2006).

Technology parameters. The labor share θ is set to .64. To calibrate the parameter ε, we use the fact that the equilibrium wage rate in our theory satisfies

$$\frac{w(h)}{h} = \text{constant} \quad h^{\varepsilon/\theta - 1},$$

Note that the elasticity of the wage rate to a change in hours of work is given by $\varepsilon/\theta - 1$. In an empirical study, Aaronson and French (2004) estimate this elasticity to be slightly above .40. We thus set $\varepsilon = 1.4 * \theta$.

Tax rates, and social security. The tax rate on consumption τ_c is set at .055 as in Conesa, Kitao, and Krueger (2009). Following Domeij and Heathcote (2004), taxes on capital income and labor income are set to $\tau_k = .40$ and $\tau_w = .27$. The social security tax rate is set to $\tau_{ss} = 0.12$, and the cap $\widehat{y}$ on social security taxation is fixed at 2.47 of average earnings in the economy ($\overline{W}$). [10]

[9] The depreciation over a yearly period is assuumed to be 4%. Because the model economy is solved in partial equilibrium, the depreciation rate does not affect any of the results in the next section of the paper.

[10] Actually, $\overline{W}$ is set at 80% of average earnings in the economy. The reason is that our model only include male workers. Using data from the CPS, we find that the average earnings among all workers in the US economy are about 80% of the average earnings of male workers.

4.2 Calibration of labor productivity

We use a GMM procedure to estimate the following annual wage process in the PSID data for college and non-college individuals:

$$\ln w_{ij} = x_j\ \kappa + \alpha_i + u_j + \lambda_j, \tag{5}$$

where x_j is a quartic polynomial in age, κ is a vector of coefficients, $\alpha_i \sim N(0, \sigma_\alpha^2)$ is a fixed effect determined at birth, $\lambda_j \sim N(0, \sigma_\lambda^2)$ is an idiosyncratic transitory shock, and u_j follows a first-order autoregression:

$$u_j = \rho u_{j-1} + \eta_j, \quad \eta_j \sim N(0, \sigma_\eta^2), u_0 = 0. \tag{6}$$

While the parameters $(\kappa, \rho, \sigma_\alpha^2, \sigma_\eta^2)$ vary across education types, this is omitted in the notation to avoid clutter. The estimated wage processes are reported in Table 4. The empirical findings show that the variance of fixed effects is quite large for both education types, with values of .06 and .08. for the non-college and the college types. Both wage processes exhibit high autocorrelation, with a value of .97 for non-college individuals and .92 for college individuals. The variance of the innovation of the autoregressive process is .013 and .029. The estimates reveal that both education types exhibit quite high transitory shocks to their wages, with variances of .097 and .079.

To calibrate the model economy, we need to find a *quarterly* stochastic process on labor productivity that is consistent with the annual wage process estimated in the data (equations (5)-(6)). To do this, we assume that labor productivity is the sum of an annual autoregressive process and a quarterly transitory shock.[11] Specifically, while the transitory shock is drawn every quarter, the persistent shock is only drawn at the first quarter of each year (age). To make these assumptions operational, we discretize all shocks by considering, for each education type, 15 values for the autoregressive shocks, 4 values for temporary shocks, and 2 values for fixed effects. The transition probabilities of the persistent shock are computed using a Tauchen routine.

The empirical literature has stressed the importance of measurement error in hours and earnings in household survey data. Moreover our empirical findings are suggestive of the importance of measurement error since the estimated variation in the transitory component

[11] We have tried an specification that allows for an autoregressive process at the quarterly level. In this case, however, we were not able to recover the stochastic process estimated in the data. When labor productivity follows an autoregressive process at the quarterly level, there is no reason to expect the logarithm of the sum of quarterly earnings to be well approximated with an autoregressive process.

of wages seems implausibly large. To get a quantitative sense of the magnitude of these variance, we have calibrated the model economy assuming that there is no measurement error in hours and earnings. We find that to match the estimated variance of transitory shocks to wages in the data the calibration requires a temporary shock process that implies that with 20% percent chance labor productivity can either increase by a factor of 15 or decrease by a factor of 15. Moreover, in the presence of such large temporary shocks, 50% of individuals in the model economy work less than 3 quarters, an implication that is grossly counterfactual.

We thus need to take seriously measurement error in the data. To this end, we assume that the transitory shock λ_j in the empirical model is the sum of a (true) temporary wage shock and measurement error in hours (m_H) and earnings (m_E), with measurement error in hours and earnings being normally distributed with mean zero. The estimated transitory variation in "observed" wages σ_λ is then the sum of the variances of transitory true wages (σ_T), measurement error in earnings σ_E^2, and measurement error in hours σ_H^2 :

$$\sigma_\lambda^2 = \sigma_T^2 + \sigma_E^2 + \sigma_H^2. \tag{7}$$

We use the implications of our theory to take a stand on the relative importance of $(\sigma_T^2, \sigma_E^2, \sigma_H^2)$ in accounting for the estimated variance in "observed" transitory wages (σ_λ^2). To this end, we assume that annual hours and earnings are measured with error in the model economy. To calibrate the variance of (true) transitory wages σ_T^2 , we note that in our theory this variance has important effects on the probability that individuals work in each quarter: The larger σ_T^2 the less likely individuals will work four quarters in a year. We thus use this statistic as a calibration target where the fraction of individuals working 4 quarters in a year is taken from the Survey of Income and Participation Program (SIPP). We note that the SIPP allows us to have more relatiable measures of labor force participation at the quarterly frequency than the PSID as it interviews individuals three times in a year (rather than once a year as in the PSID). We allow the variance of the transitory variation of wages (σ_T) to vary with age and education. For each education group, we set the transitory variation of wages at ages 25, 35, 50, 64. The transitory variation in wages at other ages are obtained by linear interporlation. We target the fraction of people working 4 quarters for four age-groups: ages 25 to 35, ages 35 to 50, ages 50 to 55 and ages 55 to 65.

To distinguish between σ_E^2 and σ_H^2, we need an additional target. This is done by comparing the variance of transitory wages in two alternative estimations of the wage process in (5)-(6): The first specification regresses observed wages while the second specification regresses wages net of the effect of hours of work on wages. Identification come from the fact that measurement error in hours and in earnings affect differently the variance of transitory earnings in the two specifications of the regression. To develop this point, we start by noticing that when wages are a non-linear function of hours, the observed wage rate is given by

$$w(h) = \frac{e^{w}\ h^{\varepsilon/\theta}\ e^{m_E}}{h\ e^{m_H}} = e^{w} h^{\varepsilon/\theta - 1} e^{m_E - m_H}, \tag{8}$$

where w is the logarithm of labor productivity, ε/θ determines the elasticity of earnings to hours of work, and (m_E, m_H) are measurement error in (log) earnings and (log) hours. In the absence of measurement error, the wage rate net of the effect of hours on wages would be uncovered by taking logs and subtracting $(\varepsilon/\theta - 1)\ln h$ from both sides of (8)

$$\ln w(h) - (\varepsilon/\theta - 1)\ln h = w.$$

In practice, though, hours are observed with error . Subtracting $(\varepsilon/\theta - 1)\ln(he^{m_H})$ from both sides of equation (8) to "clean" wages from the effect of (observed) hours gives

$$\ln w(h) - (\varepsilon/\theta - 1)\ln(he^{m_H}) = w + (\varepsilon/\theta - 1)\ln h + m_E - m_H - (\varepsilon/\theta - 1)\ln(he^{m_H}),$$

which can be re-arranged as

$$\ln w(h) - (\varepsilon/\theta - 1)\ln(he^{m_H}) = w + m_E - \varepsilon/\theta m_H$$

If w follows the empirical model in (5)-(6), we obtain the following empirical model for "clean" wages:

$$\ln w(h) - (\varepsilon/\theta - 1)\ln(he^{m_H}) = x_j\ \kappa + \alpha_i + u_j + \lambda_j + m_E + \varepsilon/\theta m_H$$

The transitory variation in "clean wages" is then given by

$$VAR(\lambda_j + m_E + \varepsilon/\theta\ m_H) = \sigma_T^2 + \sigma_E^2 + (\varepsilon/\theta)^2\sigma_H^2. \tag{9}$$

When wages are a non-linear function of hours of work $(\varepsilon/\theta > 1)$ and the wage process is estimated net of the effect of hours on wages, measurement error in hours lead to an increase in the estimated transitory variation of wages. Intuitively, the estimated transitory variation of wages increase because we are not using the "correct" hours to clean the wage data. Comparing (9) with (7), the increase in the variance of transitory wages is given by

$$\Delta VAR_T = [(\varepsilon/\theta)^2 - 1]\ \sigma_H^2.$$

For the calibrated value of $\varepsilon/\theta = 1.4$, we have that $[(\varepsilon/\theta)^2 - 1] \simeq 1$ so that $\Delta VAR_T = \sigma_H^2$. Thus, for each education type, the variance of measurement error in hours σ_H^2 is obtained as the increase in the temporary variance in wages when the wage data is clean with hours data. When we run the two specifications of the regression in the PSID data we found that the variance of temporary wages increase by .045 for non-college individuals and .031 for college individuals. We thus obtain that measurement error in hours for non-college and college individuals is given by $\sigma_H^2 = .045$ and $\sigma_H^2 = .031$. We introduce these values for measurement error in hours into the model economy. We then run the two specifications of the wage regression with model simulated data. Reassuringly, for each education type, when the regression is run with clean wages the variance of the transitory component in wages increases by an amount approximately equal to the measurement error in hours.

5 Quantitative Findings

5.1 Calibration Results

There are 29 parameters that we calibrate by solving the model economy. Table 5 shows the values for average earnings $\overline{W}$, taste for leisure φ, and discount rate β and the targets used to pin down these values. Tables 6 and 6 show the values chosen for the transitory variation in wages and the initial heterogeneity in wages (at age 25) for non-college and college individuals, respectively. As we explained before, the variance of the transitory shock on labor productivity has a first order effect on the fraction of individuals working 4 quarters which is a target in the calibration. These tables show that for all age and education groups the model replicates quite well the fraction of people working 4 quarters as well as the variance of log wages at age 25. Tables 7 and 7 report the values of the parameters characterizing the AR(1) process as well as the standard deviation of the fixed effect shock affecting labor productivity for non-college and college types. These tables also report the targeted statistics which are the estimated variance of the fixed effect and the parameters of the AR(1) process for log wages (also in Table 4). The values reported under the column Model correspond to the GMM estimation using annual model data for the baseline economy.[12]

5.2 The Facts on Labor Supply: The Performance of the Model

Figures 12-16 present the performance of the model in accounting for the facts on labor supply at the micro level. Overall, the model captures most of the salient features of labor supply. Recall that the facts on labor supply were not explicitly targeted, indicating that the features included in the analysis are important determinants of individuals' labor supply decisions.

Figure 12 displays mean annual hours over the life-cycle in the model and for various cohorts in the data. The model captures very well the pattern observed in the data.

Figure 13 displays the labor supply in the model at the intensive margin, while Figure 14 shows the dispersion of annual hours worked over the life-cycle both in the model and in the data. Overall, the model does an excellent job in accounting for the patterns observed in the data.

Finally, we investigate the persistence of annual hours worked in the model. We apply the same procedure as in the data and divide individuals into four groups: 1 − those with annual hours less than 100; 2 − those with annual hours between 100 and 1500; 3 − those with annual hours between 1500 and 2800; and 4 − those with annual hours greater than 2800. Figure 15 shows the relative size of each of these groups over the life-cycle both in the model and in the data. Two observations stand out. First, the model captures the fact that Group 3 (those working between 1500 and 2800 hours) is by far the largest group with a share which declines significantly only after the age of 55. Second, the model captures the fact that Group 1 (those working between 0 and 100 hours) is very small initially and increasing substantially only after the age of 50.

Figure 16 shows the persistence in annual hours worked both in the model and in the data. First, the model captures the fact that those who are working between 1500 and 2800 hours (Group 3) in a given year, with a very high probability will be in the same group the year after. Second, the model also captures the fact that early in the life cycle those who find themselves not working in a given year tend not to stay in the same group the year after. Later in the life cycle, however, this becomes an absorbent group − those who end up in this group tend to stay in it with a very high probability. Finally, the other two groups

[12] The values of the parameters characterizing the deterministic age-profile of labor productivity are (-3.45,0.38,$-1.31 * 10^{-}2$, $2.098 * 10^{-}4$,$-1.28 * 10^{-}6$) for non-college and (-4.71,0.48,$-1.52 * 10^{-}2$, $2.22 * 10^{-}4$,$-1.26 * 10^{-}6$) for college.

are not very big and tend to be transitory − individuals end up in those groups every now and then, but tend to quickly exit them.

5.3 The Elasticity of Labor Supply

The Frisch elasticity of labor supply describes how labor supply responds to an intertemporal change of wages that leaves the marginal utility of wealth constant. We use three alternative proceudures to compute the Frisch elasticity of labor supply in our model economy. The first procedure assumes an interior solution in the agents' labor supply decision and uses a linear approximation to the first order conditions of the maximization problem to obtain an analytical expression for the labor supply elasticity. We refer to this elasticity as the "theoretical elasticity of labor supply". The second procedure consists in simulating the model economy to generate cross-sectional data that can be used to estimate an "empirical elasticity of labor supply" using standard econometric techniques. The third procedure consists in simulating a one period (quarterly) unanticipated tax increase to estimate the aggregate labor supply response to a one time wage change. We shall refer to this third elasticity as the "macro elasticity".

5.3.1 The Theoretical Elasticity of Labor Supply

Assuming an interior solution to the first-order conditions for individuals utility maximization and using a linear approximation, it is easy to obtain an expression for the Frisch elasticity of leisure in our model economy (see appendix)

$$\eta^l = -\frac{1}{\sigma}.$$

This expression is equal to the one obtained in a model with linear wages. Hence, nonlinear wages do not affect the Frisch elasticity of labor supply *along the intensive margin.* It is standard to convert the elasticity of leisure into a labor supply elasticity by setting $\eta^h = -\frac{(1-h)}{h}\eta^l$. Hence, the theoretical elasticities of leisure and labor in the baseline economy are: $\eta^l = -.33$ and $\eta^h = .46$.

5.3.2 The Empirical Elasticity of Labor Supply

The empirical elasticity of labor supply is obtained by running the following regression on model generated data via ordinary least squares (OLS) and instrumental variables (IV):

$$\triangle \ln h_{it} = \beta_0 + \beta_1 \triangle \ln w_{it} + \varepsilon_{it}, \tag{10}$$

where the regression coefficient β_1 gives the empirical labor supply elasticity predicted by the model economy.[13] The instruments used are past wage changes (IV1) and a composite of a constant, age, age-squared, and the twice lagged log-wage (IV2). To compare results for economies with different values of the theoretical elasticity, we also simulate a model economy with $\sigma = 4$ (with all other parameters being recalibrated to match the targets used in the calibration of the baseline economy). The findings are reported on Table 8.

We find that all the estimates of the empirical elasticity of labor supply are well below the value predicted by the theoretical elasticity. The lowest estimates are obtained when using OLS and when instrumenting with lagged wage changes (IV1). The highest estimates are obtained when instrumenting with a constant, age, age-squared, and the twice lagged log-wage (IV2). Note that the IV2 instruments use information from the age-profile of wages and hours to estimate the labor supply elasticity. This procedure identifies the elasticity from anticipated life-cycle wage changes and is particularly good in correcting for measurement error in hours and wages. In the baseline economy, the empirical elasticity of labor supply obtained with IV2 is .15, well within the range of $[0, 0.5]$ in the empirical literature. Interestingly, the theoretical elasticity is equal to .46, three times the value implied by the empirical elasticity. Moreover, while the theoretical elasticity varies with σ, the empirical elasticity remains roughly constant. Another way to evaluate these results is as follows: Suppose that a researcher does not know the preference parameter σ used to generate the model data. If this researcher were to use simulated data and the expression for the theoretical elasticity to back up a value of σ, the researcher would obtain $\sigma = 9$ for all the economies considered. [14] We thus conclude that the simulated data is uninformative about the preference parameter σ. To put it differently, the empirical elasticity of labor supply is not a good calibration target for the preference parameter determining the Frisch elasticity of leisure in our model economy.

5.3.3 The Macro Elasticity of Labor Supply

The macro elasticity is obtained by computing the aggregate labor supply response to a one period unanticipated increase of the labor income tax (from .27 in the baseline economy to

[13] Note that the error term may be correlated with the contemporaneous wage change due to the wealth effect of a wage change on labor supply. To deal with this problem, the empirical literature follows an instrumental variable approach.

[14] To map the elasticity of leisure into an elasticity of labor, we use that the fraction of the time endowment allocated to work by prime-aged males in our model economy is .42.

.37). The change in the aggregate labor supply can be safely interpreted as pure "substitution effect" as the wealth effect of a one period (quarterly) tax change is small. The results from these experiments could not be more striking: The macro elasticity of labor supply in our baseline economy is 0.85, which is about six times larger than the empirical elasticity (.15), and almost twice as big as the theoretical elasticity (0.46). The large labor supply response along the extensive margin explains why the macro elasticity is much bigger than the theoretical elasticity. (Recall that the latter was derived assuming an interior solution in the labor supply decision.) Restricting attention to labor supply changes along the intensive margin, decreases the macro elasticity from .85 to .23. Our findings point that the extensive margin accounts for about 70% of the aggregate labor supply response to the tax change. We emphasize that non-linear wages play a key role in generating the large labor supply response along the extensive margin, as this response is negligible in a model with linear wages. It is interesting to point that the value of σ that would allow a standard resentative agent model of labor supply decisions (one that abstracts from incomplete markets, life-cycle, non-linear wages and the extensive margin) to match the macro elasticity predicted by our theory $\sigma = 1.8$, which is much smaller than the $\sigma = 3$ in the baseline economy and the $\sigma = 9$ implied by the simulated household-level data.

5.4 Discussion: The Role of Time Aggregation in Understanding the Disconnect

Our main finding is that there is a strong disconnect between the theoretical, the empirical, and the aggregate elasticity of labor supply: While in our Baseline Economy the macro elasticity is .85, the theoretical elasticity is .46, and the empirical elasticity is only .15. We have shown that the extensive margin plays a crucial role for understanding the high value of the macro elasticity of labor supply relative to the theoretical elasticity. We now consider two experiments that shed some light on the factors driving the low empirical elasticity of labor supply.

The Frisch elasticity of labor supply is about the substitution effect of a wage change. The empirical approach uses individual level data to estimate the response to an unanticipated wage change, keeping wealth constant. However, empirically this response is not easy to isolate. In the data, wage changes confound information about transitory and permanent (unanticipated) wage shocks, and the latter can have important wealth effects. To control for wealth effects, some empirical studies use measures of wealth or of food consumption but the measures of wealth and consumption in household surveys are far from perfect. Other empirical studies control for wealth effects by "first-differencing" the data but, again, it is not obvious that this approach works when agents face uninsurable idiosyncratic risk. Moreover, household surveys (such as the PSID) typically report wages at an annual frequency and wages (earnings) are only observed when individuals work. Thus, time aggregation together with the fact that the extensive margin can be operative within a year make observed annual wages a noisy measure of the returns to work faced by individuals during the year. Econometric theory points that regressing changes in log leisure on changes in log wages will underestimate the true elasticity if the explanatory variable is measured with noise. The experiments below show how time aggregation and an active extensive margin at the quarterly level give rise to a noisy measurement of the returns to work in the annual data, which biases downward the estimates of the elasticities of leisure and labor supply.

5.4.1 Experiment 1: Using Quarterly Data

To control for the effects of time aggregation on the estimation of empirical elasticity of labor supply, we run the regression (10) using quarterly data simulated from the model economy. Moreover, we only consider wage changes in quarters two, three, and fourth. Recall that in our calibrated model economy wage shocks in the last three quarters are temporary. By dropping the first quarter, we can better isolate the substitution effect of wage changes as temporary wage changes at the quarterly level should have small wealth effects. We run the regression with changes in "true" hours and wages (that is, we clean the model data from measurement error) because we are now interested in isolating the effects of time aggregation and the extensive margin for estimating labor supply responses. Finally, in order to incorporate labor supply responses along the extensive margin we run regression (10) on changes in log-leisure rather than changes in log hours (to avoid the log-zero problem when individuals do not work). The results are reported on Table 10.

We find that the empirical elasticity of leisure in our Baseline Economy is $-.30$ when using *quarterly* data, which is quite close to the value of the theoretical elasticity of leisure of $-.33$. To evaluate how the extensive margin affects our estimates of labor supply respones, we run the regression (10), including all individuals alive in each quarter, regardless of whether they had worked or not. In particular, when individuals do not work their leisure time is set to 100% of their time endowment and their wage rate is set to the realized value of labor productivity z in that quarter. The results is quite striking: We obtain an elasticity of leisure of $-.50$, which is quite close to the macro elasticity of leisure of $-.54$ obtained in the tax experiment. Summing up, by incorporating quarterly level data and the extensive margin we can account for almost all the difference between the empirical and the macro elasticities of labor suply. We conclude that time aggregation and the extensive margin play a crucial role in accounting for the disconnect.

5.4.2 Experiment 2: Time-Aggregation of the Returns to Work

The wage rate in annual survey data is measured as $lnw = ln\frac{AnnualEarnings}{AnnualHours}$. We now argue that (even in the absence of measurement error in hours and earnings) this wage rate gives a noisy measure of the returns to work faced by individuals that biases estimates of the empirical elasticity of labor obtained with annual data. To this end, we run regression (10) with annual model data asssuming away measurement error and using the quarterly sum of realized labor productivity as an explanatory variable ($ln\sum z$ instead of ln w). Table 11 shows that the annual elasticity of leisure increases from $-.21$ to $-.32$ when using the quarterly sum of labor-productivity instead of annual wages. Next, we show that there are some other important subtleties in aggregating the returns to work over the year. To this end, we run the regression (10) using $\sum(ln(z))$ as a regressor. While the empirical elasticity of leisure is $-.32$ when using $ln\sum z$, it is $-.38$ when using $\sum(ln(z))$. Hence, the log of the sum of labor productivities ($ln\sum z$) is a worse measure of the returns to work faced by individuals during the year than the sum of the log of quarterly labor produtivities $\sum(ln(z))$. Because there may be important wealth effects associated with change in labor productivity over the year, we isolate the substitution effect of a change in labor productivity by instrumenting he regression (10) with changes in lagged $\sum(ln(z))$. This procedure gives an empirical elasticity of leisure equal to $-.54$, which fully recovers the macro elasticity of leisure obtained in the tax experiments.

6 Conclusion

McGrattan and Rogerson (1998)

McGrattan and Rogerson (2004)

French (2005)

Osuna and Ríos-Rull (2003)

Hornstein and Prescott (1993)

Pijoan-Mas (2006)

Chang and Kim (2006)

Domeij and Flodén (2006)
Bils and Cho (1994)
Kydland and Prescott (1991)
Cho and Cooley (1994)
Rogerson and Wallenius (2006)

Table 1: Transition Matrix across Annual Hours Cells, in Percent, PSID, 1968-1996, Men.

	Ages 18-29				
			To		Relative
From	000-100	100-1500	1500-2800	2800+	Size
0-100	50.11	32.20	16.52	1.17	6.42
100-1500	7.55	46.68	42.53	3.24	22.00
1500-2800	0.81	9.73	81.34	8.11	62.19
2800+	0.32	4.15	48.78	46.75	9.39
	Ages 30-54				
			To		Relative
From	0-100	100-1500	1500-2800	2800+	Size
0-100	78.46	14.72	6.16	0.66	15.12
100-1500	10.56	37.78	47.20	4.46	8.01
1500-2800	0.60	5.48	86.02	7.90	72.86
2800+	0.26	2.61	38.78	58.35	15.12
	Ages 55-65				
			To		Relative
From	0-100	100-1500	1500-2800	2800+	Size
0-100	92.18	6.40	1.23	0.19	21.19
100-1500	31.88	42.47	23.81	1.85	13.92
1500-2800	2.22	12.80	79.44	5.54	56.17
2800+	1.37	4.61	40.52	53.51	8.72

Note - Authors' calculations from the PSID.

Table 2: Transition Probability across Annual Hours Cells, PSID, Men, High-School.

	Ages 18-29				
	To				Relative
From	000-100	100-1500	1500-2800	2800+	Size
0-100	55.03	24.12	20.56	0.29	6.90
100-1500	11.10	31.66	54.92	2.33	12.50
1500-2800	1.05	7.04	87.26	4.64	74.12
2800+	0.26	2.14	48.14	49.47	6.47

	Ages 30-54				
	To				Relative
From	0-100	100-1500	1500-2800	2800+	Size
0-100	81.37	9.85	8.43	0.36	6.26
100-1500	17.83	30.20	50.03	1.94	5.40
1500-2800	1.18	3.69	90.19	4.94	78.83
2800+	0.24	2.07	42.72	54.97	9.51

	Ages 55-65				
	To				Relative
From	0-100	100-1500	1500-2800	2800+	Size
0-100	92.35	5.30	2.35	0.0	25.55
100-1500	42.23	31.75	25.42	0.60	9.88
1500-2800	4.26	9.33	83.72	2.68	59.94
2800+	2.59	3.73	44.33	49.35	4.62

Note - Authors' calculations from the PSID.

Table 3: Transition Probability across Annual Hours Cells, PSID, 1968-1996, Men, College.

	Ages 18-29				
	To				Relative
From	000-100	100-1500	1500-2800	2800+	Size
0-100	48.75	29.99	20.14	1.13	5.65
100-1500	7.99	42.97	47.47	1.57	15.78
1500-2800	0.65	5.77	88.08	5.49	72.64
2800+	0.38	2.07	54.50	43.05	5.93
	Ages 30-54				
	To				Relative
From	0-100	100-1500	1500-2800	2800+	Size
0-100	68.47	13.82	16.90	0.81	1.65
100-1500	8.31	31.72	55.52	4.44	3.27
1500-2800	0.44	2.28	91.42	5.86	84.55
2800+	0.24	0.98	49.03	49.76	10.54
	Ages 55-65				
	To				Relative
From	0-100	100-1500	1500-2800	2800+	Size
0-100	93.69	4.48	1.83	0.0	14.13
100-1500	30.56	37.05	29.42	2.97	21.30
1500-2800	2.32	7.70	85.90	4.08	73.00
2800+	0.0	2.94	62.06	35.0	5.70

Note - Authors' calculations from the PSID.

Table 4: PSID: Stochastic Process of Hourly Wages.

	Non-college	College
Var(α)	0.063	0.083
ρ	0.972	0.921
Var(η)	0.013	0.029
Var(λ)	0.097	0.079

Notes - .

Table 5: Parameters on Preferences and Pension Formula

Parameter	Value	Variable	Target	Result
$\overline{W}$	3.3	Ratio $\overline{W}$ to Males‘average earnings	.84	.84
φ	0.65	Fraction of hours worked	42%	42%
β	0.98	Asset to income ratio	2.5	2.55

Table 6: Calibration of the Transitory shock

Non-College				
Parameter	Value	Variable	Target	Result
σ_T at age 25	0.247	fraction working 4 quarters age 25	0.72	0.70
σ_T at age 35	0.194	fraction working 4 quarters age 35	0.75	0.73
σ_T at age 50	0.163	fraction working 4 quarters age 50	0.72	0.76
σ_T at age 64	0.147	fraction working 4 quarters age 64	0.38	0.30
σ_w at age 25	0.2	variance of ln wages at age 25	0.21	0.21

College				
Parameter	Value	Variable	Target	Result
σ_T at age 25	0.2	fraction working 4 quarters age 25	0.84	0.83
σ_T at age 35	0.15	fraction working 4 quarters age 35	0.87	0.87
σ_T at age 50	0.14	fraction working 4 quarters age 50	0.85	0.85
σ_T at age 64	0.12	fraction working 4 quarters age 64	0.54	0.52
σ_w at age 25	0.205	variance of ln wages (with meas. error)	0.20	0.20

Table 7: Calibration of the Fixed Effect and Persistent Shock: Non-College

Parameter	Value	Variable	Target	Model
		Non-College		
Parameter	Value	Variable	Target	Model
σ_α	0.251	Variance of fixed component of log wages	0.063	0.041
ρ	0.971	Correlation coefficient AR component of log wages	0.972	0.0973
σ_η	0.112	Variance of innovation AR component of log wages	0.013	0.018
		College		
Parameter	Value	Variable	Target	Model
σ_α	0.288	Variance of fixed component of log wages	0.083	0.083
ρ	0.921	Correlation coefficient AR component of log wages	0.921	0.929
σ_η	0.169	Variance of innovation AR component of log wages	0.029	0.030

Table 8: The Elasticity of Labor Supply

	Theoretical	OLS	IV1 ($\Delta \ln w_{t-1}$)	IV2 (age,age^2,$\ln w_{t-2}$)
Baseline Economy	0.46	-0.26	-0.28	0.15
$\sigma = 4$	0.34	-0.26	-0.28	0.14

Notes - .

Table 9: Macro Elasticity of Labor Supply: A Tax Experiment

Elasticity:	Labor Supply	Intensive margin	Leisure
Baseline	0.85	0.23	-0.54
$\sigma = 4$	0.61	0.16	-0.39

Table 10: Elasticity of Leisure: The Extensive Margin.

	Annual	Quarterly (-first)	Quarterly (Extensive)	Theoretical	Macro
Baseline Economy	-0.21	-0.30	-0.50	-0.33	-0.54
$\sigma = 4$	-0.15	-0.21	-0.37	-0.25	-0.39

Table 11: Elasticity of Leisure: Time Aggregation.

	Theoretical	Annual	$Ln(\sum z)$	$\sum(ln(z))$	IV $\sum(ln(z))$	Macro
Baseline Economy	-0.33	-0.21	-0.32	-0.38	-0.54	-0.54
$\sigma = 4$	-0.25	-0.15	-0.26	-0.31	-0.40	-0.39

Figure 1: Mean Annual Hours Worked, 1968-1996, PSID.

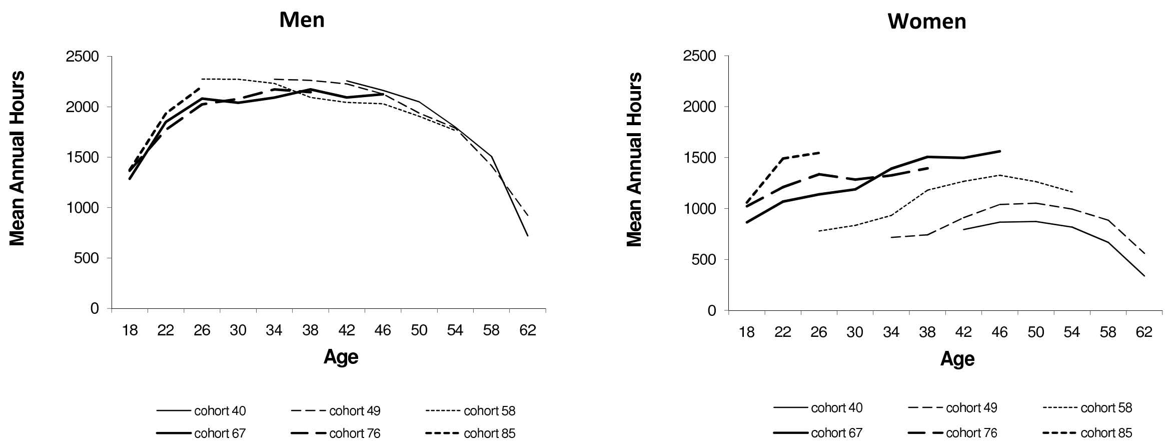


Figure 2: Mean Annual Hours Worked, Workers with Positive Hours, 1968-1996, PSID.

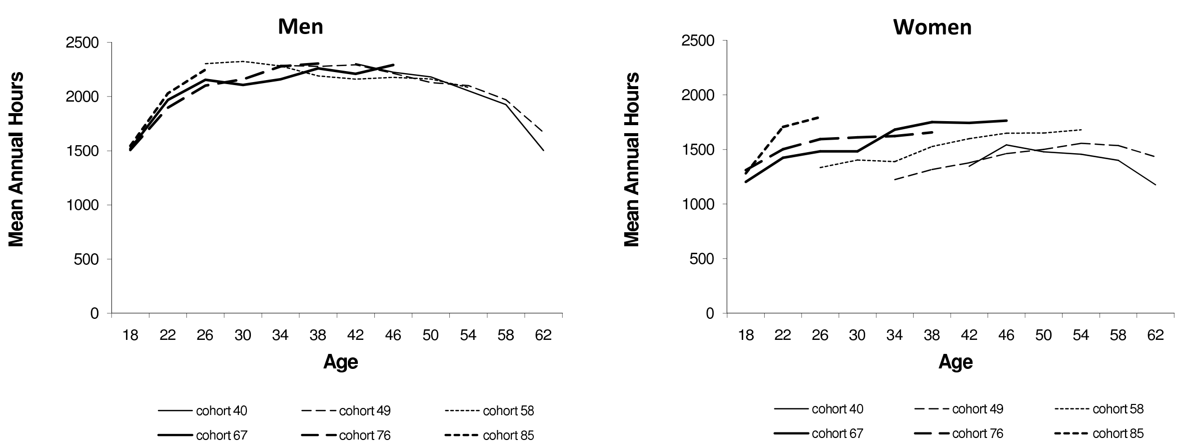


Figure 3: Participation Rate, Fraction of Workers with Positive Annual Hours, 1968-1996, PSID.

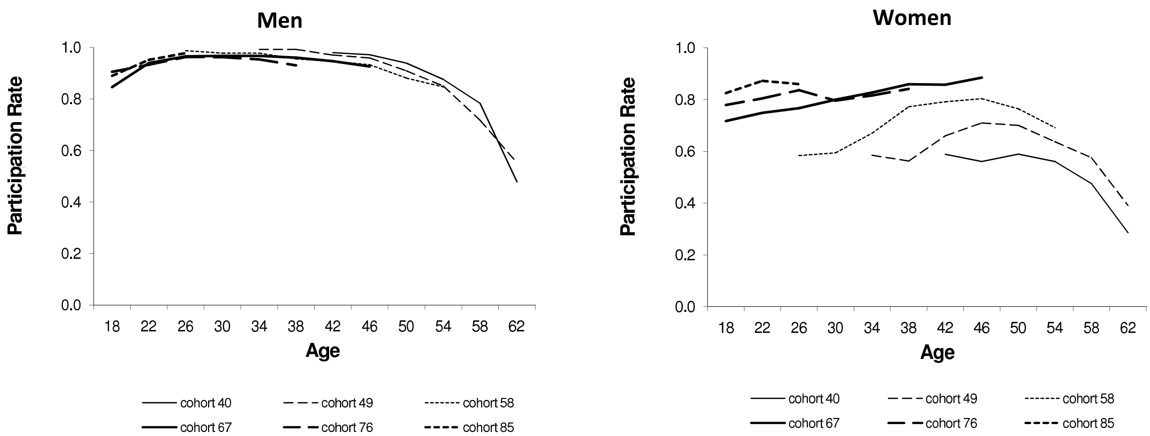


Figure 4: Coefficient of Variation of Annual Hours, 1968-1996, PSID.

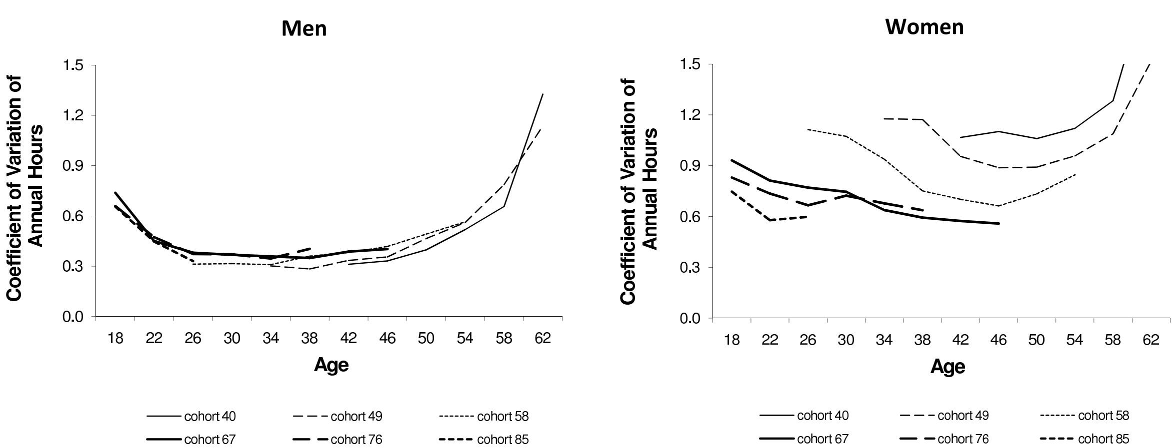
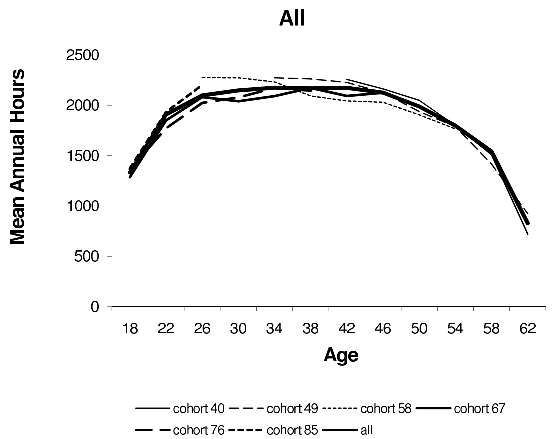


Figure 5: Mean Annual Hours Worked, 1968-1996, PSID, Men, All and by Education.



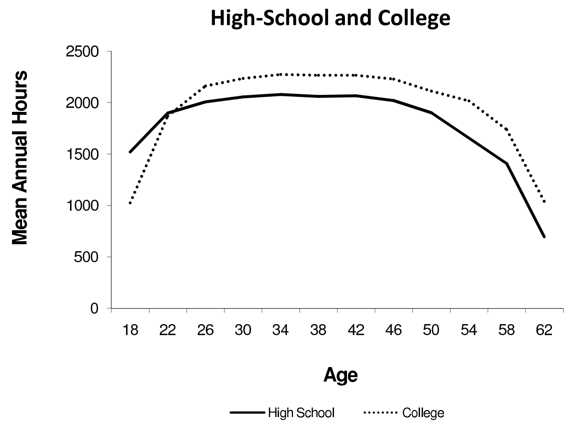
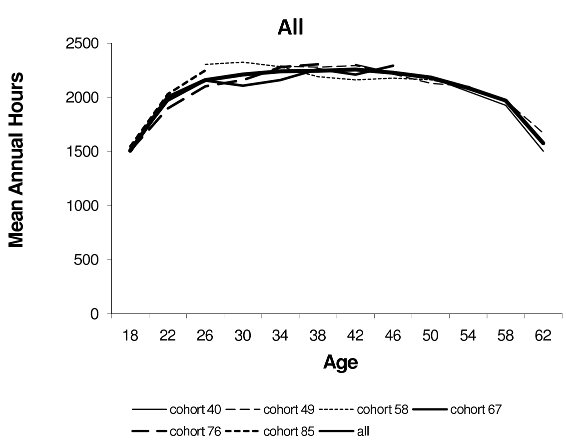


Figure 6: Mean Annual Hours Worked, Workers with Positive Hours, 1968-1996, PSID Men, All and by Education.



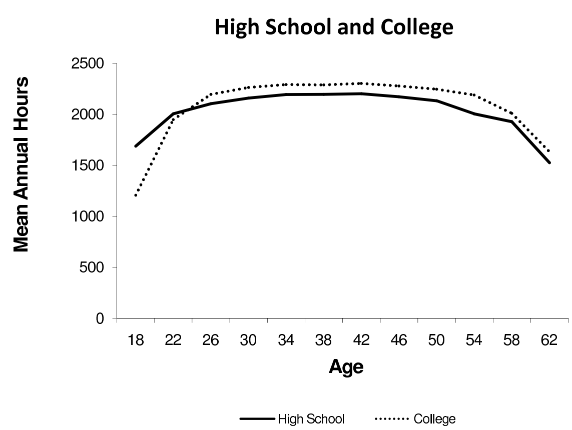


Figure 7: Participation Rate, Fraction of Workers with Positive Annual Hours, 1968-1996, PSID Men, All and by Education.

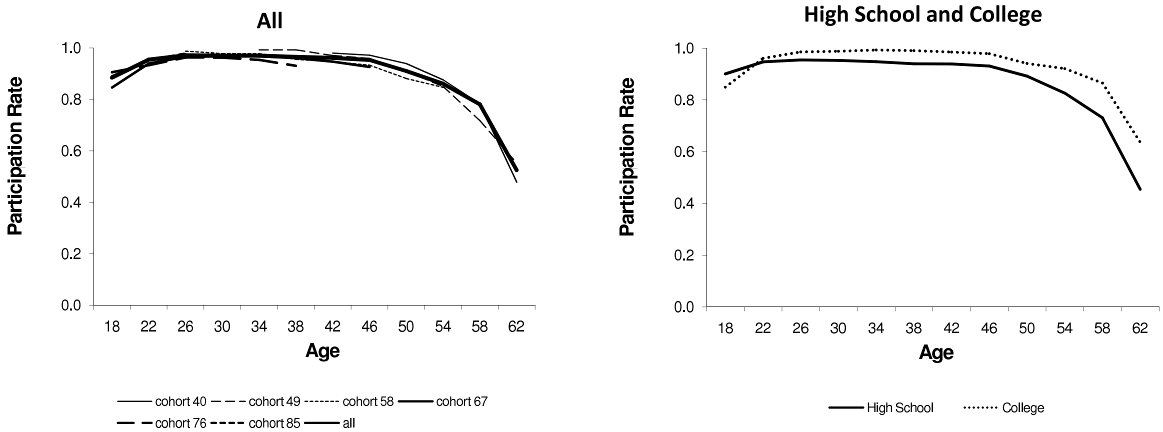


Figure 8: Coefficient of Variation of Annual Hours, 1968-1996, PSID Men, All and by Education.

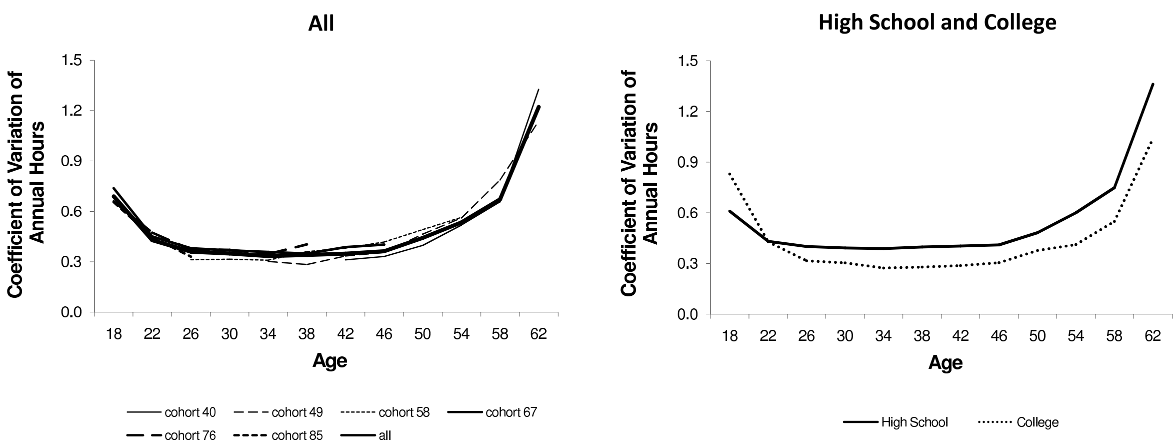


Figure 9: Persistence of Annual Hours, 1968-1996, PSID, Men.

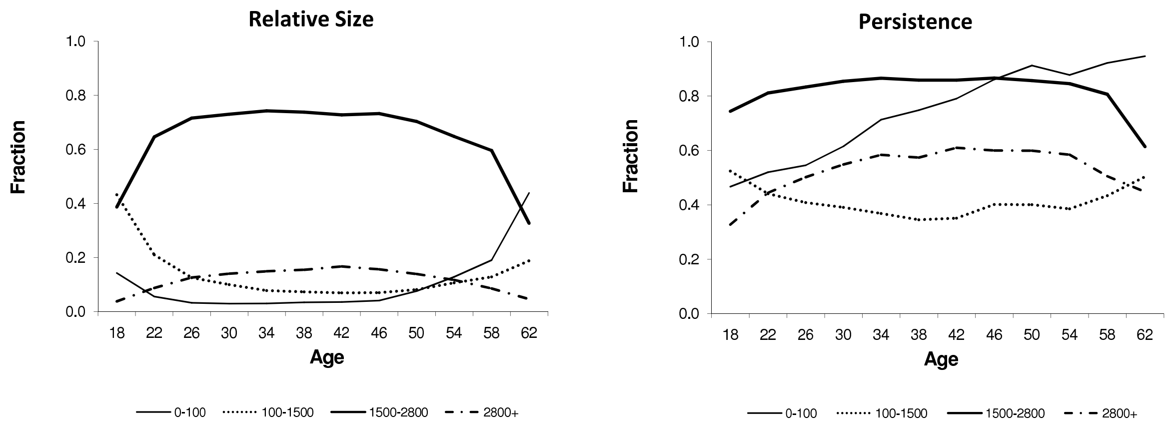


Figure 10: Persistence of Annual Hours, 1968-1996, PSID, Men, High-School.

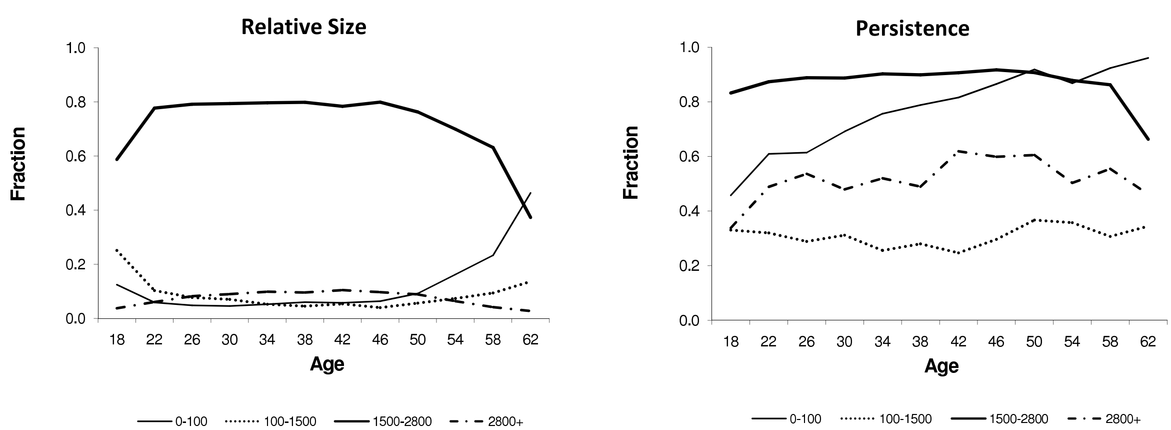


Figure 11: Persistence of Annual Hours, 1968-1996, PSID, Men, College.

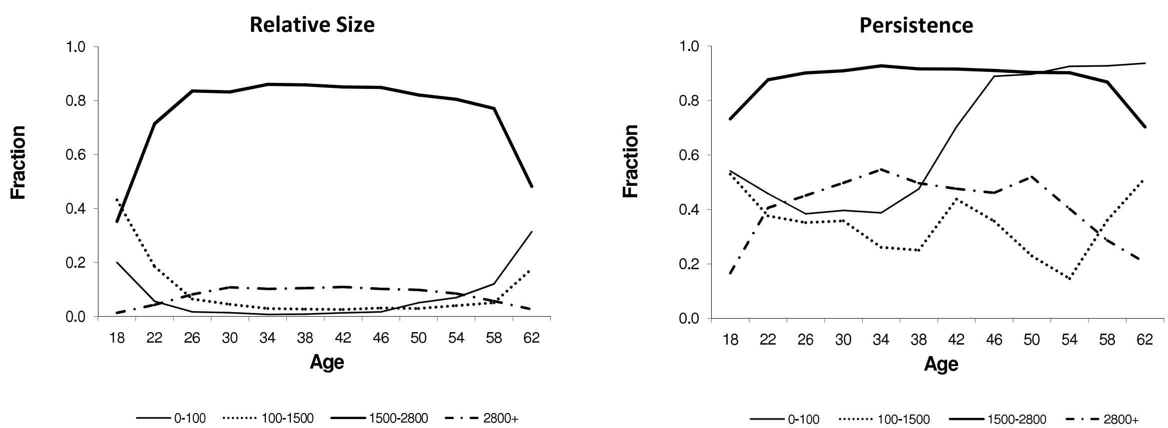
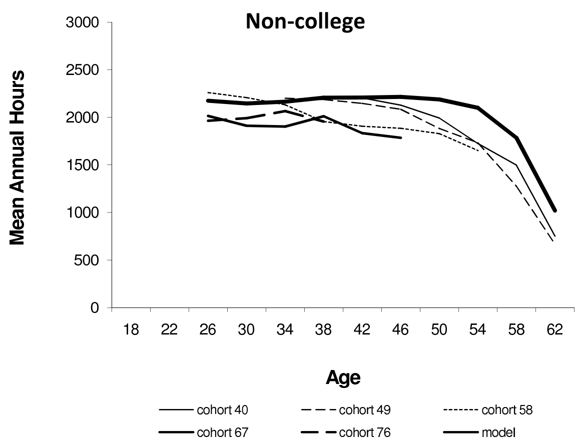


Figure 12: Mean Annual Hours Worked, Model vs. Data.



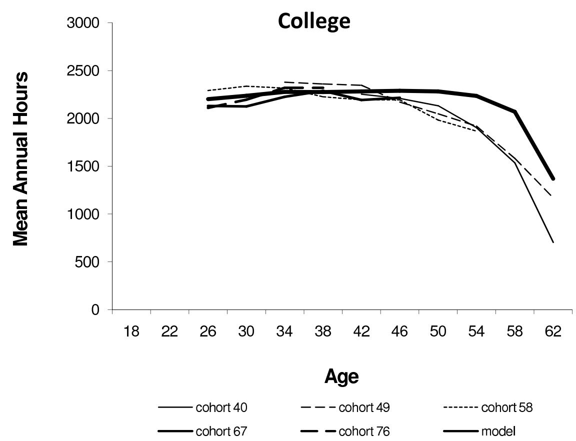
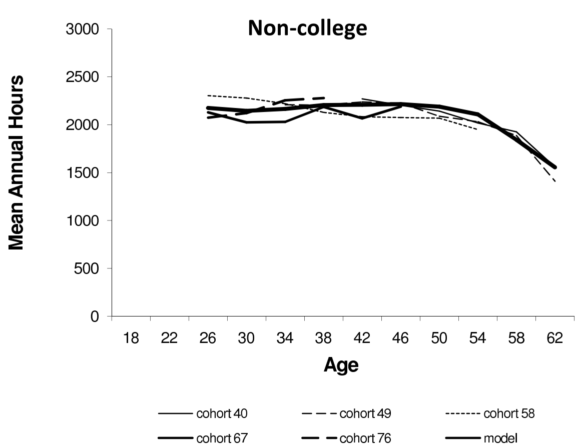


Figure 13: Mean Annual Hours Worked, Workers with Positive Hours, Model vs. Data.



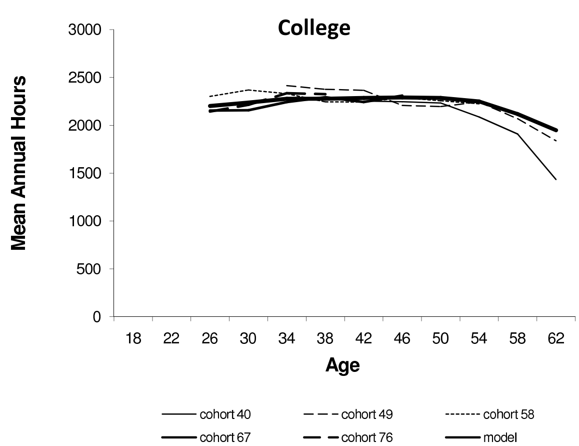


Figure 14: Coefficient of Variation of Annual Hours, Model vs. Data.

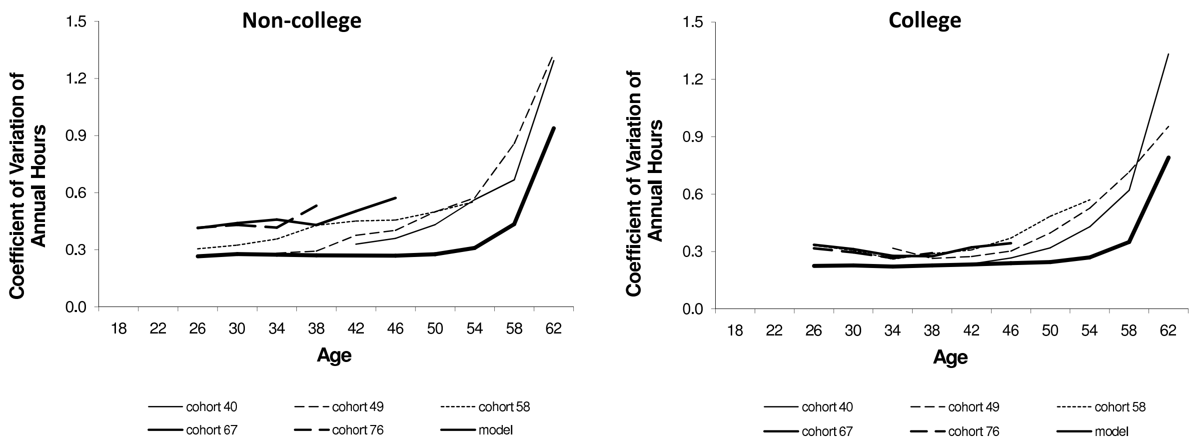


Figure 15: Annual Hours Groups, Model vs. Data.

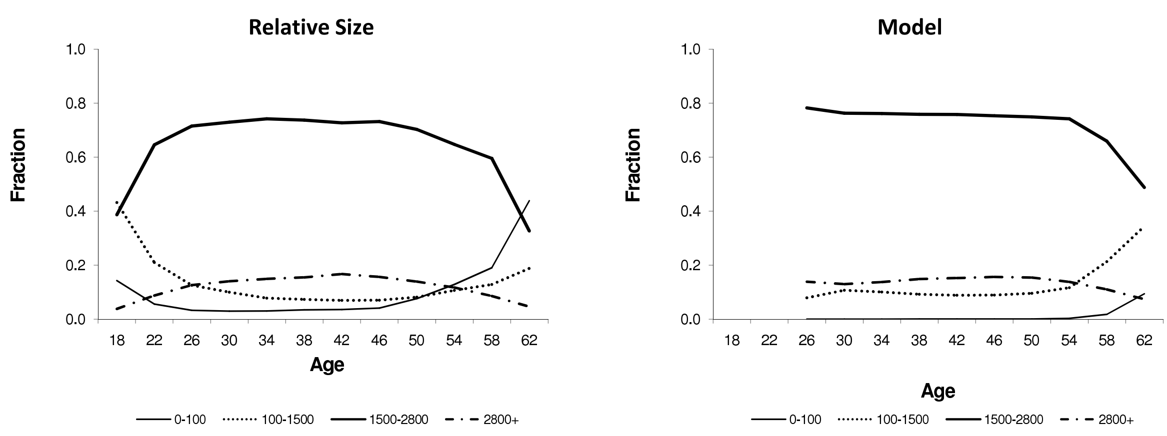
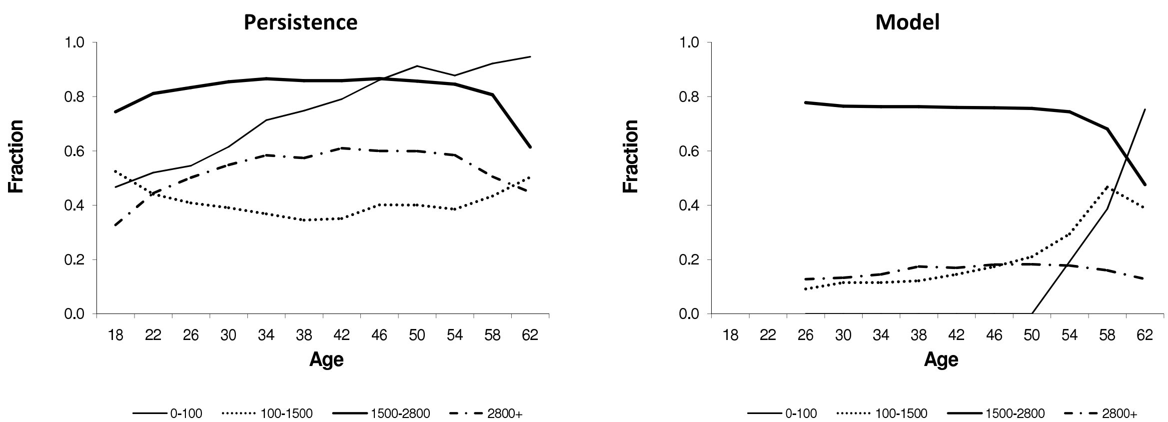


Figure 16: Persistence of Annual Hours, Model vs. Data.



References

Aaronson, D., and E. French (2004):.

Altonji, A. (1986): "Intertemporal Elasticity of Substitution in Labor Supply: Evidence from Micro Data," *Journal of Political Economy*, 94(3), S176–S215.

Badel, A., and M. Huggett (2007): "Interpreting Life-Cycle Inequality Patterns as an Efficient Allocation: Mission Impossible?," mimeo, Georgetown University.

Bils, M., and J.-O. Cho (1994): "Cyclical Factor Utilization," *Journal of Monetary Economics*, 33, 319–354.

Castaneda, A., J. Díaz-Giménez, and J.-V. Ríos-Rull (2003): "Accounting for the US Earnings and Wealth Inequality," *Journal of Political Economy*, 111(4), 818–57.

Chang, Y., and S.-B. Kim (2006): "From Individual to Aggregate Labor Supply: A Quantitative Analysis Based on a Heterogeneous Agent Macroeconomy," *International Economic Review*, 47(1), 1–27.

Cho, J.-O., and T. F. Cooley (1994): "Employment and Hours over the Business Cycle," *Journal of Economic Dynamics and Control*, 18, 411–432.

Conesa, J. C., S. Kitao, and D. Krueger (2009): "Taxing Capital? Not a Bad Idea after All," *American Economic Review*, 99(1), 25–48.

Domeij, D., and M. Flodén (2006): "The Labor-Supply Elasticity and Borrowing Constraints: Why Estimates are Biased," *Review of Economic Dynamics*, 9, 242–262.

Domeij, D., and J. Heathcote (2004): "On the Distributional Effects of Reducing Capital Taxes," *International Economic Review*, 45(2), 523–554.

French, E. (2005): "The Effects of Health, Wealth, and Wages on Labor Supply and Retirement Behavior," *Review of Economic Studies*, 72(2), 395–427.

Heathcote, J., K. Storesletten, and G. L. Violante (2004): "The Cross-Sectional Implications of Rising Wage Inequality in the United States," mimeo, New York University.

Hornstein, A., and E. C. Prescott (1993): "The Firm and the Plant in General Equilibrium Theory," in *General Equilibrium, Growth, and Trade II: The Legacy of Lionel*

McKenzie, ed. by R. Becker, M. Boldrin, R. Jones, and W. Thomson. San Diego: Academic Press.

Kaplan, G. (2007): "Inequality and the Life-Cycle," mimeo, New York University.

Kaplan, G., and G. Violante (2008): "How Much Insurance in Bewley Models?," mimeo, New York University.

Kydland, F. E., and E. C. Prescott (1982): "Time to Build and Aggregate Fluctuations," *Econometrica*, 50(6), 1345–70.

——— (1991): "Hours and Employment Variation in Business Cycle Theory," *Economic Theory*, 1, 63–81.

McGrattan, E., and R. Rogerson (1998): "Changes in Hours Worked Since 1950," *Federal Reserve Bank of Minneapolis Quarterly Review*, 22(1), 2–19.

——— (2004): "Changes in Hours Worked, 1950-2000," *Federal Reserve Bank of Minneapolis Quarterly Review*, 28(1), 14–33.

Osuna, V., and J.-V. Ríos-Rull (2003): "Implementing the 35 Hour Workweek by Means of Overtime Taxation," *Review of Economic Dynamics*, 6(1), 179–206.

Pijoan-Mas, J. (2006): "Precautionary Savings or Working Longer Hours?," *Review of Economic Dynamics*, 9, 326–352.

Prescott, E. (1986): "Theory ahead of Business Cycle Measurement," *Carnegie-Rochester Conference Series on Public Policy*, 25, 11–44.

——— (2004): "Why do Americans Work much more than Europeans?," *Quarterly Review Federal Reserve Bank of Minneapolis*, July, 2–13.

Rogerson, R., and J. Wallenius (2006): "Micro and Macro Elasticities in a Life Cycle Model with Taxes," mimeo, Arizona State University.

Storesletten, K., C. Telmer, and A. Yaron (2001): "How Important Are Idiosyncratic Shocks? Evidence from Labor Supply," *American Economic Review Papers and Proceedings*, 91, 413–417.

APPENDICES

I The PSID Dataset: Variable Description.