A Micro-founded Theory of Aggregate Labor Supply with Non-linear Wages[*]

Andrés Erosa[†]

IMDEA Social Sciences Institute

Luisa Fuster[‡]

IMDEA Social Sciences Institute

Gueorgui Kambourov[§]

University of Toronto, NBER, and RCEA

PRELIMINARY AND INCOMPLETE

This version: September 29, 2010

Abstract

In this paper we first document various facts about the labor supply decisions of male workers in the US over their life cycle. We then build a neoclassical model of labor markets with non-linear wages and heterogeneous agents that accounts for labor supply choices at the extensive margin. The key feature of our theory for delivering periods of non-participation is the non-linear mapping between hours of work and earnings, which is convex at low hours of work. We show that our heterogeneous agent model with non-linear wages can go a long way in capturing salient features of labor supply over the life-cycle for male workers. Moreover, in a quantitative experiment, we find that the aggregate response of labor supply to a one time unanticipated wage shock is

much larger than predicted by the (theoretical) Frisch-elasticity of labor supply. We show that non-linear wages play a crucial role in generating this result.

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

Keywords: Labor Supply, Heterogeneous Agents, Life Cycle.

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 model a labor supply elasticity that is much larger than the one estimated by empirical studies based on micro level data.[1] The goal of this paper is to develop a micro-founded theory of aggregate labor supply. We build a neoclassical life-cycle model of labor markets with non-linear wages and heterogeneous agents that accounts for labor supply choices at the extensive margin. The key feature of our theory for delivering periods of non-participation is the non-linear mapping between hours of work and earnings, which is convex at low hours of work. This mapping is the competitive equilibrium outcome of an economy with a production technology in which hours of work and number of workers are imperfects substitutes (see Hornstein and Prescott (1993) and Osuna and Ríos-Rull (2003)). We show that a heterogeneous agent model with non-linear wages can go along way in capturing salient features of labor supply over the life-cycle for male workers. Moreover, in a quantitative experiment, we find that the aggregate response of labor supply to a one time unanticipated wage shock is much larger than predicted by the (theoretical) Frisch-elasticity of labor supply. We show that non-linear wages play a crucial role in generating this result.

   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 Panel Study of Income Dynamics (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. We use data from the Survey of Income and Program Participation (SIPP) to document the operativeness of the extensive margin in subannual periods (quadramesters or four-month periods). Then, we develop a theory of individual labor supply that extends the neoclassical growth model in order to account for both variation in labor supply along the intensive and extensive margins. The theory models life-cycle behavior to better relate the model predictions to the data. Heterogeneity, within cohorts, is introduced by assuming that individuals are subject to uninsured labor productivity risk.

   The calibration of the model economy involves three key tasks. First, we pin down the

---

[1]Kydland and Prescott (1982) and Prescott (1986) find that aggregate labor supply is very responsive to business cycle shocks and Prescott (2004) makes the case for a large response to a change in taxes. However, MaCurdy (1981) and Altonji (1986) estimate low labor supply elasticities in the micro data.

parameter determining how hours of work affect labor productivity in the model economy using estimates by Aaronson and French (2004). While we need to calibrate a *quadramesterly* stochastic process on labor productivity, the PSID only reports data at an annual frequency and that wages are observed only for those who work. In a second step, we deal with these difficulties by using an indirect inference approach that consists in iterating in the following procedure: (i) feed a quadramesterly labor productivity process into the model, (ii) simulate the model economy and aggregate the data to an annual period, (iii) estimate the wage process in the annual model data and compare to estimates in the PSID data. We then feed a new wage process until the "same" annual wage process is obtained both in the model and in the data. The third task is to take a stand on measurement error in hours in the PSID data. We propose a novel approach to estimate measurement error that consists in comparing both in the model and in the data the variance of transitory wages in two alternative specifications of the wage process. The first specification regresses observed wages while the second specification regresses wages net of the effect of hours of work on wages estimated by Aaronson and French (2004). Identification comes from the fact that measurement error in hours has a different effect in the variance of transitory wages in the two specifications of the regression.

We calibrate alternative economies that differ in the preference parameter σ determining the intertemporal elasticity of substitution of leisure given by $1/\sigma$. While the calibration did not explicitly target the facts on labor supply, we find that when σ is between 2 and 2.5 the model captures most of the salient features about labor supply in the data. Because the economy with σ equal to 2 gives the best fit of the data, we set this economy as our baseline.[2] We find that measurement error in hours is quantitatively important and varies with the education and age of individuals (over the life cycle, it is on average slightly less than 3 percent for non-college individuals and about 2 percent for college individuals).

Next, we show that in our baseline model economy 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"

---

[2]Relative to the data, the calibrated economy with σ equal to 2.5 predicts a too flat age profile of labor hours, a too high transitory variation in wages, and the wrong sign for the covariance and correlation between changes in log hours and log wages. The economy with σ equal to 2 performs better in all these dimensions.

elasticity of labor supply. Third, we compute a "macro" labor supply elasticity by assessing the aggregate labor supply response to a one period (quadramester) unanticipated wage change. The results from these experiments could not be more striking: While the empirical elasticity of labor supply is 0.36 (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 two and four (with values of .61 and 1.27, respectively).

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 60% of the aggregate labor supply response to the temporary wage change and that non-linear wages play a crucial role in generating the large labor supply response along the extensive margin. This result is supported by the empirical evidence in Kimmel and Kniesner (1998): Using SIPP data on males, these authors estimate fixed effects labor supply models separating the extensive and intensive margins. They find that the extensive margin accounts for almost 70% of the wage elasticity of labor supply, with an estimate of the wage elasticity of aggregate labor supply of 1.25.

We find that time aggregation, together with an operative extensive margin at the quadramester 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 quadramester, there are no traces of these low wage shocks in the annual wage data (since the low wage shocks in a quadramester are unobserved in the annual wage data when individuals do not work). To assess the role of time aggregation, we construct in our simulated data an annual measure of the returns to work that uses information on quadramesterly level productivity (where it is important that we *sum the logs* of quadramesterly productivity rather than use *log of the sum* of quadramesterly productivity) and we find that a regression on annual level data now recovers the macro elasticity.

Our paper contributes to a literature that considers extensions of the neoclassical model of the labor market to account for choices at the extensive margin. 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 in a model of indivisible labor, 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. More recently, Rogerson and Wallenius (2009) develop a theory of non-linear wages that provides some qualitative insights showing that the aggregate labor supply response to permanent tax changes is unrelated to the theoretical elasticity of labor supply implied by preference parameters. Domeij and Flodén (2006) and Pijoan-Mas (2006) study labor supply decisions in an infinitely-lived framework with incomplete markets.[3]

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 standard econometric estimates of labor supply elasticities are biased downwards when agents accumulate human capital. Domeij and Flodén (2006)) make a similar point in the context of a model with liquidity constraints. Our paper shows that time aggregation combined with an active extensive margin can also bias downwards empirical estimates of the labor supply elasticity. Relative to Imai and Keane (2004) and Domeij and Flodén (2006)), a distinctive feature of our theory is that temporary wage changes can lead to labor supply responses that are much larger than implied by the intertemporal elasticity of substitution parameter in preferences. Moreover, this implication of the theory is supported by the evidence in Kimmel and Kniesner (1998) (see also Heckman (1993) and Bils, Chang, and Kim (2009)).

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

---

[3]French (2005).

documented on labor supply and evaluates aggregate labor supply responses to temporary and permanent wage changes.

2 Empirical Facts

2.1 The Data

We use the Michigan Panel Study of Income Dynamics (PSID) for the period 1968-1997 in order to compute all annual statistics.[4] 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.[5] Appendix I provides a detailed description of the construction of the dataset 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, the age-18 group on the graphs includes individuals between the age of 18 and 21, while the 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.[6] We use PSID sample weights in the analysis.

As the data suggests, cohort effects are not significant in the case of men. Figures A-1-A-3 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

---

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

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

[6]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.

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, we do not take out cohort effects for men.[7]

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, and the dispersion of annual hours.

- 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.

- The labor supply behavior of high school and college graduates is different enough to warrant a separate analysis for each of these groups.[8]

2.2   Facts on the Life-Cycle Labor Supply of Men

2.2.1   Average Annual Hours over the Life-Cycle

Figure 1 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 2 and 3 illustrate the intensive and extensive margins at the annual level of the labor supply of men over the life-cycle. Between ages 30 to 46 working hours are quite

---

[7]Furthermore, as it will become clear later, the analysis will not directly target the facts on labor supply.
[8]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 that the current partition is a sensible one.

constant and average annual hours are about 2,200 for non-college and 2,300 for college graduates. The extensive margin at the annual level 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 4 displays the dispersion of annual hours over the life-cycle as measured by the coefficient of variation of annual hours. This figure illustrates 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 coefficient of variation of hours is between 0.25 and 0.90. Indeed, the inequality in hours worked is comparable to the inequality of wages or consumption observed 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 life-cycle 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.[9]

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.[10] We then construct transition matrices where cell ij indicates the fraction of all

---

[9]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.

[10]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

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 A-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 5 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 A-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 A-1 further shows that, between 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 A-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

---

documented here.

men.

2.2.4 Lifetime Labor Supply

The dispersion in lifetime labor supply is another useful statistic which is closely related to the persistence in an individual's labor supply over time. Due to the nature of the PSID dataset, we do not observe individuals throughout all their life − some of them have already been in the labor market for some time when the survey starts in 1968 while those who enter the labor market in 1968 at the age of 18 are only in their 40s in 1997. Nevertheless, we can learn a lot even if we follow individuals for shorter periods. We choose to follow individuals for periods of 10 years at different stages in their life-cycle: ages 25-34, 35-44, 45-54, and 55-65. We drop all individuals who have a missing observation during the relevant ten years and sum the hours worked for each individual during the whole ten years. Then we compute the dispersion in this cumulative measure of hours worked. Considering two extreme examples is useful for illustrating how to interpret the results. Consider a particular group, e.g. the group between the ages of 35 and 44, and suppose that all individuals work the same number of hours throughout the whole period as at the beginning at the age of 35. In that case, the coefficient of variation of the cumulative hours worked throughout the whole period would be the same as the coefficient of variation (cross-sectionally) at the age of 35 (or any other age in the period). Alternative, suppose that individual hours fluctuate a lot over the period and those who work a lot in one year work very little the year after that. In that case, workers would end up working quite similar cumulative hours over the period, and the coefficient of variation of the cumulative hours worked throughout the whole period would be quite small and substantially lower than the coefficient of variation (cross-sectionally) at the age of 35 (or any other age in the period).

Table 1 reports the coefficient of variation of the cumulative hours worked for the four age groups defined above. The results indicate that hours are quite persistent, especially over the ages of 25 and 55. This analysis provides us with two important findings. First, the dispersion in cumulative hours is quite substantial, indicating that individuals tend to be quite persistent in their labor supply behavior. This is consistent with the mobility matrices discussed in section 2.2.3. Second, the dispersion of cumulative hours is smaller than the cross-sectional dispersion at any age in the 25-65 interval. This implies that workers do sometimes change their hours worked. This is also consistent with the mobility matrices discussed in section 2.2.3 since − as seen in the middle panel of Table A-1 for those between the ages of 30 and 54 − the diagonal elements of the mobility matrices are not zero and we

do observe workers who switch across the hours categories.

2.3 The Survey of Income and Program Participation (SIPP)

We also use data from the Survey of Income and Program Participation (SIPP). The SIPP interviews individuals three times a year (rather than once a year as in the PSID) and allows us to compute other labor market statistics of interest at a lower frequency, such as a quadramester (a 4-month period). We use the 1990 SIPP Panel. Figure 6 shows the distribution of hours within a quadramester, for a few age-education groups. Figure 7 graphs the fraction of individuals working three quadramesters in a year.

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 balanced 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 cost, 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

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

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 $\varepsilon > \theta$ the wage rate increases 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 implies 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 denotes 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 quadramester, for computational simplicity we compute average quadramesterly earnings over the 35×3 highest quadramesterly earnings as follows

$$\overline{w}_{j+1} \;=\; \overline{w}_j + zw(h_j)/(35 \times 3) \text{ for } j \leq 35 \times 3, \tag{1}$$

$$\overline{w}_{j+1} \;=\; \overline{w}_j + \max\{0, (\min\{zw(h_j), \widehat{y}\} - \overline{w}_j)/(35 \times 3)\} \text{ for } j > 35 \times 3, \tag{2}$$

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 compact way setting:

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

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 $(\overline{W})$.[11] The social security benefit is given by

$$b(\overline{w}) = \begin{cases} 0.90 \times \overline{w} \text{ for } \overline{w} < .2\overline{W}, \\ 0.90 \ \times 0.2\overline{W} + 0.33 \times (\ \overline{w} - 0.2\overline{W}) \text{ for } \overline{w} \in (.2\overline{W}, 1.24\overline{W}], \\ 0.90 \ \times 0.2\overline{W} + 0.33 \times (1.24\overline{W} - 0.2\overline{W}) + 0.15 \times (\ \overline{w} - 1.24\overline{W}) \text{ for } \overline{w} > 1.24\overline{W}, \end{cases} \tag{4}$$

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.

---

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

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

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

The value of a person that has not retired is

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

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.

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 through an indirect inference approach. In the estimation, moreover, we exploit that our theory of non-linear wages provides a very natural way to identify measurement error in hours in the data. In order to model the variation in employment within a year, the calibration sets the model period to 1 quadramester (a 4-month period). This choice allows us to use employment data, such as the fraction of individuals working all three quadramesters in a year, from the Survey of Income and Participation Program (SIPP), which interviews individuals three times in a

16

year (rather than once a year as in the PSID). The SIPP, however, is longitudinally fairly short to allow us to estimate the stochastic process for wages. As a result, we use the PSID for this purpose. In calibrating a *quadramesterly* 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 follow an indirect inference approach (see Smith (1990), Gourieroux, Monfort, and Renault (1993), and Guvenen and Smith (2010)):

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 *quadramesterly* labor productivity process into the model economy.

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

5. Aggregate the *quadramesterly* 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 *quadramesterly* 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 to one quadramester (4 months). The model economy is solved in partial equilibrium for a fixed interest rate. The quadramesterly interest rate is chosen so

that the implied annual rate of return on capital (net of depreciation) is 4%.[12]

Preference parameters, time endowments, and mortality rates. Following Kaplan and Violante (2008), the discount factor 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 Osuna and Ríos-Rull (2003) and Prescott (2004), 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 annually about 42% of their available time. The curvature parameter on leisure (σ) is set at 2 implying a Frisch-elasticity of leisure of -0.5. In the sensitivity analysis we also calibrate an economy with σ equal to 2.5. 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

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

Note that the elasticity of the wage rate to a change in hours of work is given by $\frac{\varepsilon}{\theta} - 1$. 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 $\tau_h = .27$. 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 ($\overline{W}$).[13]

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:

---

[12]The depreciation over a yearly period is assumed 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.

[13]Actually, $\overline{W}$ is set at 80% of average earnings in the economy. The reason is that our model only includes 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.

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

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 2. The empirical findings show that the variance of fixed effects is quite large for both education types, with values of .097 and .072. for the non-college and the college types. Both wage processes exhibit high autocorrelation, with a value of .94 for non-college individuals and .97 for college individuals. The variance of the innovation of the autoregressive process is .019 and .021 respectively. The estimates reveal that both education types exhibit transitory shocks to wages with quite high variances as reported in Figure 11.

To calibrate the model economy, we need to find a *quadramesterly* 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 quadramesterly transitory shock.[14] Specifically, while the transitory shock is drawn every quadramester, the persistent shock is only drawn at the first quadramester 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 of wages seems implausibly large (see Figure 12).

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

---

[14]We have tried an specification that allows for an autoregressive process at the quadramester 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 quadramester level, there is no reason to expect the logarithm of the sum of quadramesterly earnings to be well approximated with an autoregressive process.

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 all three quadramesters in a year: The larger σ_T^2 the less likely individuals will work during all periods in a year. We thus use this statistic as a calibration target where the fraction of individuals working 3 quadramesters in a year is taken from the Survey of Income and Participation Program (SIPP).[15] Figure 7 shows that this fraction is roughly constant for prime-age males (age 30 to 50) but that it decreases substantially after age 50. To mimic the data in a simple way, for each education group the process for transitory shocks is parameterized with two values σ_{T50}^2 and σ_{T64}^2, where the variance of transitory shocks is assumed to be equal to σ_{T50}^2 up to age 50 and that it then increases linearly up to the value σ_{T64}^2.[16]

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 comes 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

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

---

[15]We note that the SIPP allows us to have more reliable measures of labor force participation at the quadramesterly frequency than the PSID as it interviews individuals three times in a year (rather than once a year as in the PSID).

[16]The value of σ_{T50}^2 is 0.0179 for non-college and 0.0148 for college while σ_{T64}^2 takes the value 0.0151 for non-college and 0.0097 for college.

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 leads to an increase in the estimated transitory variation of wages. Intuitively, the estimated transitory variation of wages increases 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 transitory variance in wages when the wage data is "cleaned" with hours data. We then introduce the estimates for measurement error in hours into the model economy and run the two specifications of the wage regression with model simulated data.

Reassuringly, as discussed below, 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 estimated in the data.

5 Quantitative Findings

While below we present in detail the calibration results for our baseline economy with $\sigma = 2$, we also calibrate two alternative economies. The first alternative economy was calibrated to the same targets as the baseline economy but setting a value of $\sigma = 2.5$. The second alternative economy has $\sigma = 2$ but differs from the baseline in that it features linear wages and fixed costs of work. To keep homotheticity, the fixed cost of work is formulated in terms of time rather than in terms of goods or utility. Since the baseline economy best matches the facts, the analysis below mostly focuses on this economy.

5.1 Calibration Results

There are 29 parameters that we calibrate by solving the model economy. Table 3 shows the values and the calibration targets for three of these parameters: the average earnings $\overline{W}$, taste for leisure φ, and discount factor β. For each education group, we use an indirect inference approach to pin down a quartic polynomial for the wage age-profile (5 parameters), a stochastic process of wages (4 parameters giving variance of fixed effects, persistence and variance of innovations, transitory shock), and the variance of measurement error in hours and earnings (2 parameters).

5.1.1 Wages: Age-Profile and Stochastic Process

Because in our baseline economy there is an active extensive margin in labor supply decisions, people who work are a non-random selection of the population. Hence, we can't mechanically plug an age-profile for wages into our model but need to solve for it. Nonetheless, Figure 9 shows that the baseline economy matches almost exactly the age-profile of wages for both education groups.[17] Table 4 reports 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. This table also reports the targeted statistics which are

---

[17]The values of the parameters characterizing the deterministic age-profile of labor productivity are $(-1.85, 0.22, -6.98 * 10^{-2}, 1.04 * 10^{-4}, -6.18 * 10^{-7})$ for non-college and $(-4.21, 0.45, -1.41 * 10^{-2}, 2.04 * 10^{-4}, -1.15 * 10^{-6})$ for college.

the estimated variance of the fixed effect and the parameters of the AR(1) process for log wages (also in Table 2). The values reported under the column Model correspond to the GMM estimation using annual model data for the baseline economy.

As explained in the calibration procedure, the process for transitory shocks is parameterized with two values σ^2_{T50} and σ^2_{T64}, where the variance of transitory shocks is assumed to be equal to σ^2_{T50} up to age 50 and that it then changes linearly to the value σ^2_{T64}. The targets are the fractions of prime-aged males (age 35-50) and males aged 50-64 that work all three quadramesters in a year. The calibration results in σ^2_{T50} equal to 0.0179 for non-college and 0.0148 for college individuals, while σ^2_{T64} takes the value 0.0151 for non-college and 0.0097 for college. Hence, to match the fact that in the data non-college individuals are less likely to work three quadramesters in a year, the calibration implies that non-college individuals are subject to larger true temporary shocks than college individuals. The model replicates reasonably well the fraction of people working 3 quadramesters in a year but for young non-college individuals the calibration tends to overpredict the fraction of individuals working all periods in a year (see Figure 10).

5.1.2 Measurement Error

To estimate measurement error in hours our calibration procedure compares the variance of transitory wages in two alternative estimations of the wage process both in actual and in model data. The first specification estimates the wage process using data on observed wages, the second specification estimates a process for wages net of the effects of hours worked on wages. Regarding the first specification, Figure 11 shows that the model matches well the age-profile for the variance of the transitory component of residual log wages in the data for both education groups. Figure 12 shows that the model also matches well the variance of the transitory component for "clean" wages. Measurement error in hours is obtained as the difference between the variance of transitory wages across the two specifications for the wage process. The results for measurement error in hours are reported in Figure 13. We find that measurement error in hours is higher for non-college than for college individuals. The age-profile of measurement error has a slight U-shaped for non-college individuals: It takes a value of around 3 percent for very young individuals, it is below 3 percent for prime-aged males and it increases to 5 percent when individuals are close to the retirement age. For college individuals, measurement error is about 2 percent for most of the life-cycle, with a mild increase to 4 percent prior to retirement.

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

Figures 14-18 present the performance of the model in accounting for the facts on labor supply documented in Section 2. 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.

Age-Profile of Hours of Work Figure 14 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 though it does not fully account for the decline in hours late in the life-cycle. The model matches well the decline in working hours among individuals with positive hours of work as shown in Figure 15.

Dispersion in Hours of Work Figure 16 shows the dispersion of annual hours worked over the life-cycle both in the model and in the data. For both education types, the model captures the fact that the dispersion in working hours is flat for prime-age males and that it increases substantially late in the life-cycle. The baseline economy underpredicts the dispersion in hours worked but this should not be surprising as the model abstracts from many factors that could lead to heterogeneity in working hours across individuals. In considering what these factors maybe, it is suggestive that the baseline economy underpredicts the heterogeneity in lifetime labor supply (see Table 5). While in the data the coefficient of variation in lifetime labor supply (labor supply over a ten year period) for individuals aged 35 to 45 and aged 45 to 55 is .26 and .37, this statistic takes values of .11 and .16 in the baseline economy for the two age groups considered. Hence, it seems that the theory abstracts from some factors leading to persistent differences in working hours across individuals in the U.S., such as heterogeneity in health, preferences, or demographics. This observation is supported by the findings of Bils, Chang, and Kim (2009) who model and calibrate permanent differences in tastes for work across individuals to match micro facts in the US economy.

Persistence in Annual Hours Worked 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 17 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 mimics the observation that the size of Group 1 (those working between 0 and 100 hours) in the data is very small for individuals younger than 50 but that it rises substantially late in the life-cycle.

Figure 18 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 mimics the observation 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 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.

Distribution of Hours Figure 19 compares the distribution of hours in a quadramester both in the model and in the data for two education groups and three age groups. The model mimics the facts that the distribution of hours is highly concentrated around 600 hours and has a spike at zero. Moreover, the fraction of individuals working zero hours is highest for non-college individuals and it increases with age. The model underpredicts the fraction of individuals working zero hours.

Change in Log Hours and Log Wages: Covariance and Correlation Table 6 shows that both in the baseline economy and in the data the covariance between changes in log hours and changes in log wages have an inverted U-shape and a negative sign. The baseline economy matches these data remarkably well. Table 7 shows that the baseline economy is also successful in predicting the negative correlation between the change in log hours and the change in log wages in the data, though the quantitative fit is not as good as in the case as with the covariances between changes in log wages and log hours. It is interesting to compare these predictions with the ones for the two alternative calibrated model economies. Recall that the first alternative economy has a value of $\sigma = 2.5$ and the second alternative economy has $\sigma = 2$, linear wages, and an active extensive margin due to fixed costs of work. Two observations are of note: The economy with $\sigma = 2.5$ fits the data worse than the baseline economy. The economy with fixed costs of work predicts the wrong sign for the covariances

and correlations in changes in hours and wages.

5.3 The Frisch-Elasticity of Labor Supply

We now assess how labor supply responds to wage changes in our baseline economy. Following a large literature, we focus on the Frisch-elasticity of labor supply. This elasticity describes how labor supply responds to an intertemporal change of wages that leaves the marginal utility of wealth constant. We use three alternative procedures 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 in terms of the preference parameter σ. We refer to this elasticity as the "theoretical elasticity of labor supply" or the "intensive-margin elasticity of labor supply." The second procedure consists in simulating the model economy to generate cross-sectional data and estimate an "empirical elasticity of labor supply" with standard econometric techniques. The third procedure consists in simulating a one period (quadramester) unanticipated decrease in the aggregate wage and compute the aggregate labor supply response in order to obtain an "aggregate elasticity of labor supply."

Our findings are quite striking. First, we find that the Frisch elasticity of labor supply estimated with conventional econometric techniques is well below the value implied by the theoretical elasticity and well within the range of values estimated in the empirical literature. Thus, conventional econometric techniques applied to model simulated data fail to recover the value of the key preference parameter determining labor supply decisions. While this finding is similar in spirit to the results in Domeij and Flodén (2006) and Imai and Keane (2004), the mechanism underlying our result differs from the one in these papers. Second, we find that the elasticity of labor supply to a temporary-aggregate wage change is much higher than implied by the theoretical elasticity and is consistent with empirical evidence reported in Kimmel and Kniesner (1998). This result follows from modeling non-linear wages and does not hold in the frameworks of Imai and Keane (2004) and Domeij and Flodén (2006).

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, non-linear 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 intensive-margin elasticities of leisure and labor in the baseline economy are: $\eta^l = -.5$ and $\eta^h = .61$.[18]

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.[19] 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 $\sigma = 2.5$ (with all other parameters being recalibrated to match the targets used in the calibration of the baseline economy). The findings are reported in Table 8.

We find that all the estimates of the empirical elasticity of labor supply are well below the value predicted by the intensive-margin 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 .36, well within the range of $[0, 0.5]$ in the empirical literature. Interestingly, the theoretical elasticity is equal to .61, almost twice the value

---

[18]To obtain a value for η^h we compute the median $\frac{(1-h)}{h}$ across individuals who work in the model economy.

[19]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.

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 $\sigma = 3.4$.[20] 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, a point recently emphasized by Rogerson and Wallenius (2009).

5.3.3 The Elasticity of Aggregate Labor Supply

The aggregate elasticity is obtained by computing the aggregate labor supply response to a one period unanticipated decrease in the wage rate of two percent. The change in the aggregate labor supply can be safely interpreted as pure "substitution effect" as the wealth effect of a one period wage change is small. The results from these experiments could not be more striking: The aggregate elasticity of labor supply in our baseline economy is 1.27, which is 3.5 times larger than the empirical elasticity (.36), and twice as big as the theoretical elasticity (0.61). The large labor supply response along the extensive margin explains why the aggregate 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 aggregate elasticity from 1.27 to 0.58. Hence, the extensive margin accounts for about 54% of the aggregate labor supply response to a temporary wage change. This result is supported by the empirical evidence in Kimmel and Kniesner (1998): Using SIPP data on males, these authors estimate fixed effects labor supply models separating the extensive and intensive margins. They find that the extensive margin accounts for almost 70% of the wage elasticity of labor supply, with an estimate of the wage elasticity of labor supply of 1.25 and a wage elasticity of employment of 0.86.

It is interesting to compare our findings to those of Imai and Keane (2004). In a structural estimation exercise on NLSY data that assumes human capital accumulation, Imai and Keane estimate a very large value for the intertemporal elasticity of substitution (i.e.s) of labor - roughly about 4. The key to their large estimate is that the incentives to supply labor in

---

[20]To map the elasticity of leisure into an elasticity of labor, we use that the median value of $(1 - h)/h$ in our model economy is 1.22.

their framework are driven by the sum of the wage rate and the returns to human capital accumulation. When the returns to human capital are large, labor supply responds little to wage changes leading to high estimates of the i.e.s. Notice, however, that this reasoning also implies that labor supply should not be very responsive to aggregate wage shocks. Indeed, Imai and Keane (2004) simulate the effects of a one period change in the wage rate of 2% and find the average labor supply change for individuals aged 30 to 50 is about 1.5%, a much smaller response that the 8% predicted by the estimated i.e.s. of 4. On the contrary, in our paper the labor supply response to a temporary wage change is more than twice the value predicted by the i.e.s parameter. In our framework with non-linear wages, temporary wage changes have a large effect on the extensive margin leading to labor supply responses larger than implied by the i.e.s parameter. Moreover, our findings provide an explanation for the evidence in Kimmel and Kniesner (1998) of a highly wage-elastic extensive margin at subannual periods in the SIPP data.

Domeij and Flodén (2006) argue that standard econometric estimates of the elasticity of labor supply may be biased downwards when there are liquidity constraints. The idea is that individuals who are liquidity constrained may not be able to reduce their labor supply when they are hit by a negative temporary wage shock. Since consumption smoothing can only be achieved by an increase in labor supply, the labor-supply response of liquidity constrained individuals is thus smaller or of the opposite sign that what is predicted by an analysis that ignores such constraints. Note, again, that this reasoning also implies that labor supply should not be very responsive to aggregate wage shocks. In fact, this intuition is confirmed with an experiment where we simulate a one period unanticipated wage change in an economy with linear wages.[21] We find that with linear wages there is no response at the extensive margin and the aggregate elasticity is quite close to the theoretical elasticity.

5.4   Discussion on Non-linear Wages and the Aggregate Elasticity

We have seen that a key implication of our theory is that when wages are non-linear the aggregate elasticity is substantially larger than the theoretical elasticity because the latter only describes labor supply responses along the intensive margin, thereby neglecting labor supply responses along the extensive margin. Hence, the importance of the disconnect between the theoretical and the aggregate elasticities depends on the wage-elasticity of employment. We now examine some key factors affecting labor supply decisions.

---

[21]This economy features incomplete markets and borrowing constraints so that the main difference with the economy considered by Domeij and Flodén (2006) is that we model the life-cycle.

Age Table 9 reports the elasticity of labor supply to a temporary-unanticipated wage change of 2 percent for different age groups. The elasticity of labor supply increases steeply with age: It rises from 1.0 for individuals aged 25-35, to 1.98 for individuals aged 55-64. While the response along the intensive margin is roughly flat over the life-cycle, the wage-elasticity of employment rises from .38 for individuals aged 26-35 to 1.56 for people aged 55-64. The employment of old individuals is very responsive to temporary wage changes because they, on average, have a buffer stock of savings that allows them to smooth consumption well in response to an unanticipated wage shock. On the contrary, young individuals are less responsive in their labor supply because they are poorer and they need to build a stock of precautionary savings to insure income risk over the life-cycle.

Transitory versus Permanent Changes in Wages To evaluate the compensated elasticity to a permanent wage change, we simulate an increase in the labor income tax from .27 in the baseline economy to .37. In this experiment, tax proceeds are rebated with a lump sum transfer to working-age individuals. The amount of the transfer is education specific so that there is no income redistribution across education types. Overall, we find that the elasticities for both the intensive and extensive margins are reduced by a half relative to the case of a temporary wage change (see Table 9). Now the labor supply elasticity is .65 and the employment elasticity is .35, which should be compared to the elasticities of 1.27 and .69 to a temporary wage change. Hence, not surprisingly, individuals respond more strongly to a temporary wage change than to a permanent compensated-wage change. More interestingly, the age-profile of the wage-elasticity to the permanent change in wages has a U-shape. It starts at .70 for young individuals, decreases to .56 for individuals aged 55-64 and it increases to .72 for people aged 55-64. Young individuals respond strongly to the tax increase because the lump sum transfer helps them to smooth consumption. This effect is stronger for college than non-college individuals since college individuals face a steep age-profile of labor productivity. As a result, the aggregate elasticity for the age-group 25-35 is higher for college than for non-college individuals (.81 versus .65).

Intertemporal Elasticity of Substitution The parameter σ affects more importantly labor supply responses when wage changes are temporary rather than permanent. As σ varies from 2 to 2.5, the labor supply elasticity to a temporary wage change decreases from 1.27 to .88 and the elasticity to a permanent wage change decreases from .65 to .51 (see Table 10). The parameter σ has more important effects on labor supply responses along the

extensive margin than the intensive margin both in the case of permanent and temporary wage changes.

Non-Linear Wages vs. Fixed Costs of Work It is interesting to compare labor supply responses in our baseline economy to those of an alternative theory of the extensive margin. We thus evaluate aggregate labor supply responses in a model with linear wages in which the extensive margin is active because of fixed costs of work.[22] We find that the economy with fixed costs of work has a much lower labor supply elasticity both at the intensive and extensive margins. While the labor supply elasticity to a temporary wage change is 1.27 in the baseline economy, it is 0.59 − less than a half − in the economy with fixed costs of work. The elasticity at the extensive margin is reduced by more than a half (from .69 to .26). The response to a permanent compensated-wage change is also much smaller for the economy with fixed costs of work, especially at the extensive margin (from .35 to .18) (see Table 10). In understanding these results note that the incentives to work long hours are much weaker in the economy with fixed costs of work than in the economy with non-linear wages. First, in the economy with fixed costs of work individuals are less willing to work long hours because the concavity of the utility function implies that the marginal utility of leisure decreases more steeply with working hours than in the economy with no fixed costs of work. Second, contrary to the non-linear wage economy, the hourly wage rate does not rise with working hours in the economy with fixed costs of work. As a result, the elasticity of labor supply along the intensive margin is lower in the economy with fixed costs of work than in the economy with non-linear wages (.43 versus .61). To understand the low labor supply response at the extensive margin, note that the more costly is for individuals to work long hours the more costly is for them to take periods off work.[23] Summing up, the economy with fixed time costs of work exhibits low labor supply responses both along the intensive and extensive margins. We conclude that the non-linear wage economy represents

---

[22]Note that in the absence of fixed costs of work, the extensive margin is not active in the presence of linear wages. To keep homotheticity, the fixed cost of work is formulated in terms of time rather than goods or utility. Otherwise, to be consistent with the evidence the fixed costs of work would have to change across individuals in the income distribution, over time, and across countries. This alternative economy is calibrated to the targets used in the calibration of the baseline economy. We find that, in order to match the calibration target for the fraction of people working 3 quadramesters we need the variance of the transitory wage shocks to be roughly 5 times the variance in the baseline economy with non-linear wages. As a result, the calibrated economy with fixed costs of work matches all the calibration targets but the variance of transitory wages, which is higher than in the data.

[23]This also explains why the calibration of the economy with fixed costs of work requires large transitory shocks to wages.

a parsimonious theory of labor supply decisions with an active extensive margin.

Heterogeneity and Non-linear Wages In a model of indivisible labor, Chang and Kim (2006) find that increases in the variance of the cross-sectional distribution of wages can imply a large reduction in the aggregate elasticity of labor supply by reducing the mass of individuals close to their reservation wage. As discussed by Rogerson and Wallenius (2010), this effect is less prominent in a non-linear wage economy as a lower aggregate response on the extensive margin can be counterbalanced by a higher response at the intensive margin. Moreover, we note that the homotheticity of our baseline economy implies that an increase in the variance of fixed effects (permanent heterogeneity in wages) should not have consequences for the elasticity of labor supply along the extensive margin. Finally, we emphasize that aggregation in our model is disciplined by a host of micro facts.

5.5 Discussion on Time Aggregation and the Low Empirical Elasticity

We have shown that the extensive margin plays a crucial role in generating the high value of the aggregate elasticity of labor supply relative to the theoretical elasticity. We now consider an experiment that shows that time aggregation may play an important role in generating a low empirical elasticity of labor supply.

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 within the year 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.

The wage rate in annual survey data is measured as $ln(w) = ln\frac{Annual Earnings}{Annual Hours}$. 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 which biases estimates of the empirical elasticity of labor obtained with annual data. To this end, we run a regression (10) with annual model data assuming away measurement error and using the quarterly sum of realized labor productivity as an explanatory variable ($ln \sum z$ instead of ln w). 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). Table 11 shows that the annual elasticity of leisure increases from $-.37$ to $-.52$ 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 $-.52$ when using $ln \sum z$, it is $-.57$ 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 productivities $\sum(ln(z))$. 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 $-.88$, which fully recovers the aggregate elasticity of leisure obtained in the temporary wage experiment.

6   Conclusion

Table 1: The Coefficient of Variation in Lifetime and Cross-sectional Hours, PSID.

Age	Lifetime	Cross-sectional
25-34	0.27	0.36
35-44	0.26	0.34
45-54	0.37	0.41
55-65	0.64	0.86

Note - Authors' calculations from the PSID.

Table 2: PSID: Stochastic Process of Hourly Wages.

	Non-college	College
$\mathrm{Var}(\alpha)$	0.097	0.072
ρ	0.940	0.977
$\mathrm{Var}(\eta)$	0.019	0.021

Table 3: Parameters on Preferences and Pension Formula

Parameter	Value	Variable	Target	Result
$\overline{W}$	3.3	Ratio $\overline{W}$ to Males average earnings	.80	.82
φ	1.0	Fraction of hours worked	42%	42%
β	0.9815	Asset to income ratio	2.5	2.6

Table 4: Calibration of the Fixed Effect and Persistent Shock

Non-College

Parameter	Value	Variable	Target	Model
σ_α	0.283	Variance of fixed component of log wages	0.10	0.09
ρ	0.937	Correlation coefficient AR component of log wages	0.94	0.94
σ_η	0.138	Variance of innovation AR component of log wages	0.02	0.02

College

Parameter	Value	Variable	Target	Model
σ_α	0.236	Variance of fixed component of log wages	0.07	0.07
ρ	0.972	Correlation coefficient AR component of log wages	0.97	0.98
σ_η	0.117	Variance of innovation AR component of log wages	0.02	0.02

Table 5: The Coefficient of Variation in Lifetime and Cross-sectional Hours, Data vs. Model.

Age	Cross-sectional		Lifetime	
	Data	Model	Data	Model
25-35	0.36	0.20	0.27	0.08
35-45	0.34	0.22	0.26	0.11
45-55	0.41	0.25	0.37	0.15
55-65	0.86	0.47	0.64	0.36

Table 6: The Covariance Between the Change in Log Hours and the Change in Log Wages, by Age Groups.

Age	25-34	35-44	45-54	55-65
Data	-0.04	-0.03	-0.03	-0.04
$\sigma = 2.0$: Baseline	-0.02	-0.01	-0.01	-0.03
$\sigma = 2.0$: Fixed cost	0.01	0.03	0.02	0.01
$\sigma = 2.5$	0.00	0.02	0.01	0.00

Table 7: The Correlation Between the Change in Log Hours and the Change in Log Wages, by Age Groups.

Age	25-34	35-44	45-54	55-65
Data	-0.26	-0.27	-0.24	-0.24
$\sigma = 2.0$: Baseline	-0.13	-0.06	-0.06	-0.10
$\sigma = 2.0$: Fixed cost	0.07	0.15	0.11	0.02
$\sigma = 2.5$	0.01	0.11	0.07	0.00

Table 8: The Elasticity of Labor Supply

	Theoretical	OLS	IV1 $(\Delta \ln w_{t-1})$	IV2 $(age, age^2, \ln w_{t-2})$
$\sigma = 2.0$: Baseline	0.61	-0.07	-0.08	0.36
$\sigma = 2.0$: Fixed cost	0.44	0.09	0.13	-0.07
$\sigma = 2.5$	0.49	0.06	0.09	0.15

Table 9: Aggregate Labor Supply Elasticity by Age: Baseline Economy

	Temporary wage decrease			Permanent wage decrease		
Elasticities:	Total	Employment	Intensive	Total	Employment	Intensive
26-35	1.0	0.38	0.62	0.70	0.26	0.43
35-45	1.17	0.53	0.64	0.62	0.31	0.21
45-55	1.35	0.72	0.63	0.56	0.31	0.25
55-64	1.98	1.56	0.42	0.72	0.53	0.18
All	1.27	0.69	0.58	0.65	0.35	0.29

Table 10: Aggregate Labor Supply Elasticity:

Elasticities:	Temporary wage decrease			Permanent wage decrease		
	Total	Employment	Intensive	Total	Employment	Intensive
$\sigma = 2.0$: Baseline	1.27	0.69	0.58	0.65	0.35	0.29
$\sigma = 2.0$: Fixed cost	0.59	0.26	0.33	0.45	0.18	0.27
$\sigma = 2.5$	0.88	0.42	0.46	0.51	0.21	0.30

Table 11: Elasticity of Leisure: Time Aggregation, No Measurement Error.

	Theoretical	Annual	$Ln(\sum z)$	$\sum(ln(z))$	IV $\sum(ln(z))$	Macro
$\sigma = 2.0$: Baseline	-0.50	-0.37	-0.52	-0.57	-0.88	-0.85
$\sigma = 2.0$: Fixed cost	-0.50	-0.27	-0.34	-0.39	-0.46	-0.43
$\sigma = 2.5$	-0.40	-0.30	-0.45	-0.50	-0.63	-0.60

Figure 1: Mean Annual Hours Worked, 1968-1996, PSID, Men, All and by Education.
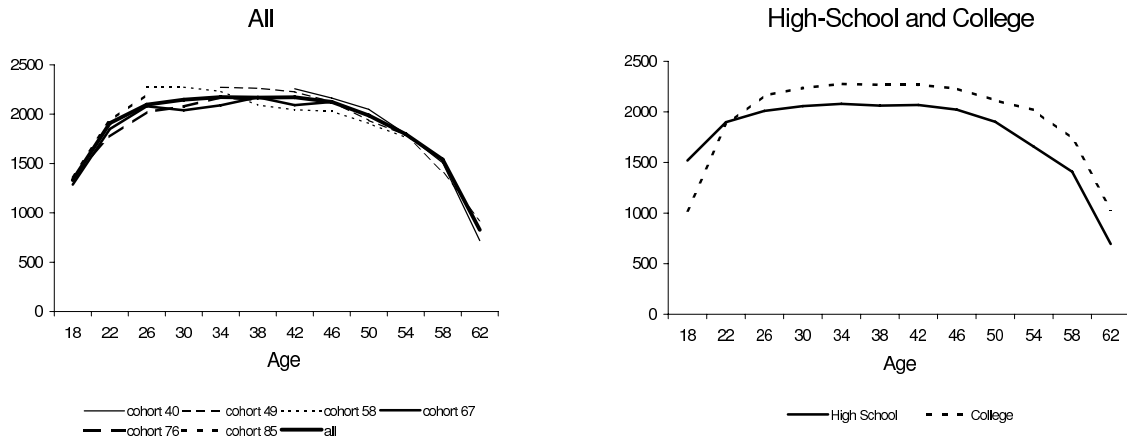
Figure 2: Mean Annual Hours Worked, Workers with Positive Hours, 1968-1996, PSID Men, All and by Education.
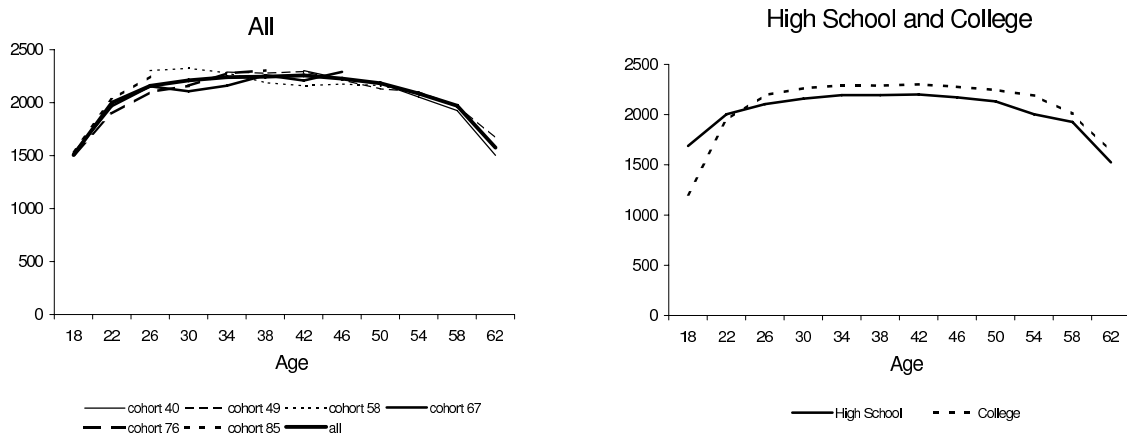
Figure 3: Participation Rate, Fraction of Workers with Positive Annual Hours, 1968-1996, PSID Men, All and by Education.
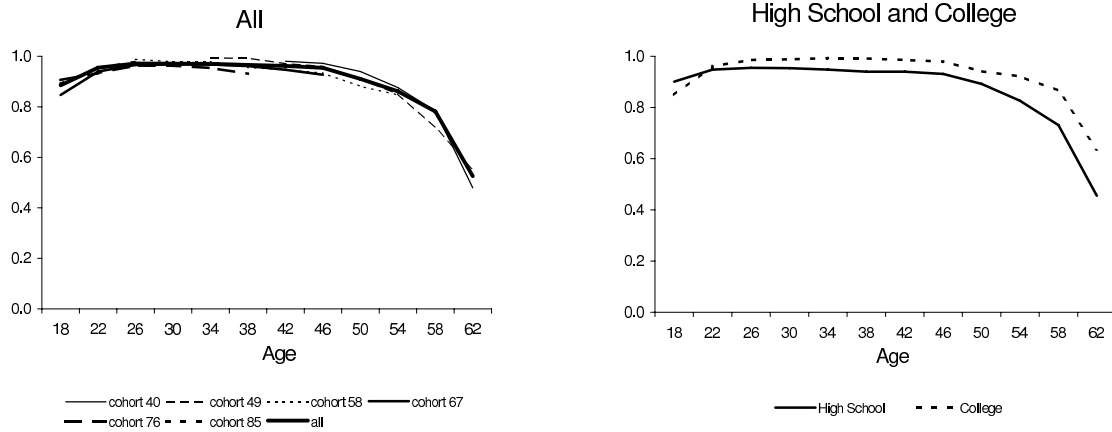
Figure 4: Coefficient of Variation of Annual Hours, 1968-1996, PSID Men, All and by Education.
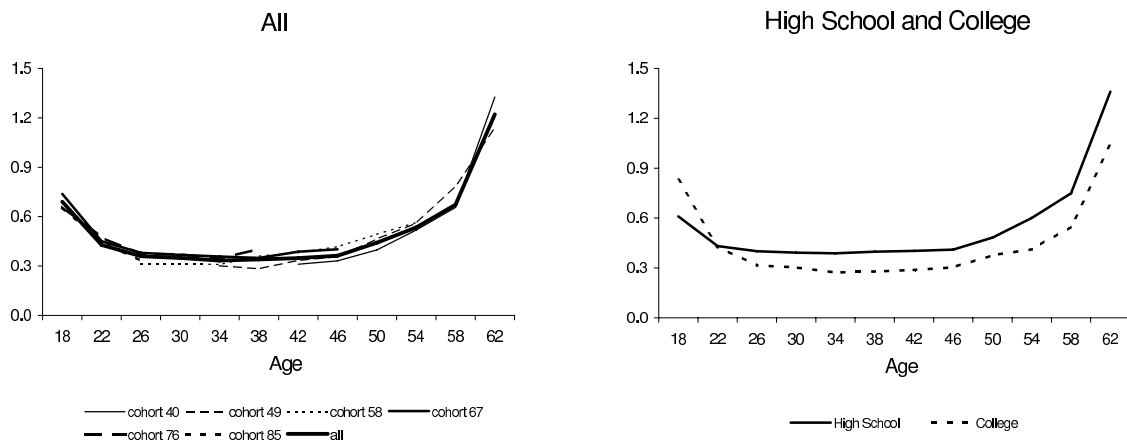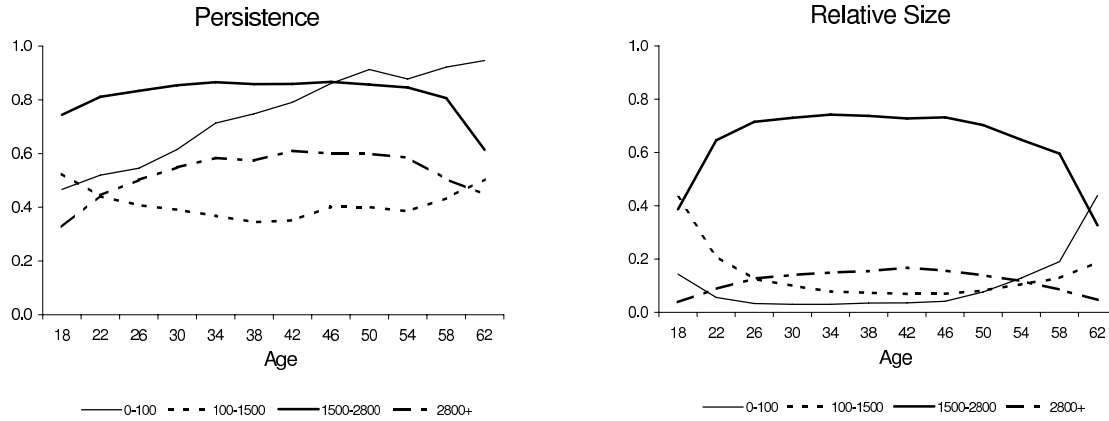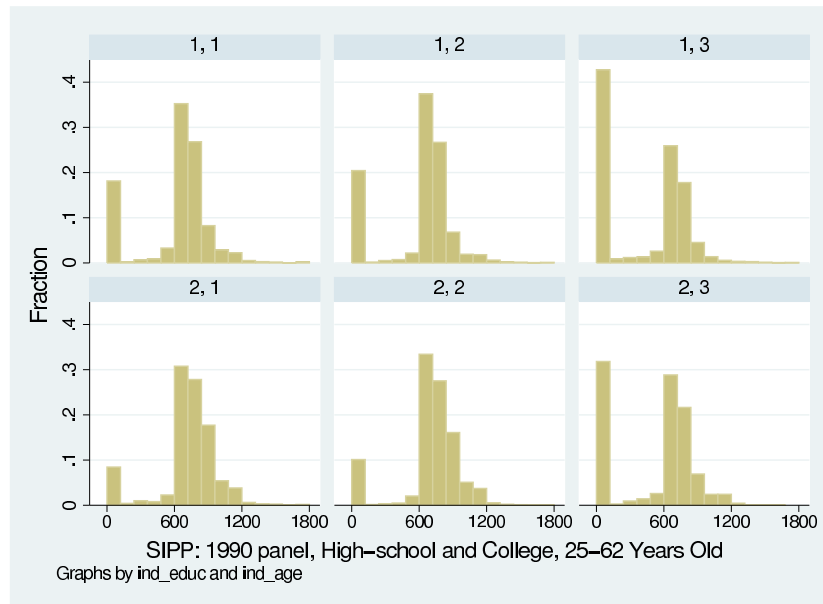
Figure 5: Persistence of Annual Hours, 1968-1996, PSID, Men.

Figure 6: The Distribution of Hours within a 4-month Period, SIPP: 1990 Panel.

Notes: The graphs show the distribution of hours in a 4-month period for age and education groups. Group ij represents a particular age (i) and education (j) group. i: 1. 25-40, 2. 41-55, 3. 56-61; j: 1. non-college, 2. college.

Figure 7: The Fraction of Individuals Working Three Quadramesters in a Year, by Education, SIPP: 1990 Panel.
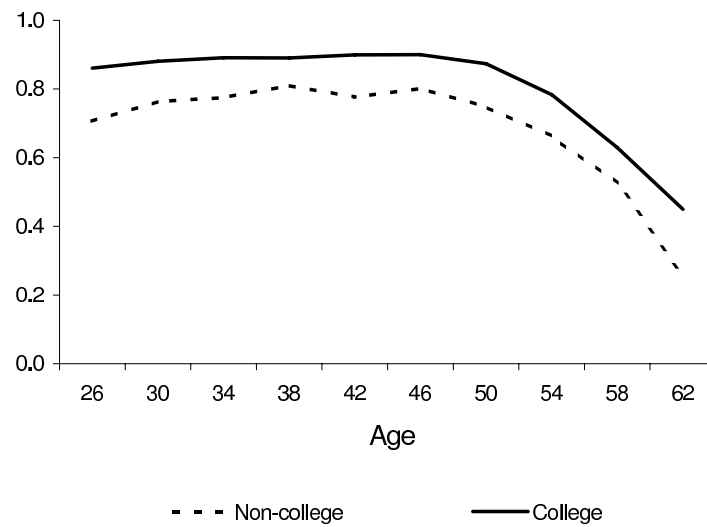
Figure 8: Mean Annual Hours, PSID, Men: 30-45 Years Old, by Education and Fixed Effects.
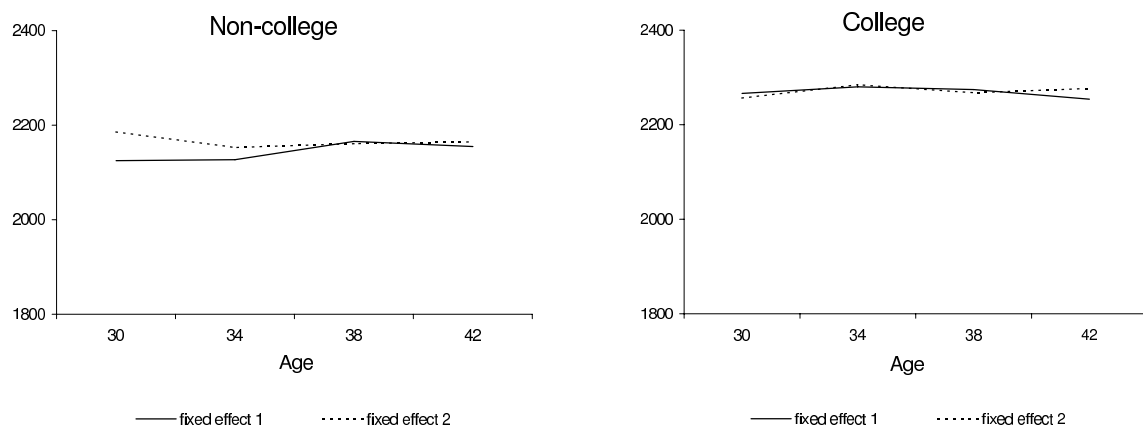
Figure 9: The Life-cycle Deterministic Component of Wages, by Education, Model vs. Data.
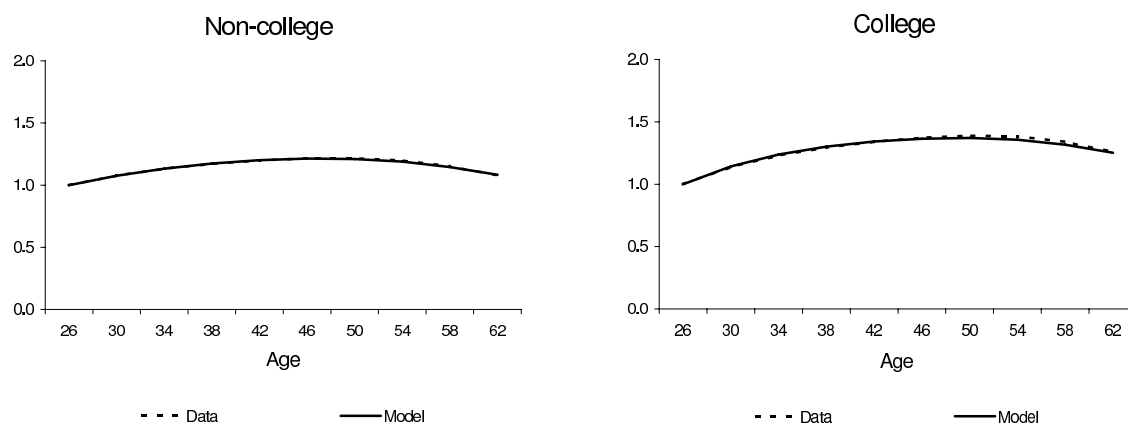
Note: On the graphs, by education, wages have been normalized to 1 for the first age group.

Figure 10: The Fraction of Individuals Working Three Quadramesters in a Year, by Education, Model vs. Data.
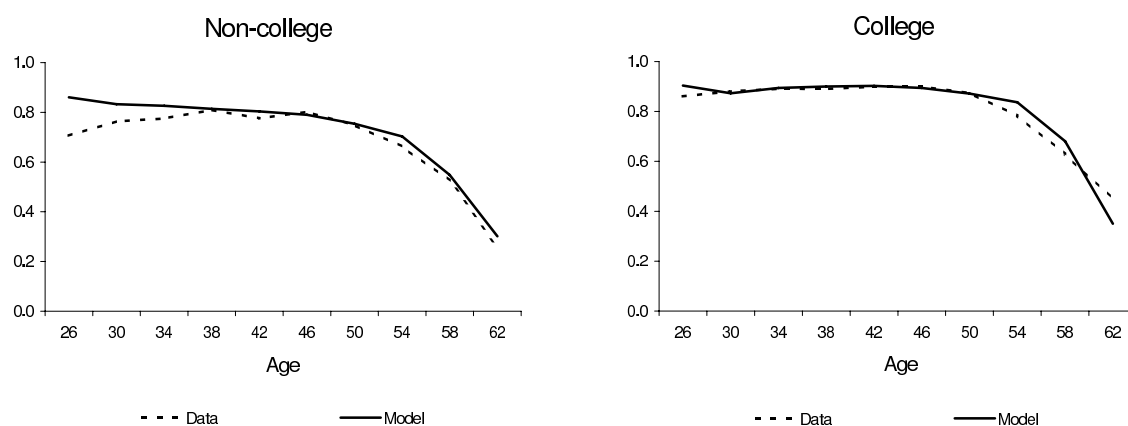
Figure 11: Variance of the Transitory Component of Residual Log Wages, by Education, Model vs. Data.
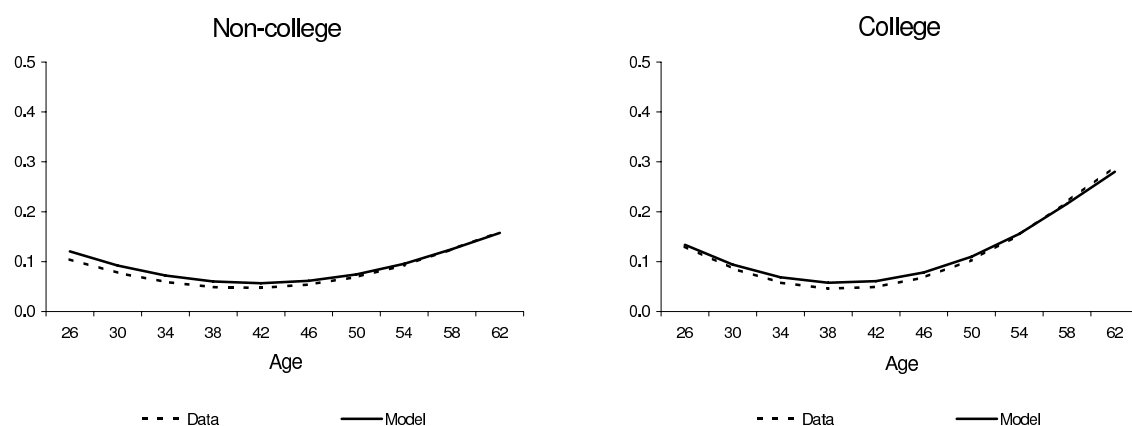
Non-college

College

Age

Age

- - - Data ——— Model

- - - Data ——— Model

Figure 12: Variance of the Transitory Component of Residual Log Wages, Net of the Effect of Hours Worked, by Education, Model vs. Data.

Non-college
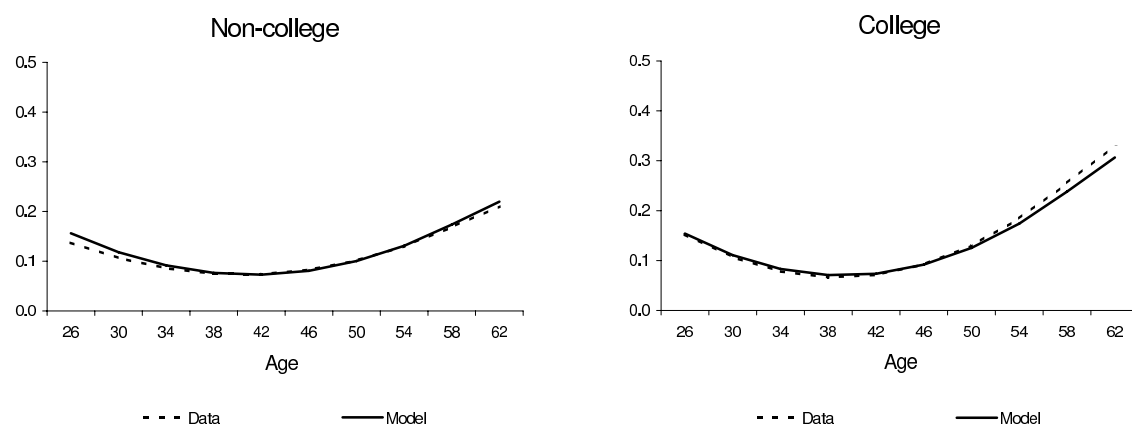
College

Age

Age

- - - Data ——— Model

- - - Data ——— Model

Figure 13: The Variance of Measurement Error in Hours Worked, by Education, Model vs. Data.
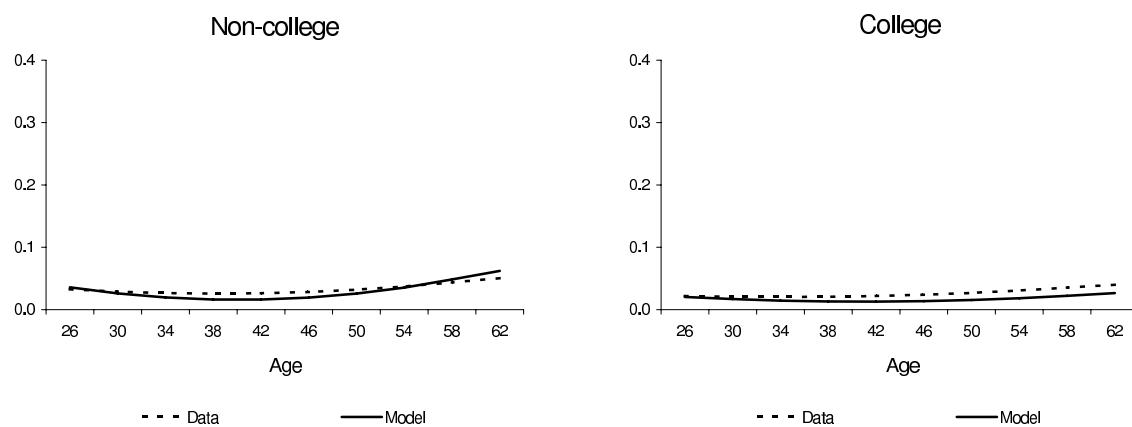
Non-college

College

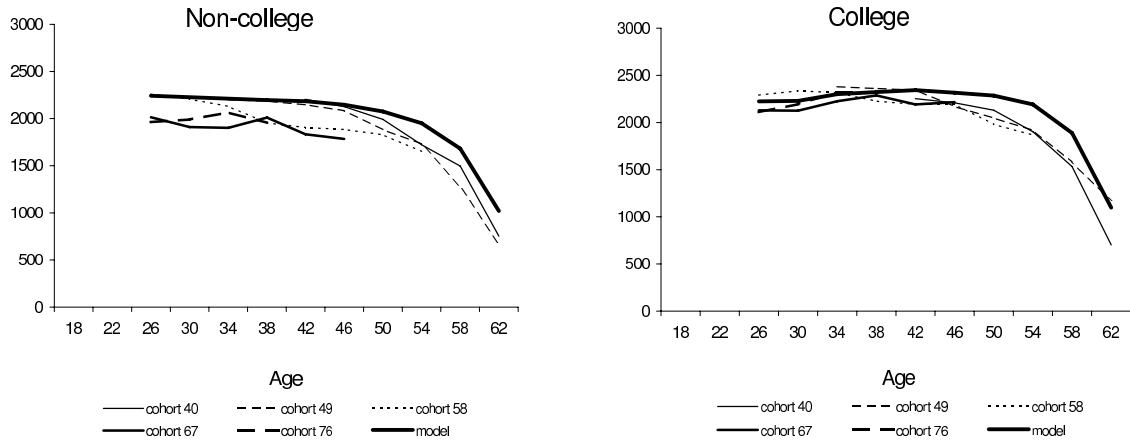Figure 14: Mean Annual Hours Worked, Model vs. Data.

Figure 15: Mean Annual Hours Worked, Workers with Positive Hours, Model vs. Data.
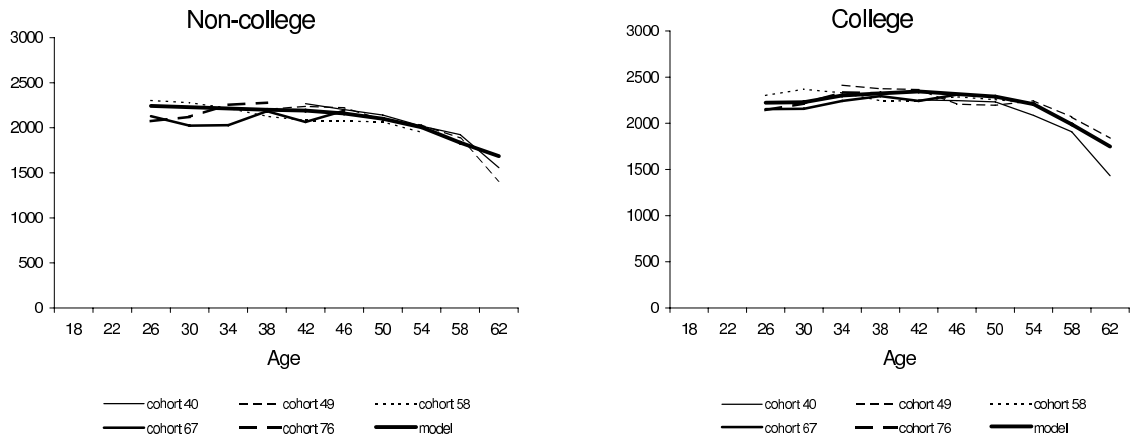
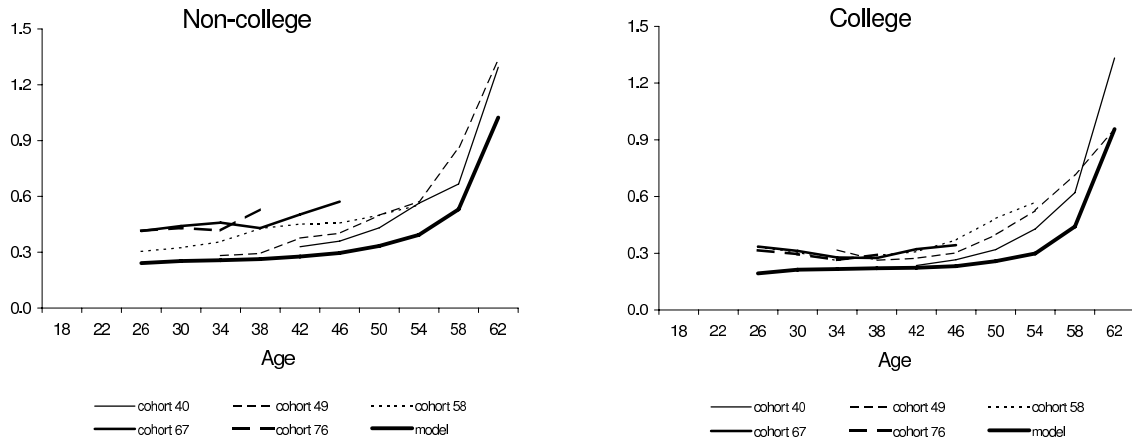Figure 16: Coefficient of Variation of Annual Hours, Model vs. Data.

Figure 17: Annual Hours Groups, Model vs. Data.

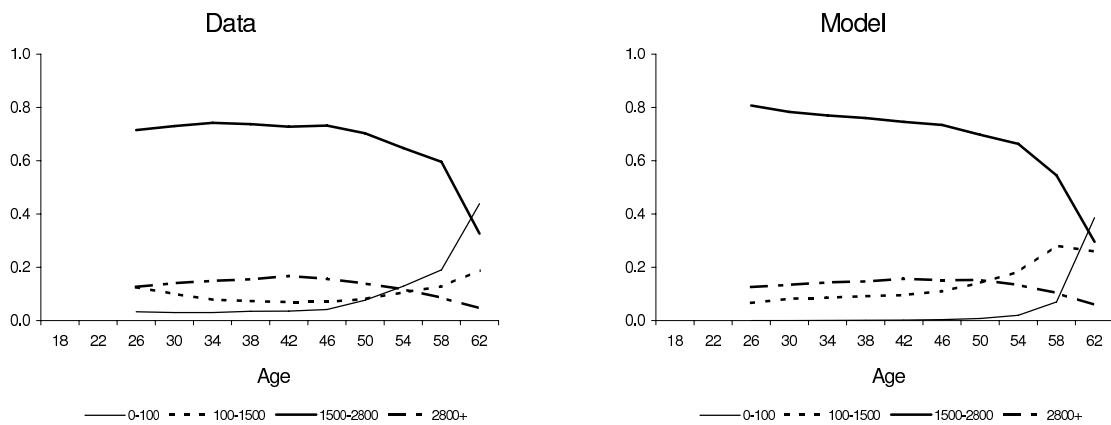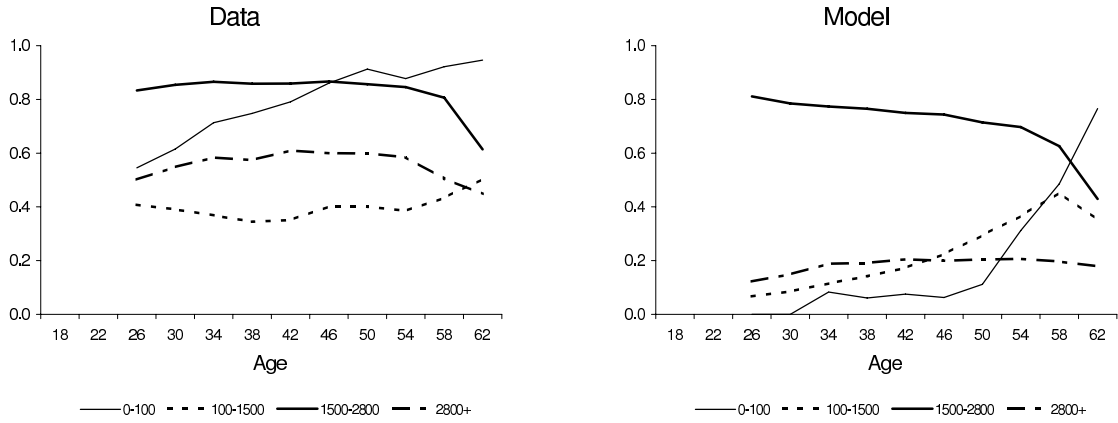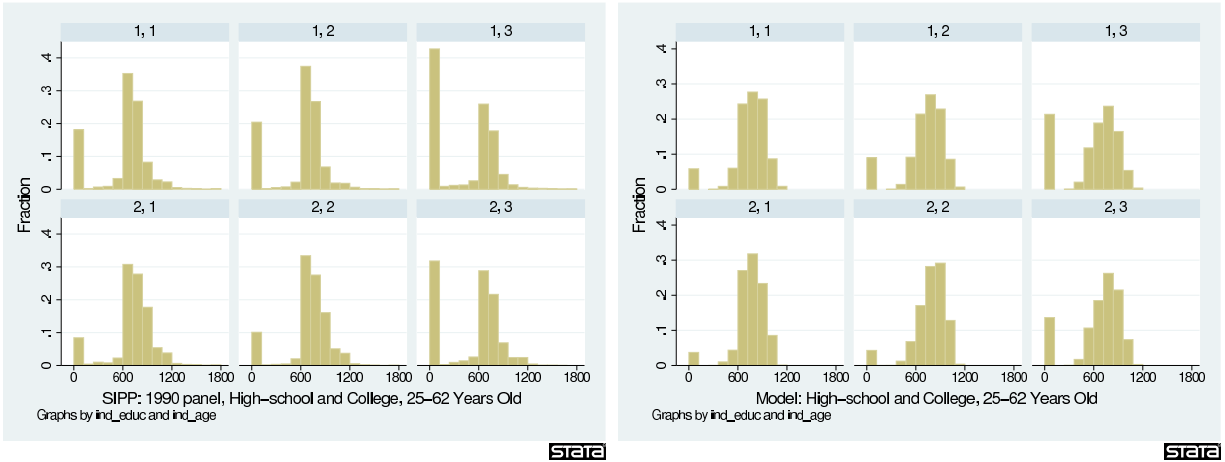Figure 18: Persistence of Annual Hours, Model vs. Data.

Figure 19: The Distribution of Hours within a 4-month Period, Model vs. Data.

Notes: The graphs show the distribution of hours in a 4-month period for age and education groups, in the data and in the model. Group ij represents a particular age (i) and education (j) group. i: 1. 25-40, 2. 41-55, 3. 56-61; j: 1. non-college, 2. college.

References

AARONSON, D., AND E. FRENCH (2004): "The Effect of Part Time Work on Wages: Evidence from Social Security Rules," *Journal of Labor Economics*, 22(2), 329–252.

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.

BHATTACHARYA, J., AND D. LAKDAWALLA (2006): "Does Medicare Benefit the Poor?," *Journal of Public Economics*, 90(1-2), 277–292.

BILS, M., Y. CHANG, AND S.-B. KIM (2009): "Comparative Advantage and Unemployment," *RCER Working Papers, University of Rochester*, 547.

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.

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.

GOURIEROUX, C., A. MONFORT, AND E. RENAULT (1993): "Indirect Inference," *Journal of Applied Econometrics*, 8(S), S85–112.

GUVENEN, F., AND A. SMITH (2010): "Inferring Labor Income Risk from Economic Choices: An Indirect Inference Approach," *Staff Report Federal Reserve Bank of Minneapolis*, 450.

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.

HECKMAN, J. J. (1993): "What Has Been Learned about Labor Supply in the Past Twenty Years?," *American Economic Review*, 83(2), 116–121.

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.

IMAI, S., AND M. P. KEANE (2004): "Intertemporal Labor Supply and Human Capital Accumulation," *International Economic Review*, 45(2), 601–641.

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.

KIMMEL, J., AND T. J. KNIESNER (1998): "New Evidence on Labor Supply: Employment versus Hours Elasticities by Sex and Marital Status," *Journal of Monetary Economics*, 42, 289–301.

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

MACURDY, T. E. (1981): "An Empirical Model of Labor Supply in a Life-Cycle Setting," *Journal of Political Economy*, 89(6), 1059–1085.

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. (1988): "Indivisible Labor, Lotteries and Equilibrium," *Journal of Monetary Economics*, 21, 3–16.

Rogerson, R., and J. Wallenius (2009): "Micro and Macro Elasticities in a Life Cycle Model with Taxes," *Journal of Economic Theory*, forthcoming.

Smith, A. (1990): "Three Essays on the Solution and Estimation of Dynamic Macroeconomic Models," *PhD Dissertation, Duke 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.

I   The PSID Dataset: Description.

II   The SIPP Dataset: Description.

III Tables.

Table A-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 A-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 A-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.

IV Figures.

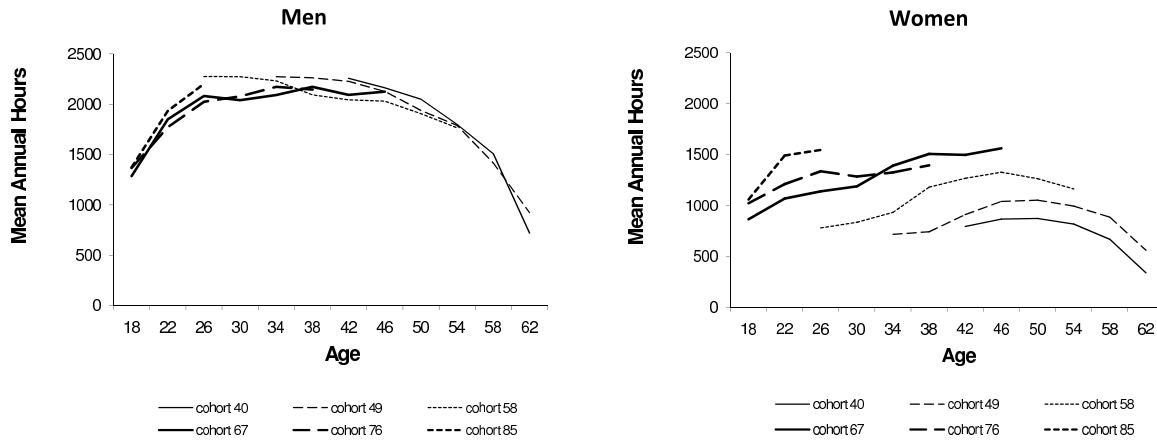Figure A-1: Mean Annual Hours Worked, 1968-1996, PSID.

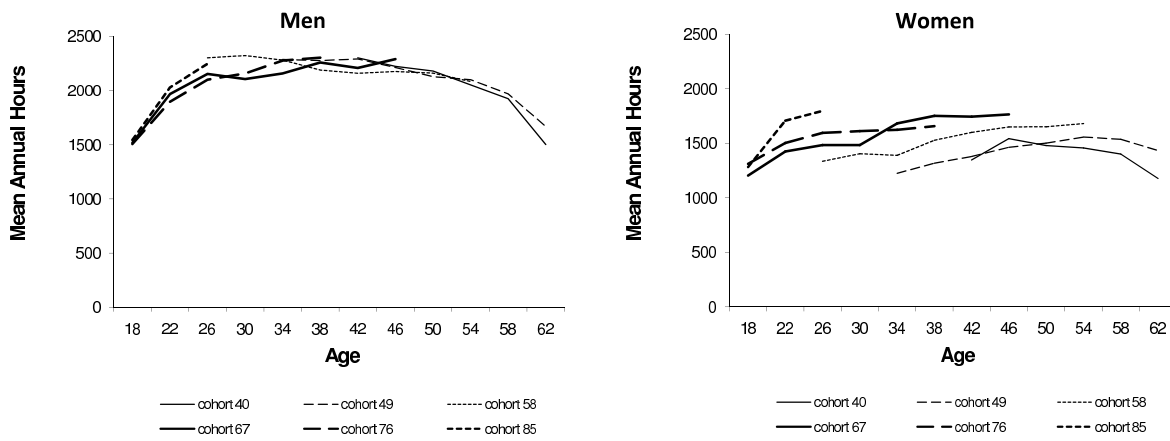Figure A-2: Mean Annual Hours Worked, Workers with Positive Hours, 1968-1996, PSID.

Figure A-3: Coefficient of Variation of Annual Hours, 1968-1996, PSID.

Men

Women