Revisiting Wage, Earnings, and Hours Profiles*

Peter Rupert† Giulio Zanella‡

February 15, 2010

Abstract. This paper consists of two parts. In the first part we use the PSID (which is now long enough to allow the observation of the entire working life) to document empirical life cycle profiles (wage, earnings, hours) of a cohort of workers followed from when they enter the labor market in 1967 to when they retire in 2006. We find an empirical pattern that is inconsistent with the human capital model, where the three profiles are hump-shaped. In particular, in this model the wage and hours profiles necessarily have the same shape if human capital depreciates and the real interest rate is not too high relative to the subjective discount rate. We find instead that in our data the wage profile does not decline while hours per worker drop sharply beginning at age 50, when many workers seem to start a smooth transition into retirement by working progressively fewer hours. The decline in earnings is therefore driven by labor supply. This pattern is confirmed by the CPS. We show that a model in which the value of leisure/disutility of work starts growing after age 50 reconciles theory and data. In the second part we study an OLG model that incorporates two main forces that may be driving such shift in preferences: health deterioration and appearance of grandchildren. The model produces conditions under which unhealthy workers and grandparents reduce labor supply. Under certain conditions the average effects of such shocks can be estimated in a straightforward way without having to resort to full structural estimation. We illustrate such identification conditions and show that while the health channel has an insignificant effect, intergenerational transfers of time (grandparenting) play a major role in reducing labor supply along the intensive margin towards the end of the working life.

JEL Classification Codes: E24; J13; J22; J24; J26

Keywords: wage profile, earnings profile, labor supply, life cycle, synthetic cohorts, human capital, intergenerational transfers, grandparents, child care, value of leisure.

*We are grateful to Alberto Dalmazzo, Eric French, Peter Kuhn, and Richard Rogerson, for insightful comments, as well as to seminar participants at the universities of Aarhus, Chicago Booth, Leicester, Rome–La Sapienza, and Siena. Bonnie Queen provided outstanding research assistance.

†University of California–Santa Barbara. E-mail: rupert@econ.ucsb.edu.

‡University of Bologna E-mail: giulio.anella@unibo.it

1 Introduction

What is the shape of the life cycle profiles of hourly wages, hours of market work, and labor earnings? This question may sound uninteresting at first because of the established consensus that these profiles are all hump-shaped, reflecting the process of optimal investment in human capital. The human capital model initiated by Ben-Porath (1967) and further developed in the 1970s by Ghez and Becker (1975), Blinder and Weiss (1976), Ryder, Stafford, and Stephen (1976), Heckman (1976), and Rosen (1976) posits that the wage rate is equal to the unit return on the individual stock of human capital. The wage rate initially grows because investment (e.g. via reduced compensation in return for on-the-job training) causes this stock of skills to increase. Then, as investment in new human capital optimally falls below depreciation, the wage rate starts declining, tracing a hump-shaped profile. Hours and earnings follow a similar process. Figure 1 below, reproduced from Weiss' (1986) survey, illustrates. This model was meant to explain the stylized facts summarized by Weiss (1986):

> The major stylized facts which the theory attempts to explain are: a life cycle earnings profile which is increasing at early ages and is declining towards the end of the working period. A wage profile which tends to increase over the life cycle with a weak tendency for wage reduction towards the end of the working period. An hours of work life cycle profile which is increasing at early ages and declining at older ages, with the peak occurring earlier than in the earnings or wage profiles. (p. 603)

At the time these empirical regularities were isolated, researchers could only rely on cross-sectional data or short panels. This limitation forced them to infer the shape of life cycle profiles from synthetic cohorts[1] or repeated cross sections from administrative, such as social security, data. Figure 2 illustrates that if one constructs life cycle profiles using synthetic cohorts from a single cross-sections of the Panel Study of Income Dynamics (PSID), these match the theoretical profiles reported in Figure 1.[2]

[1] By synthetic cohort we mean the time series constructed from a cross-section at, say, time t using the $a + j$ years olds at time t as a counterfactual for the a years olds at time $t + j$.

[2] The wage and earnings profiles in Figure 2 result from "tilting" the raw cross sectional profiles to allow for 1.5% real wage growth every year. In other words, the wage of an individual a years old in the cross section is multiplied by $(1.015)^{a-23}$, where 23 is the age of the youngest worker in the cross section. This correction makes the wage and earnings profiles much less hump-shaped than they would otherwise be. The interpolating lines are the predicted values from a regression on a quadratic function of age. This was the standard representation before Murphy and Welch (1990) suggested a quartic function of age should be used instead.

Figure 1. Theoretical life cycle profiles in the human capital model.
[Reproduced from Weiss (1986)]

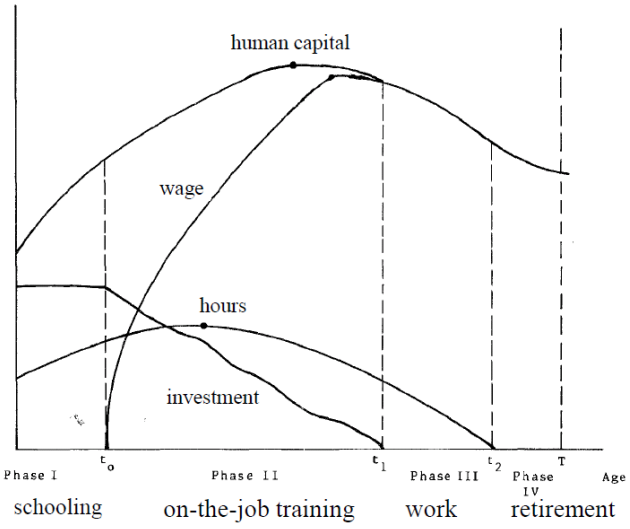
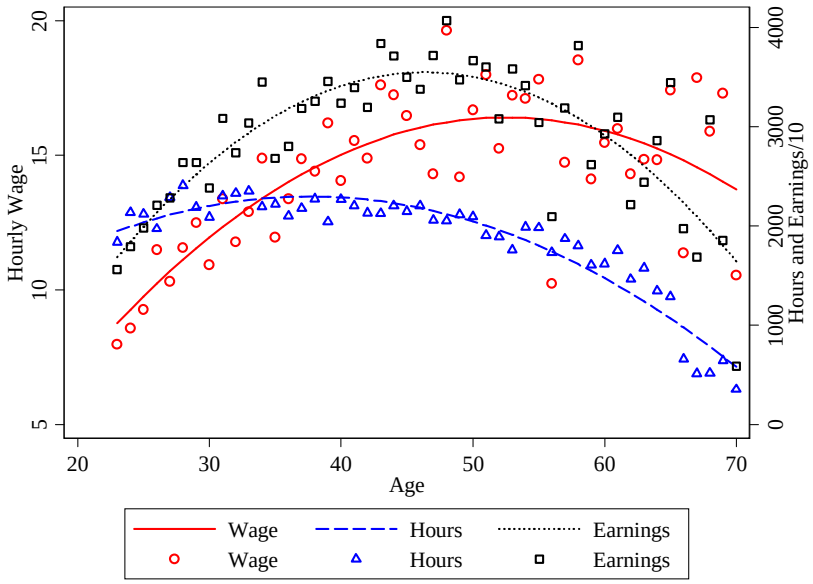


Figure 2. Empirical life cycle profiles from the 1970 synthetic cohort in the PSID
[Profiles are "tilted" to allow for real wage growth – see footnote 2]



These days we can get a more accurate empirical picture of life cycle profiles by analyzing panel data. In this paper we exploit a unique feature of the PSID: this panel dataset now covers 40 years and so it spans the entire working life of the cohort who entered the labor market in 1967. Therefore, we can construct actual life cycle profiles by averaging individuals in the cohort at each point in time and see whether the theoretical prediction of the human capital model hold in panel data like they do in single cross sections. Furthermore, we can do this for younger cohorts to control for cohort and time effects on the shape of the empirical profiles.

This paper consists of two parts. In the first part we document such empirical profiles in the core sample of the PSID (1967-2006) and to show that they are inconsistent with the benchmark human capital model. By "benchmark human capital model" we mean a human capital model in which the value of leisure is constant, human capital depreciates—as we illustrate in detail later in the paper. In particular, even after correcting for selection into employment, the wage rate shows no tendency to fall towards the end of the working life—at least not before age 65—contrary to cross-sectional stylized facts.[3] On the other hand, hours of market work of employed individuals decline sharply after age 50. The observed decline in earnings after age 50 is therefore completely driven by an adjustment of labor supply. To overcome the difficulties generated by small sample size and attrition we show that this pattern is roughly confirmed by the March supplement of the CPS (1978-2008). This overall pattern agrees with what Mincer (1974) had already noticed for weekly earnings analyzing census data:

> There is no visible decline at these later ages in weekly earnings. Apparently, declines in weeks worked per year are the main factor in the decline of annual earnings during preretirement years. (p. 70)

This is inconsistent with the benchmark human capital model because, as we show below, in this model the wage and hours profiles necessarily have the same shape if the real interest rate is not too high relative to the subjective discount rate. Even in its modern, fully-fledged general equilibrium version the human capital model generates hump-shaped life cycle profiles when calibrated with realistic parameters (Hugget, Ventura, and Yuron, 2009). Therefore, we need a different model to accomodate the evidence. We can reconcile these findings in a model in which the value of leisure increases after age 50.

In the second part of the paper we suggest an OLG model that incorporates two main forces that may be driving such shift in preferences: health deterioration

[3] This had already been noticed by Thornton, Rodgers, and Brookshire (1997), who report findings from longitudinal data constructed from social security and census datasets.

and appearance of grandchildren. The model produces conditions under which unhealthy workers and grandparents reduce labor supply. We quantify empirically the role of these two channels in the observed decline in hours. We find that health has a negligible effect, while grandparenting explains roughly 40% of the reduced labor supply during preretirement years. We argue that becoming grandparents is an exogenous event after conditioning on a few observable that affect the probability of such event. Results derived under this assumption are robust to using single-parent families (as a proxy of unexpected pregnancies) as instrument for the event of becoming grandparents.

Cardia and Ng (2003) have documented that such time transfers are substantial and have a positive effect on the labor supply of young parents. Our model is an extension of their model that highlights two additional facts. First intergenerational time transfers may also depress the labor supply of older individuals—young grandparents—long before retirement. Second, in equilibrium such transfers are driven to a good extent by grandparents deriving utility from time spent with grandchildren.

We do not take a stand on why the wage profile is not declining before age 65. It could be that workers are not becoming less productive as they approach retirement, or that they are becoming less productive but their wage rate is above productivity—for instance because efficient wage profiles a la Lazear (1979, 1981) are in place. In either case, since labor supply responds to the wage rate (regardless of whether this coincides with true productivity or not) the puzzle of a sharp decline in ours while hourly wage are not declining and income/wealth effects are modest is present under both scenarios—with a caveat to be discussed later.

The paper is organized as follows. Section 2 describes the dataset and documents the empirical life cycle profiles in the PSID. Section 3 shows that the empirical pattern we uncover cannot be replicated in the benchmark human capital framework and sketches a model that instead is able to. In this model, deteriorating health and becoming grandparents generate arguably exogenous shocks that reduce labor supply. Section 4 evaluates the contribution of these two channels in a regression framework.

2 Data

2.1 Data sources and data selection criteria

After the release of the 2007 wave in August 2009, the PSID offers the unique opportunity to observe in a panel individual life cycle profiles of wages, hours, and earnings for 39 years (from 1967 to 2006), i.e. the entire working life of the cohort who entered the labor market in 1967. We compute weighted averages of wage

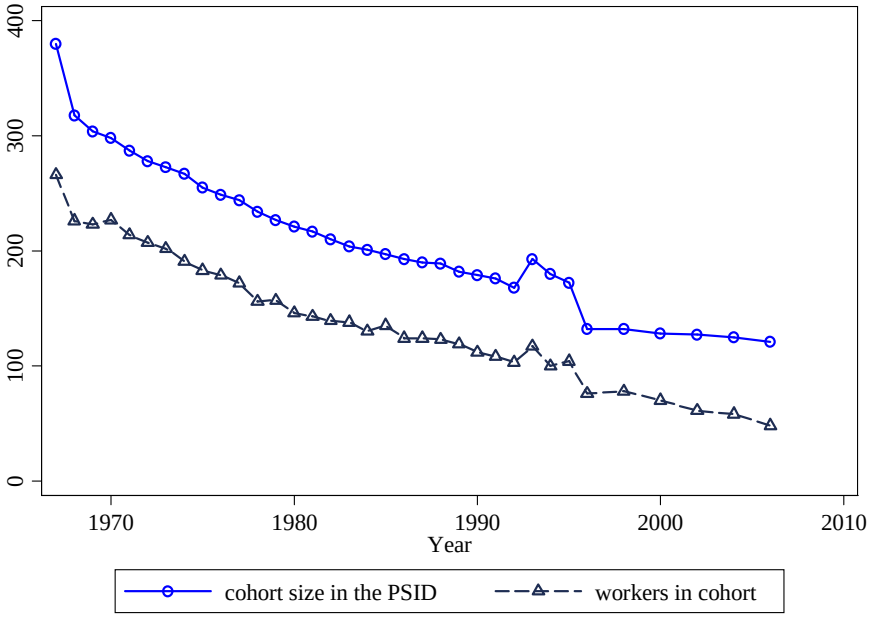
rates, hours, and earnings for a given cohort in a given year to trace the life cycle profiles of individuals in that cohort. Zeros are not averaged: this means we have a selection problem we need to deal with.

We used the following data selection criteria. First, we discard the SEO sample and work with the core sample only. Sampling weights applied to the core sample produce statistics that are representative of the US population. We use the PSID family-level data file and so have longitudinal information of interest mostly for "household heads", i.e. men. We discard the few women who are household heads to have an accurate picture of at least male workers. In order to have a sufficient number of observations to compute meaningful averages, we define cohorts based on 5-year bins and use the central value of the bin to keep track of the age of the cohort. For instance, men who where between 21 and 25 years old in 1967 (i.e. at the outset of the survey) are the "23 years old" cohort, those who where between 26 and 31 years old in 1967 are the "28 years old" cohort, and so on. Self-employed individuals are excluded from the analysis both because they are more subject to measurement errors of variables such as the wage rate and because the latter is a mix of labor and capital income.

The wage rate is computed in the PSID as the ratio between self-reported labor income and self reported hours of market work. This makes some people skeptic about the reliability of such variable. However, panel regressions (not reported) of earnings on hours produce an elasticity of exactly 1. This elasticity answers the following question: what's the percentage variation in reported earnings following a 1% increase in reported hours? Since these two variables are independently measured, a unit elasticity suggests that measurement errors in hours (and correspondingly in wages) are negligible—unless there is a proportional measurement error in earnings, which is unlikely. In other words, when people choose to work less this is fully reflected in their earnings, so that the implied hourly wage is a good measure of the unobserved wage rate. All nominal quantities are expressed in constant 1982 dollars using the CPI-urban. We exclude individuals whose implied nominal wage rate in a given year is below the Federal minimum wage in force that year. Furthermore, in order to eliminate outliers we trim the distribution of wages and hours at or above the top 1% and at or below the bottom 1% every year. This eliminates one or two observations every year.

In addition to the main labor market variables of interest (wages, hours, and earnings), we observe a number of covariates we will use to condition life cycle profiles and to make inference: age, education, health, non-labor income, number of children, family size, geographic mobility. Focusing on a specific cohort in a panel that included about 4,000 observation at its outset in 1967 means working with relatively few individuals. Furthermore, attrition is likely to be severe on a 40-years horizon. Figure 3 illustrates these facts for the 23 years old cohort.

Figure 3. Number of individuals and number of workers
cohort: individuals 21-25 years old in 1967



This picture shows that the cohort of men aged 21-25 years old in 1967 comprised almost 400 individuals at the outset of the panel, almost 300 of which worked. Attrition is substantial but it seems unrelated to employment, at least until mid-1990s.[4] Notice that after mid 1990s there is very little attrition in this cohort while the number of workers declines steadily due to the process of retirement. The last data point is 2006: the cohort is 60-64 years old and we will be averaging about 60 observations for which wages and hours are observed.

To overcome these difficulties we collected data for this same cohort from the March supplement of the CPS to compute the same life cycle profiles from a much larger dataset (the rotation group provides labor market information for about 100,000 individuals every year) not affected by attrition—and using the same selection criteria described above. The disadvantage of course is that the CPS is not a panel. Yet, large repeated cross sections provide an approximation that is good enough for the purposes of checking the robustness of a relatively small (but long) panel.

[4]The increase in sample size in 1993 is due to the reappeareance of "recalls", individuals who where present in 1967, had been lost, and recontacted that year.

2.2 Empirical profiles

Figure 1 shows the wage profile for the oldest cohort (21-25 years old in 1967). The wage profile is increasing and steep until 40 year old. Real wage growth is then slower but there is no tendency to decline at retirement age (the cohort is 60-64 years old in 2006). This pattern is confirmed by the CPS[5] and contradicts the first stylized fact discussed above, according to which we should observe a decline in the profile. Figure 5 shows this same profile for workers with and without college degree. It appears that the profile of low-skilled individuals is in fact slightly declining. However, it is not declining for skilled workers, which is still inconsistent with the first stylized fact.

Since we are averaging only individuals who work, the suspicion is that the unconventional shape of this profile may be an artifact of selection into employment. In particular, if less productive individuals leave the labor force earlier (e.g. because they have more fatiguing jobs), then we are overestimating the average wage rate at the end of the working life. We can gain some quantitative intuition into this bias by employing a consistent predictor of the wage rate of individuals who are not employed. We do this using a standard two-stage Heckman procedure, i.e. under the assumption of joint normality of the errors in the selection and wage equations. At first stage we estimate the probability of being employed using a random effects Probit estimator. The variables used to predict employment are age, education, health, family size, non-labor income, and two dummies for being 62 and 65 years old. Since these are the two points in the working life of Americans at which they an incentive to leave the labor force under the current social security rules, these dummies provide an instrument for the retirement process. At second stage, we consistently predict the wage rate of individuals who left their jobs using the coefficients from a regression of hourly wage on the same variables used at first stage (except the two age dummies), individual dummies, and the correction term—the ratio between the normal pdf and cdf evaluated at the predicted probability of being employed. Then we compute the wage profile by including these predicted wages in the averages. The result is a counterfactual wage profile, i.e. the profile we would have observed had those workers who retired remain employed. Actual and counterfactual wage profiles are plotted in Figure 6. This illustrates that while the presumption that mean wages are overestimated towards the end of the working life, the bias is not dramatic. In particular, even after the correction the shape of the wage profile is at odds with the first stylized fact.

[5] Notice that the fundamental difference between the PSID (a panel) and the March supplement of the CPS (repeated cross-sections, effectively) is reflected in the different volatily of the two profiles.

Figure 4. Wage profile, cohort of men 21-25 years old in 1967

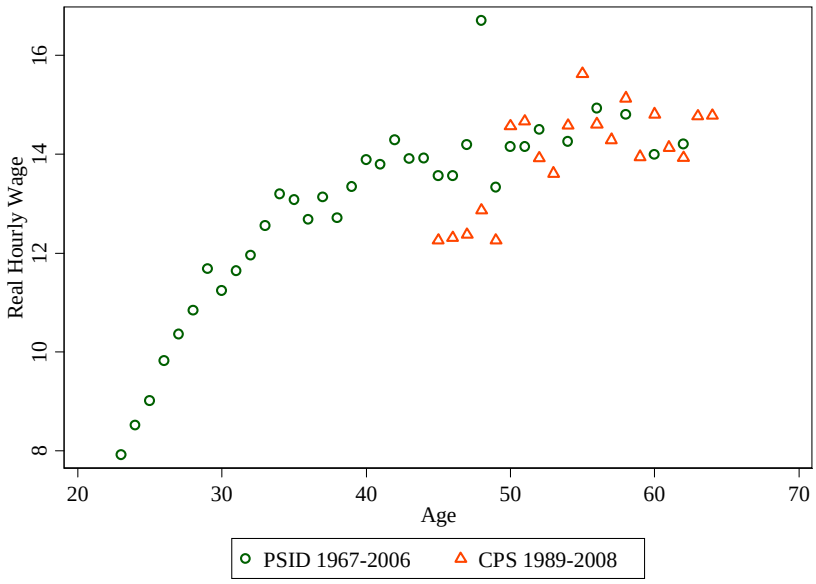


Figure 5. Wage profile by education

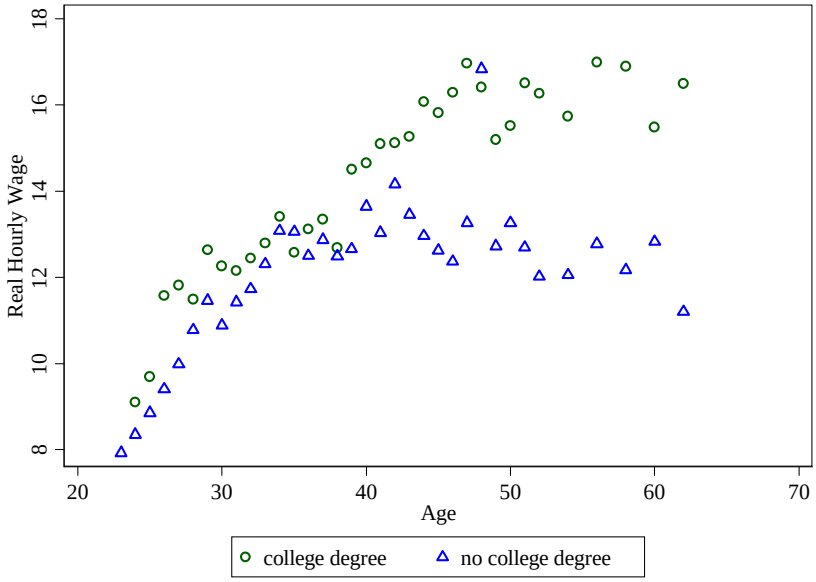


Figure 6. Wage profile, effect of selection-correction.

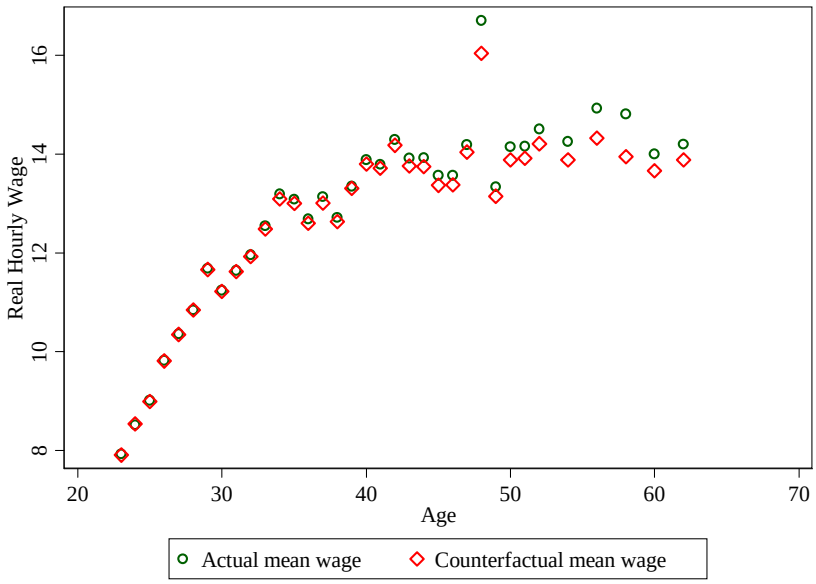


Figure 7. Wage profile, cohort of men 26-30 years old in 1967

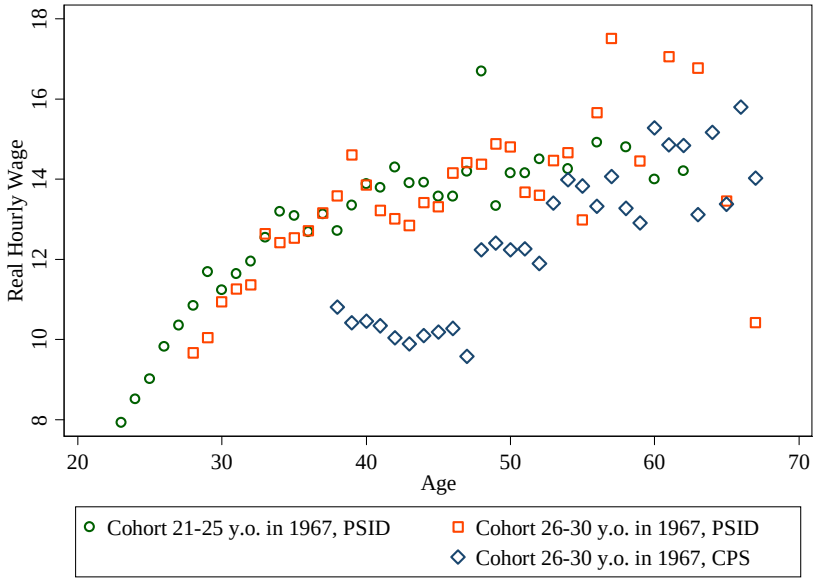


Figure 7 compares the wage profiles of this cohort with the profile of the older cohort of men who were 26-30 years old in 1967 and therefore are 65-69 years old at the last observation in 2006. We also report estimates from the CPS for this older cohort, like we do in Figure 4 for the younger cohort. The profile for the older cohort in Figure 7 shows no obvious tendency to decline, except at the last data point: the drop in real wage when the cohort is 67 years old, in 2006, is clear. However, this last data point is obtained from averaging a mere 25 men who are still working after age 65 (out of 400 observations in 1967 for this older cohort) and so this is likely to be a poor estimate. Estimates from the CPS (the diamonds in the Figure) confirm that this profile is not declining.

Figure 8 reports the earnings profile for the cohort that is 21-25 years old in 1967. This profiles has instead the traditional hump-shape, confirmed by CPS estimates: labor income starts declining visibly after age 50. Figure 9 shows the profile conditional on education, which confirms this pattern.

Declining earnings and non-declining hourly wages imply that earnings fall because of a reduction in hours of market work. The hours profile in Figure 10 illustrates this pattern: workers begin reducing labor supply after age 50. In the PSID, the cumulative reduction between ages 51 and 62 with respect to average hours worked between ages 30 and 50 is about 350 hours, or about 15%. Figure 11 shows that this happens for both college graduates and non college graduates. This conforms to the more general notion of retirement advocated by Heckman (1976):

> Retirement can more generally be defined as a period with few hours of work supplied to the market. (p. S15)

Summing up, the picture that emerges from these empirical profiles is the following. The wage profile is increasing over the life cycle but is not declining at the end of the working life. The earnings profile increases over the life cycle and declines after age 50. The hours profile is increasing at early ages, then is roughly constant before declining sharply after age 50.

This pattern suggests that labor supply drives the decline in earnings: men choose to work less after age 50. The remainder of the paper is concerned with understanding what determines this choice. We cannot rule out the alternative interpretation that the decline in hours reflects the demand side. For instance, if workers are paid above their marginal product when they are old—as in Lazear (1979; 1981)—the employer may have an incentive to induce them to work fewer hours. Even if this is the case workers are accepting to work fewer hours at a given wage rate, i.e. are still choosing to reduce their labor supply in equilibrium.

Figure 8. Earnings profile, cohort of men 21-25 years old in 1967

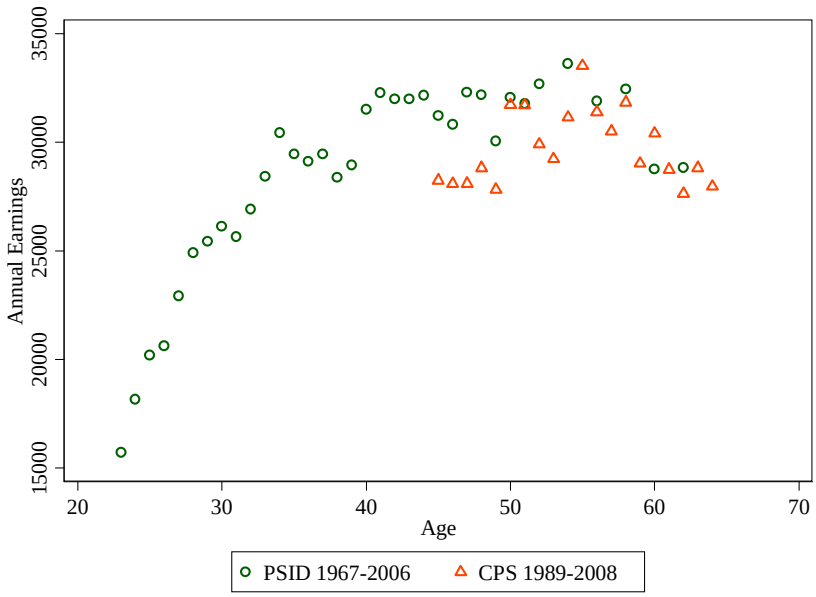


Figure 9. Earnings profile by education

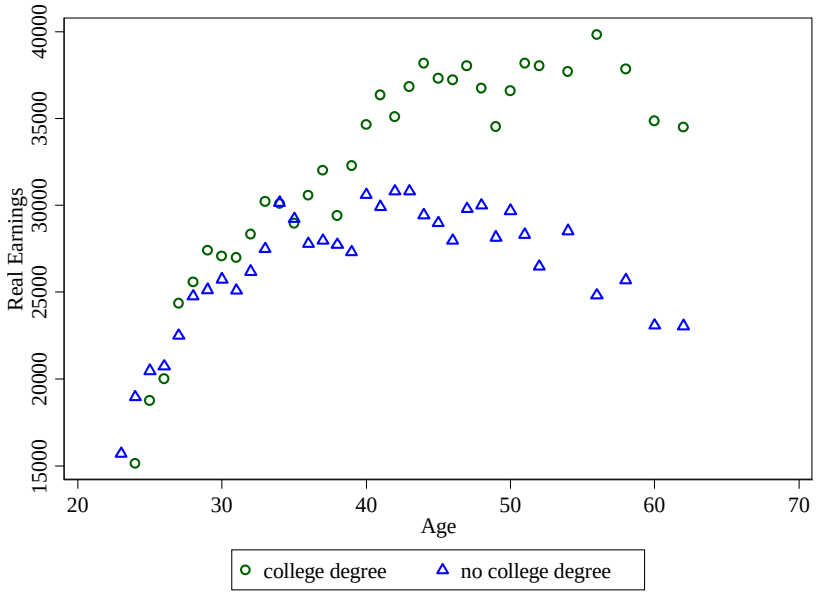


Figure 10. Hours profile, cohort of men 21-25 years old in 1967

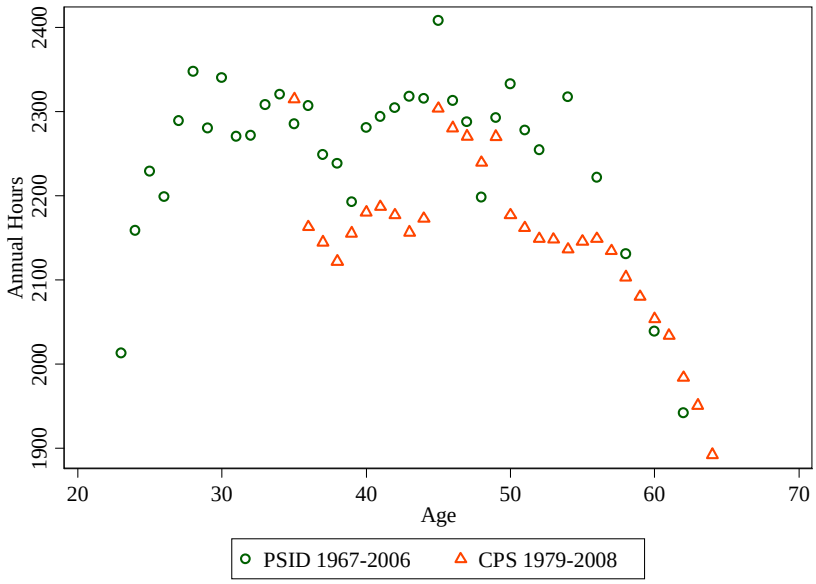
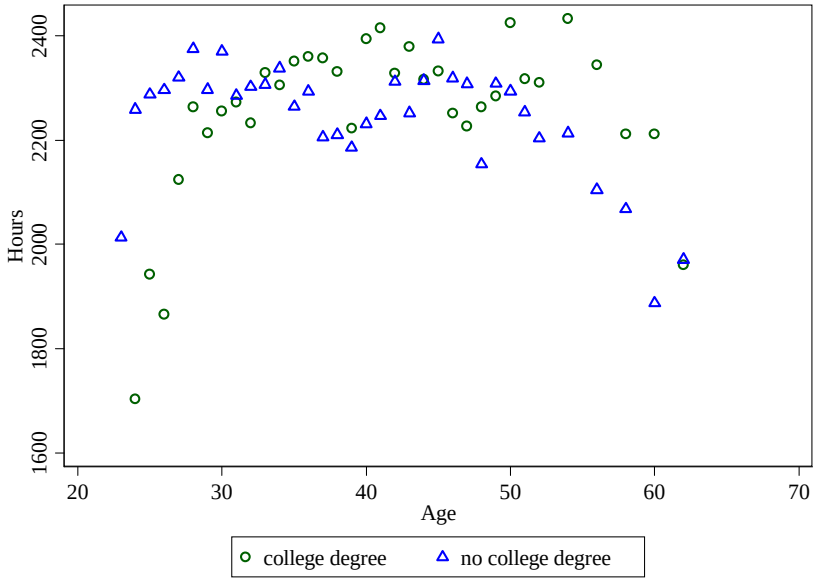


Figure 11.Hours profile by education



3 Competing models

3.1 Benchmark human capital model

The benchmark human capital model cannot reproduce this pattern. Consider the following version of the model. In each period t there are overlapping generations alive, indexed by age, a. Each individuals enters the labor market at age $a = 1$, lives for A years afterwards, and chooses consumption (c), hours of work (h), investment (x) in human capital or skills (s) and assets (k) to maximize discounted lifetime utility. Assets yield a real return i. Investment in human capital takes the form of on-the-job training and is costly because the hourly wage is correspondingly reduced. So investment x is interpreted as the proportion of hours spent on the workplace training rather than producing (and so is constrained to the unit interval), in which case the wage rate is the product of the return on the human capital stock and the fraction of hours spent producing. The learning production function is concave in investment and increasing in both labor supply and the pre-existing stock of human capital. Each unit of human capital, which is not tradable, has a net return r that is independent of age and depreciates at a constant rate δ. In other words, each individual solves the following problem:

$$\begin{aligned} \max_{\{c_a,h_a,x_a,k_{a+1},s_{a+1}\}} \quad & : \quad \sum_{a=1}^{A} \beta^a \left[v(c_a) - \gamma \frac{h_a^{1+\eta}}{1+\eta} \right] \\ \text{s.t.} \quad & : \quad c_a + k_{a+1} = k_a(1+i_t) + r_t s_a (1-x_a) h_a \\ & : \quad s_{a+1} = s_a + (x_a^{\theta} h_a - \delta) s_a, \end{aligned} \tag{1}$$

where $r_t s_a(1-x_a)$ is the wage rate, $x_a^{\theta} h_a s_a$ (with $\theta < 1$) is the learning production function, and γ reflects the value of leisure (i.e. any alternative use of time other than work). At the optimum, investment, human capital, labor supply, and the wage rate grow at the following gross rates:

$$\frac{x_{a+1,t+1}}{x_{a,t}} = \left[\beta(1+i_t)\frac{r_t}{r_{t+1}}\frac{\mu_{a+1,t+1}}{\mu_{a,t}}\right]^{\frac{1}{1-\theta}} \quad (2)$$

$$\frac{s_{a+1,t+1}}{s_{a,t}} = 1 + x_a^{\theta} h_a - \delta \quad (3)$$

$$\frac{h_{a+1,t+1}}{h_{a,t}} = \left[\frac{1}{\beta(1+i_t)}\frac{r_{t+1}}{r_t}\frac{s_{a+1}}{s_a}\frac{1+x_{a+1}(\theta^{-1}-1)}{1+x_a(\theta^{-1}-1)}\right]^{\frac{1}{\eta}} \quad (4)$$

$$\frac{w_{a+1,t+1}}{w_{a,t}} = \frac{r_{t+1}}{r_t}\frac{s_{a+1}}{s_a}\frac{1-x_{a+1}}{1-x_a}, \quad (5)$$

where $\mu_{a,t}$ is the multiplier on the law of motion of human capital, i.e. the shadow price of skills. Since human capital is not tradable and the residual stock at age A is valueless, an additional unit is valued less and less as the individual age and so the shadow price of skills monotonically decreases with age. Therefore, unless the unit return on human capital decreases very fast, equation (2) indicates that investment decreases with age—a standard property of the human capital model. As a consequence, human capital first increases (until $x_a^{\theta} h_a > \delta$) then decreases resulting in a single-peaked concave profile. This hump-shape transfers to the hours and wage profiles, equations (4) and (5), unless the return on human capital grows at a sufficiently high rate. The typical profiles are those illustrated in Figure 1. By inspecting equations (4) and (5) one immediately sees that in this model a non-decreasing wage profile implies a non-decreasing hours profile due to the substitution effect.[6]

3.2 Age-dependent value of leisure

As a preliminary step towards the model we put forward, consider the following, simpler life cycle model in which human capital is left in the background, the value of leisure is age-dependent, and the wage rate is exogenous and never decreases (all the notation is as above):

$$\begin{aligned} \max_{\{c_a, h_a, k_{a+1}\}} \quad & : \quad \sum_{a=1}^{A} \beta^a \left[v(c_a) - \gamma_a \frac{h_a^{1+\eta}}{1+\eta} \right] \\ \text{s.t.} \quad & : \quad c_a + k_{a+1} = k_a(1+i) + w_{a,t} h_a, \end{aligned} \quad (6)$$

[6] Income effects are captured by the term $1/\beta(1+i_t)$, which is the optimal gross rate of growth of the marginal utility of wealth from the Euler equation. If assets are accumulated optimally, then income effects cannot drive hours down unless the interest rate is increasing, which is not the case in the in the period of interest in the US.

with $w_{a+1,t+1} > w_{a,t}$ for each a and t. The dynamics of hours in this model is given by:

$$\frac{h_{a+1,t+1}}{h_{a,t}} = \left[\frac{1}{\beta(1+i_t)}\frac{w_{a,t+1}}{w_{a,t}}\frac{\gamma_a}{\gamma_{a+1}}\right]^{\frac{1}{\eta}}. \tag{7}$$

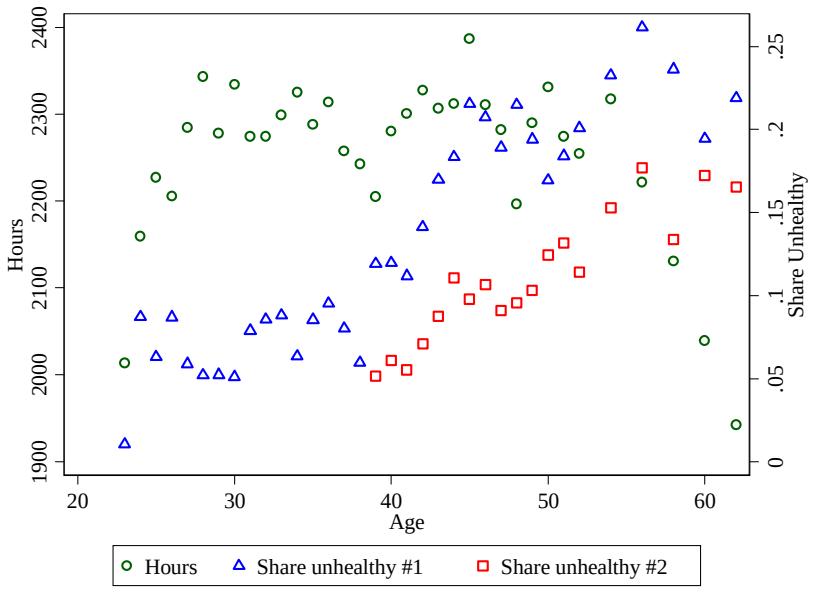
Therefore, this model reproduces a declining hours profile in the presence of a non-decreasing hourly wage—with the same caveat on wealth effects—if the value of leisure/disutility of work increases with age. Except for the dependence of γ on age and the identity between hours of work and labor services, this is the discrete-time version of the model in Rogerson and Wallenius (2009). We emphasize this connection because their model assumes that wages start decreasing at mid-career so to generate a declining hours of work profile and eventually retirement. Of course nothing would change if Rogerson and Wallenius assumed that it is instead the value of leisure that is hump-shaped. The empirical patterns in the PSID suggests this is in fact the case.

In particular, this simpler model reproduces the empirical pattern in the PSID if the value of leisure increases sharply after age 50. We consider two forces that may generate such increase.

3.2.1 Health

First, the deterioration of physical and mental health. Except in few circumstances, work is not a pleasant activity and after 20-30 years workers may well desire or be forced to reduce hours of work. Evidence from survey data shows that stress levels at work peak precisely at 50-55 years of age and decrease before retirement [Griffiths *et al*. (2009)]. There are two variables that measure health in the PSID. First, the following generic question has been asked is all waves: "Do you have any physical or nervous condition that limits the type of work or the amount of work you can do?", with the answer being either yes or no. Second, a question to self-report own health status is asked since 1984: "Would you say your health in general is excellent, very good, good, fair, or poor?" If we define an unhealthy individual as someone who is not in good health, i.e. someone who responded "yes" to the first question (label this health measure #1) and either "fair" or "poor" to the second (label this health measure #2) and plot the share of unhealthy individuals and hours of market work as a function of age in our sample, we obtain Figure 12.

Figure 12. Health and hours

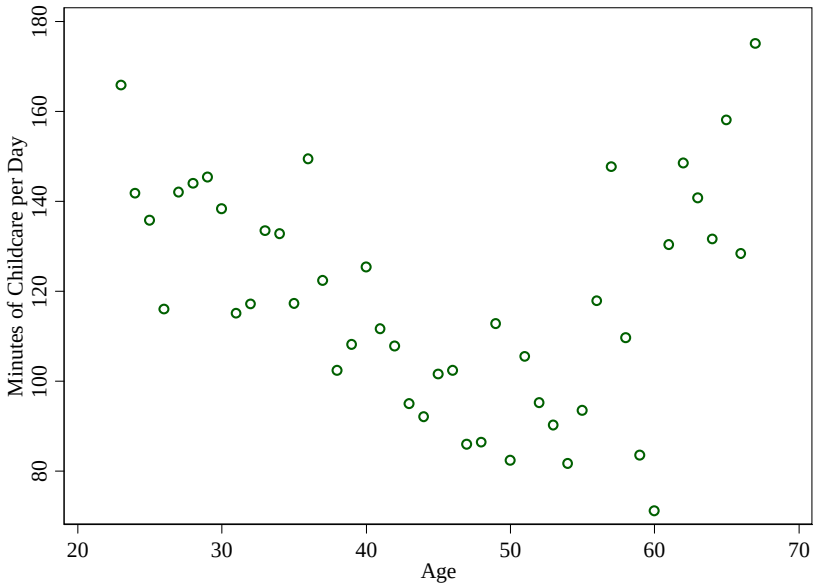


This picture shows that the two health variables give the same information. In particular, health deteriorates after age 40, reaches the bottom around age 50 when hours of market work also decline, and improves afterwards. The improvement before retirement is likely to be overestimated as unhealthy individuals are more likely to disappear from the survey. In what follows we use the second measure of health in the PSID, following what is a standard choice in the literature [e.g. French (2005)].

3.2.2 Gandparenting

Second, intergenerational transfers of time in the form of grandparenting. Figure 13 reports from the 2008 cross-section of the American Time Use Survey the amount of minutes per day individuals spend taking care of children. If we adopt the synthetic cohort method, this picture suggests that individuals spend a considerable amount of time taking care of children as young parents, then decrease this amount as their children grow-up but then start spending again a lot of time taking care of children when they become grandparents from age 50 onward.

Figure 13. Child care and hours



This picture is consistent with data from the Health and Retirement Study: in 2008, on average, a couple of sampled American adults over 50 provided 500 hours of grandparenting. This is equivalent to about 13 38-hours workweeks. Such large transfers are likely to have substantial effects on labor market behavior of senior workers during the preretirement years. Cardia and Ng (2003) study a model in which intergenerational time transfers in the form of grandparenting, motivated by altruism towards own children, are substantial and may have first-order effects on labor supply. However, they seem to miss the offsetting negative labor supply effect on senior workers. We next suggest a model similar to theirs that makes such effect explicit and produces implications we can bring to data.

3.3 An OLG model

We follow by and large the model of Cardia and Ng (2003), with a few important extensions we discuss below. Consider a world in which overlapping generations are born at each date t and live for three periods: they are children in period 0 and live with their parents, young in period 1 when they form their own household, and old in period 2. Since children make no economic choice in this model, we will ignore period 0 and employ a two-period setting. Each individual belongs to

a one-parent family and a household. A parent can have more than one child, but we define families in vertical line. That is, a family is composed of a grandparent, a parent, and children. A household, instead, is composed of a parent and children living with her. Grandparents do not live with their grandchildren. A parent is altruistic in that she cares about the welfare of her children. Both young and old individuals value average household consumption and leisure, supply labor, and enjoy spending time with children who belong to their family. Childrearing requires a fixed amount of time per child, and child care is also available on the market with exogenous and infinitely elastic supply. Therefore, time spent with children is both an input to childrearing and home production of a good (time spent with family children) for which there is no market.

The only difference between the two generations is that the young can have children of their own while the old cannot. Therefore, the only way old individuals can spend time with children is by supplying time to the childrearing process, i.e. grandparenting. Since childrearing time can be purchased by parents on the market, grandparenting is a form of intergenerational transfer. The other form is a standard asset transfer. Bequests and inter-vivos transfers of assets are indistinguishable in our model.

In what follows we will use superscripts to indicate birth date and subscripts to indicate life cycle period. Denoting consumption per household member by c_j^t, $j \in \{1,2\}$, leisure by l_j^t, and total time spent with children by parents and grandparents by q_1^t and q_2^t, respectively, preferences of an individual born at date t are given by:

$$U^t = u_1(c_1^t, l_1^t, q_1^t) + \beta u_2(c_2^t, l_2^t, q_2^t) + \alpha\beta U^{t+1}, \tag{8}$$

where u_1 and u_2 are strictly concave functions that are separable and increasing in all their arguments, $\beta \in (0,1)$ is the discount factor and $\alpha \in [0,1]$ the degree of intergenerational altruism. This formulation implicitly assumes that a parent cares about the quality of children—U^{t+1} is the welfare of the representative child—rather than their quantity. Denote by n^t the number of own children for an individual born at time t. When old, this individual will have $n^t n^{t+1}$ grandchildren. Therefore, total time spent with children at each point in the life cycle can be written as

$$q_1^t = x_1^t n^t, \tag{9}$$
$$q_2^t = x_2^t n^t n^{t+1}, \tag{10}$$

where x_1^t and x_2^t denote per-child parenting time and per-grandchild grandparent-

ing time, respectively. We make the following, additional, assumption on preferences:

Assumption 1 (health weakly deteriorates): $\left|\frac{\partial u_2}{\partial h_2^t}\right| \geq \left|\frac{\partial u_1}{\partial h_1^t}\right|$ for all $h_2^t = h_1^t$

In other words, work is more painful when old than when young.

It takes a fixed amount $\overline{h}$ of time to raise a child. If the sum of time inputs by parents and grandparents falls short of this amount, the difference must be purchased on the market. Therefore, denoting the market price of childrearing by p^t, first period total assets by k_1^t, second period per-child assets by k_2^t, the wage rate by w_j^t, and labor supply by h_j^t, the first period budget constraint is:

$$(1+n^t)c_1^t + (\overline{h} - x_1^t - x_2^{t-1})n^t p^t + k_1^t \leq w_1^t h_1^t + k_2^{t-1}. \tag{11}$$

Grandparents must pay a transportation cost ϕ^t per unit of time transfer. This cost is a friction that captures the spatial separation between grandparents and grandsons. Therefore, the second period budget constraint is:

$$c_2^t + n^t k_2^t + \phi^t q_2^t \leq w_2^t h_2^t + (1+r^t)k_1^t. \tag{12}$$

We also impose a non-negativity constraint on assets in period 2—the old cannot pass their debt over to their children:

$$k_2^t \geq 0. \tag{13}$$

Finally, normalizing the time endowment to 1, the time constraints are:

$$l_1^t + h_1^t + q_1^t = 1, \tag{14}$$
$$l_2^t + h_2^t + q_2^t = 1. \tag{15}$$

We impose the additional constraint

$$x_1^t + x_2^{t-1} \leq \overline{h} \tag{16}$$

to ensure that time spent with children has no value outside the family. That is, parents and grandparents cannot sell child care on the market. This assumption, and more generally the absence of a model for the production side of the economy,

makes this a partial equilibrium framework. Although a general equilibrium model would make our analysis richer and more rigorous it would be more than we need to reconcile life cycle theory and life cycle empirical profiles—which is the goal of this part of the paper.

The problem is to maximize (8) with respect to c_1^t, c_2^t, l_1^t, l_2^t, n^t, x_1^t, x_2^t, k_1^t, and k_2^t subject to (9) through (16), as well as to the optimal response of one's children to the time transfer:

$$n_{t+1} = n_{t+1}(x_2^t). \tag{17}$$

This final constraint recognizes the strategic interaction between parents and children: the latter respond optimally to time transfers from the former. Parents, in turn, will choose the per-grandchild time transfer responding optimally to the fertility choice of their children. Therefore, n_{t+1} and x_2^t are simultaneously determined in a Nash equilibrium.

Our model differs from the grandparenting model of Cardia and Ng (2003) in several respects. First, and most important, in their model the old have zero labor supply by assumption. Instead we want the old to trade-off labor supply and grandparenting before fully retiring from the labor market, consistently with the evidence discussed above.

Second, in their model time spent with children has no intrinsic utility for the old: the only motive for time transfers from them to the young is intergenerational altruism. As we show below, if the old can work a strictly positive amount of hours and have no direct preference for grandparenting, then it is less likely to observe grandparenting in equilibrium. Under certain conditions (the marginal value of child care to parents is below the marginal value of labor income and leisure for the old) direct preferences for grandparenting are necessary to observe any time transfer in equilibrium.

Third, the number of children is exogenous in their model and endogenous in ours.

On the other hand, we simplify along other lines and ignore important aspects of the Cardia and Ng model, such as old age care and richer substitution patterns between market and non-market inputs in the production of child care.

We are primarily interested in interior solutions because we want to explain the reduction in hours worked along the intensive margin only in a world in which both parents and grandparents spend time with family children. However, we will briefly discuss corner solutions below. In what follows we denote by u_{ya}^t the first partial derivative of the utility function with respect to variable y in period a for an individual born at date t, and by u_{yza}^t the corresponding cross-partial derivative

with respect to y and z. the first-order conditions for an interior optimum boil down to the following relations, all expressed in the form [*marginal cost*] = [*marginal benefit*]

$$\begin{aligned}
\text{Work when young} \quad &: \quad u_{l1}^t = \frac{w_1^t u_{c1}^t}{1+n_t} && (18)\\
\text{Work when old} \quad &: \quad u_{l2}^t = w_2^t u_{c2}^t && (19)\\
\text{Savings (Euler)} \quad &: \quad u_{c1}^t = \beta(1+r_t)(1+n_t)u_{c2}^t && (20)\\
\text{Asset transfers} \quad &: \quad n^t u_{c2}^t = \alpha u_{c1}^{t+1} && (21)\\
\text{Children} \quad &: \quad x_1^t u_{l1}^t + \beta n^{t+1} x_2^t u_{l2}^t + \\
&\quad + [c_1^t + (\overline{h} - x_1^t - x_2^{t-1})p^t + x_1^t w_1^t] u_{c1}^t \\
&\quad + \beta n^{t+1}(\phi^t + w_2^t) x_2^t u_{c2}^t = \\
&\quad x_1^t u_{q1}^t + \beta n^{t+1} x_2^t u_{q2}^t && (22)\\
\text{Parenting} \quad &: \quad u_{l1}^t + w_1^t u_{c1}^t = u_{q1}^t + p^t u_{c1}^t && (23)\\
\text{Grandparenting} \quad &: \quad u_{l2}^t + (w_2^t + \phi^t) u_{c2}^t = \alpha p^t u_{c1}^{t+1} + u_{q2}^t && (24)
\end{aligned}$$

Despite the seemingly cumbersome notation these equations are self-explanatory, as they reflect simple cost-benefit calculations at the margin.

We first need to establish existence and uniqueness of an equilibrium for the number of children in this model, so that the OLG structure makes sense in the first place.

Proposition 1 *An equilibrium number of children n^{t*} exists and is unique.*

Proof. See Appendix. ■

Such equilibrium number of children is part of the Nash equilibrium in the interaction between the old and the young generations.

Lemma 2 $n^{t+1} = 0 \Longrightarrow x_2^t = 0$.

Proof. See Appendix. ■

Proposition 3 *A Nash equilibrium in pure strategies (n^{t*}, x_2^{t-1*}) exists and in unique.*

Proof. See Appendix. ■

Next, we establish that worsening health leads a senior worker to reduce labor supply if health deteriorates faster than the marginal value of labor income. If the weak inequality in Assumption 1 is strict then the senior worker is said to be *unhealthy*. If instead it holds with equality then the senior worker is said to be *healthy*.

Proposition 4 *At the equilibrium, an unhealthy senior worker reduces labor supply, h_2^{t*}, relative to the case in which the worker remains healthy if health deteriorates at a rate above $(w_2^t/w_1^t)(u_{c2}^t/u_{c1}^t)(1+n^t)$.*

Proof. See Appendix. ■

Finally, we establish that in this model senior workers who are grandparents work less than their non-grandparents counterpart.

Proposition 5 *At the equilibrium, a senior worker reduces labor supply, h_2^{t*}, after grandchildren are born relative to the case in which grandchildren are not born if the marginal utility of time spent with children when there are no grandchildren is larger than $u_{l2}^t + (w_2^t + \phi^t - p^t n^t)u_{c2}^t$*

Proof. See Appendix. ■

Propositions 4 and 5 are the model implications we want to bring to data in order to explain the fall in labor supply after age 50.

4 Empirical analysis

Our final goal is to go as far as we can in understanding the determinants of the decline in hours after age 50 despite a non-declining wage profile. In particular, based on the model we suggested, we want to evaluate the role of grandparenting and health deterioration.

Define period 2 in the model as the portion of the life cycle after age 49, $\{a : a \geq 50\}$. Also, define $b_{i,a}$ as an indicator that takes value 1 if individual i is in bad health at age a and zero otherwise, and $g_{i,a}$ as an indicator that takes value 1 if individual i is grandparent at age a and zero otherwise. Then, introducing individual characteristics $X_{i,a}$, proposition 4 and 5 suggest the following restrictions on the average effects of bad health and grandparenting:

$$\mathbb{E}[h^t_{i,a}|X^t_{i,a}, b_{i,a} = 1] - \mathbb{E}[h^t_{i,a}|X^t_{i,a}, b_{i,a} = 0] < 0, \quad (25)$$

$$\mathbb{E}[h^t_{i,a}|X^t_{i,a}, g_{i,a} = 1] - \mathbb{E}[h^t_{i,a}|X^t_{i,a}, g_{i,a} = 0] < 0. \quad (26)$$

These average effects can be evaluated by estimating the following model (we can omit the cohort index t because we are focusing on a single cohort):

$$h_{i,a} = \beta_0 + \beta_1 g_{i,a} + \beta_2 b_{i,a} + \beta_3 x_{i,a} + \theta_i + \varepsilon_{i,a}, \quad (27)$$

where $x_{i,a}$ are observable covariates and θ_i individual fixed effects. Identification of causal effects is achieved under the following assumptions. As for health shocks, health deterioration can be considered exogenous to the extent that there is no correlation between working a lot (e.g. above the median) early in the life cycle and poor health later on. There is no such correlation in our data.

As for having grandchildren, this is of course more tricky. Conditional on having children, having grandchildren is at least partly an exogenous event because it is determined by fertility choices of someone else. Of course individuals who have a strong desire to become grandparents can exert pressure on their children to induce them to have children themselves. This is not much of a problem, because we can condition on fixed effects and wash out the unobserved heterogeneity in preferences. What is more problematic is the fact that a couple may decide to have children in anticipation of grandparents reducing labor supply. Although in a Nash equilibrium this cannot happen (would-be grandparents cannot commit, before becoming grandparents, to work less after their grandchildren are born), this remains a potential problem. In order to check how severe the ensuing bias may be we use unexpected pregnancies as an instrument for the event of being a grandparent. Unexpected pregnancies are proxied in our data by single-parent households at the time a child is born. That is, we assume that if a grandchild is born from a parent that lives alone, then a senior worker becomes grandparent exogenously.

We compare individuals who have children and who are older than 49 years old—the reason is that we want to explain the decline in hours after age 50. Table 1 compares the characteristics of the two groups. The P-value of the t-test of equality of means shows that the two groups are indeed different. Less educated and relatively low-income individuals are more likely to be grandparents. This clearly reflects differential patterns of fertility across the two groups.

Table 1. Characteristics of "treated" and "control" group.

	With grandchildren (n=147)		Without grandchildren (n=344)		P-value
Variable	Mean	Std. Dev.	Mean	Std. Dev.	
age	56.09	3.90	55.15	3.63	0.01
highschool	0.41	0.49	0.39	0.49	0.75
collegemore	0.29	0.45	0.42	0.49	0.01
familysize	2.30	0.74	2.44	0.91	0.09
moved	0.08	0.27	0.08	0.28	0.92
health	2.67	0.94	2.09	0.93	0.00
familyincome	40920.8	17104.8	54334.1	37943.5	0.00

Our results are reported in Table 2. Fixed effects estimates consistently (across specifications) deliver the result that grandfathers reduce labor supply by about 300 hours on average between 50 and 62 years old, about 8 workweeks. On the other hand, an unhealthy individual reduces (insignificantly) labor supply by about 100 hours relative to a healthy individual, about 3 workweeks. The IV estimate, using having a son or daughter that is single parent as instrument for being grandparent as described above, confirms this result, except that the standard error explodes due to the limited variability of the instrument. Second, in order to check the robustness of our results we use being father of a single mother (something that is highly correlated with unexpected pregnancies) as an instrument for the event of becoming grandparent. In addition to this check, the robustness of our results to inclusion or exclusion of covariates is a telltale sign of the substantial exogeneity of becoming grandparents.

What do these results suggest about the decline in hours? Since 1/20 of individuals in our sample experience worse health after age 50, worsening health explains a mere 1.5% of the 350 hours decline in labor supply (100*.05/350). On the other hand, since 2/5 have grandchildren after age 50, grandparenting explains as much as 35% of such decline (300*.40/350).

Table 2. Results

	Dependent variable: hours		
	(1)	(2)	(3)
	FE	FE	FE
grandparent	-314.562**		-315.709**
	[115.681]		[116.734]
disability		-80.223	-83.046
		[75.264]	[76.600]
age	135.988	108.467	143.311
	[156.743]	[158.720]	[157.859]
age2	-145.377	-124.624	-151.116
	[141.010]	[142.908]	[141.869]
Constant	-782.875	17.753	-998.421
	[4,333.460]	[4,386.912]	[4,366.595]
	(4)	(6)	(5)
	FE	IV	OLS
grandparent	-279.371*	-233.947	39.231
	[116.255]	[949.723]	[49.670]
disability	-114.119	-114.109	-227.945**
	[76.630]	[83.588]	[72.304]
age	144.936	139.422	112.93
	[157.459]	[185.163]	[191.723]
age2	-155.168	-150.956	-122.708
	[141.180]	[156.607]	[172.336]
familysize	-60.388	-60.532	31.269
	[33.978]	[44.765]	[34.083]
children	23.864	21.974	-193.645**
	[63.325]	[83.605]	[62.445]
familyincome	0.003*	0.003*	0.005**
	[0.001]	[0.001]	[0.001]
moved	-70.47	-70.361	-106.903
	[74.531]	[75.389]	[86.159]
Constant	-963.245	-802.696	-546.493
	[4,373.960]	[5,249.339]	[5,301.504]
Observations	585	585	585
Individuals	112	112	112
R-squared	0.1	0.09	0.1

5 Conclusion

Empirical profiles from panel data do not seem to fit widely accepted stylized facts and are inconsistent with the human capital model with depreciation and constant value of leisure. Care of grandchildren goes a long way towards explaining the progressive reduction of hours after age 50. Health is quantitatively less important. To the extent that grandparenting is driven by preferences for time spent with children the reduction in labor supply during preretirement years is efficient.

6 Appendix: Proofs

Proposition 6 *An equilibrium number of children n^{t*} exists and is unique.*

Proof. Given parenting and grandparenting time, the marginal cost (*MC*) of having children, i.e. the LHS of (22), is increasing in n^t due to the concavity of the utility function in leisure:

$$MC'(n^t) = -(x_1^t)^2 u_{ll1}^t - \beta n^{t+1}(x_2^t)^2 u_{ll2}^t > 0.$$

Similarly, given parenting and grandparenting time, the marginal benefit (*MB*) of having children, i.e. the RHS of (22), is decreasing in n^t due to the concavity of the utility function in time spent with children:

$$MB'(n^t) = (x_1^t)^2 u_{qq1}^t + \beta n^{t+1}(x_2^t)^2 u_{qq2}^t < 0.$$

Therefore, since marginal cost and benefit of children are positive, the marginal cost and marginal benefits curve cross once in the positive orthant and determine a unique equilibrium number of children, $n^{t*} : MC(n^{t*}) = MB(n^{t*})$. ■

Such equilibrium number of children is part of the Nash equilibrium in the interaction between the old and the young generations.

Lemma 7 $n^{t+1} = 0 \Longrightarrow x_2^t = 0.$

Proof. This is true by definition: time spent with grandchildren is zero if a senior worker has no grandchildren, i.e. no transfer of time can take place. ■

Proposition 8 *A Nash equilibrium in pure strategies (n^{t*}, x_2^{t-1*}) exists and in unique.*

Proof. The marginal benefit of having children (RHS of equation (22)) does not depend directly on the time transfer from the old, x_2^{t-1}. On the other hand, the marginal benefit of having children (LHS of equation (22)) increases directly with x_2^{t-1}. Therefore, the child's best response function,

$$n^t(x_2^{t-1}),$$

is strictly increasing. The concavity of the utility function implies that this best response function is itself concave. Furthermore, equation (22) and constraint (16) imply that marginal cost and benefit are strictly positive at $x_2^{t-1} = 0$ so that $n^t(0) > 0$.

Applying the implicit function theorem to equation (24), one sees that the parent's best response function,

$$x_2^{t-1}(n^t),$$

is strictly increasing. The concavity of the utility function implies that this best response function is convex.

Therefore, Lemma 2 implies the two best response functions cross once and only once. ■

Next, we establish that worsening health leads a senior worker to reduce labor supply if health deteriorates faster than the marginal value of labor income. If the weak inequality in Assumption 1 is strict then the senior worker is said to be *unhealthy*. If instead it holds with equality then the senior worker is said to be *healthy*.

Proposition 9 *At the equilibrium, an unhealthy senior worker reduces labor supply, h_2^{t*}, relative to the case in which the worker remains healthy if health deteriorates at a rate above $(w_2^t/w_1^t)(u_{c2}^t/u_{c1}^t)(1+n^t)$.*

Proof. The proof follows trivially from Assumption 1 and conditions (18) and (19). Since $u_{la}^t = -\dfrac{\partial u_a}{\partial h_a^t}$ for $a = 1, 2$, the rate at which health deteriorates is the rate at which the marginal utility of leisure increases keeping labor supply constant:

$$\frac{u_{l2}^t|_{h_2^t = h_1^t}}{u_{l1}^t},$$

which is positive for an unhealthy worker. At the optimum it must be:

$$\frac{u_{l2}^t}{u_{l1}^t} = \frac{w_2^t}{w_1^t}\frac{u_{c2}^t}{u_{c1}^t}(1+n^t).$$

Therefore, the unhealthy worker reduces labor supply if

$$\frac{u_{l2}^t|_{h_2^t=h_1^t}}{u_{l1}^t} > \frac{w_2^t}{w_1^t}\frac{u_{c2}^t}{u_{c1}^t}(1+n^t).$$

■

Finally, we establish that in this model senior workers who are grandparents work less than their non-grandparents counterpart.

Proposition 10 *At the equilibrium, a senior worker reduces labor supply, h_2^{t*}, after grandchildren are born relative to the case in which grandchildren are not born if the marginal utility of time spent with children when there are no grandchildren is larger than $u_{l2}^t + (w_2^t + \phi^t - p^t n^t)u_{c2}^t$*

Proof. Equation (24) and (21) suggests that if the benefit of grandparenting is larger than its cost in terms of leisure and consumption,

$$u_{q2}^t|_{q_2^t=0} > u_{l2}^t + (w_2^t + \phi^t - p^t n^t)u_{c2}^t,$$

then a grandparent will transfers time:

$$n^{t+1} > 0 \Longrightarrow x_2^t > 0.$$

Convexity of preferences together with the time constraint imply that such transfer is realized by reducing, at an interior equilibrium, both leisure and labor supply. Therefore, using Lemma 2:

$$\begin{aligned} l_2^{t*}|_{n^{t+1}>0} - l_2^{t*}|_{n^{t+1}=0} &< 0, \\ h_2^{t*}|_{n^{t+1}>0} - h_2^{t*}|_{n^{t+1}=0} &< 0. \end{aligned}$$

■