Human Capital and the State-Dependent Intertemporal Elasticity of Substitution

Kjersti Næss Torstensen*

PRELIMINARY

August 6, 2013

Abstract

This paper offers new insights on the intertemporal elasticity of substitution (*i.e.s.*) both at a theoretical and an empirical level. I first show that within a standard life cycle model of labor supply featuring human capital accumulation, the *i.e.s.* is a state-dependent variable which strongly depends on the return to human capital accumulation. I identify two important sources of heterogeneity w.r.t. return to human capital accumulation: age and education. Second, I argue that the average *i.e.s.* is low and comparable with micro estimates, even in the presence of human capital. Estimating the life cycle model for two different education groups; workers with and without a college degree, I find that the average *i.e.s.* in the sample is 0.34. However, the average *i.e.s.* hide important heterogeneity: for college graduates the *i.e.s.* more than doubles over the life cycle, whereas it increases by about 66 percent for non-college graduates. Moreover, young non-college graduates have 35 percent higher *i.e.s.* than young college graduates. Finally, I find that the welfare costs of permanent tax changes are much lower than previously found in models with human capital.

JEL Classification: E24, J24

Keywords: Intertemporal elasticity of substitution, Human capital, Life cycle, Labor supply

____

* European University Institute. E-mail kjersti.torstensen@eui.eu.

1 Introduction

Over the last decades substantial effort has been devoted to estimating the intertemporal elasticity of substitution for labor supply (*i.e.s.*). The *i.e.s.* plays a crucial role in both the public finance literature and in the real business cycle literature, as it describes how individuals shift labor supply across periods in response to tax changes and wage changes. Moreover, it governs the life cycle labor supply decision. The *i.e.s.* describes both how labor supply changes along the expected wage profile, and how labor supply responds to unexpected shifts in the wage profile.

In this paper I show, both theoretically and empirically, that the intertemporal elasticity of substitution of labor with respect to wage changes is not constant across the population nor is it constant throughout an individual's working-life. In particular, the *i.e.s.* is a state-dependent variable that depends crucially on worker's return to labor market experience, henceforth referred to as human capital. When workers accumulate human capital through labor supply, return to labor is the sum of contemporaneous wages and discounted future gains in wages. The larger is the (discounted) future gains in wages relative to contemporaneous wages, the smaller is the response of labor supply to changes in contemporaneous wages. Thus, workers with a high return to human capital accumulation will have a lower intertemporal elasticity of substitution compared to workers with a low return to human capital accumulation.

I identify two important sources of heterogeneity with respect to the return to human capital accumulation; age and education. Age, and in particular years left until retirement, is important as it determines how long workers will benefit from an increase in human capital. Moreover, it is plausible that production of human capital becomes more difficult with age. As a result, the *i.e.s.* will increase with age. Looking at cohort averages of real wage profiles taken from the Current Population Survey (CPS), from age 23 to age 45, where wages typically starts to level off, it seems evident that there are large differences in the return to human capital accumulation across education groups: real wages increases by 27 percent for individuals with less than High School, 42 percent for High School graduate, 70 percent for workers with some College education, and finally 137 percent for College graduates.[1] As the *i.e.s.* is decreasing in the return to human capital accumulation, it will decrease with education.

A second contribution of the paper is to draw attention on the sample used to estimate the *i.e.s.*. There is an important distinction to be made between voluntary and involuntary quits. This is because involuntary quits are not a choice by the worker, and should therefore not be included in the measure of labor supply when estimating the *i.e.s.*. I show that the majority of unemployment spells are involuntary; The share of voluntary transitions from employment to unemployment are less than 20 percent for young college graduates and about 15 percent for young workers without a college degree, according to monthly data on labor market transitions taken from the CPS dataset. From age 30 and onwards, the shares of voluntary quits are even lower. Moreover, one must be

[1] The importance of education depends on one reasonable assumption, namely that changes in real wages reflects changes in the value of human capital, either due to a change in the stock of human capital or in the rental rate on human capital. Then, differences in the growth of real wages reflects differences in the return to human capital accumulation.

cautious with respect to changes in labor supply due to sample selection, in particular changes in the share of students. I show that most of the sharp increase in the extensive margin, measured as no. of weeks worked during a year, between age 20 and age 25-30 is due to a significant reduction in weeks of unemployment between age 20-30 and an increase in the no. of weeks worked due to students leaving university and starting to work throughout the year. Whether or not these changes are included in the measure of labor supply matter significantly for the estimate of the *i.e.s.*. This is because the main changes in labor supply comes from the extensive margin and occurs early in life, when the increase in wages is strongest. Ideally, one should then control for all workers who are involuntary unemployed and all workers who switch from being a student to working during a survey year. However, when this is not possible, I argue that it is better to only include the intensive margin as a measure of optimal labor supply. This implies that the *i.e.s.* estimates refers to the intertemporal elasticity of the intensive-margin only.[2]

To quantify the importance of the state-dependency of the *i.e.s.* and the empirical measure of labor supply, I estimate a life cycle model with endogenous labor supply along the intensive margin, human capital accumulation through learning by doing and wage uncertainty, using cohort data on white males from CPS for workers with and without a college degree. As the *i.e.s.* is a state-dependent variable, it cannot be estimated. However, given the estimated model parameters, I use the simulated data and compute the *i.e.s.* over the life cycle for both education groups.

First, I find that the *i.e.s.* is low, and comparable with micro estimates, even in the presence of human capital accumulation. A vast micro literature has estimated the *i.e.s.* assuming that wages are exogenous to the hours decision, and typically find an estimate of the *i.e.s.* in the range $(0-0.5)$. (See e.g. MaCurdy (1981), Browning, Deaton and Irish 1985; Altonji (1986), Ham 1986; Abowd and Card 1987; Hansen and Wright 1992; Card 1994; Reilly 1994; Millard, Scott and Sensier 1999; Ham and Reilly (2002); French (2005)).[3] However, as first noted by Heckman (1976), the assumption of exogenous wages may bias the estimate of the *i.e.s.* downwards. Imai and Keane (2004) and Wallenius (2011) estimate life cycle models with human capital using micro data, and find estimates of the *i.e.s.* of 3.8 and 0.7-1.3, respectively.[4] Both Imai and Keane and Wallenius target annual labor supply, which implicitly assumes that all non-employment spells are voluntary.[5] The use of labor supply measure is important: using a similar methodology but targeting only the intensive margin of labor supply, I find an estimate of the *i.e.s.* for soon-to-retire-workers of 0.45.[6] This

[2] See e.g. Chetty et al. (2011) for a discussion of the difference between estimates of the *i.e.s.* for the intensive and extensive margin.

[3] See Pencavel (1986) for a survey of this literature.

[4] The papameter estimated by Imai and Keane (2004) and Wallenius (2011) refer to the *i.e.s.* with respect to changes in the total return to labor, and not only wages. Alternatively, it is the *i.e.s.* for the workers with zero return to human capital accumulation, which in their models is the workers soon-to-retire.

[5] In order to minimize the effect of students, Imai and Keane (2004) use only observations for workers the year after they last reported to be students. However, as the labor supply is asked in retrospect "hours worked last calender year", Imai and Keane still include a large share of students in their sample and will have an increase in their labor supply measure as students graduate and start working.

[6] The intertemporal elasticity of substutution estimated in the model represents two measures: (*i*) the elasticity with respect to changes in the total return to labor for all workers (labeled "opportunity cost of time" by Keane (2011)), and (*ii*) the elasticity with respect to wage changes for individuals with zero return to human capital accumulation,

estimate is however likely to be biased downwards, as some of the changes in the extensive margin are voluntary and related to wage changes. But, given the large share of involuntary unemployment it seems reasonable to assume that the upward bias on the *i.e.s.* of targeting annual hours is considerably larger than the downward bias of only targeting the intensive margin. Moreover, it serves as an important benchmark for the *i.e.s.* of the intensive margin in model with human capital accumulation.

Second, I find that there is a large heterogeneity across workers with respect to the return on human capital, and thus a large heterogeneity *w.r.t.* the intertemporal elasticity of substitution. Whereas workers soon-to-retire has an *i.e.s.* of about 0.45, the estimate of the *i.e.s.* for workers aged 25 is only 0.27 and 0.20 for non-college graduates and college graduates, respectively. Thus, the estimate of the intertemporal elasticity of substitution of labor for college graduates aged 25 is less than half of the estimate for workers soon-to-retire. And, young non-college workers have 35 percent higher *i.e.s.* than young college educated. Using population weights calculated from the CPS, and a share of 1/3 college graduates, the average *i.e.s.* in the population is 0.34. Thus, the average elasticity is well below the estimate of the *i.e.s.* for workers soon-to-retire, which is identical to the estimate of the (inverse) curvature of the marginal disutility of labor. An estimate of the average *i.e.s.* of 0.34 is within the range of what is obtained in the micro literature, using a comparable sample with respect to age and education composition.

The structural estimates of the intertemporal elasticity of substitution by different age and education groups are novel. However, previous studies have argued that the response of labor supply to temporary wage fluctuations would differ across age, see Shaw (1989), Imai and Keane (2004) and Keane (2011). Shaw estimates a life cycle model with human capital and argues that the elasticity is increasing with age. However, due to the choice of utility function she cannot recover the *i.e.s.*. Imai and Keane compute the uncompensated response of hours to a temporary shock to wages, and find that it increases in age. However, they do not control for the income effect which is also age-dependent and likely to be stronger than in standard models as transitory wage changes have persistent effects on wages through changes in human capital accumulation, see Keane (2011) for a detailed discussion of the "persistence effect". Moreover, neither of these studies consider educational differences.[7]

Imai and Keane (2004) and Wallenius (2011) find that the estimation bias of the excluding human capital is large. Imai and Keane provides several estimates using different reduced form regressions on simulated data, excluding human capital: Comparing results using the same age group in both the structural and the reduced form estimation, they find that estimate of the *i.e.s.* falls from 3.8 in the structural estimation to between $0 - 0.5$ using reduced form regressions. Wallenius finds that

which (in the model) is workers who have few periods left in the labor market. I will refer to this measure as the *i.e.s.* for soon-to-retire-workers.

[7]Moreover, for the interpretation of the utility parameters estimated, there is an important distiction to be made between the finding that the uncompensated response in hours increases with age and the finding that the *i.e.s.* increases with age. When the *i.e.s.* is state-dependent, the average *i.e.s.* in the sample considered is no longer given by the (inverse) curvature of the utility function w.r.t. to labor. This is important to take into account when comparing estimates of the *i.e.s.* across different models.

the estimate more than doubles when taking human capital into account. Although the estimated intertemporal elasticity of substitution is estimated to be small in this paper, I still find that the bias of excluding human capital is large in a simple Monte Carlo experiment on simulated data. Ignoring human capital yields estimates of the *i.e.s.* close to zero. Not surprisingly, I also find that the bias of excluding human capital varies significantly with the sample considered: The bias of excluding human capital increases in both the share of young workers and in the share of college graduates, as the return to human capital is decreasing in age and increasing in education. Including only older workers, the bias is only 7%, compared to a bias of 400% when using the whole sample. This implies that in order to obtain unbiased estimates of the *i.e.s.* using reduced from techniques, one should use data on older workers. However, if one is interested in the average *i.e.s.*in the total population, or the *i.e.s.* for particular age or education groups, such an estimate is of limited use. At best, it as an upper bound on the *i.e.s.* for the total population. Thus, to fully take into account the fact that wages are endogenously related to previous labor supply when estimating the *i.e.s.*, one must use a structural model and then solve for the *i.e.s.* of the group(s) of interest.

The state-dependency of the *i.e.s.* and the difference in estimation bias across different samples in terms of age and education indicate that one should be cautious when comparing estimates of the intertemporal elasticity of substitution obtained from different samples with respect to e.g. the age and education distribution. Moreover, it provides a theoretical explanation for why studies find that the labor supply elasticity increases with the wealth distribution (see e.g. Ziliak and Kniesner (1999)) and increases with age and decreases with education (see e.g. Blau and Kahn (2007)). Analyzing the life cycle labor supply effects of taxes, Ziliak and Kniesner find that the intertemporal elasticity of substitution increases by about 39 percent from the lowest wealth quartile to the top quartile.[8] This is consistent with the findings in this paper, as wealth and the *i.e.s.* both increase over the life cycle. Looking at the wealth quartiles of the simulated data from the model, which is however not matched to real data, I find that the *i.e.s.* increases by 32 percent from the lowest to the top quartile. Blau and Kahn estimate labor supply elasticities for married women, and analyze the elasticity separately for different age groups and different education groups. They find that the intertemporal elasticity of substitution increases sharply with age, and falls with education, which is consistent with a model of human capital accumulation of the sort presented in this paper.[9]

Keane (2011) finds that welfare costs associated with permanent tax changes are large, once the effects of human capital are properly taken into account. Whereas human capital dampens the effect of transitory tax changes for young workers, because it does not affect the return to human capital, Keane finds that human capital amplifies the effect of permanent tax changes through a "snowball-effect". A decline in current labor supply reduces (expected) future wages through

[8]Note that increase in the *i.e.s.* over wealth quartiles most likely is too large. I find that the negative bias of excluding human capital is largest for young workers, which typically also are wealth-poor workers. Thus, as the estimation bias shrinks with age, it most likely also shrinks with wealth quartiles. Thus, the increase in estimates of the *i.e.s.* over wealth quartiles is overstated.

[9]Note that the estimates of Blau Kahn are however likely to be biased to zero, as they do not take into account that wages are not exogenous to the labor supply decision. And, in light of the findings in this paper, the estimation bias for younger workers and higher educated workers are larger.

a lower production of human capital. Lower future wages reduces future labor supply, and thus further lowers future wages. Combined with a high intertemporal elasticity of substitution, Keane finds that this snowball-effect is large, and the welfare costs of permanent tax changes are large, and more than 4 to 7 times larger than what is found in conventional models excluding the effects of human capital. Performing a similar exercise as Keane, but using the model estimated below, I find that human capital do amplify the effect of permanent tax changes, but the amplification effect is small. In response to a 5 % permanent tax increase, life time consumption is reduced by 1.9 percent and 2 percent for non-college workers and college workers, respectively. In a model without human capital, the corresponding consumption loss is 1.5 percent for both education groups. The amplification effect of human capital is much smaller due to both a much lower intertemporal elasticity of substitution, but also due to a better model fit. As shown by Wallenius (2011), the out-of-sample fit of the model by Imai and Keane, which is above age 36, is poor. Hours and wages fall to sharply towards zero. With decreasing returns to scale, the level of hours matters for the snowball effect: the lower is the labor supply, the larger is the production of human capital per unit of labor supply, and the larger is the corresponding reduction in the production of human capital due to a tax change.

The remainder of the paper is organized as follows. Section 2 presents the life cycle model. Section 3 describes the data used in the estimation, and section 4 explains the estimation strategy and present the results. In section 5, the estimation bias of the *i.e.s.* in discussed in more detail, and section 6 provides some robustness checks. Section 7 examines the welfare costs of tax changes in the model estimated in section 4, and finally, section 7 concludes.

2 Model

The model is a discrete-time life cycle model for a finitely lived agent facing human capital uncertainty. I use the deterministic model of Wallenius (2011), and add wage/human capital uncertainty in the same way as Imai and Keane (2004). The agent is endowed with one unit of time at each date, and retirement is exogenously imposed at a given age. For simplicity, I omit age subscripts on the variables since there is a one to one relationship between time, t, and age, a. The agents are subject to an idiosyncratic shock to the stock of human capital at the beginning of each period, and must have non-negative assets at death.

The agent has separable preferences over consumption, c, and hours worked, n, given by

$$\mathbb{E}_t \sum_{t=0}^{T} \beta^t \left[U(c_t) - v(n_t)\right], \tag{1}$$

where β is the discount factor and $\mathbb{E}_t$ is the expectation operator, where the expectation is taken with respect to future human capital shocks. U is a concave function, v is a convex function, and they are both continuous and twice differentiable. A working age agent faces an intertemporal budget constraint

$$c_t + k_{t+1} \leq (1+r)k_t + w_t n_t,$$

where k_t is the asset holdings at time t, r is the interest rate on assets, and w_t is the hourly wage at time t.[10] The observed wage per hour at age t is defined as the product of the stochastic human capital stock $\varepsilon_t h_t$ and the rental rate on human capital p_t^h :

$$w_t = p_t^h \varepsilon_t h_t, \tag{2}$$

where the shock to the human capital stock, ε_t, is assumed to be a mean one, *i.i.d.* stochastic process with a log-normal distribution,

$$\ln(\varepsilon_t) \sim N\left(-\frac{1}{2}\sigma^2, \sigma^2\right).$$

The shock to the human capital stock may be thought of as a shock to the depreciation rate of human capital. Not that although the shock is *i.i.d.*, the effect on future wages is persistent through the human capital channel: a high shock today increases labor supply and human capital accumulation, and thus increases future wages.

The choice of introducing uncertainty into the model may have important effects on the estimate of the *i.e.s.*. If workers are risk averse, the human capital shock will decrease the expected utility weighted return of human capital, as low realizations of the shock gets a higher weight than high realizations. This will decrease the return to labor. However, if asset markets are incomplete, labor supply may act as an insurance devise against idiosyncratic risk which increases the return to labor. First, in order to smooth consumption, workers may increase labor supply in response to a low human capital shock today.[11] Second, as workers accumulate human capital while working, labor supply also acts as an insurance devise against low shocks to human capital in the future. As young workers are typically also poor workers who need insurance the most in order to smooth consumption, they have a stronger incentive to increase labor supply today in response to a bad shock and to insure against future wage shocks. This increases the expected utility weighted return of human capital. Moreover, if asset markets are incomplete, uncertainty about future wages will lead to precautionary savings for poor workers. Thus, the return to labor will increase for these workers, as their marginal utility of consumption increases. As young workers are typically poor workers, they have a stronger precautionary savings effect. Thus, the effect due to incomplete markets increases the return to labor, and more so for young workers. The total effect is unclear: whether or not uncertainty makes the life cycle profile for the total return to labor, and thus optimal labor supply profile, steeper or flatter is determined by the parameters. A steeper (flatter) profile for the optimal labor supply results in a lower (higher) estimate of the *i.e.s.*, compared to the model without uncertainty.

[10] The price of consumption is normalized to 1.

[11] Pijoan-Mas (2006) finds that households use work effort extensively as a mechanism to smooth consumption when facing uninsurable idiosyncratic productivity risk.

Following Imai and Keane (2004) and Wallenius (2011) I assume that the rental rate on human capital is constant for all ages and normalizes it to one for all time periods. This implies that the wage rate equals the current stock of human capital, $w_t = \varepsilon_t h_t$.

I assume that human capital accumulation is limited to learning-by-doing, or equivalently, is a by-product of market work. The more time a worker spends in the market, the higher is his accumulation of human capital. Human capital evolves according to

$$h_{t+1} = (1-\delta)\varepsilon_t h_t + G(n_t, \varepsilon_t h_t, a_t), \quad 0 \leq \delta \leq 1, \tag{3}$$

where δ is the rate of depreciation and $G(n_t, \varepsilon_t h_t, a_t)$ is the production function of human capital, which is concave, twice differentiable and increasing in n_t and $\varepsilon_t h_t$. Age is included in the function to allow for different human capital production at different ages.

Workers maximize life time utility with respect to consumption, labor supply, asset holdings and human capital accumulation. The value function is given by

$$V_t(k_t, h_t, \varepsilon_t) = \max_{c_t, k_{t+1}, n_t, h_{t+1}} \left\{ u(c_t) - v(n_t) + \beta \mathbb{E}_t V_{t+1}(k_{t+1}, h_{t+1}, \varepsilon_{t+1}) \right\},$$

and maximized subject to

$$\begin{aligned}
c_t &\leq \begin{cases} (1+r)k_t - k_{t+1} + \varepsilon_t h_t n_t, \ t < \hat{T} & \\ (1+r)k_t - k_{t+1} + \bar{P}, & \text{if} \ \ t \geq \hat{T} \end{cases}, && (4)\\
h_{t+1} &= (1-\delta)\varepsilon_t h_t + G(n_t, \varepsilon_t h_t, a_t), \ \ t < \hat{T} && (5)\\
k_{T+1} &= 0, && (6)\\
0 &\leq n_t \leq 1, \ \forall t < \hat{T}, \ n_t = 0 \ \forall t \geq \hat{T} && (7)\\
c_t &\geq 0, \ k_0, h_0 \text{ given} && (8)
\end{aligned}$$

where $\bar{P}$ is the pension benefit, and $\hat{T}$ is the exogenously imposed retirement age.

Let λ_t denote the Lagrange multiplier on the age t budget constraint, and μ_t the multiplier on the age t human capital accumulation constraint. The first order conditions are

$$\begin{aligned}
c_t &: \ u'(c_t) = \lambda_t, \ \ \forall t && (9)\\
k_{t+1} &: \ \lambda_t = \beta \mathbb{E}_t \frac{\partial V_{t+1}(k_{t+1}, h_{t+1}, \varepsilon_{t+1})}{\partial k_{t+1}}, \ t < \hat{T} && (10)\\
n_t &: \ v'(n_t) = \lambda_t \varepsilon_t h_t + \mu_t \frac{\partial G(n_t, \varepsilon_t h_{t,} a_t)}{\partial n_t}, \quad t < \hat{T}, \ \varepsilon_t = 1 && (11)\\
h_{t+1} &: \ \mu_t = \beta \mathbb{E}_t \frac{\partial V_{t+1}(k_{t+1}, h_{t+1}, \varepsilon_{t+1})}{\partial h_{t+1}}, \ t < \hat{T} && (12)
\end{aligned}$$

Using the envelope conditions to substitute for the derivatives of the value function yields the following equations

$$
\begin{aligned}
u'(c_t) &= \beta(1+r)\mathbb{E}_t u'(c_{t+1}), && (13)\\
v'(n_t) &= u'(c_t)\varepsilon_t h_t + \mu_t \frac{\partial G(n_t, \varepsilon_t h_t, a_t)}{\partial n_t}, && (14)\\
\mu_t &= \beta \mathbb{E}_t \left(\left((1-\delta)\varepsilon_{t+1} + \frac{\partial G(n_{t+1}, \varepsilon_{t+1} h_{t+1}, a_{t+1})}{\partial h_{t+1}} \right) \mu_{t+1} + \varepsilon_{t+1} u'(c_{t+1}) n_{t+1} \right). && (15)
\end{aligned}
$$

Equation (13) is the familiar consumption Euler equation. Equation (14) equates the marginal disutility of labor to its marginal benefit. The first term on the *r.h.s.* is the effect on current utility from higher labor income. The second term is the learning effect; working an additional hour increases human capital and hence future wages. Multiplying the learning effect with the shadow price on skill, μ, gives the marginal utility of increased human capital. Equation (15) is the law of motion for the shadow price of skill.

Equation (14) demonstrate why the *i.e.s.* is no longer a constant, but a state-dependent variable: Due to human capital accumulation, the return to labor is a function of both the contemporaneous wage and discounted future gains in wages. As transitory wage shocks only affects the contemporaneous term, workers with a high future gains in wages (in relative terms) will respond less to transitory shocks compared to workers with low gains in future wages.

Let's define the *i.e.s.* as the percentage change in hours over time in response to an exogenous unanticipated temporary change in the rental rate on human capital. And, assume that the disutility of labor and the production function of human capital takes the following functional forms

$$
\begin{aligned}
v(n_t) &= b\frac{n_t^{\frac{1}{\gamma}+1}}{\frac{1}{\gamma}+1}, \qquad b \geq 0,\ \gamma > 0,\\
G(n_t, \varepsilon_t h_t, a_t) &= A\exp(-d a_t)\varepsilon_t h_t n_t^{\alpha}, \quad 0 \leq \alpha \leq 1,\ A, d \geq 0.
\end{aligned}
$$

Then, the *i.e.s.* is defined as[12]

$$
\frac{d\left(\frac{n_t}{n_{t+1}}\right)}{dp_t^h}\frac{p_t^h}{\frac{n_t}{n_{t+1}}} = \gamma \frac{1}{\left(b n_t^{\frac{1}{\gamma}-1} - \gamma \mu_t \frac{\partial^2 G_t}{\partial n_t^2}\right)} \left(u'(c_t)\varepsilon_t h_t + \frac{\partial G_t}{\partial n_t}\frac{\partial \mu_t}{\partial p_t^h}\right)\frac{p_t^h}{n_t}\left(1 - \frac{\partial n_{t+1}}{\partial h_{t+1}}\frac{\partial G_t}{\partial n_t}\frac{n_t}{n_{t+1}}\right). \quad (16)
$$

Without human capital accumulation $\frac{\partial G_t}{\partial n_t}$ is zero, and the *i.e.s.* is equal to γ. With human capital the *i.e.s.* possibly depends on all state variables; human capital, asset position, the shock to the human capital stock and age, and all parameters of the model.

Let's analyze how the *i.e.s.* will evolve over the life cycle. For all parameter values considered below, the product $\frac{\partial G_t}{\partial n_t}\frac{\partial \mu_t}{\partial p_t^h}$ is very close to zero, and so is $\frac{\partial n_{t+1}}{\partial h_{t+1}}\frac{\partial G_t}{\partial n_t}\frac{n_t}{n_{t+1}}$. This implies that the

[12] For a more detalied calculation of the *i.e.s.*, see the appendix.

importance of the state-dependency of the *i.e.s.* does not come from that fact that it alters also future labor supply decisions through the human capital effect. Instead, the importance of human capital in making the *i.e.s.* state-dependent is to change the "pass-through" of wage changes. As wages is only a faction of the return to labor, and this fraction changes with age, so does the labor supply elasticity. In order to simplify the expression, I therefore impose that $\frac{\partial G_t}{\partial n_t}\frac{\partial \mu_t}{\partial p_t^h} = \frac{\partial n_{t+1}}{\partial h_{t+1}}\frac{\partial G_t}{\partial n_t}\frac{n_t}{n_{t+1}} = 0$, such that the expression becomes

$$\frac{d\left(\frac{n_t}{n_{t+1}}\right)}{dp_t^h}\frac{p_t^h}{\frac{n_t}{n_{t+1}}} = \gamma\frac{1}{\left(1+\frac{1}{u'(c_t)p_t^h\varepsilon_t h_t}\mu_t\frac{\partial G_t}{\partial n_t}\left(1-\gamma(\alpha-1)\right)\right)}. \tag{17}$$

As shown below, the shadow price on skills, μ_t, is decreasing over the life cycle as the working life horizon shrinks, contributing to an increase in the intertemporal substitutability of labor over the life cycle. The same is true for the slope of the production function of human capital, $\frac{\partial G_t}{\partial n_t}$, both due to decreasing returns to scale in the production of human capital and because learning becomes more difficult with age ($d > 0$). Moreover, due to consumption smoothing, the increase in wages is larger than the increase in consumption over the life cycle, so that $\frac{1}{u'(c_t)\varepsilon_t h_t p_t^h}$ is decreasing over the life cycle. Again, this contribute to an increasing *i.e.s.* over the life cycle. Hence, the *i.e.s.* will be increasing over the life cycle, as contemporaneous wages becomes relatively more important for the workers, and will approach γ, which is the *i.e.s.* of the workers in the last period before retirement (when $\mu_{\hat{T}} = 0$).

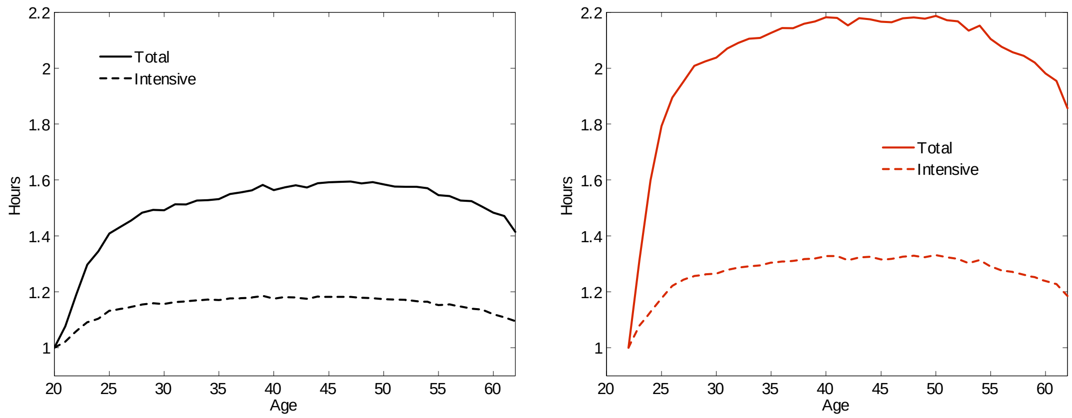
The larger is the fall in the shadow price on skills over the life cycle, the larger is the increase in the *i.e.s.* over the life cycle, and the further apart is the average intertemporal elasticity of substitution of labor from the value of γ. As more education is associated with a higher return to human capital for young workers, and thus larger fall in the shadow price on skills over the life cycle, the *i.e.s.* should increase more for higher educated workers. The state-dependency of the *i.e.s.* i further explored below, in section X.

Note that the response of labor supply to a human capital shock, ε_t, would be different than the response to changes in the rental rate, as ε_t changes the whole future path of expected wages and then also the expected return on current human capital accumulation.

3 Data

The choice of using data from the CPS is motivated by the findings in Wallenius (2011); she shows that the life cycle profile implied by the estimates found in Imai and Keane (2004), using a life cycle profile based on individuals aged 20-36 years, have poor out of sample fit and stress the importance of using long life cycle profiles, which is possible using the CPS. I use monthly CPS data from 1979 to 2008, and construct an average life cycle profile for wages and labor supply for individuals aged 25 to 60 years.[13] As shown below, there are large differences in hours and wages across education

[13] The data are from CEPR, Center for Economic and Policy Research. 2006. CPS ORG Uniform Extracts, Version 1.5 . Washington, DC



No College degree   College degree

Figure 1. Life cycle patterns for age averages of different measures of hours for worker with and without a college degree, normalized to 1 at age 20 and age 23.

levels. The main educational difference is between those with and those without a college degree. I therefore choose to split my sample in workers with a college degree and workers without a college degree. To make my results comparable with most of the literature, I limit my sample to white males.[14] I use only data on individuals which are in the labor force and have well defined hours and wage data.[15]

Unlike both Imai and Keane (2004) and Wallenius (2011), I start my sample at age 25 instead of at age 20, and I use monthly data on "usual weekly hours" instead of annual data on "usual weekly hours" multiplied by "weeks worked". This is motivated from the desire to find the best measure of labor supply, when estimating the *i.e.s.*. The life cycle profile for "usual weekly hours", labeled "intensive", and "usual weekly hours x weeks worked", labeled "total", are plotted in figure 1 for workers with and without a college degree.[16] Given that wages exhibits its largest increase in the period between age 20 and age 30, the choice of labor supply measure matters for the estimate of the *i.e.s.*: the larger is the increase in the optimal labor supply, the stronger is the relation between wages and hours, and the higher is the estimate of the intertemporal elasticity of substitution. The natural question to ask is then: what is behind this sharp increase in the extensive margin, i.e. "weeks worked", plotted in figure 2.

The large increase in the no. of weeks worked is due to a decline in the average number of weeks workers are unemployed, see figure 3(a), and a sharp increase in the participation rate, *i.e.* a sharp decline in the no of weeks out of the labor force, see figure 3(b).

Let's first analyze the nature of unemployment. Using monthly CPS data on labor market transitions from 2001 to 2007, which span one business cycle defined by the NBER, I compute the share

[14] I also exclude individuals that are students, have served in the military and self employed workers.

[15] I include only individuals with usual weekly hours above 4 hours and below 99 hours.

[16] Data from: Miriam King, Steven Ruggles, J. Trent Alexander, Sarah Flood, Katie Genadek, Matthew B. Schroeder, Brandon Trampe, and Rebecca Vick. Integrated Public Use Microdata Series, Current Population Survey: Version 3.0., Minneapolis: University of Minnesota, 2010.

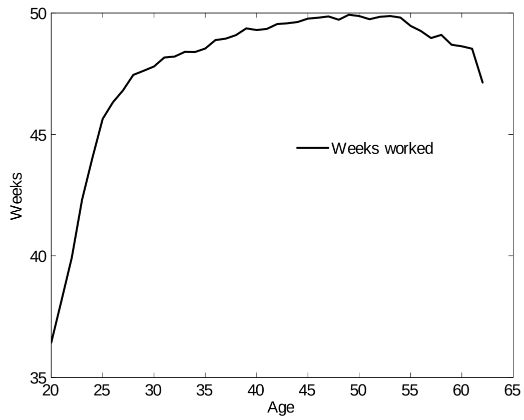
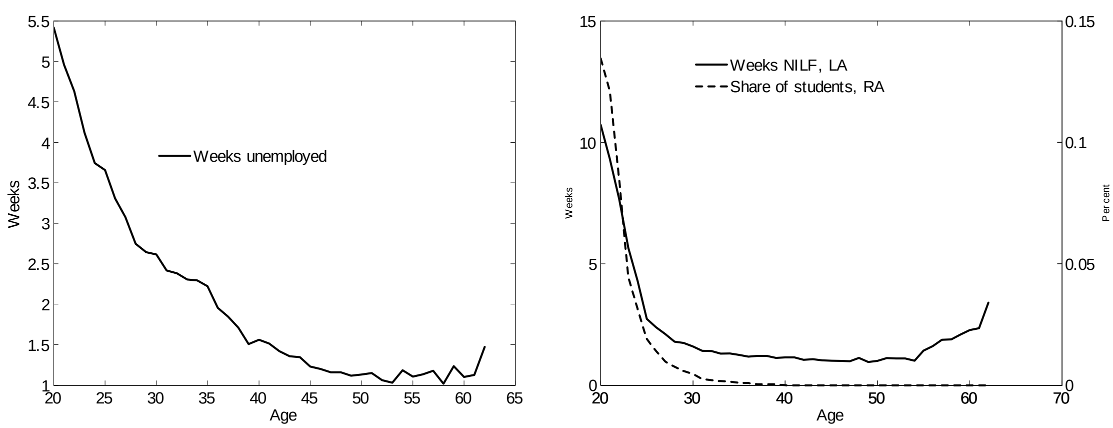


Figure 2. Weeks worked, age averages.



(a) (b)

Figure 3.
(a) Weeks unemployed
(b) Weeks NILF and the share of students, age averages. Left axis (LA) measures weeks not in the labor force, and the right axis (RA) measures the share of students.

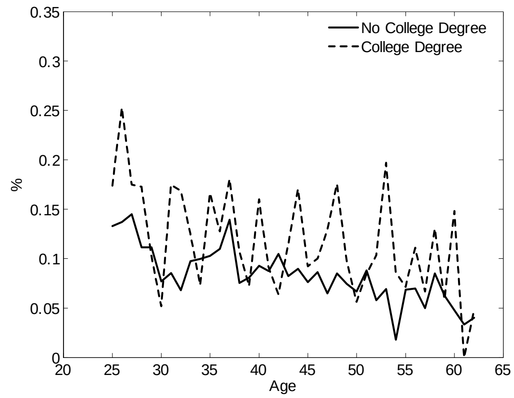


Figure 4. Job separation probability, share of voluntary quits. Age averages.

of voluntary and involuntary employment to unemployment transitions and the job finding probability of workers with and without a college degree.[17,18,19] The fraction of voluntary unemployed are slightly higher for the workers with a college degree, but the fraction are small for all ages, see figure 4.[20] Given that I cannot distinguish between voluntary and involuntary unemployment in data used in the estimation below, I argue that changes in the extensive margin due to unemployment should not be included in the measure of optimal labor supply.

Second, the sharp decrease in weeks out of the labor force goes hand in hand with a sharp decrease in the share of students in the sample, see 3 (b). This indicates that a part of the increase in the extensive margin is due to sample selection issues: students aged 20-25 work on average 21 weeks opposed to the average of non-students of 43 weeks, thus as the share of students decline the average weeks worked in the sample increases significantly. The choice of years of education, and thus the choice of low labor supply during studies, is linked to the expected wage increase over the life cycle, not just the expected wage in the years shortly after leaving university. Thus, the choice to increase labor supply the year after leaving university is not directly linked to the wage increase experienced in the same period, and should not be used in the estimation of the *i.e.s.* Due to the nature of the annual data I cannot remove students from the sample.[21] In order to minimize the sample selection issue due to students entering the labor force, I therefore start the sample at age 25, when most individuals have finished their studies.

Excluding the weeks workers are unemployed from the measure of labor supply, due to the finding

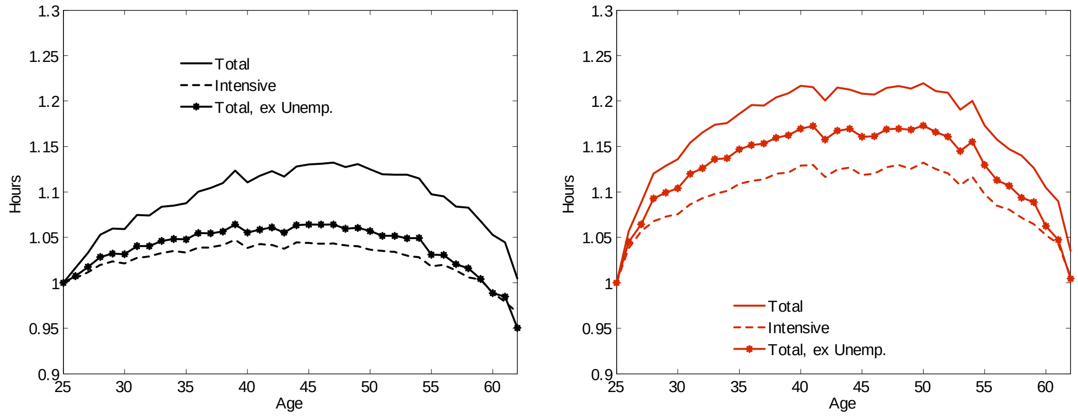
[17] Voluntary unemployed is defined as those who report that they are unemployed because they "quit job". The involuntary are those who are unemployed due to "temporary job ended", "lost job" or "on layoff".

[18] The data are collected from the NBER. I match individual transitions based on household identifiers, age, sex, race, individual line number and months in sample. Due to limitations on availiable data I cannot calculate these measures for the whole time period from 1979 and onwards.

[19] I use data that spans one business cycle due to Shimer (2007) findings that there are substantial fluctuations in the job finding probability at business cycle frequensies.

[20] I also find that there is no significant difference between the job finding probability of the two groups. Results are availiable on request.

[21] In the CPS, the student status refers to the time of the interview, whereas data on labor supply refer to the year preceding the interview. Thus, I cannot detect students that left university during the last year, and possibly had low labor supply during that year.



No College degree College degree

Figure 5. Life cycle patterns for age averages of different measures of hours for worker with and without a college degree. Normalized to 1 at age 25.

that most unemployment is involuntary, reduces the measure of labor supply towards the measure of the intensive margin, see figure 5.[22] For non-college-graduates the two measures are almost identical, and for college graduates the measure of total labor supply excluding unemployment lies inbetween the two other measures. However, once starting at age 25, the three measures of labor supply are very similar. Looking at the increase in wage dispersion over time, Lemieux (2006) finds that "... The main problem with the March CPS is that it poorly measures the wages of workers paid by the hour (the majority of the work force)", and suggests using monthly data on ORG CPS instead. To be robust towards measurement errors in the wage data I therefore use monthly data, and thus restrict the labor supply measure to only include the intensive margin. Thus, the intertemporal elasticity of substitution estimated refers only to the intensive margin. However, given that most unemployment spells are involuntary and given that the measure of total labor supply excluding unemployment is so similar to the intensive margin, I do not believe that *i.e.s.* of the total labor supply is similar to the elasticity obtained here.

In the model, the period time endowment is assumed to be equal to one. As is standard in the literature, I assume that the maximum time allotted to work is 14 hours a day 7 days a week. The fraction of hours worked by an employed individual in a given period is hence

$$hours = \frac{'usual\ weekly\ hours'}{14 * 7}.$$

The hourly wage is computed as

$$wage = \frac{'usual\ weekly\ earnings'}{'usual\ weekly\ hours'},$$

and is deflated by the CPI to make it comparable across time.

[22] The measure of total labor supply, exclusive changes in unemployment is calculated as follows $Total, exUnemp = \frac{"weeks\ worked" * "usual\ weekly\ hours"}{(52 - "weeks\ unemployed") * 14 * 7}$

The life cycle profiles for hours and wages are constructed from an average cohort. Unfortunately, the monthly data starts in 1979, and I thus do not have enough years to cover a full life cycle, from 25 to 62 years, for one cohort. I therefore take the average of several cohorts. In order to have a minimum of 5 observations at each age, I use 17 consecutive cohorts starting with those who were 21 years old in 1979, and continuing with the those who were 24 in 1979, and so forth up to those who were 37 in 1979. I take the average growth rate by age of the three variables, and combine these with the initial condition given by the average outcome of individuals aged 25 years. The profile obtained for hours and wages are very robust to changes in number of cohorts once I end my sample at age 60. I therefore limit my sample to individuals aged 25 to 60 years old.

In addition to the age effect, there is a significant time effect in the data.[23] Given that I merge different cohorts, and that the focus of this paper is on the age effect (life cycle effect), I want to remove the time effect from the data. I adjust the growth rates of all variables for all cohorts for the education specific average growth in the sample

$$\Delta\tilde{x}_{a,i,t} = \Delta x_{a,i,t} - \Delta\bar{x}_{i,t},\ \forall\ t = 1980 : 2008,\ a = 21 : 62,\ i = 1, 2,$$

where $\Delta x_{a,i,t}$ is the change in e.g. the average wage for individuals with age a, education level i, at time t, and $\Delta\bar{x}_{i,t}$ is the change in the age average of wages for individuals with education level i at time t.[24] Assuming that the rental rate on human capital evolves according to the education specific average growth rate of wages, the growth adjustment of wages can be seen as removing the changes in the rental rate on human capital over time, and hence making the data more comparable with the model.

Due to a limited sample size and possibly sampling errors, the life cycle profiles are jagged. Since I do not want the estimates to be influenced by these issues, I choose to smooth the data.[25] The life cycle profiles of wages, hours for smoothed and unsmoothed data and the variance of wages of unsmoothed data are plotted in figure 6. Both wages and hours are hump-shaped over the life cycle for both education groups.

The variance of wages are similar for both education groups at age 25, but the variance increases much less over the life cycle for non-college graduates. This is consistent with the modeling choice of a complementarity between labor and human capital in the production of human capital, which increases the variance of human capital in the level of human capital.

[23] There is potentially cohort effects present in the data, as I merge several cohort. However, analysing life cycle profiles of inequality in wages and hours Heathcote et al. (2005) find no evidence that cohort effects are important. I therefore abstract from cohort effects.

[24] I control for changes in age distributions over time, by using the same age-weighting each year.

[25] In particular I smooth the data using a Hodrick-Prescott filter with a smoothing parameter of 75 for real wages and the unemployment rate and 30 for hours.

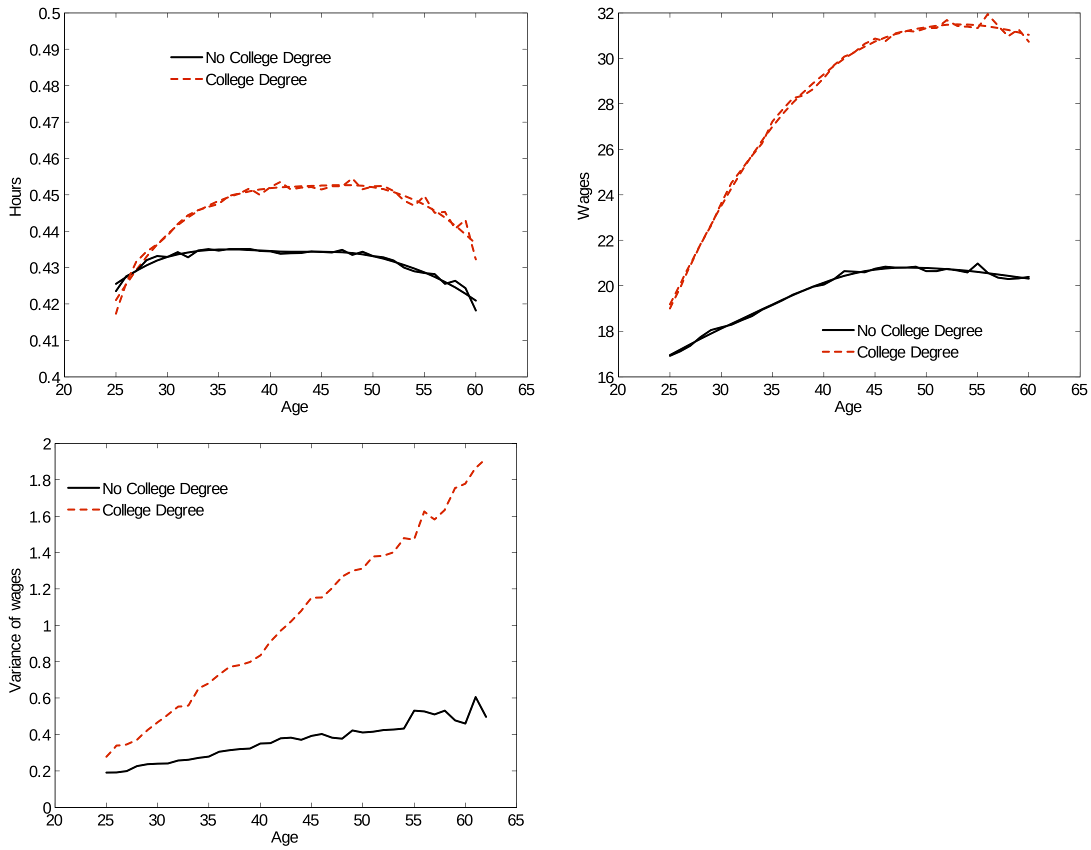


Figure 6. Wages, hours and the variance of wages. Age averages.

4 Estimation

The model is estimated using simulated method of moments, where I minimize the sum of squared percentage deviation of averages of hours and wages by age for both education groups.[26,27] The model period is set to one year.

In order to solve the model, I must specify functional forms, initial level of human capital and assets and the exogenous age of retirement and death. All individuals enter the labor force at age 25, retire at age 62, and after 15 years in retirement they exogenously die at age 77.[28] I further assume that all workers are identical in the first period, in particular the initial human capital level is set to one, $h_{25} = 1$, and the initial level of assets is set to zero, $k_{25} = 0$. Both education groups are assumed to have the same preferences, but differs in the production function of human capital. I use the following functional forms for preferences and the human capital production function, $G(\cdot)$

$$
\begin{aligned}
u(c_t) &= \ln(c_t), && (18)\\
v(n_t) &= b\frac{n_t^{\frac{1}{\gamma}+1}}{\frac{1}{\gamma}+1}, \qquad b \geq 0,\ \gamma > 0, && (19)\\
G^i(n_t, \varepsilon_t h_t, a_t) &= A^i \exp(-d^i a_t)\varepsilon_t h_t n_t^{\alpha^i}, \qquad 0 \leq \alpha \leq 1,\ A, d \geq 0,\ i \in \{NDC, CD\}. && (20)
\end{aligned}
$$

The functions for preferences and the production function for human capital are taken from Wallenius (2011). Preferences are assumed to be separable and consistent with a balanced growth path. The choice of utility function may have implications for the estimation result. In particular, since the intertemporal elasticity of substitution of consumption and labor are linked, the choice of log utility in consumption may affect the estimate of the inter-temporal elasticity of substitution of labor. The possibility of a higher risk aversion is explored in section X.

As discussed above, it is common to refer to γ as the intertemporal elasticity of substitution of labor, also when workers accumulate human capital, see e.g. Imai and Keane (2004) and Wallenius (2011). The inclusion of human capital makes the return to labor intertemporal and depending on years left to retirement and other state variables, depending on the specification of the production function of human capital. This will make the *i.e.s.* state dependent. γ will only refer to the intertemporal elasticity of substitution for those workers who have no future gain of human capital accumulation, *i.e.* those workers who will retire next period.

The production function of human capital is a Cobb-Douglas with labor and human capital as inputs, where $A \exp(-da_t)$ is the productivity parameter. The choice of complementarity between

[26] The simulated moments are constructed in the following way: I solve the model for a particular parameter guess, simulate the life cycle profile for 50000 individuals, and then compute average hours and wage profiles at a given age. I then compute the age averages of the two variables.

[27] In the weighting matrix I put double weight on the moments of labor supply, in order to obtain a good match for both variables.

[28] The age of death in the model is motivated by the life expectancy at age 20 for white males in the US, which is 76.8 years (Source: U.S. National Center for Health Statistics, National Vital Statistics Reports (NVSR), Deaths: Preliminary Data for 2008, Vol. 59, No. 2, December 2010.)

Table 1. Parameter estimates deterministic model, targeting age 25-60.

			NCD	CD
γ	0.3985	A	0.0950	0.2097
b	23.8421	d	0.0179	0.0313
β	0.9593	α	0.3109	0.3602
SSE	0.0123			

Note: NCD refers to no college degree, whereas CD refers to college degree. SSE refers to sum of squared errors.

labor and human capital in the production of human capital is due to the finding by e.g. Imai and Keane (2004) that the marginal effect of an hour on future wages is increasing significantly with current wages. If d is positive, human capital accumulation will be more demanding with age.[29]

The pension benefit is set to 40% of the average income for the average individual over the life cycle.[30] In the model, only the product $\beta(1 + r)$ matters for the labor supply. Hence, I cannot identify both, and I therefor set $r = 0.04$. In order to limit the number of parameters to estimate, I follow Wallenius (2011) and assume that the annual rate of depreciation, δ, is 3 per cent. This is in line with micro evidence on the wage costs of labor market intermittency provided by Kim and Polachek (1994), which find that the annual rate of depreciation is between 2-5 percent when controlling for unobserved heterogeneity and endogeneity issues. I further assume that the preference parameters, β, b and γ, are the same across educations groups, whereas the parameters of the human capital production function d, A, and α differs across education level.

In addition I need to specify the distribution of the shock to human capital. Due to computational constraints, I cannot estimate the shock process of the human capital shock. I therefore first estimate the model without any uncertainty, and then estimate the model using the distribution of the human capital shock estimated by Imai and Keane (2004). The first model, the model without uncertainty, is comparable to Wallenius (2011), whereas the model with uncertainty is similar to Imai and Keane's model. Hence, by comparing the results obtained below I can assess the effect of uncertainty on the estimate of the *i.e.s.*, and by comparing my results to the ones obtained by Wallenius and Imai and Keane, I can assess the difference in results stemming from a different target for labor supply.

4.1 Results, no shock

Adjusting the labor supply profile to take into account that most of the changes in the extensive margin are involuntary or unrelated to the changes in wages gives an estimate of γ equal to 0.40, which is significantly lower than the estimates obtained by Wallenius (2011) $(0.8 - 1.3)$ using the same model, see table 1. By using total annual hours as a measure for optimal labor supply, Wallenius significantly overestimate the intertemporal substitutability of labor supply.

[29] This is a simpler production function than the one studied by Imai and Keane (2004). However, Wallenius (2009) shows that the results in Imai and Keane also applies to this production function.

[30] According to French and Jones (2012), Social security replaces about 40% of pre-retirement earnings.

The model fit for hours and wages for the set of parameters estimated for the two education groups are plotted in figure 7. Apart from predicting to high labor supply among young college graduates, the model fits the targets well.

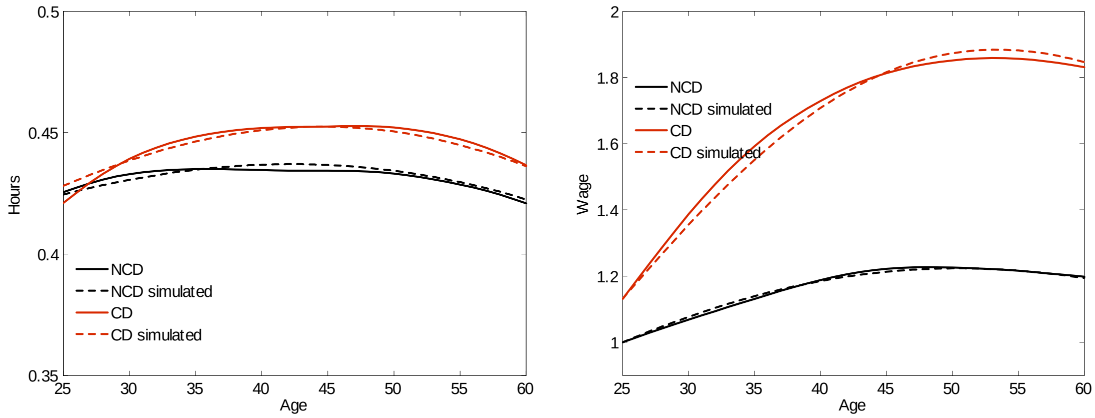


Figure 7. Model fit for hours and wages, targeting age 25-60.

As shown above, the *i.e.s.* is state-dependent. Using equation 16, one can calculate the average *i.e.s.* by age and education, using the simulated data. Figure 8 plots the average *i.e.s.* by age and education. Although the *i.e.s.* for workers close to retirement is equal to the estimate of γ, 0.40, the *i.e.s.* for young college graduates is as low as 0.21, *i.e.* only about half the value of γ. For workers without a college degree, the estimate of the *i.e.s.* for workers aged 25 is 0.28 and it increases by 41 percent over the life cycle.[31] This implies that the average intertemporal substitutability in the population is much lower than the estimate of γ obtained above. Using a 1/3 share of college graduates, and the population weights from CPS, the average *i.e.s.* in the sample is 0.32, which is 20 percent lower than γ. Unfortunatly, Wallenius (2011) does not report how the *i.e.s.* varies with age in her model, so I cannot compare.

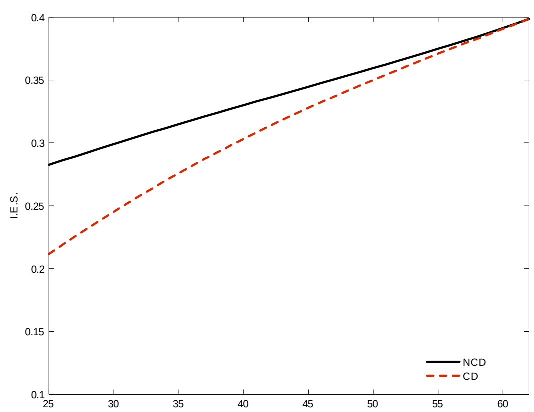


Figure 8. Intertemporal elasticity of substitution over the life cycle. Age averages of simulated data.

[31] For more details on the calculation, see appendix A.

The difference in the intertemporal elasticity of substitution comes mainly from large differences in the return on human capital, see figure 9. Non-college graduates have 33 percent higher *i.e.s.* than college graduates, and 42 percent lower return on human capital than college graduates at age 25. As one approaches retirement, the gap between the return on human capital narrows, as does the gap between the intertemporal elasticities of substitution.

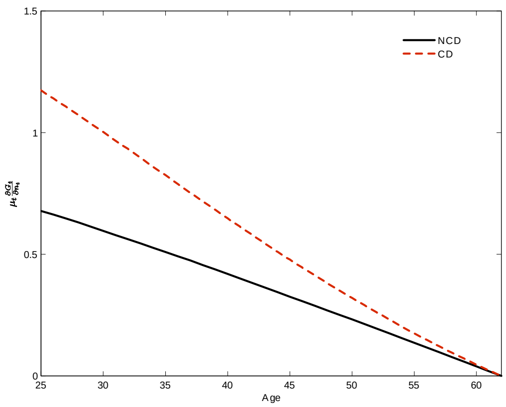


Figure 9. Returns to human capital, $\mu_t \frac{\partial G_t}{\partial n_t}$, for non-college graduates and college graduates,

4.2 Results, human capital shock

As shown above, the variance of human capital increases significantly over the life cycle, supporting the model choice of a stochastic human capital stock. Moreover, introducing uncertainty may influence the estimate of the *i.e.s.* due to the insurance role human capital then plays. As explained above, depending on the parameters, adding uncertainty may either increase or decrease the estimate of the *i.e.s.* compared to the model without uncertainty.

Due to computational constraints, I am not able to estimate the distribution of the shock. I therefore use the distribution for the same shock estimated in Imai and Keane (2004), using NLSY-data for the cohort aged 20 in 1979:

$$\ln(\varepsilon_t) \sim N\left(-\frac{1}{2}0.05781^2, 0.05781^2\right),\ t > 25$$

Further, I set the standard deviation of the initial shock, ε_{25}, so that I match the education specific distribution of wages in the data. The estimate of γ increases from 0.40 to 0.45 when uncertainty is introduced into the model, see table 2. This is however considerably lower than Imai and Keane (2004)'s estimate of 3.8, obtained using a similar model.[32]

The model fit for hours and wages are still good, see figure 10. Even though I do not estimate the distribution of the human capital shock, and I do not target the variance of the wages in the estimation, the model explains most of the variance for wages for both education groups. However,

[32] The difference in model between the one above and Imai and Keane (2004) should not influence the estimate of the *i.e.s.* significantly; using the same deterministic model as above, Wallenius (2011) obtains similar estimate to Imai and Keane (2004) when using the same age group.

Table 2. Parameter estimates stochastic model, targeting age 25-60.

			NCD	CD
γ	0.4496	A	0.0980	0.2411
b	18.6758	d	0.0188	0.0360
β	0.9592	α	0.3109	0.3171
SSE	0.0133			

Note: NCD refers to no college degree, whereas CD refers to college degree. SSE refers to sum of squared errors.

at older ages the variance levels off in the model whereas it continues to grow for college educated in the data. The variance levels off as the variance of the production of human capital decreases. With d positive, production of human capital is more difficult with age, irrespective of human capital level and labor supply. As a result, production of new human capital is very low for all older workers which together with a positive rate of depreciation of human capital slow down the increase in the wage dispersion.

Due to human capital, even a small wage shock can create a large increase in the variance of wages, and due to different production function of human capital across age groups the same shock gives rise to very different profiles for the variance of wages for the two groups.

Again, using equation 16, one can calculate the average *i.e.s.* across age and education. Introducing wage uncertainty leads to an even larger increase in the *i.e.s.* over the life cycle; for non-college graduates the *i.e.s.* increases by about 66 percent whereas it increases by 120 percent for college graduates, see figure 11. The results indicate that it is important to be cautious when comparing estimates of the *i.e.s.* obtained from different samples with respect to age and education, as there is no reason to expect that they should find the same estimate of the intertemporal elasticity of substitution.

Using population weights from the CPS and a 1/3 share of graduate students, the average *i.e.s.* in the population is 0.34. This is 24 per cent lower than the estimate of γ. Again, interpreting the estimate of γ as an estimate of the *i.e.s.*, significantly overstates the average substitutability of labor in the population. Interestingly, the only state variable that seems to matter is age: the variance of the *i.e.s.* for a given age and education level is practically zero, although the variance of e.g. consumption and human capital differs significantly.

Shaw (1989) was the first to note that the *i.e.s.* is increasing with age due to human capital accumulation, but she does not consider other state variables, and she cannot provide quantitative estimates for different ages due to her choice of preferences. Thus, I cannot compare my findings with her's. Using a similar model Imai and Keane (2004) find that the average (uncompensated) response in hours to a 2 percent unexpected wage increase grows exponentially with age; whereas hours increase by 0.6 percent for workers in their early twenties, hours increases by nearly 4 percent at age 60 and about 5.5 percent at age 65. But, as Imai and Keane do not control for the income effect, which also is age-dependent, it is difficult to compare the two measure. The income effect is particularly important in their model as temporary wage changes have persistent effects on future

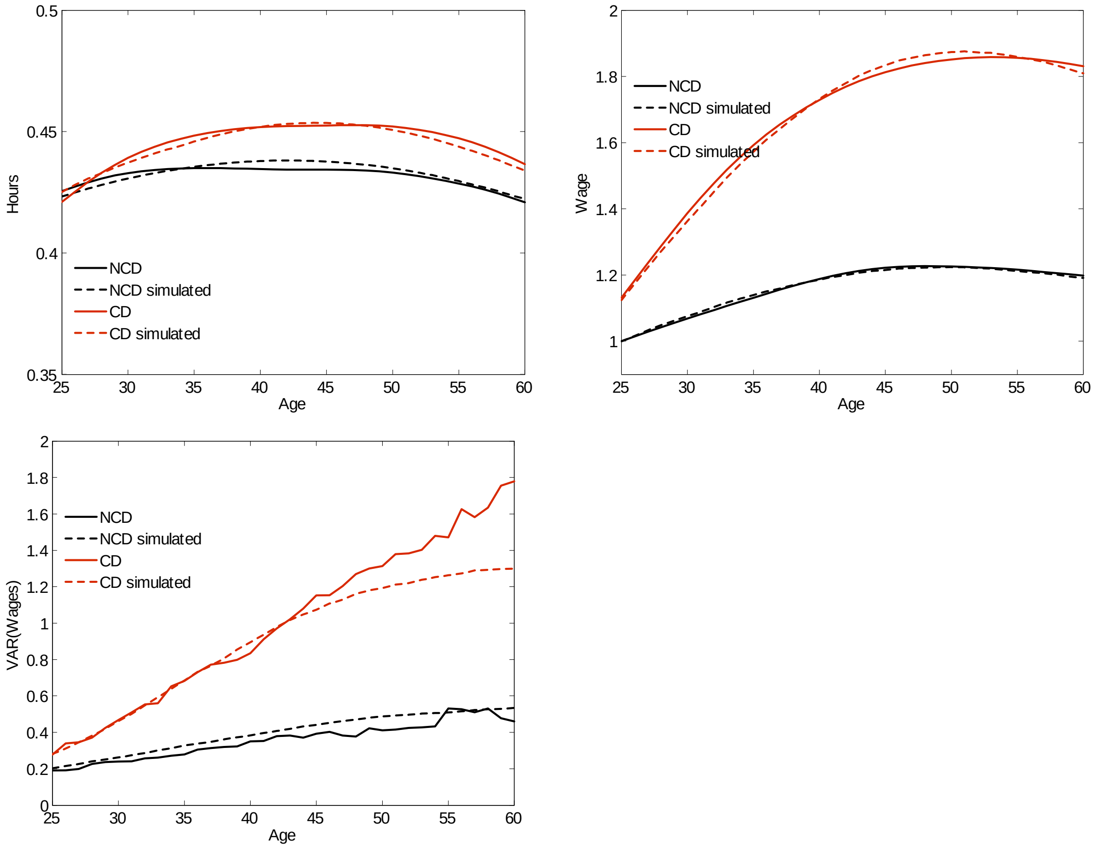


Figure 10. Model fit for hours, wages and the variance of wages with stochastic human capital, targeting age 25-60.

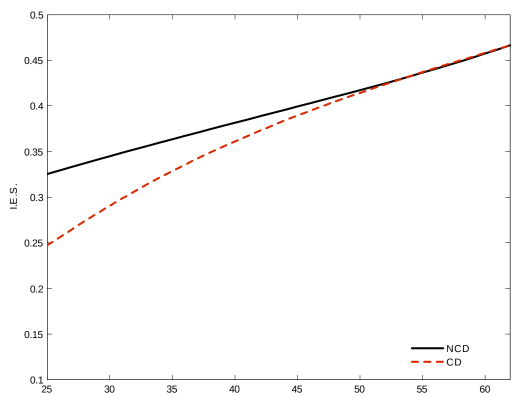


Figure 11. Intertemporal elasticity of substitution over the life cycle with stochastic human capital. Age averages of simulated data.

wages. With a higher estimate of the *i.e.s.*, labor supply reacts stronger, which affects the production of human capital and thus future wages more. Thus, transitory shocks does not only affect current labor income but also future labor income. Moreover, Imai and Keane do not distinguish between education groups, although they estimate different production functions for different educational groups. However, compared to Imai and Keane I find both a much smaller response of hours at all ages, and also, most likely, a much smaller increase in the *i.e.s.* over the life cycle.

Domeij and Floden (2006) argue that ignoring liquidity constraints will bias the estimate of the *i.e.s.* downwards. Once they exclude individuals likely to be liquidity constraint in their sample, their estimate of the *i.e.s.* more than doubles. Human capital accumulation is an alternative source of an estimation bias, but has also some implications for the bias found when ignoring liquidity constraints. If liquidity constraint individuals are mainly young workers, which I find has the lowest *i.e.s.* in the population, when excluding these workers the average *i.e.s.* in the population will increase due to the sample selection effect, and not only due to the importance of the liquidity constraint.

The state-dependency of the *i.e.s.* introduces some interesting policy implications with respect to e.g. labor taxation; young workers and in particular young college graduates will respond much less to tax changes than older workers. This implies that economies with a high share old workers will have a more tax-sensitive workforce relative to economies with a younger workforce. Also the distribution of education and or occupation may matter; the ability to accumulate human capital through learning by doing, and the value of human capital may differ significantly across education groups and occupations. Thus, economies with workers with high gains from human capital accumulation, here college graduates, will be less tax-sensitive, than economies with workers which have low gains from human capital accumulation.

5 Bias

Imai and Keane (2004) and Wallenius (2011) find that the estimation bias of the excluding human capital is large; Imai and Keane provides several estimates using different reduced form regressions, excluding human capital, on the simulated data: Comparing results using the same age group in the estimation, they find that estimate of the *i.e.s.* falls from 3.8 in the structural estimation to between $0 - 0.5$ using reduced form regressions. Wallenius finds that the estimate more than doubles when taking human capital into account. As I obtain a much lower intertemporal elasticity of substitution, it is interesting to see how this affects the estimation bias of excluding human capital.

To account for the bias, I perform a Monte Carlo experiment. I estimate a model without human capital accumulation using the simulated data from the stochastic model with human capital accumulation. More precisely, I estimate the following equation

$$\ln n_{t,i} = B_1 + B_2 \ln c_{t,i} + B_3 \ln w_{t,i} + B_3 \ln \varepsilon_{t,i}, \tag{21}$$

which is simply the natural logarithm of the first order condition of labor supply in the model above, without human capital accumulation. $\ln \varepsilon_{t,i}$ is, as assumed above, an *i.i.d.* normally distributed shock. B_3 (and $-B_2$) is an unbiased estimator of γ (=*i.e.s.)*, given that the assumption of no human capital accumulation is true. Using (simulated) individual panel data, B_3 is identified from the co-variation of wages and hours over the life cycle.

Table 3. Bias by age groups.

Sample	"true" *i.e.s.*	B_3	Bias $\left(\frac{\text{"true" } i.e.s.}{B_3}\right)$	Bias II $\left(\frac{\gamma}{B_3}\right)$
$25 - 62$	0.34	0.07	5.18	6.53
$35 - 62$	0.38	0.12	3.33	3.90
$45 - 62$	0.41	0.33	1.24	1.36
$55 - 62$	0.43	0.39	1.11	1.15
$60 - 62$	0.44	0.42	1.07	1.08

When examining the estimation bias of excluding human capital, that is it comparing estimates of the *i.e.s.* obtained from a structural model with human capital with estimates from reduced form regressions, is important to take the state-dependency of the *i.e.s.* into account. In models with human capital, the average *i.e.s.* in the sample is no longer given by any of the preference parameters, but is given by an expression of several parameters and state-variables (see equation (16)). Ignoring the state-dependency of the *i.e.s.*, and comparing the estimate of γ, which is the *i.e.s.* for workers soon-to-retire who have the highest elasticity in the population, with an estimate of the *i.e.s.* from a reduced form regression, which is a (weighted) average of the population considered, one will overstate the bias of excluding human capital significantly if the *i.e.s.* differs across the population.

Using equation (16) I can calculate the average *i.e.s.* in the simulated data for various age groups, and for the two education groups. I label this measure the ' "true" *i.e.s.'*. As shown above, the intertemporal elasticity of substitution is increasing in age and the increase is steeper for college graduates. I thus expect that the bias is increasing in both the share of young workers and the share of college graduates.

I first estimate equation (21) for the various age groups, using a 1/3 share of college graduates, and calculate the bias of the estimate of the *i.e.s.*. For comparison, I also present the estimation bias with respect to the estimate of γ, which is overstating the true bias, but is the bias Imai and Keane (2004) and Wallenius (2011) report. The results are given in table 3. As expected, the OLS estimates of the *i.e.s.* is increasing in the share of old workers, see column 5 of table 3. Similarly, the bias is decreasing in the share of old workers, as the return on human capital decreases with age, see column 6 of table 3. When the sample consists of only workers close to retirement, age $60 - 62$, there is almost no bias. This is confirming and explaining the the finding by Imai and Keane using IV techniques on simulated data: "It seems that if older individuals are heavily represented in the data, then the elasticity estimates tend to increase..." For all age groups the "true bias" is smaller than the bias found using γ as an estimate of the *i.e.s.*. However, for older age groups the difference between the two measures of the bias shrinks as the average *i.e.s.* in the population converges to γ.

Second, I estimate equation (21) for each education group separately, and calculate the bias. Again, as expected, the bias is larger for college educated than for non-college educated, see table 4. This is because college graduates have a steeper learning curve, and thus a higher return for human capital accumulation. Even old college graduates have a significant return from human

Table 4. Bias by education

Sample	NCD	CD
$25-62$	$3,72$	$5,45$
$35-62$	$2,16$	$7,18$
$45-62$	$1,16$	$1,52$
$55-62$	$1,01$	$1,26$
$60-62$	$0,99$	$1,16$

capital accumulation, and the bias is as high as 16 percent for the age group $60-62$, whereas there is practically no bias for non-college graduates in the same age group. Thus, to obtain an unbiased estimate of the *i.e.s.* using reduced form regressions, one should use only older workers and preferably low educated workers. However, as shown in this paper, the results obtained from such a sample is not representative for the whole population but do serve as an upper bound on the intertemporal elasticity of substitution in the population.

6 Robustness

Above, I assumed that both education groups had common preferences. This is however possibly a strict assumption, as the two education groups differs significantly with respect to wage growth and wage dispersion, they may also differ significantly with respect to disutility of working. I therefore estimate the stochastic model without the assumption of common preferences. Interestingly, the estimates for γ are very similar across education groups, and are similar to the results with common preferences although slightly lower, see table 5.[33] As the estimates of γ are so similar across education groups, it is natural to ask why the estimate differs from the estimate with common preferences. The reason is because the estimates of b and β differs across the two groups, which also affect the labor supply profile over the life cycle.

As emphasized above, the estimate of γ is of limited interest when it comes to comparing the intertemporal elasticity of substitution across the two model specifications, as this is also influenced by the other parameters of the model. I therefore use equation (16) to compute the *i.e.s.* using the simulated data, and compare the evolution of the *i.e.s.* across the two models. Without the assumption of common preferences, the life cycle profile of the *i.e.s.* shifts downwards for both education groups, see figure 12. The increase in the *i.e.s.* is common across the two models; for non-college graduates the *i.e.s.* increases more with common preference, 66 percent versus 61 percent with uncommon preferences, whereas for college graduates the i.e.s. increases more with uncommon preferences, 143 percent versus 120 percent.

In the estimation above, I have assumed log-utility. However, when facing uncertainty the degree of risk aversion may affects the incentives to insure against this risk, and thus affect labor supply decisions. As the need for insurance is largest for young workers, higher risk aversion leads to higher labor supply when young, and thus the assumption of log-utility may bias the estimates of

[33] The change in results for the deterministic model is very similar, and are available upon request.

Table 5. Parameter estimates for the stochastic model with education specific preferences, targeting age 25-60.

	NCD	CD
γ	0.4022	0.4099
b	23.7480	22.8892
β	0.9605	0.9557
A	0.0904	0.2570
d	0.0165	0.0365
α	0.3073	0.3782
SSE	0.0009	0.0070

Note: NCD refers to no college degree, whereas CD refers to college degree. SSE refers to sum of squared errors.

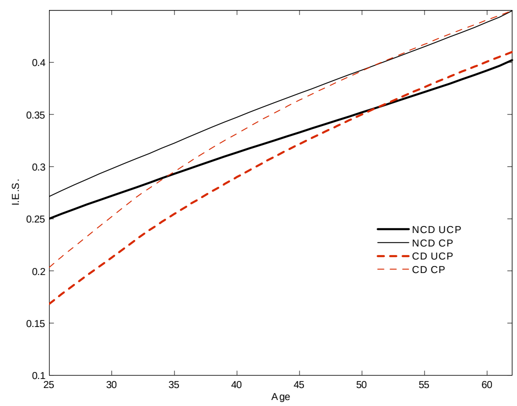


Figure 12. Comparing the intertemporal elasticity of substitution with and without common preferences. Age averages of simulated data. UCP and CP refers to uncommon preferences and common preferences, repectively.

γ downwards if the degree of risk aversion is considerably higher. In order to test the sensitivity of my results with respect to risk aversion, I estimate the stochastic model using a CRRA utility function

$$u(c_t) = \frac{c_t^{1-\sigma}}{1-\sigma},$$

with a relative risk parameter, σ, of 2. The results are presented in table 6. As expected, the estimate of γ increases, although the increase is small. Notice that the fit, measure as the sum of squared errors, are considerably worse than above. Again, to assess the changes in the intertemporal elasticity of substitution of a higher risk aversion, I compute the life cycle path of the *i.e.s.*. For non-college graduates, the *i.e.s.* increases by 60 percent which is similar to the increase for log-utility. For college graduates it increases by 160 percent, compared to 120 percent increase with log-utility. Thus, it seems that the assumption of log-utility biases the estimate of γ and the increase in the *i.e.s.* for college graduates slightly downwards. However, as the fit with higher risk aversion is considerably worse, it is difficult to fully compare the estimates of the *i.e.s.* as they depend on the simulated data which is much more similar to the actual data with log-utility.

Table 6. Parameter estimates for the stochastic model, with a relative risk aversion of 2. Targeting age 25-60.

			NCD	CD
γ	0.5026	A	0.1028	0.1787
b	32.6282	d	0.0247	0.0239
β	0.9596	α	0.1058	0.4430
SSE	0.1901			

Note: NCD refers to no college degree, whereas CD refers to college degree. SSE refers to sum of squared errors.

7 Labor income taxation - Welfare analysis

When workers accumulate human capital the impact of transitory wage shocks on hours are dampend, and more so for workers with a high return to human capital accumulation. Thus, also the impact of transitory tax changes are smaller in the model estimated above, compared to a standard model without human capital. In standard models without human capital, the response in hours to transitory wage changes exceeds the response in hours to permanent wage changes. Then, the *i.e.s.*/Frisch elasticity represents an upper bound on the response of hours to permanent tax changes. However, as Keane (2011) emphasizes, the *i.e.s.* is not the best suited parameter to evaluate the effects of permanent tax policies in models with human capital. First, with human capital the effect of permanent tax changes may exceed the effect of transitory tax changes if the return to human capital accumulation is large, or if the income effect is sufficiently small. Second, in contranst to transitory tax changes, where human capital dampens the effect, with permanent tax changes the role of human capital is to amplify the effect of taxes in the long-run. As Keane puts it, human capital creates a "snowball effect"; increasing taxes permanently decreases today's labor supply, which reduces human capital accumulation. A reduction in human capital in the next period causes wages to fall more than the tax increase and thus further reduces labor supply. Thus, using the estimated *i.e.s.* for different age and education groups as an upper bound on the response of hours to permanent changes in taxes may yield too low welfare costs assosiated tax changes.

Keane (2011) finds that welfare costs of permanent tax changes in a model with human capital are large, and 4 to 7 times larger than in a model with no human capital accumulation. This is both due to the fact that human capital makes permanent tax changes exceed transitory shocks for the youngest and the oldest worker, and because Keane finds that the snowball-effect is very large. I revisit these findings using the model estimated in section 4. There are two key differences between the model here and the models used by Keane (2011). First, Keane uses an estimate of the intertemporal elasticity of substitution for soon-to-retire workers of 3.8, which is the *i.e.s.*

estimated by Imai and Keane (2004). This is a very high estimate of the *i.e.s.* compared to the rest of the literature, and about 10 times larger than what I obtain. The welfare costs of permanent tax changes depends crucially on the size of the snowball effect, which strongly depends on the intertemporal elasticiy of substitution. Second, Imai and Keane estimates their model using data on workers between age 20 and age 36. However, as found by Wallenius (2011), the out-of sample performance of the model is poor. In particular, hours declines too sharply after the age of 40, and so does wages. As the production function of human capital features decreasing returns to scale, a too low level of labor supply is associated with too large changes in human capital accumulation from a change in hours. The model fit of hours and wages in the model estimated above is good for all ages and for both education groups, and I take the good fit of hours and wages as a support of reasonable production function estimates.

Before I turn to the welfare analysis of permanent tax changes, I look at the response of both transitory and permanent shocks in the model estimated above, and compare the findings with Keane (2011). I use the benchmark model estimated in section 4, which assumes log utility. Thus, my results are only directly comparable with parts of the results obtained by Keane, as he mostly uses a lower than unity risk aversion. The effects on labor supply of an unexpected tax increase of 5% occurring at different ages are presented in table 7. The first 4 columns looks at the response of hours for college and non-college workers to a 5% temporary tax increase. In the uncompensated case, the tax income is just vaste, whereas in the compensated case, tax income is distributed back to the workers in a lump-sum form on a yearly basis. In standard models without human capital, the change in hours to a transitory shock exceeds the change in hours to a permanent shock. With human capital however, changes in hours due to permanent tax changes can exceed the effect of transitory shocks. This is the case if the return to human capital accumulation is sufficiently strong, or if the income effect is small. Let's first compare the transitory and permanent uncompensated tax increase, column 1-2 and 5-6. As expected, with log utility the permanent uncompensated effect is close to zero, except for older ages where the horizon of the tax increase is closer to the transitory increase. Thus, for the uncompensated increase, with log utility, the standard result holds that transitory changes have a larger effect than permanent. With compensated tax increases, permanent tax changes have a larger effect than transitory changes for young workers. And, the effect is strongest for college educated; for non-college educated permanent tax changes have a larger effect on labor supply than transitory until the age of 30, whereas for college educated the effect of permanent tax changes exceed the effect of transitory until the age of 40. This is qualitatively similar to what Keane (2011) finds, which reports the average of all workers. However, I find a smaller difference than Keane at young ages, and I do not find that the permanet effect is larger for older workers as Keane does. Here, the effect of permanent tax changes approaches the effect of transitory changes as workers approaches retirement, but it never exceeds. Thus, using a much smaller estimate of the *i.e.s.*, the result that permanent tax changes may exceed transitory changes is maintained, but I find the magnitude of this effect to be smaller than what Keane finds.

When introducing a permanent tax change not only does the immediate response of the economy

Table 7. Percentage change in labor supply in response to a 5 percent tax increase at the indicated age.

	Transitory (unanticipated)				Permanent (unanticipated)			
	Uncompensated		Compensated		Uncompensated			Compensated
	NCD	CD	NCD	CD	NCD	CD	NCD	CD
Age	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
25	-1.49%	-1.11%	-1.52%	-1.12%	-0.06%	-0.18%	-1.58%	-1.55%
30	-1.58%	-1.31%	-1.62%	-1.35%	-0.03%	-0.05%	-1.57%	-1.54%
35	-1.65%	-1.49%	-1.71%	-1.54%	-0.05%	0.01%	-1.58%	-1.56%
40	-1.74%	-1.63%	-1.81%	-1.70%	-0.10%	-0.00%	-1.60%	-1.57%
45	-1.81%	-1.76%	-1.90%	-1.85%	-0.21%	-0.09%	-1.64%	-1.60%
50	-1.88%	-1.86%	-1.99%	-1.98%	-0.42%	-0.29%	-1.71%	-1.67%
55	-1.94%	-1.94%	-2.08%	-2.08%	-0.79%	-0.69%	-1.83%	-1.79%
60	-1.98%	-1.99%	-2.15%	-2.16%	-1.48%	-1.44%	-2.04%	-2.03%

Note: Response in hours to a 5% unexpected tax change at the indicated age. In the compensated tax increase, tax income is distributed back in a lump-sum form, on a yearly basis. NCD refers to workers without a college degree, whereas CD refers to workers with a college degree.

matters, but also, and possibly more important, are the long run effects of a tax change. To evaluate the long run changes, I compute the changes in labor supply over the life cycle of a permanent 5% tax increase at age 25. Then, future wages are not only effected by the tax increase directly, but also indirectly through lower human capital accumulation due to lower labor supply. This is the snowball effect, which Keane (2011) finds to be very important. The change in hours and pre-tax wages over the life cycle for the log-utility model is given in table 8. For comparison, I have also included the changes in hours and wages in the Imai and Keane (2004) model (which assumes a risk aversion of 0.74), taken from Keane (2011). The simulation by Keane suggests that individuals who face a 5 percentage points higher labor tax throughout their working life have 11 percent lower wages and 30 percent lower labor supply at the age of 65. This suggests that one should see very large cross-country differences in labor supply for older workers who face different labor income taxe rates.

Using the model estimated above, I find that pre-tax wages falls by much less compared to the Imai and Keane model. This is due to a lower response in hours, which is due to a much lower estimate of the intertemporal elasticity of substitution, and also due to the model fit. In the model by Imai and Keane simulated hours after the age of 40 are substantially lower than in the data. With decreasing returns to scale, this translates into a too large change in human capital production, as the production per unit is substantially larger than for higher levels of labor supply. Thus, it seems that the large snowball effect found by Keane hinges on a large estimate of the *i.e.s.*.

By comparing the change in hours for older non-college and college workers in table 8 with the ones in the last two columns of table 7, it is evident that the change in hours are smaller when the tax change occurred at the beginning of the working life. The snowball effect is apparently not large enough to outweigh the income effect. Although workers are compensated for their tax payments, they are not compensated for the loss of future income due to lower human capital accumulation. Part of the decrease in wages are compensated through a higher labor supply, thus reducing the long-run labor supply effect of a permanent tax change. This is in sharp contrast to the findings by Keane, which finds that due to the large snowball effect the reduction in hours at age 60 of a permanent tax increase that started at age 25 is double the effect of a transitory tax increase at age 60.

Table 8. Effects of permanent tax increase on labor supply and wages at different ages. A comparison of the response of the model estimated in section X with the model of Imai and Keane (2004).

	Benchmark model				Imai and Keane (2004)	
	NCD		CD			
Age	Hours	Wage*	Hours	Wage*	Hours	Wage*
25	-1.58%		-1.55%		-2.7%	
30	-1.61%	-0.12%	-1.63%	-0.17%	-2.9%	-0.4%
35	-1.64%	-0.22%	-1.67%	-0.33%	-3.2%	-0.7%
40	-1.66%	-0.33%	-1.69%	-0.46%	-3.8%	-1.0%
45	-1.69%	-0.42%	-1.72%	-0.58%	-5.1%	-1.3%
50	-1.72%	-0.51%	-1.74%	-0.69%	-7.9%	-2.0%
55	-1.75%	-0.59%	-1.78%	-0.77%	-13.3%	-3.6%
60	-1.77%	-0.66%	-1.82%	-0.84%	-19.3%	-7.5%
62/65*	-1.79%	-0.70%	-1.84%	-0.86%	-29.2%	-11.6%

Note: The tax increase is a permanent 5% increase, unexpectedly introduced at age 25. Tax income is distributed back in a lump-sum form, on a yearly basis. The effects from the model by Imai and Keane (2004) are taken from Keane (2011).

* Reduction in pre-tax wages.

Given the much smaller snow-ball effect of human capital, I expect to find much lower welfare costs of tax changes compared to Keane (2011). I use two measures to evaluate the long-run welfare costs of taxes. Since I have a dynamic model from age 25 to death, it is natural to look at the change in the life time utility at age 25, given by equation 1. I also look at the change in life time consumption, which is a commonly used. I keep the assumption that taxes are distributed back lump-sum. In response to a 5 percentage point permanent tax increase, life time consumption, ΔC, is reduced by 1.9 percent and 2 percent for non-college workers and college workers, respectively, see table 9. Looking at change in life time utility, ΔV, which also takes into account the increase in utility from leisure due to reduced labor supply, the change is tiny: -0.03 percent for non-college workers and -0.14 percent for college workers. When comparing with a model with no human capital accumulation, the welfare losses increases. Looking at life time consumption, the increase in welfare loss is 26 percent and 33 percent for non-college and college workers, respectively. Note that the in the model without human capital, there is no difference in the loss of consumption between the two education groups. The increase in welfare loss is however significantly lower than what Keane (2011) finds, where human capital increases the loss by 4-7 times.

For the sake of comparison with Keane, the welfare costs reported are the long-run costs, and do not include costs along the transition. However, I expect the changes along the transition to be similar to the long-run costs, as the response of hours in table 7 and 8 are very similar. More importantly however, is the failure to take into account the welfare costs of a reduction in human capital accumulation. However, this requires a general equilibrium model. But, as I find the reduction in human capital of a 5 percentage point permanent tax increase to be relativly modest, I expect that the welfare costs of the reduction in human capital to be modest.

Table 9. Effects of permanent tax increase on labor supply and wages at different ages. A comparison of the response of the model estimated in section X with the model of Imai and Keane (2004).

	Benchmark model		No Human Capital
	NCD	CD	($A = 0$, $\delta = 0$)
ΔC	-1.9%	-2.0%	-1.5%
ΔV	-0.03%	-0.14%	-0.04%

Note: The tax income is distributed back in a lump-sum form, on a yearly basis. ΔC is the change in life-time consumption, discounted to age 25. ΔV is the change in the life-time utility, at age 25.

8 Conclusion

In this paper I argue that the average *i.e.s.* is low, and comparable with micro estimates, even in the presence of human capital accumulation. This is due to the findings that (i) a large share of the increase in the extensive margin early in the working life should not be included as changes in optimal labor supply (ii) human capital accumulation implies a state-dependent intertemporal elasticity of substitution, which makes the average elasticity lower than the estimated utility parameter. First, when estimating the *i.e.s.* there is an important distinction to be made between voluntary and involuntary quits. This is because involuntary quits are not a choice by the worker, and should therefore not be included in the measure of labor supply when estimating the *i.e.s.*. Using monthly data on labor market transitions taken from the CPS I show that the majority of unemployment spells are involuntary. Moreover, I find that most of the changes in the extensive margin of labor supply occurs early in life, and is due to either involuntary unemployment or due to a sample effect caused by students leaving university and starting to work throughout the year. Second, when workers accumulate human capital the *i.e.s.* is no longer only a function of the preference parameters, but a state-dependent variable that crucially depends on the return to human capital accumulation. I identify two import sources of heterogeneity with respect to return to human capital accumulation; age and education. I show that the *i.e.s.* is increasing in age and decreasing in education.

Estimating a life cycle model with endogenous labor supply along the intensive margin, human capital accumulation through learning by doing and wage uncertainty, using cohort data on white males between age 25 and age 62 from CPS for workers with and without a college degree. I target only the intensive margin of labor supply, and find an estimate of the *i.e.s.* for soon-to-retire-workers of 0.45. However, there are significant differences across age and education; the estimate of the *i.e.s.* for workers aged 25 is 0.27 and 0.20 for non-college graduates and college graduates, respectively. Second, at young ages non-college workers have 35 percent higher *i.e.s.* than young college educated. Using population weights calculated from the CPS, and a share of 1/3 college graduates, the average *i.e.s.* in the population is 0.34. Thus, the average elasticity is well below the estimate of the *i.e.s.* for workers soon-to-retire, which is identical to the estimate of the (inverse) curvature of the marginal disutility of labor. The increase in the *i.e.s.* over the life cycle is however significantly lower than what Imai and Keane finds in a similar model, when looking at uncompensated changes in labor supply.

I perform robustness checks on two of my key assumptions, namely the assumption of common preferences across education groups and log-utility of consumption. Relaxing these assumptions yields small changes in the estimates, and the key conclusions are not altered; the estimate of the average intertemporal elasticity of substitution is low, and the *i.e.s.* increases significantly over the life cycle.

Finally, I revisit the findings by Keane (2011), which argue that once the effects of human capital is taken into account the long-run welfare costs of tax changes are large, and more than 4-7 times larger than in conventional models. This is because human capital amplifies the long-run effect of taxes through a "snowball-effect"; lower labor supply today's reduces human capital accumulation, and thus reduces future wages which again reduces future labor supply. Using the model estimated above, I find that the importance of the snowball-effect hinges on the high estimate of the *i.e.s.* used by Keane in the analysis (3.8); the welfare costs of a permanent 5% tax increase in the model estimated above, with an average *i.e.s.* of 0.34, implies a consumption loss of 1.9 percent and 2 percent for non-college and college workers, compared to a loss of 1.5 percent loss for both education groups in the same model without human capital. So, although human capital do increase the loss of taxation, the difference is much smaller than previously found. The analysis is however, as Keane's, a partial equilibrium analysis, which may understate the importance of human capital, as the productivity costs of reduced human capital accumulation is not taken into account.

The state-dependency of the *i.e.s.* and the difference in estimation bias across different samples in terms of age and education indicate that one should be cautious when comparing estimates of the intertemporal elasticity of substitution obtained from different samples with respect to e.g. the age and education distribution. Moreover, the state-dependency of the *i.e.s.* has important policy implications, as the substitutability of labor depends on age and ability to accumulate human capital. Thus, both the age structure of the economy and the education/occupation structure of the economy may matter significantly for the effect of e.g. tax changes on average labor supply. These effects are however not explored here, but are topics for future work.

A State-dependent i.e.s.

I define the intertemporal elasticity of substitution as the percentage change in the intertemporal labor supply to a one percent temporary increase in the rental rate on human capital, p_t^h, ($p_j^h = 0, \forall j \neq t$), keeping marginal utility of consumption fixed:

$$i.e.s. = \frac{d\left(\frac{n_t}{n_{t+1}}\right)}{dp_t^h} \frac{p_t^h}{\frac{n_t}{n_{t+1}}} \tag{A.1}$$

Rewriting this yields

$$i.e.s. = \frac{dn_t}{dp_t^h}\frac{p_t^h}{n_t} - \left(\frac{\partial n_{t+1}}{\partial k_{t+1}}\frac{dk_{t+1}}{dp_t^h} + \frac{\partial n_{t+1}}{\partial h_{t+1}}\frac{dh_{t+1}}{dp_t^h}\right)\frac{p_t^h}{n_{t+1}}.$$

Without human capital, the last term in parenthesis would have been zero and the intertemporal elasticity of substitution would have been identical to the Frisch elasticity. However, with human capital, also future labor supply responds to changes in today's rental rate, as this affects future wages through today's production of human capital.

I the following I assume, as in the model, that preferences are separable in consumption and leisure, and in time, and that production of human capital is linear in the capital stock, $\frac{\partial^2 G_{t+1}}{\partial h_{t+1}^2} = 0$. Given the functional forms, there are no analytical solutions for the policy functions. I therefore take the total derivative of the first order condition for labor, equation (14), imposing that $dh_t = dk_t = 0$:

$$\frac{dn_t}{dp_t^h} = \frac{1}{\left(b\frac{1}{\gamma}n_t^{\frac{1}{\gamma}-1} - \mu_t\frac{\partial^2 G_t}{\partial n_t^2}\right)}\left(u'(c_t)\varepsilon_t h_t + \frac{\partial G_t}{\partial n_t}\frac{d\mu_t}{dp_t^h}\right). \tag{A.2}$$

The derivative of the shadow price on skills with respect to an increase in the current price on skills, is given by

$$\frac{d\mu_t}{dp_t^h} = \beta E_t\left[\left(1 - \delta + \frac{\partial G_{t+1}}{\partial h_{t+1}}\right)\frac{d\mu_{t+1}}{dp_t^h} + \frac{bn_{t+1}^{\frac{1}{\gamma}}}{h_{t+1}}\frac{dn_{t+1}}{dp_t^h}\right], \tag{A.3}$$

where I have used the first order condition for n_{t+1} and the linearity of G_{t+1}in h, so that $h_t\frac{\partial^2 G_{t+1}}{\partial h_{t+1} n_{t+1}} = \frac{\partial G_{t+1}}{\partial n_{t+1}}$.

The total derivative of future variables depends only on human capital, and not on future capital stocks. The proof for this is the following: Using the budget constraints, and the law of motion for human capital, we know that

$$h_{t+2}\left(1 - \frac{dk_{t+2}}{dk_{t+1}}\right)\frac{dn_{t+2}}{dk_{t+2}} = \left(\frac{dk_{t+3}}{dh_{t+2}} - n_{t+2}\right)\frac{dG_{t+1}}{dn_{t+1}}\frac{dn_{t+1}}{dk_{t+1}}. \tag{A.4}$$

Then, if for any $t : \frac{dn_t}{dk_t} = 0$, $\frac{dn_t}{dk_t} = 0\ \forall t$. Looking at the derivative of the first order condition for labor with respect to capital in period T, $\frac{dn_T}{dk_T}$, this is zero, as $\mu_T = 0$,and given that $u'(c_t) = 0$ by assumption. Moreover, from the first order condition for labor supply in any period

$$\frac{dn_{t+2}}{dk_{t+2}} = \frac{1}{b\frac{1}{\gamma}n_t^{\frac{1}{\gamma}-1} - \mu_{t+2}\frac{\partial^2 G_{t+2}}{\partial n_{t+2}^2}}\frac{\partial G_{t+2}}{\partial n_{t+2}}\frac{d\mu_{t+2}}{dk_{t+2}}, \tag{A.5}$$

we see that as $\frac{dn_t}{dk_t} = 0\ \forall t$, so is also $\frac{d\mu_t}{dk_t} = 0, \forall t$. So, labor supply responds only to human capital (and the rental rate). Hence $\frac{dn_{t+1}}{dp_t^h} = \frac{\partial n_{t+1}}{\partial h_{t+1}}\frac{dh_{t+1}}{dp_t^h}$, and $\frac{d\mu_t}{dp_t^h}$ simplifies to

$$\frac{d\mu_t}{dp_t^h} = \beta E_t\left[\left(1-\delta+\frac{\partial G_{t+1}}{\partial h_{t+1}}\right)\frac{\partial \mu_{t+1}}{\partial h_{t+1}} + \frac{bn_{t+1}^{\frac{1}{\gamma}}}{h_{t+1}}\frac{\partial n_{t+1}}{\partial h_{t+1}}\right]\frac{dh_{t+1}}{dp_t^h}. \tag{A.6}$$

Moreover, the intertemporal elasticity of substitution simplifies to

$$i.e.s. = \frac{dn_t}{dp_t^h}\frac{p_t^h}{n_t}\left(1 - \frac{\partial n_{t+1}}{\partial h_{t+1}}\frac{\partial G_t}{\partial n_t}\frac{n_t}{n_{t+1}}\right) \tag{A.7}$$

Now, one only needs to find $\frac{dn_{t+1}}{dh_{t+1}}$ as a function of $\frac{\partial \mu_{t+1}}{\partial h_{t+1}}$, so that $\frac{d\mu_t}{dp_t^h}$ can be solved recursively. This is obtained from the FOC of labor

$$\frac{dn_{t+1}}{dh_{t+1}} = \frac{1}{\left(b\frac{1}{\gamma}n_{t+1}^{\frac{1}{\gamma}-1} - \mu_{t+1}\frac{\partial^2 G_{t+1}}{\partial n_{t+1}^2}\right)}\left(\frac{bn_{t+1}^{\frac{1}{\gamma}}}{h_{t+1}} + \frac{\partial G_{t+1}}{\partial n_{t+1}}\frac{d\mu_{t+1}}{dh_{t+1}}\right) \tag{A.8}$$

Then, the *i.e.s.* is obtained from solving the following system of equations recursively

$$\begin{aligned}
ies_t &= \frac{dn_t}{dp_t^h}\frac{p_t^h}{n_t}\left(1 - \frac{\partial n_{t+1}}{\partial h_{t+1}}\frac{\partial G_t}{\partial n_t}\frac{n_t}{n_{t+1}}\right)\\
\frac{dn_t}{dp_t^h}\frac{p_t^h}{n_t} &= \gamma\frac{1}{\left(\Upsilon_t - \gamma\left(\frac{\partial G_t}{\partial n_t}\right)^2\Psi_t\right)}u'(c_t)\varepsilon_t h_t\frac{p_t^h}{n_t},\\
\frac{dn_t}{dh_t}\frac{h_t}{n_t} &= \frac{dn_t}{dp_t^h}\frac{p_t^h}{n_t}\frac{1}{u'(c_t)\varepsilon_t}\left(\frac{bn_t^{\frac{1}{\gamma}}}{h_t} + \frac{\partial G_t}{\partial n_t}\Psi\frac{h_{t+1}}{h_t}\right),\\
\Psi_t &\equiv \beta E_t\left\{\frac{n_{t+1}}{h_{t+1}}\left[\frac{h_{t+2}}{h_{t+1}}\frac{1}{\frac{\partial G_{t+1}}{\partial n_{t+1}}}\left[\Upsilon_{t+1}\frac{dn_{t+1}}{dh_{t+1}}\frac{h_{t+1}}{n_{t+1}} - bn_{t+1}^{\frac{1}{\gamma}-1}\right] + \frac{bn_{t+1}^{\frac{1}{\gamma}}}{h_{t+1}}\frac{dn_{t+1}}{dh_{t+1}}\frac{h_{t+1}}{n_{t+1}}\right]\right\}\\
\Upsilon_t &\equiv bn_t^{\frac{1}{\gamma}-1} - \gamma\mu_t\frac{\partial^2 G_t}{\partial n_t^2}.
\end{aligned}$$

In the estimated models above $\gamma\left(\frac{\partial G_t}{\partial n_t}\right)^2\Psi_t \approx 0$, and $\frac{\partial n_{t+1}}{\partial h_{t+1}}\frac{\partial G_t}{\partial n_t}\frac{n_t}{n_{t+1}} \approx 0$, so the *i.e.s.* simplifies to

$$ies_t \approx \gamma\frac{1}{bn_t^{\frac{1}{\gamma}-1} - \gamma\mu_t\frac{\partial^2 G_t}{\partial n_t^2}}u'(c_t)\varepsilon_t h_t\frac{p_t^h}{n_t} \tag{A.9}$$

References

Altonji, J. G. (1986). Intertemporal substitution in labor supply: Evidence from micro data. *Journal of Political Economy*, 94(3):S176–S215.

Blau, F. D. and Kahn, L. M. (2007). Changes in the labor supply behavior of married women: 1980-2000. *Journal of Labor Economics*, 25:393–438.

Chetty, R., Guren, A., Manoli, D., and Weber, A. (2011). Are micro and macro labor supply elasticities consistent? a review of evidence on the intensive and extensive margins. *American Economic Review*, 101(3):471–75.

Domeij, D. and Floden, M. (2006). The labor-supply elasticity and borrowing constraints: Why estimates are biased. *Review of Economic Dynamics*, 9(2):242–262.

French, E. (2005). The effects of health, wealth, and wages on labour supply and retirement behaviour. *Review of Economic Studies*, 72(2):395–427.

French, E. and Jones, J. (2012). Public pensions and labor supply over the life cycle. *International Tax and Public Finance*, 19(2):268–287.

Ham, J. C. and Reilly, K. T. (2002). Testing intertemporal substitution, implicit contracts, and hours restriction models of the labor market using micro data. *American Economic Review*, 92(4):905–927.

Heathcote, J., Storesletten, K., and Violante, G. L. (2005). Two views of inequality over the life cycle. *Journal of the European Economic Association*, 3(2-3):765–775.

Heckman, J. J. (1976). A life-cycle model of earnings, learning, and consumption. *Journal of Political Economy*, 84(4):S11–44.

Imai, S. and Keane, M. P. (2004). Intertemporal labor supply and human capital accumulation. *International Economic Review*, 45(2):601–641.

Keane, M. (2011). Income taxation in a life cycle model with human capital. Working Papers 201117, ARC Centre of Excellence in Population Ageing Research (CEPAR), Australian School of Business, University of New South Wales.

Kim, M.-K. and Polachek, S. W. (1994). Panel estimates of male-female earnings functions. *Journal of Human Resources*, 29(2):406–428.

Lemieux, T. (2006). Increasing residual wage inequality: Composition effects, noisy data, or rising demand for skill? *American Economic Review*, 96(3):461–498.

MaCurdy, T. E. (1981). An empirical model of labor supply in a life-cycle setting. *Journal of Political Economy*, 89(6):1059–85.

Pencavel, J. (1986). Labor supply of men: A survey. In Ashenfelter, O. and Layard, R., editors, *Handbook of Labor Economics*, volume 1 of *Handbook of Labor Economics*, chapter 1, pages 3–102. Elsevier.

Pijoan-Mas, J. (2006). Precautionary savings or working longer hours? *Review of Economic Dynamics*, 9(2):326–352.

Shaw, K. L. (1989). Life-cycle labor supply with human capital accumulation. *International Economic Review*, 30(2):431–56.

Wallenius, J. (2009). Human capital accumulation and the intertemporal elasticity of substitution of labor? Unpublished mimeo.

Wallenius, J. (2011). Human capital accumulation and the intertemporal elasticity of substitution of labor: How large is the bias? *Review of Economic Dynamics*.

Ziliak, J. P. and Kniesner, T. J. (1999). Estimating life cycle labor supply tax effects. *Journal of Political Economy*, 107(2):326–359.