Accepted Date: 03/11/2015

Article Type: Conference Article

Labour Supply: the Roles of Human Capital and the Extensive Margin

Short Title: Human Capital and the Extensive Margin

Abstract: We specify and estimate a life-cycle labour supply model that expands on earlier work by simultaneously including human capital accumulation, saving and bequests, an active extensive margin, a realistic specification of the progressive tax structure and the Social Security system, and an accounting for private pensions and health expenditures. By incorporating all these features, we develop new insights into how taxes affect life-cycle labour supply. For instance, we find that labour supply elasticities vary in important ways with age, education and the tax structure itself. We also show how human capital affects elasticities differently on the intensive vs. extensive margins.

Keywords: Human Capital, Labour Supply, Taxation, Life-Cycle Model, Hicks elasticity, Frisch elasticity, Extensive Margin, Social Security

JEL Codes: D91, E24, H31, J22, J24

Acknowledgements: We thank the editor Kjell Salvanes and an anonymous referee for very helpful comments. This is the second of two papers that formed the basis of Keane's Sargan Lecture to the Royal Economic Society on March 30, 2015. Wasi's research on this project has

been supported by the Alfred P. Sloan Foundation for the University of Michigan's Census-Enhanced HRS (CenHRS) Project. Keane's work on this project has been supported by Australian Research Council grants FF0561843 and FL110100247, by the AFTS Secretariat of the Australian Treasury and the ARC Centre of Excellence in Population Aging Research (grant CE110001029) at UNSW. The views expressed are entirely our own. Corresponding author: Michael Keane, Nuffield College, Oxford, OX1 1NF, UK.
Email: michael.keane@nuffield.ox.ac.uk.

In this paper we specify and estimate a life-cycle labour supply model that incorporates many key features of the (U.S.) economic environment that have not heretofore been unified in a single framework.[1] These features include: human capital accumulation, an active extensive margin, saving and bequests, a realistic specification of the Social Security system, an accounting for private pensions and health expenditures, and a realistic specification of the progressive tax structure. By accounting for all these features, we develop new insights into how taxes affect life-cycle labour supply. For instance, we find that labour supply elasticities vary in important ways with age, education and the tax structure itself.

Our work was originally motivated by the well-known controversy over the magnitude of labour supply elasticities. The conventional wisdom among economists, at least until recently, was that Frisch and Hicks elasticities are quite small, at least for men (see, e.g., MaCurdy (1981), Browning et al (1985), Altonji (1986) and Blundell and Walker (1986)). Recently, this consensus has come under attack from two directions:

A line of work starting with Imai and Keane (2004), and represented recently by Keane (2015), argues that the failure of most of the male labour supply literature to account for human capital has led to severe downward bias in Frisch and Hicks elasticity estimates.

At the same time, a line of research dating back to Rogerson (1988) argues that failure to account for the extensive margin (participation decisions) led the male labour supply literature to understate labour supply elasticities. Some important papers in this line of research are Kimmel and Kniesner (1998), French (2005), Prescott et al (2009) and Rogerson and Wallenius (2009).[2]

[1] While we use US data, our basic modelling approach could be applied to any developed country by substituting in country specific tax and social security rules.

[2] We won't be discussing female labour supply in this paper. This is not because we view it as any less important. Rather, as Keane (2011) noted in an extensive review of the literature, there appears to be a rather broad consensus that female labour supply is quite elastic. So there is really no "female labour supply controversy" to speak of. We conjecture the reason the female literature tends to find large elasticities is not that women differ from men in their

Thus, in the past decade, much of the economics profession has shifted toward a view that labour supply elasticities are larger than was previously thought. But, as noted by Keane and Rogerson (2012), there is not yet an agreement on why: Is it human capital, the extensive margin or some other factor that led prior work to understate elasticities? Clearly, a better understanding of the *mechanism*s that drive labour supply responses is important for many reasons, such as assessing welfare effects of tax changes and determining the optimal design of the tax system.

The main goal of the present paper is to develop and estimate a unified model of life-cycle labour supply in which both the extensive margin and human capital mechanisms are operative. This is an important exercise for several reasons:

First, even amongst economists who believe labour supply elasticities are large, there has emerged some controversy as to whether the human capital or extensive margin models provide a better account of the data – see, e.g., Wallenius (2011). Second, it is important to understand how the two mechanisms interact to affect labour supply behavior. For example, does the introduction of human capital primarily affect elasticities on the intensive or extensive margin?

Third, while both the extensive margin and human capital mechanisms generate labour supply elasticities that vary by age, the exact patterns differ greatly. As Keane (2015) explains, models with endogenous human capital (and interior solutions) predict that labour supply elasticities should grow with age.[3] In contrast, models with an extensive margin (and no human capital) predict a U-shape for elasticities.[4] Does one or the other pattern predominate in the data, implying that one or the other mechanism is of primary importance? Or do the two mechanisms interact to generate a more complex pattern than either would generate on its own?

Fourth, the two types of model also make different predictions for effects of permanent vs. transitory tax changes. The extensive margin model implies transitory tax effects are greater than permanent tax effects (consistent with conventional wisdom). But the with human capital permanent tax effects can be larger than transitory (see Keane (2015) for further discussion).

---

preferences and/or constraints (although they certainly do). Rather, it is because the female literature has long viewed the data through the lens of models that emphasize the extensive margin and/or human capital investment. See, for example, Heckman and MaCurdy (1982) and Eckstein and Wolpin (1989).

[3] The basic point is that for young workers the opportunity cost of time is much greater than the wage, because youth have substantial returns to work experience. As a result, young workers will be will relatively insensitive to the direct wage component of compensation.

[4] The basic point is that, given fixed costs of work, it is not optimal to work a low level of hours. Instead, workers jump from zero hours to substantial positive hours when the offer wage passes the reservation wage. Hence, both young and old workers tend to be close to indifferent between not working at all and working substantial positive hours – the young because their offer wages tend to be low, the old because their reservation wages tend to be high.

Thus, while both human capital and extensive margin models generate labour supply elasticities that are substantially greater than conventional wisdom suggests, they do so in very different ways. As a result, it is of considerable interest to develop a model that contains both mechanisms, and to quantitatively assess labour supply behavior in this unified framework.

The reason that life-cycle labour supply models have not previously incorporated human capital, saving and the extensive margin is purely computational. Imai and Keane (2004) noted that it was very difficult to estimate a model with both human capital and assets as state variables, while also allowing for choice of hours and the possibility of corner solutions. But given the present level of computer speed, the problem is feasible. In fact, we can go further:

We are also able to incorporate another important feature of the economic environment: Social Security retirement benefits and "retirement" decisions. Computationally this is a major extension, as the accumulated level of retirement benefits is an additional continuous state variable (along with assets and human capital). Furthermore, "retirement" is an additional decision variable. That is because, under US Social Security rules, "retirement" is not a decision to stop working, but rather a decision to start collecting benefits.[5] Henceforth, we refer to this decision as "claiming benefits." Additional important features of the economic environment that we account for (albeit in a very simple way) are private pension benefits and medical costs.

Both static and life-cycle labour supply models have been criticized in the past for the assumption that workers can adjust hours continuously in response to changes in wages/taxes (see, e.g., Moffitt (1984), Dickens and Lundberg (1985), van Soest et al (1990), Dagsvik and Strøm (1992), Aaberge et al (1995)). In the data, it is well-known that hours tend to "bunch up" at a few at a few key points (e.g., 35, 40, 45 hours per week). An inability to adjust hours smoothly between such points (due to demand side/technological constraints or adjustment costs) may well dampen labour supply responses to wage/tax changes.

In this paper we address this issue in two ways: First, we incorporate a search structure where job offers, and involuntary separations, arrive probabilistically. Thus, involuntary job losses may account for part of the variation in hours and employment. This is important, as it means the model does not "force" all of the observed variation in hours/employment to be explained as optimal responses to wage/tax changes. Second, we assume that workers can only

---

[5] From age 62 onward one can work and collect Social Security simultaneously. The tax treatment of benefits and earnings in this case is rather complex, and we incorporate that in our model.

choose from among six discrete levels of hours, thus capturing the well-known bunching of hours in the data. Given this feature, the model can capture the idea that hours may be unresponsive to small changes in after-tax wages.

A final important feature of the model is that we incorporate progressive taxation. The life-cycle labor supply literature, from MaCurdy (1981) onwards, has mostly ignored progressive taxes (exceptions are Blomquist (1985) and Ziliak and Kniesner (1999)). Altonji (1986) argued that this should not create too great a problem, as year-to-year tax changes tend to be too small to have much impact on year-to-year changes in after-tax-wages (see Keane (2011)).

Nevertheless, the literature on static labour supply models has shown that effects of tax changes can be very different in an environment with a progressive tax system than with a flat-rate tax. In particular, as was noted by Hausman (1981, 1985), under a progressive tax structure uncompensated tax changes will have effects similar to those of compensated changes. This is because they alter "virtual" non-labour income in a way that approximates Hicks compensation (see also MaCurdy (1992), Blomquist (1983)). Dynamic structural models of labour supply have ignored this mechanism. Thus, a key open question is the extent to which a progressive tax structure amplifies Marshallian elasticities in a dynamic setting.

Given our estimated model, we conduct two types of simulations: First, we construct a set of hypothetical tax experiments that generate the "textbook" Marshall and Hicks elasticities implied by the model, i.e. compensated/uncompensated effects of permanent/transitory changes in the tax rate on labour earnings. Second, we perform a set of more realistic tax experiments that involve plausible changes in the U.S. progressive tax structure.

As we'll see, the response of labour supply to changes in the progressive tax structure can be very different from what one might predict based solely on knowledge of the Hicks and Marshall elasticities. The point that elasticities are not invariant parameters, but depend on the tax structure and the nature of tax changes, was stressed in the static setting by Hausman (1981), Blomquist (1983), and other early contributors to the structural labour supply literature.

Finally, we perform additional simulations to gain a better understanding of how human capital affects labour supply elasticities on the intensive vs. extensive margin. Here, we consider versions of our model where returns to human capital are shut down or reduced.

The paper is organized as follows: In Section 1 we present our life-cycle labour supply model. Section 2 describes the solution. In Section 3 we describe the data and estimation method.

We combine data from the Current Population Survey (CPS), Health and Retirement Survey (HRS), Consumer Expenditure Survey (CEX) and Medical Expenditure Panel Survey (MEPS), as no one of these data sets contains all the information we require. Estimation is by method of simulated moments (MSM). We present our estimation results in Sections 4, and discuss model fit in Section 5. We report our tax change experiments in Section 6, while Section 7 reports experiments where we change the importance of endogenous human capital (and examine how this effects intensive vs. extensive margin elasticities). Section 8 concludes.

1. A Model of Labor Supply, Human Capital Accumulation and Retirement

The decision period in our model is annual. Let *t* denote age and period. Agents start making labour supply and savings decisions after they leave school. We solve and estimate the model separately for dropouts, high school graduates and college graduates. The initial decision making age is 16, 18 and 22 for these three types of workers, respectively.

Agents' choice set depends on their age. From the time of entering the labour market until age 61, agents make annual decisions about consumption c_t and work hours h_t. Beginning at age 62, application for Social Security (SS) retirement benefits is an additional option. Note that one may elect to receive SS benefits but still continue to work, so receipt of SS benefits and retirement from the labour force are not necessarily linked.[6]

We assume that by age 75 all agents must claim SS and must also retire from the labour force. Thus, from age 75 until death agents only choose consumption. Mortality is probabilistic, but to simplify solution of the model we assume a terminal period when all agents must die (age 93 for college graduates, age 90 for high school graduates and dropouts). This allows us to obtain the solution by backsolving.

1.1. *Utility Function*

We assume the same per-period utility function as in MaCurdy (1981) and Imai and Keane (2004). Suppressing the individual subscript *i* we have:

$$u(c_t, h_t) = \frac{c_t^{a_1}}{a_1} - b_n \frac{h_t^{a_2}}{a_2} \tag{1}$$

Here $a_1 < 1$ and $a_2 > 1$ are parameters, and b_n is a type *n* specific taste for work, where *n* = 1,2.

[6] If one delays the take-up of SS benefits until after age 62, then benefits increase according to a formula we discuss later. If one works while receiving SS then benefits may be taxed according to a formula we also discuss later.

Agents also get utility from bequests, according to the function:

$$B(A_{T+1}) = \begin{cases} 3ln(A_{T+1} + \phi) - 1 - 3ln(\phi) & \text{if } A_{T+1} > 0 \\ \left(\frac{A_{T+1}+\phi}{\phi}\right)^3 & \text{otherwise} \end{cases} \quad (2)$$

Higher values of ϕ reduce the marginal utility of bequests.[7] This is the same function used by Imai and Keane (2004), who chose it because it is continuously differentiable and concave in assets, even at A_{T+1}=0. It strongly discourages negative bequests, but does not rule them out.

1.2. *The Wage Process and Human Capital Investment*

The wage rate w_t is determined by human capital. Human capital at age t is denoted k_t, and it evolves with age according to the process:

$$k_{t+1} = g(k_t, h_t, t)\varepsilon_{t+1} \qquad \Rightarrow \qquad lnk_{t+1} = lng(k_t, h_t, t) + ln\varepsilon_{t+1} \quad (3)$$

$$\ln g(k_t, h_t, t) = \lambda_0 + \lambda_1 lnk_t + \lambda_2 \max\left(h_t - \bar{h}, 0\right) + \lambda_3 \max\left((h_t - \bar{h})^2, 0\right) \quad (4)$$
$$+\lambda_4(t - t_0(e)) + \lambda_5(t - t_0(e))^2$$

Equation (3) says that human capital at t+1 depends on lagged human capital, hours worked at age t, and age t, as well as a shock ε_{t+1}. The function $g(\cdot)$ governs the deterministic part of the process mapping current human capital, work hours and age into next period's. According to equation (4), the increment to human capital is a quadratic in hours of work. However, hours must exceed the threshold $\bar{h}$ to be productive in producing human capital.

Equation (4) also allows for pure age effects on human capital. Indeed if $\lambda_2 = \lambda_3 = 0$ our model reduces to a standard exogenous wage-path model. This is important, as it means we do not force observed wage growth to be explained by work experience. In the quadratic in age, the term $t_0(e)$ = 16, 18 or 22 is the age of entering the labour force for dropout, high school and college workers respectively. Finally, $0 < \lambda_1 < 1$ captures depreciation.

The shocks to the human capital production process are given by:

[7] We normalize ϕ to the value of 30,000 used by Imai and Keane (2004).

$$\log(\varepsilon_t) \sim N\left(-\frac{1}{2}\sigma^2, \sigma\right) \quad (5)$$

Agents face a discrete choice of six possible levels of annual hours:

$$h_t \in [0, 500, 1000, 1500, 2000, 2500] \quad (6)$$

Wages are given by:

$$w_t = \begin{cases} k_t & if\ h_t \geq 1500 \\ .85k_t & if\ h_t < 1500 \end{cases} \quad (7)$$

Thus, the wage is equal to human capital if the agent works 1500 hours or more. There is a 15% penalty for working less than 1500 hours. We do not attempt to estimate the part-time penalty, but rather calibrate it based on consensus estimates in the literature (see Keane (2011)).

1.3. *Fixed Costs of Work*

Crucially, the model includes fixed costs of work. This is important for generating corners solutions where workers choose not to work. Specifically, we have that pre-tax wage income (E_t) is given by:

$$E_t = w_t * \max\{h_t - fc_h, 0\} - fc_m \cdot I(h_t > 0) \quad (8)$$

Here fc_m is a monetary fixed cost of work, and fc_h is a time fixed cost of work (also known as a start-up time). This non-linear wage schedule is similar to the idea in Rogerson and Wallenius (2009) and Prescott, Rogerson and Wallenius (2009), and is central to their story for why labor supply elasticities should be much higher on the extensive margin than on the intensive margin. Indeed, a key motivation of our work is to develop a model that includes both human capital investment, as in (3)-(4) and an operative extensive margin, which is generated by (8).

1.4. *Job Offer Probabilities*

In each period, an agent has a probability of receiving a job offer. Let $pjob_t = 1$ if a person has a job offer and 0 otherwise. The offer probability is specified as a flexible function of lagged participation and age. In particular, we specify a logit with latent index $pjob_t^*$ given by:

$$pjob_t^* = m^1 + m^{21}(t - t') * I(t \leq t') + m^{22}(t-30)I(t<30) + m^{23}(t-30)I(t>30)$$
$$+m^{24}(t-50)I(t>50) + m^{25}(t-59)I(t>59)$$
$$+m^{30}(1-p_{t-1}) + m^{31}(1-p_{t-1})I(t>30)$$
$$+m^{32}(1-p_{t-1})(t-40)I(t>40) + m^{33}(1-p_{t-1})(t-59)I(t>59) \qquad (9)$$

where $p_{t-1} = I(h_{t-1} > 0)$ is an indicator for lagged work. $pjob_t = 1 \; iff \; pjob_t^* + \zeta_{offer,t} > 0$, where $\zeta_{offer,t}$ is drawn from logistic distribution. Define $J_t(p_{t-1}, t) = Pr[pjob_t = 1 | p_{t-1}, t\,]$.

The first set of terms in (9), with coefficients m^{21} through m^{25}, define a spline in age with notches at *t'*, 30, 40, 50 and 59.[8] The second group of terms, with coefficients m^{30} through m^{33}, allows the probability of receiving an offer to depend on lagged employment status, in a way that varies with age. For example, we would expect workers in their 60s who were not employed in the previous period to have a difficult time finding a job, would be captured by $m^{33} < 0$. Recall that we estimate the model separately for dropouts, high school graduates and college graduates, so the parameters of the job offer function will differ freely between the three groups.

Our main motivation for including job offers is to avoid the strong assumption that all fluctuations in hours of work can be explained as voluntary responses to wage changes. Such an assumption may bias labour supply elasticity estimates, particularly if probabilities of job offers and involuntary separations are state dependent.

1.5. *Social Security Benefits*

The US Social Security system works as follows: People become eligible to start collecting SS "retirement benefits" at age 62. They can delay, with (roughly) actuarially fair adjustments, until age 70. One can keep working while receiving SS benefits, so "claiming SS" and "retirement" are two distinct decisions. Let ss_t be a 1/0 indicator for claiming SS benefits.

Social Security benefits are based on "Average Indexed Monthly Benefits" (AIME). This is a person's average earnings over their top 35 years. Let $AIME_t$ denote the value of this state variable at age *t*. Structural modelling of SS is extremely challenging, because, in order to keep

[8] The first notch *t'* is set at age 23 for dropouts and high school workers, and 26 for college graduates.

track of how the $AIME_t$ state variable evolves over time, one must save all 35 previous highest earnings, which is obviously infeasible. Instead, we use a method due to French (2005), Blau (2008) and French and Jones (2011) to approximate the evolution of $AIME_t$.

To explain the approximation, consider first the case of a person who has not yet worked a full 35 years. Let $N(t)$ denote the number of years the person has worked as of age t.

$$AIME_t = \frac{\sum_{\tau=t_0(e)}^{t} w_\tau h_\tau p_\tau}{35*12} = \frac{\sum_{\tau=t_0(e)}^{t-1} w_{\tau-1} h_{\tau-1} p_{\tau-1}}{35*12} + \frac{w_t h_t p_t}{35*12} = AIME_{t-1} + \frac{w_t h_t p_t}{35*12} \qquad (10)$$

Thus, if $N(t) \leq 35$, one can update $AIME_t$ just by adding $(w_t h_t p_t/35 \cdot 12)$ to $AIME_{t\text{-}1.}$ However, once $N(t) > 35$ the problem is more difficult, as we need to drop the minimum earnings year and then add $(w_t h_t p_t/35 \cdot 12)$. Specifically, we have:

$$AIME_t = AIME_{t-1} - \min_{\tau, p_\tau > 0} \frac{w_\tau h_\tau p_\tau}{35*12} + \frac{w_t h_t p_t}{35*12} \qquad (11)$$

The difficulty is that we need to have all 35 values of $w_\tau h_\tau$ for $p_\tau > 0$ saved in memory in order to know which one is the minimum.[9] The approximation due to French (2005), Blau (2008) and French and Jones (2011) replaces (11) with:

$$AIME_t = AIME_{t-1} + \max\{0, \frac{w_t h_t - AIME_{t-1} * 12}{35*12}\} \qquad (12)$$

In words, *AIME* is updated by letting earnings at age t replace an *average* year of previous earnings, provided $w_t h_t$ is above average. Otherwise *AIME* is left unchanged. In general, this will tend to understate growth in AIME, because we drop an average year rather than the lowest year. But in practice the approximation appears to be quite accurate.[10]

Social Security benefits are determined by applying a highly progressive tax structure to *AIME* to obtain what is known as the "primary insurance amount" (PIA). This is the SS benefit

[9] This is an extremely formidable computational problem (35 state variables!), and there is no simple trick that can solve it. For example, one might conjecture it would be good enough to keep track of just two state variables, *AIME* and the lowest previous earnings level. Indeed, this would be enough information to update *AIME* once. But then, after the lowest year to is dropped, one finds one needs to know the second lowest year, as it become the new lowest year. And so the proposed solution unravels.

[10] There is a maximum annual contribution amount (see Appendix A), but we ignore it here to conserve on notation.

that a person receives if he elects to begin receiving SS at the "normal" retirement age. The PIA formula differs by birth year (see Appendix A). In 2015 the formula was:

$$\begin{aligned} \text{PIA} = \; & 0.90 \text{ of AIME up to } \$826 \\ & +0.32 \text{ of AIME from } \$826 \text{ to } \$4{,}980 \\ & +0.15 \text{ of AIME over } \$4{,}980 \end{aligned} \tag{13}$$

A person's year of birth also determines the normal retirement age, which we denote by t_R^N. Normal retirement age was 65 for those born up until 1937, but will gradually increase to 67 for those born in 1960 or later. The benefit is reduced for early retirement and increased for late retirement. So, letting t_R denote actual retirement age, we have:

$$SSinc^* = f(PIA,\; t_R\,, t_R^N\,) \tag{14}$$

If one delays retirement benefits increase at a (roughly) actuarially fair rate up until age 70.[11]

A final complication is the taxation of earnings while receiving SS benefits. This is known as the "Earnings Test." A full description of the rules is complicated (see Appendix B). Basically, if one earns above a threshold, a fraction of benefits is withheld, but is credited back at a later age. This was age 70 until 2000 (covering most of our sample), but it was changed to the normal retirement age thereafter. Letting *SSinc* denote SS benefits net of withholding, we have:

$$SSinc_t = SSinc^* - \theta \cdot BW(w_t h_t, t, SSinc^*, t_B) \tag{15}$$

where $BW(\cdot)$ is the amount of benefits withheld, t_B is the person's birth year and θ is a parameter. The return of this "tax" (at age 70 or the normal retirement age) is approximately actuarially fair, conditional on the person expecting to live to 80+ and a fairly low discount rate. We estimate the

[11] Obviously, modeling SS is a major source of complication. It adds two state variables (AIME and retirement age) and a decision variable (claiming benefits). But the goal of our paper is not to model SS *per se*. We incorporate SS because we expect it may have important effects on life-cycle labour supply and savings. Thus, ignoring SS may lead to severe bias in estimating labor supply elasticities. For example, suppose that wages fall for men in their 60s due to depreciation of human capital. SS retirement benefits become available at the same time. A model that ignored SS could attribute the whole drop in labour force participation for men in their 60s to falling wages, and thus exaggerate labour supply elasticities. This is why we feel the added complexity of including SS is worthwhile.

parameter $\theta \in (0,1)$ which captures how people *perceive* the tax/refund. If $\theta = 0$ it is perceived as actuarially fair, and the tax is ignored. But if $\theta = 1$ then agents treat $BW(\cdot)$ as a straight tax.

1.6. *Private Pensions*

Aside from Social Security, another important consideration is private pensions. In contrast to SS, private pension rules are person specific, making it difficult if not impossible to capture their complexity. Instead we adopt a simple reduced form specification. We assume that, from age 55, there is a probability of receiving a private pension that depends on age, education, and lagged pension status ($dpen_{t-1}$). The probability is given by a logit model with latent index:

$$dpen_t^* = q_1 + q_2 t + q_3 t^2 + q_4 t^3 + q_5 HS + q_6(HS * t) + q_7(HS * t^2) \tag{16}$$

$$+q_8 college + q_9(college * t) + q_{10}(college * t^2) + q_{11} dpen_{t-1}$$

We estimate (16) on the HRS, and report the results in Appendix C. If a person gets a pension, we assume they get the mean conditional on their education level. The values, denoted Pen_t, are \$8992 for dropouts, \$14,617 for high school graduates and \$23,565 for college graduates.

1.7. *Unemployment Benefits*

Agents in our model may receive an unemployment benefit (ub_t) if they do not work. We assume the following simple benefit rule:

$$ub_t = \begin{cases} B & if \quad h_t = 0, h_{t-1} > 0, ss_t = 0 \\ 0 & \qquad otherwise \end{cases} \tag{17}$$

where B is a parameter to be estimated. While we refer to B as an "unemployment benefit," it is net of any non-pecuniary rewards from employment (see Keane and Wolpin (1997)). Thus, we would expect B to understate actual employment benefit levels.

1.8. *Medical Expenditures*

Our model also accounts in a simple way for medical expenditures (med_t). We feel this is important for two reasons. First, expected medical costs are an important aspect of life-cycle planning. Second, in order to fit asset paths at older ages it is quite important to account for medical costs, particularly as the US Medicare system covers only about 50% of the medical costs of people who are 65+. Thus, we fit the following simple function to out-of-pocket expenditures of males in the 1996-2009 Medical Expenditure Panel Survey (MEPS):

$$med_t = c_1 + c_2 t + c_3 t^2 + c_4 I(t \geq 65) \quad (18)$$

We report the estimates of this function in Appendix D.

1.9. *The Progressive Tax Structure*

An important feature of our model is that we account for the progressive tax structure. In order to calculate a person's tax liability, we need to calculate his total taxable income (TI_t):

$$TI_t = max\{w_t h_t + Cap_t + tSSinc_t + Pen_t - med_t - SD, 0\} \quad (19)$$

Here Cap_t is capital income, $tSSinc_t$ is taxable Social Security income, and SD is the "standard deduction." Capital income is given by:

$$Cap_t = \max\left\{\frac{rA}{1+r}, 0\right\} \quad (20)$$

We assume that if a person has negative assets then he has no taxable asset income.

Note that by entering gross earnings $w_t h_t$ in (19), rather than earnings net of fixed costs E_t, we assume the fixed costs of work in (8) are not tax deductible. In reality some fraction of fixed costs may be deductible, but it simplifies the analysis to abstract from this. Conversely, equation (19) assumes that medical expenses are deductible. In practice they are only deductible if they are sufficiently large, but it again simplifies the analysis to abstract from this.[12]

The taxable part of Social Security is based on a rather involved formula that basically takes half of $SSinc_t$, adds this to total income from other sources, and, based on the result, determines what part of $SSinc_t$ is taxable. We describe the algorithm in Appendix B.

We assume the standard deduction (SD) is equal to the 1999 level of \$7050.[13] Of course, some people itemize, but this is a close approximation to the mean deduction claimed by people

[12] We experimented both with dropping med_t from (19) and with replacing $w_t h_t$ with E_t, and found that our results were not sensitive to these changes.

[13] Specifically, this is the sum of the \$4300 standard deduction and the \$2750 personal exemption.

in our CPS data.[14] Finally, we approximate the progressive tax structure by using the CPS to fit the equation:

$$ln(Tax_t) = -3.9543 + 1.2263 \cdot ln\,(TI_t) \tag{21}$$

The function provides an excellent fit to the data, as we show in Appendix E.

1.10. *The Inter-temporal Budget Constraint*

We are now in a position to write the law of motion for assets, which is:

$$A_{t+1} = (1 + r)(A_t + E_t + SSinc_t + Pen_t + ub_t - c_t - med_t - Tax_t) \tag{22}$$

Here E_t is earnings net of fixed costs of work, which is given by equation (8). The interest rate r is set at .05. This discount factor is set to $\beta = 1/(1+r) = .9524$.

Note that we abstract from any special treatment of capital income. As can be seen from (19) to (22), we assume it is taxed at the same rate as labour and other forms of income. The cost of an extra unit of consumption today in terms of next period assets is $1 + r[1 - \tau(TI_t)]$ where the marginal tax rate $\tau(TI_t) \equiv \partial Tax_t/\partial TI_t$ is a function of taxable income. With a progressive tax structure, higher income leads to a lower after-tax interest rate, so the price of consumption is lower (Blomquist (1985)). As a result, *ceteris paribus*, agents would like to consume more at ages when income is higher, which of course is during the prime earning years.

1.11. *Survival Probabilities*

Let π_t denote the mortality rate at age t so $(1-\pi_t)$ is the probability of survival from age t to age t+1. We obtain survival probabilities for males from the US Life Tables for 2000 reported in Arias (2002). We also use the results in Brown, Liebman and Pollet (2002) to adjust these mortality rates for education, obtaining specific rates for dropouts, high school graduates and college graduates. For computational reasons, we assume that all college workers die with certainty at age 93, while all dropout and high school people die with certainty at age 90. This enables us to backsolve our model from a particular age.

2. Solution of the Model

[14] In order to make this comparison we used US Census Bureau calculations of taxable incomes of CPS respondents. Combining this with the total income data we can back out an estimate of deductions for each respondent.

In our model both the choice set and the state variables change at three points in time. These are age 55, when agents may start to receive private pensions, age 62, when the may start to receive Social Security and age 75, when we assume that all agents retire and must start to receive Social Security (if they have not already done so). Thus, we describe the solution of the model in five stages. Recall that we specify three different working lives for the three education groups: Dropout: ages 16-90; High school: ages 18-90; College: ages 22-93.

Stage A. Age (16,18, or 22)-54: State variables = $\{A_t, k_t, AIME_{t-1}, p_{t-1}\}$

From the initial age of entry into the labour market through age 54 there are three sources of uncertainty: wage shocks, job offers and mortality. Thus we have expected value functions:

$$E_\varepsilon E_{pjob_t} V(A_t, k_t, AIME_{t-1}, p_{t-1}, \varepsilon) = E_\varepsilon \{J_t(p_{t-1}, t) V(A_t, k_t, AIME_{t-1}, \varepsilon | pjob_t = 1) + (1 - J_t(p_{t-1}, t)) V(A_t, k_t, AIME_{t-1}, \varepsilon | pjob_t = 0)\}$$

where:

$$V(A_t, k_t, AIME_{t-1}, \varepsilon | pjob_t = 1) = \max_{c_t, h_t} \{u(c_t, h_t) + \beta[\pi_t EV_{t+1}(A_{t+1}, k_{t+1}, AIME_t) + (1 - \pi_t)\delta B(A_{t+1})]\}$$

$$V(A_t, k_t, AIME_{t-1}, \varepsilon | pjob_t = 0) = \max_{c_t} \{u(c_t, 0) + \beta[\pi_t EV_{t+1}(A_{t+1}, k_{t+1}, AIME_t) + (1 - \pi_t)\delta B(A_{t+1})]\}$$

Stage B. Age 55-61: State variables = $\{A_t, k_t, AIME_{t-1}, p_{t-1}, dpen_{t-1}\}$

At age 55 agents may begin to receive private pensions. A pension arrives stochastically with a probability denoted by $ppen_t$. This probability is determined by equation (16). Thus, there are four sources of uncertainty: wage shocks, job offers, pension receipt and mortality. We have:

$$E_\varepsilon E_{dpen_t} E_{pjob_t} V(A_t, k_t, AIME_{t-1}, p_{t-1}, dpen_{t-1}, \varepsilon) = E_\varepsilon \{ppen_t [J_t(p_{t-1}, t) V(A_t, k_t, AIME_{t-1}, \varepsilon | pjob_t = 1, dpen_t = 1) + (1 - J_t(p_{t-1}, t)) V(A_t, k_t, AIME_{t-1}, \varepsilon | pjob_t = 0, dpen_t = 1)]$$

$$+(1-ppen_t)[J_t(p_{t-1},t)V(A_t,k_t,AIME_{t-1},\varepsilon|pjob_t=1,dpen_t=0) \\ +(1-J_t(p_{t-1},t))V(A_t,k_t,AIME_{t-1},\varepsilon|pjob_t=0,dpen_t=0)]\}$$

where:

$$V(A_t,k_t,AIME_{t-1},\varepsilon|pjob_t=1,dpen_t) \\ =\max_{c_t,h_t}\{u(c_t,\ h_t)+\beta[\pi_t EV_{t+1}(A_{t+1},k_{t+1},AIME_t)+(1-\pi_t)\delta B(A_{t+1})]\}$$

$$V(A_t,k_t,AIME_{t-1},\varepsilon|pjob_t=0,dpen_t) \\ =\max_{c_t}\{u(c_t,\ 0)+\beta[\pi_t EV_{t+1}(A_{t+1},k_{t+1},AIME_t)+(1-\pi_t)\delta B(A_{t+1})]\}$$

Stage C. Ages 62-74: State variables = $\{A_t, k_t, AIME_{t-1}, p_{t-1}, dpen_{t-1}, ss_{t-1}, age_{ss}\}$
During this age range the agent is choosing each year whether to apply for SS benefits. He is also making choices about labour supply and consumption. And there continue to be four sources of uncertainty (as in stage B). This is by far the most computationally difficult stage of our model. We consider four cases:

(i) Those who haven't applied for SS benefits and who receive a job offer have a choice set with 12 discrete options (6 levels of hours and whether to apply for SS) along with consumption :

$$V(A_t,k_t,AIME_{t-1},\varepsilon|pjob_t=1,dpen_t,ss_{t-1}=0) \\ =\max_{c_t,h_t,ss_t}\{u(c_t,\ h_t)\ +\beta[\pi_t EV_{t+1}(A_{t+1},k_{t+1},AIME_t,dpen_t)+(1-\pi_t)\delta B(A_{t+1})]\}$$

(ii) Those who haven't applied for SS benefits and do not receive a job offer have a choice set with only two discrete options (whether to apply for SS). If they do not apply they receive the unemployment benefit, but if they do apply they do not:

$$V(A_t,k_t,AIME_{t-1},\varepsilon|pjob_t=0,dpen_t,ss_{t-1}=0) \\ =\max_{c_t,,ss_t}\{u(c_t|h_t=0)\ +\beta[\pi_t EV_{t+1}(A_{t+1},k_{t+1},AIME_t,dpen_t)+(1-\pi_t)\delta B(A_{t+1})]\}$$

(iii) Once a person has applied for Social Security, they can still work if they receive a job offer, giving 6 discrete labour supply choices along with consumption:

$$V(A_t, k_t, AIME_{t-1}, age_{ss}, \varepsilon | pjob_t = 1, dpen_t, ss_{t-1} = 1)$$
$$= \max_{c_t,\, h_t}\{u(c_t,\ h_t)\ + \beta[\pi_t EV_{t+1}(A_{t+1}, k_{t+1}, AIME_t, age_{ss}, dpen_t) + (1-\pi_t)\delta B(A_{t+1})]\,\}$$

Here age_{ss} $\in$[62,75] is the age when the person applied for Social Security. Once the person applies age_{ss} becomes a time-invariant state variable that affects the person's benefit level.

(iv) Those who have already applied for Social Security and who do not receive a job offer only choose consumption:

$$V(A_t, k_t, AIME_{t-1}, age_{ss}, \varepsilon | pjob_t = 0, dpen_t, ss_{t-1} = 1)$$
$$= \max_{c_t}\{u(c_t | h_t = 0)\ + \beta[\pi_t EV_{t+1}(A_{t+1}, k_{t+1}, AIME_t, age_{ss}, dpen_t) + (1-\pi_t)\delta B(A_{t+1})]\,\}$$

Stage D. Ages 75-89 (or 92): State variables $\{A_t, AIME_{74}, dpen_{t-1}, age_{ss}\}$

We assume agents must retire and start to collect SS benefits at age 75, if they have not already done so. So from age 75 onward the only choice is consumption, and only one source of uncertainty is whether one receives a private pension. The value function is given by:

$$V(A_t, AIME_{74}, age_{ss}, dpen_{t-1})$$
$$= E_{dpen_t}\Big[\max_{c_t}\{u(c_t)\ + \beta[\pi_t EV_{t+1}(A_{t+1}, AIME_{74}, age_{ss}, dpen_t)$$
$$+ (1-\pi_t)\delta B(A_{t+1})]\,\}\Big]$$

Stage E. Terminal Period: Age 90 (or 93): State variable $\{A_{T+1}\}$

We assume all agents have a terminal value function given by (2), the bequest function that depends on terminal assets and parameter δ. We backsolve the model from this terminal stage.

3. Estimation and Data construction

Having solved the model as discussed in section 2, we can simulate a sample of artificial agents, and estimate the model by method of simulated moments. That is, we seek to minimize:

$$\sum_k \sum_t \left(\frac{x_{k,t,model} - x_{k,t,data}}{se_{k,t,data}} \right)^2$$

where $x_{k,t,model}$ and $x_{k,t,data}$ are statistics on variable k at age t taken from the simulated and actual data, respectively, and $se_{k,t,data}$ is the standard error from actual data. The model is estimated separately for each of the three education groups. Recall that we assume there are two discrete types of taste for work within each education group (see equation (1)).

We fit the model to eleven types of moments (k=1,…,11), which we now describe:

1) Average employment rate at each age (i.e., from t=16, 18 or 22 to t=74).
2) Average annual hours conditional on work (t=16, 18 or 22 to t=74).

3) Median full-time hourly wage (t=16, 18 or 22 to t=74).

The data for moments 1 to 3 are taken from the CPS 1996-2005 (male household heads or male spouse of head). The annual hours variable is based on questions about usual hours worked and weeks worked last year. If someone works less than 250 hours a year, we treat him as not working. Wages are converted to 1999 dollars, and "full-time" in 3) is defined as working more than 35 hours a week and 40 weeks a year. Observations with an hourly wage below \$1.52 or above \$152 are excluded.

4) Average consumption (t=16, 18 or 22 to t=89 or 92).

Data is taken from the CEX for 2002-2006. We define consumption as total household expenditure net of medical costs. As the CEX unit of observation is a household, we apply a household equivalence scale to adjust to the individual level. We use the square root scale. Some cells with small sample sizes are grouped together, specifically t = 16-18 and t = 71-75, 76-80, 81-89 (or 92). This leaves, e.g., 66 moments for high school types.

5) Percent who Apply for Social Security benefits (t=62,..,75)

Data is taken from the HRS for 1992-2012. We use male respondents who were ≤ 55 and not yet retired when they entered the HRS study, who were born in 1937 or earlier, who report SS retirement income, and who don't report Social Security disability income. Our simulation assumes a birth year of 1934, so we chose the actual data to include men who were born near that date so they faced similar SS rules. Cells with small sample sizes are grouped together, specifically t = 68-69, t =70-71 and t = 72-75, giving 9 moments.

6) Standard deviation of hours (t=16, 18 or 22 to t=74).

7) Standard deviation of hours conditional on work (t=16, 18 or 22 to t=74).

8) Standard deviation of log of the hourly wage rate (t=16, 18 or 22 to t=74).

Moments 6 to 8 are based on the same CPS data used for k = 1 to 3. The data are grouped into 12 five-year intervals, so items 6-8 contribute 12 moments each. Before constructing the standard deviations of hours, the CPS data on hours are grouped into the same six discrete categories used in the model. That is 0-250 is classified as 0, while 251-750 is classified as 500, and so on, until 2251+ is classified as 2500. Work by Keane and Wolpin (1997) and Imai and Keane (2004), among others, finds that roughly half the standard deviation of reported (log) wages is due to measurement error. To account for this, we divide the standard deviation of the log hourly wage by two.

9) Transition rate: Probability of working conditional on working last year

10) Transition Rate: Probability of working conditional on not working last year

The transition rates in 9) and 10) are from CPS data from 1996 to 2005 (as in k =1 to 3). We use the CPS questions about the preceding week and the last year. For "last year," "employed" is based on the same definition as in k=1 to 3. For "preceding week," "employed" includes those who report any paid work, including those who have jobs but were not at work. The data are grouped into 12 five-year intervals to reduce noise.

11) Standard deviation of consumption

The data is a subset of that used in k=4. We use only consumer units observed for 9+ months to calculate annual consumption. For units observed for less than a year, we scale up consumption to 12 months. Based on prior estimates, we divide the standard deviation by two to account for measurement error. The data are grouped in 14 five-year intervals.

In summary, we have selected moments that capture means/medians, standard deviations and transition rates for the key variables in our model (employment, hours, wages, consumption and retirement). The model is fit to 326, 320 and 302 data moments for the dropout, high school and college groups, respectively. As we discussed in Section 1, we also use HRS data to fit the process for private pensions, MEPS data to fit the process for medical expenditures, and CPS data to fit the tax function. These data are described in more detail in Appendices C, D and E.

4. Parameter Estimates

The parameter estimates are reported in Table 1. The model contains 27 free parameters that we estimate separately for each of the three education groups. We view this as a very parsimonious structure, given the amount of data the model attempts to fit. Given the nature of our model, the parameter estimates are of secondary interest relative to the policy simulations. But here we highlight some of the more important estimation results.

Given the utility function in (1), if we were to abstract from human capital, fixed costs of work, progressive taxes and other complications, the model would generate the following (age invariant) elasticities: Frisch = $1/(a_2-1)$, Hicks = $1/(a_2-a_1)$ and Marshall = $a_1/(a_2-a_1)$. Our estimates of a_1 and a_2 are similar for all three education groups, with $a_1 \approx 0.25$ and $a_2 \approx 1.5$. These values imply that Frisch ≈ 2, Hicks ≈ 0.80 and Marshall ≈ 0.20. In Section 6 we will compare these figures with the actual elasticities implied by the model.

In the human capital process, the main point of interest is the parameter λ_2, the coefficient on work hours, which is 0.0039 for high school dropouts, 0.0048 for high school graduates, and 0.0059 for college graduates. Thus, the returns to work experience are greater for more educated workers. Our estimates imply small effects of age itself on human capital. The annual depreciate rate of human capital (1-λ_1) is 9.4% for college types, and 8% for high school and dropout types.

As for fixed costs of work, the time fixed cost (fc_h) ranges from 107 hours for the dropouts to 127 hours for the college graduates. This means that one must work at least that many hours per year before earning any income. The monetary fixed cost of work is estimated to be about $700 to $800 per year.

There are two notable aspects of the job offer probability function. First, the intercept increases substantially with education (from 2.48 for dropouts to 2.96 for college graduates). Thus, *ceteris paribus*, more educated workers have a higher probability of receiving an offer. Second, the coefficients (m^{30}) on the indicator for lagged unemployment (1-p_{t-1}) are large and negative, indicating the chance of receiving an offer is much lower if one was not employed in the previous period. Thus, in our model, the shadow price of time exceeds the after-tax wage for three reasons: (i) working increases human capital, (ii) working increases the probability of a job offer in the next period, and (iii) working leads to accrual of SS benefits.

Finally, the parameter θ that captures perceptions of the "earnings test" is close to zero for college and high school workers, implying they do not view it as a tax. But for dropouts we obtain θ =0.40, implying they do not believe the refund is even close to being actuarially fair. This may be because the dropouts have a shorter life expectancy, or are more concerned about short-run liquidity constraints.

5. Data Patterns and Model Fit

In this Section we describe the fit of the model. Along the way, we also describe some key patterns in the data. In Figure 1 we see that the model provides a good fit to variation in average hours of work, both in terms of the life-cycle paths and the level differences across education groups. Notice that for college workers, hours rise rather quickly over the first ten years of the working like (e.g., from about 1600 at age 22 to 2200 at age 32), remain fairly flat until roughly age 50, and then begin a decline that starts slowly at first but accelerates in the 60s. The high school and drop out groups follow very similar patterns, except that (a) the peak level of hours is lower for the less educated, and (b) the decline in hours starts at a younger age for the less educated (e.g., at roughly age 40 for the dropouts).

Figures 2 and 3 decompose life-cycle hours paths into extensive (i.e., employment) and intensive (i.e., hours given employment) margins. The model provides quite a good fit to both variables. A striking pattern is that hours conditional on employment are very flat over most of the life-cycle. They do rise substantially over about the first 10 years of the working life (e.g., from about 1750 at age 22 to 2250 at age 32 for college workers). But hours are then remarkably flat until about age 55. There is a decline from roughly 55 to 65, followed by another flattening out. Strikingly, men who continue to work at 65+ still tend to work about 1500 to 1750 hours.

Thus, the steep drop in average hours at ages 50+ that we see in Figure 1 is driven primarily by the decline in participation that we see in Figure 2. This pattern is what we would expect in a model with fixed costs of work and a nonlinear earnings schedule, as emphasized by Prescott et al (2009), Rogerson and Wallenius (2009) and Wallenius (2011).

Figure 4 reports how the model fits the median full-time wage rate by age and education group. A key feature of the data, which the model captures well, is the much steeper wage-age profile for more educated workers, suggesting greater returns to work experience. As Imai and Keane (2004) note, this has important implications for labour supply elasticities (see below).

Figure 5 reports how the model fits average consumption. This is consumption for the individual men in our sample, adjusted for family composition via an equivalence scale (Section 3). Aside from the level differences between education groups, the most striking pattern is that college workers have a pronounced hump shape in consumption over the life-cycle, while the path for dropouts is very flat. The model captures these features well. High school workers have a more modestly humped consumption profile than college graduates. One failure of the model is that it generates excessively smooth consumption for high school workers over the life-cycle.[15]

Figure 6 plots model predictions for the age of claiming Social Security. For high school and dropout workers the fit is quite good. We capture the large spike in claims at age 62 (the first age of eligibility) and we capture that nearly everyone claims by 65. A small weakness of the model is that we generate that almost no one delays past 65, while in the data a few percent do. In the data, college workers tend to claim Social Security later. The model captures this pattern, but it exaggerates it. Specifically, we fail to capture the spike in college workers claiming Social Security at age 62 (about 33% in the data vs. 9% in the model), and we exaggerate the spike at

[15] We have been unable to determine why this problem occurs, but we did not regard it as serious enough to warrant complicating the model to try to correct it.

65 (40% vs. 30%).[16] It is important to emphasize, however, that claiming Social Security is not equivalent to leaving the labour force, and Figure 2 shows we capture the timing of labour force exit for all three education groups quite accurately.

Figure 7 shows how the model fits the standard deviation of hours of work. Recall that the model divides work hours into six discrete categories, so the standard deviation summarizes the dispersion of hours across these categories. There are three notable patterns, (i) dispersion generally increases with education, (ii) dispersion grows with age, and (iii) the peak in dispersion is later for more educated workers. The model captures these patterns well.[17]

In Figure 8 we see two notable patterns for the standard deviation of log wages: (a) it is greater by about .05 log points for college workers than for high school and dropout workers, and (ii) after growing very slowly over most of the working life, wage dispersion starts to grow sharply at about age 55+. The model misses this sharp growth in dispersion at older ages, and it over-estimates the standard deviation for high school workers by about .05 log points.

Figure 9 plots the standard deviation of consumption. Overall, the model provides a good fit, both to the inverse U-shaped life-cycle paths and to level differences across education groups. But it somewhat overstates the degree of dispersion for dropouts.

Finally, Figures 10 and 11 report the fit to two key transition rates, the work-to-work rate and the non-work to work rate. In Figure 10 we see that the work-to-work transition rate is very high for all three education groups from labour market entry through to about age 45 (i.e., about 90% to 95%). But from the late 40s through to the 70s it begins a sharp decline. For instance, by age 65 the work-to-work transition rate falls into the 60% to 70% range for all three education groups. The model captures this decline qualitatively, but it fails to capture its magnitude. Of course, the number of older people who work is fairly small, so the variances of the moments involving this transition rate are high. Thus the model does not put much weigh on these errors.

[16] We did not see missing the peak at age 62 in Social Security claims for college workers as a serious enough problem to warrant complicating the model to try to capture this detail.

[17] The standard deviation of hours *conditional* on working follows a "bathtub" shape over the life-cycle, dropping for roughly 10 years after labour force entry, staying flat until about age 55, and then growing sharply at older ages. There is also a clear ranking by education, with the standard deviation at the bottom of the "bathtub" ranging from about 500 hours for dropouts to 400 hours for college types. The model captures these shapes and level differences (across groups) very well. But it systematically understates the standard deviation for all groups by about 100 hours. So, in contrast to most prior work (that has largely ignored nonlinear earnings schedules and assumed continuous hours choice), our model errs on the side of making the hours distribution slightly too concentrated (rather than too dispersed).

In contrast, the fit to non-work to work transitions that we see in Figure 11 strikes us as very impressive, especially at older ages. This statistic does not have the problem of being based on small sample sizes at older ages. An obvious feature of the data is that the probability a non-working individual transitions to work is much greater for more educated workers. The probability that a non-working individual transitions to work also declines steeply with age, particularly after about age 40. The model captures these patterns quite well.

6. Tax Policy Experiments

In this Section we use the model to simulate several different tax policy changes.

6.1. *Simple Flat Tax Experiments*

First, we consider the introduction of a new 5% flat rate tax on all labour income. This means we modify the tax function in (21) to be:

$$Tax_t = exp(-3.9543 + 1.2263 \cdot ln\,(TI_t)) + 0.05 \cdot w_t h_t \tag{23}$$

Of course, this is not a (politically) realistic change in the tax rule. We consider (23) because it allows us to calculate Frisch, Hicks and Marshall elasticities for our model. A key motivation of our work is to develop an understanding of how these elasticities behave in a model with both human capital investment and an active extensive margin.

We consider 3 types of tax change: (1) transitory, (2) permanent uncompensated, and (3) permanent compensated. These generate the Frisch, Marshall and Hicks elasticities, respectively. We assume the transitory tax changes are unanticipated and uncompensated.[18] We will consider both permanent tax changes that occur at the very start of the working life (regime shifts), and permanent tax changes that occur as surprises at later ages. We describe how we calculate the compensation needed to obtain the Hicks elasticity in Appendix F.

6.1.1. *The Frisch elasticity*

Figure 12 plots our calculations of the Frisch elasticity by age. The most striking aspect of the figure is simply that the Frisch varies substantially with age. In the standard life-cycle model (MaCurdy (1981) the Frisch elasticity is a constant. And, given our estimated preference

[18] This gives a good approximation to the Frisch elasticity. But it is a bit of an underestimate as even transitory tax changes have small income effects.

parameters, that constant should be $1/(a_2\text{-}1) \approx 2$. Keane (2015) shows that if human capital is added to the standard model the Frisch is dampened at young ages and increases with age. But in our model, which also includes the extensive margin and other features, the Frisch varies with age in a much more complex way. Some key features of Figure 12 are worth noting:

First, the Frisch elasticity is much lower for more educated workers. Second, at young ages the Frisch oscillates quite substantially for all three groups. Third, from roughly age 25 to 60 the Frisch is fairly stable for dropouts, hovering in the 1.0 to 1.3 range. In contrast, it exhibits a pronounced U-shape for high school and college workers. For example, for high school types it falls from 0.90 at age 25 to only 0.46 at age 35 back up to 0.94 at age 60. Fourth, at ages 60+ the Frisch goes well above 1.0 for all three groups.

6.1.2. *The Hicks and Marshall elasticities*

Next, we turn to Hicks and Marshall elasticities. Consider a permanent 5% tax increase that is in effect for a person's entire life (i.e., a tax regime shift). We are interested in the effect on total lifetime hours. This may be called a "long-run" effect, because it allows for human capital investment to adjust over the life-cycle. The uncompensated (Marshallian) elasticities are 0.33, 0.19 and 0.22 for dropout, high school and college workers, respectively. The compensated (Hicks) elasticities are 1.01, 0.65 and 0.74. To put these figures in perspective, the Imai and Keane (2004) model implies a Hicks elasticity of 1.30, averaged over all education groups. The values we obtain here are smaller, but still much larger than most of the prior literature (e.g., the Keane (2010) survey finds an average Hicks elasticity of 0.31).

Next, we look at how the Hicks and Marshall elasticities vary with age. In Figure 13, we consider the same permanent 5% tax increase considered above, but now we plot the Hicks (compensated) response by age. For all three education groups, the lifetime hours reduction is concentrated at young and especially older ages. As Keane (2015) discusses, the life-cycle model with human capital implies that the Hicks elasticity should rise with age (as human capital concerns become less important), while the extensive margin model predicts a large Hicks elasticities for both young and old workers (who are close to indifferent between working and not working). The pattern we see in Figure 13 appears to be a compromise between the two models. For high school and college workers the Hicks elasticity is a bit larger for young workers than for prime-aged (30-55) workers, but the increase at age 55+ is very dramatic, just as in Imai and Keane (2004). But for dropouts, for whom returns to human capital are always small, we see a more symmetric U-shape that looks very much like the patterns predicted by the extensive margin models of Prescott et al (2009) and Rogerson and Wallenius (2009).

The Marshallian elasticities over the life cycle are plotted in Figure 14. We see the same basic pattern as for the Hicks elasticity. What is notable is that, given that the already modest Marshallian elasticities of lifetime hours are concentrated among young and especially older workers, the Marshallian elasticities for males in the 30 to 50 age range are really quite small. Thus, nothing in our results contradicts the standard finding that uncompensated labour supply elasticities for prime age men are very small.

6.1.3. *Surprise tax changes*

Next, we consider surprise permanent tax changes that may occur part-way into the working life. The top panel of Table 2 reports the short-run impact of tax changes that occur at various different ages. By "short-run" we mean the impact in the year the tax change occurs. This holds human capital fixed, as it has not had time to adjust to the changed environment. We find average (over all ages) short-run Hicks elasticities of 1.31, 0.69 and 0.65 for dropouts, high school and college workers respectively. In the CPS 1996-2005 data we used in estimation, the percentage from dropout, HS, and college graduates are 22%, 55% and 23%. So our estimates imply a short-run Hicks elasticity of roughly 0.82 in the population.

One reason it is interesting to examine the immediate or "short-run" impact of permanent tax changes is that this creates a point of comparison with the growing literature on tax reform experiments. For instance, Chetty (2012) pools estimates from many existing studies, most using the short-run effects of tax reforms as the source of identification, and obtains a Hicks elasticity of 0.58. This is a below our estimate, but that is not surprising given that many of the studies he examines focus on prime age males, who have relatively low elasticities.

Keane (2015) notes that, in a model with human capital, it is theoretically possible for the Hicks elasticity to exceed the Frisch. An interesting result in the top panel of Table 2 is that this happens at every age for college workers. For example, at age 25, the short-run elasticity with respect to a compensated permanent tax change is 1.56, while that with respect to a transitory tax change is 1.09. This illustrates how the desire to continue to acquire human capital dampens the labour supply response of college workers.

The bottom panel of Table 2 reports the long-run impact of tax changes that occur at several different ages. Given human capital accumulation, we would expect these to be larger than short-run impacts (as a higher tax rate reduces human capital accumulation). For example, the Imai and Keane (2004) model implies an average (over all ages) of the short-run Hicks elasticity of 0.70 rising to 1.30 in the long run. But the situation here is more complex, because the Hicks elasticity varies in a more complex way over the life cycle, and can sometimes be very large at young ages. For example, as we see in Table 2, short-run Hicks elasticities are very large for college workers at around age 25. Thus, the short-run elasticity with respect to a permanent tax increase is 1.56, while the long run elasticity is only 0.75. However, once we get past age 30, the general pattern that long run effects are greater than short run effects does hold.

6.2. *Changes in the Baseline Tax Structure*

We now consider more realistic tax experiments that involve changing parameters of the baseline tax function (21). As the tax structure is progressive, the rate depends on the taxable income level. For example the tax on \$20k of taxable income is \$3606 (or 18%), while that on \$100k is \$25,950 (or 26%). This gives an idea of the degree of progressivity implied by (21).

In our first experiment, we increase the constant term in (21) from -3.9543 to -3.7048. This corresponds to an across-the-board 28% increase in the tax rate on taxable income.[19] Labour supply elasticities are typically reported with respect to percentage changes in the after-tax rate (1-τ), not the tax rate (τ). A δ% increase in the tax rate corresponds to a $-\delta\cdot\tau/(1-\tau)$% decrease in the after-tax rate.[20] Thus, the percent change in τ is invariant to the level of income, but, under a progressive structure, the percent change in (1-τ) is not. To deal with this, we report elasticities evaluated at average values of (1-τ) for simulated workers. Appendix G describes three methods of averaging that generate similar results, so here we just give results from our preferred method.

Our model implies uncompensated (i.e., Marshallian) elasticities of lifetime total hours of 0.58, 0.39 and 0.35 for dropout, high school and college workers, respectively. It is interesting that these values are roughly twice as large as what we found in the flat-tax experiment (0.33, 0.19, 0.22). The reason is the mechanism emphasized by Hausman (1981, 1985). In a world with progressive taxation, a decrease in the after-tax rate on taxable income generates an increase in

[19] Under the experiment the tax on \$20k of taxable income increases to \$4628 (i.e., from 18% to 23.1%), while that on \$100k increases to \$33, 290 (i.e., from 26.1% to 33.3%).
[20] For example, if τ = 20% then the 28% increase in tax rates in our experiment corresponds to a 7% decline in the after-tax rate, while if τ = 30% the decline is 12%.

virtual non-labour income (by pivoting the budget constraint). This induces an income effect that augments the Marshallian elasticity. To our knowledge, this is the first time this effect has been shown to be quantitatively important in a dynamic setting.

The compensated (Hicks) elasticities are 1.17, 0.82 and 0.87 for dropout, high school and college workers, respectively. These figures are 16% to 26% greater than what we found in the flat-tax experiment (1.01, 0.65, 0.74), and much greater than typical values of the Hicks elasticity in the literature. Recall that our utility function estimates imply Marshall and Hicks elasticities of roughly 0.20 and 0.80 in a flat-tax, linear earnings function world with no human capital. Thus, our results suggest that accounting for these features of the economic environment magnifies the Marshall and Hicks elasticities in an economically significant way.

Next we change the slope coefficient in (21) to simulate an increase the progressivity of the tax structure. Specifically, we increase the slope coefficient in (21) from 1.2263 to 1.2534. At the low taxable income level of $10,000 this has the same effect as our previous increase in the intercept. That is, it raises the tax rate by 28% and lowers the after-tax rate by 5%. However, at the higher taxable income level of $100k we now raise the tax rate by 37% and reduce the after-tax rate by 13%. The marginal tax rate at $100k increases from 26% to 35%.

Our model implies uncompensated (i.e., Marshallian) elasticities of lifetime total hours of 0.71, 0.47 and 0.43 for dropout, high school and college workers, respectively. These are more than 20% greater than elasticities we obtained by changing the intercept, which increased rates by 28% across the board, holding the degree of progressivity fixed. The compensated (Hicks) elasticities that we obtain are 1.34, 0.93 and 1.03 for dropout, high school and college workers, respectively. These are roughly 15% greater than what we obtained holding progressivity fixed.

We also tried increasing the degree of progressivity even further by setting the slope equal to 1.2658. This raises the tax rate at $100k from 26% to 41%. Given this tax structure, Marshallian elasticities increase to 0.75, 0.52 and 0.50 for the three education groups, while Hicks elasticities increase to the very substantial levels of 1.40, 1.00 and 1.26, respectively.

These experiments illustrate the point stressed by Keane and Rogerson (2012), and earlier by Hausman (1981), Blomquist (1983), that labour supply elasticities are not invariant parameters, but depend on the tax structure and the nature of tax changes. We see how labour supply becomes more elastic under a more progressive tax structure, and that increases in progressivity generate more substantial declines in labour supply than we would expect given knowledge of the utility function parameters alone.[21]

6.3. *Labour Supply on the Intensive vs. Extensive Margins*

In this section, we decompose the elasticity of total hours into elasticities on the intensive vs. extensive margins. This decomposition has been the subject of considerable interest in the literature – see, e.g., Kimmel and Kniesner (1998) and Erosa, Fuster and Kambourov (2014). Figure 15 shows the results of the introduction of a compensated 15% flat rate tax on labour earnings. That is, in equation (23) we set the coefficient on $w_t h_t$ equal to 0.15.[22]

For dropouts, the participation elasticity has a pronounced U-shape. It is over 1.5 in the teens, drops to a trough of about 0.50 in the 30s, and then rises above 1.0 at 55+. In contrast, the elasticity of hours (conditional on work) is quite flat over the whole life-cycle, hovering in the vicinity of 0.70 (except for values over 1.0 at 16 to 17). These patterns are broadly consistent with the predictions of the extensive margin models of Prescott et al (2009) and Rogerson and Wallenius (2009). At both young and old ages, workers are close to the margin of indifference between working and not working (as productivity is low and/or health is poor). Furthermore, human capital accumulation is not an important factor for these workers (see Figure 4), so it has little impact on the life-cycle elasticity patterns.

But the elasticity patterns for high school and college workers are very different. For them, extensive margin elasticities are trivially small at young ages, and reach essentially zero just a few years into the working life. But their participation elasticities begin to grow sharply in the 50s, and surpass 1.0 in the early 60s. The intensive margin elasticities of high school and college workers also grow sharply at 60+. These patterns are broadly consistent with the human capital model described by Keane (2015). Given that human capital is important for both high school and college workers (see Figure 4), their labour supply is relatively insensitive to the after-tax wage rate at young ages. But once they reach and 50s and 60s, the returns to human capital investment become negligible, so their labour supply elasticities increase markedly.

7. The Impact of Human Capital on the Intensive vs. Extensive Margins of Labour Supply

[21] Recall that our utility function parameters would imply Marshall and Hicks elasticities of only 0.20 and 0.80 in simple flat-tax world without human capital and the other complicating aspects of our model.

[22] There is no unique decomposition of extensive vs. intensive margin elasticities: We can always construct a small enough tax increase that no one ceases to work, or a large enough increase that everyone ceases to work. We tried several different changes, and found the extensive margin elasticity tends to grow with the size of the tax increase. For increases smaller than 15% we find essentially no participation response for young high school and college workers.

In this section we ask how human capital affects labour supply behaviour in a model with an active extensive margin. From Imai and Keane (2004) and Keane (2015) we already have a good idea of how human capital affects elasticities in a model with only an intensive margin. But here we can answer additional questions: For instance, does human capital primarily affect elasticities on the intensive or extensive margins? Or is it important for both?

In order to address these questions we conduct two experiments. In each case we start from the baseline model for college graduates. In one experiment we replace the parameters of the human capital production function for college graduates with those for high school drop outs.[23] In a second experiment we shut down the endogenous human capital channel entirely. That is, we set the returns to experience to zero, and let wages follow a fifth order polynomial in age. This polynomial provides a very good fit to the life-cycle wage path. However, entirely eliminating returns to experience significantly reduces lifetime labor supply, so we recalibrate the taste for leisure parameters so that the model still matches total hours of work over the life-cycle.[24] We will refer to these two experiments as the "low HC" and "no HC" scenarios.

Figure 16 shows how the experiments affect the life-cycle hours profile. In both experiments, there is little change in the employment profile. Thus, we only show the plot for total hours. Note that, when human capital is eliminated from the model ("no HC"), the life-cycle hours profile becomes much steeper at young ages, and declines more slowly at older ages (thus providing a poor fit to the true life-cycle hours profile). Bear in mind that we have kept the shape of the wage-age profile essentially unchanged. Thus, if we want to restore the fit of the model to the hours profile, we would need to reduce the responsiveness of hours to wages by increasing parameter a_2. Of course this would reduce the Frisch, Hicks and Marshall elasticities.

This is another way of illustrating the argument in Imai and Keane (2004) that failure to account for human capital leads to downward bias in labour supply elasticities. But we can go further: From our results we see that <u>the main effect of ignoring human capital on *estimation* will be downward bias in intensive margin elasticities</u>. This is because ignoring human capital distorts the shape of the hours profile, but has little effect on the participation profile.

Next we consider how reducing the importance of human capital affects labour supply

[23] These parameters are λ_1 through λ_5 and $\bar{h}$ (see Table 1C). We leave the skill endowment λ_0 unchanged.

[24] That is, we set the experience coefficients $\lambda_2 = \lambda_3 = 0$ and also set $\bar{h}=0$. We augment the second order polynomial in age in equation (4) with terms up to the fifth order, which we denote by λ_6 to λ_8. We then re-estimate λ_1 and λ_4 to λ_8 to fit the wage data. We then recalibrate b_1 and b_2 to match total lifetime hours. But we do not re-estimate other parameters, as our goal is to analyze the effect of eliminating endogenous human capital holding other factors fixed.

elasticities *conditional* on parameter estimates. It is worth emphasizing that in this exercise we do not recalibrate the "low HC" and "no HC" models in an attempt to improve their fit to the shape of the life-cycle hours profile, beyond the simple level adjustment in taste for leisure noted earlier. This because we want to examine the *ceteris paribus* effect of reducing the importance of human capital on labour supply elasticities.[25]

First we consider the Hicks elasticity on the extensive (participation) margin. As we see in Figure 17, for college workers the extensive margin elasticities are very small at young ages, and reach essentially zero just a few years into the working life, but they begin to grow sharply at 55+. However, this pattern changes substantially when we reduce the importance of human capital. The extensive margin elasticity grows substantially at young ages (<30) and it is reduced at older ages (55+). Thus, without human capital, the extensive margin elasticity clearly has the "bathtub" shape discussed in Keane (2015) as a prediction of extensive margin models.

These results are consistent with the idea that, without human capital, both young and old workers are close to the margin of indifference between working and not working, so both young and old workers have high elasticities (generating the bathtub shape). But with human capital, the price of time is well above the current wage for young workers, so they are no longer close to the margin of indifference between working and not working. As a result, returns to human capital investment substantially reduce extensive margin elasticities at young ages.[26]

Now consider the Hicks elasticity on the intensive margin. Interestingly, as we see in Figure 18, the complete elimination of human capital accumulation has a negligible effect until age 45. Then, at older ages, the intensive margin elasticity is reduced in the model without human capital, just as the extensive margin elasticities were reduced at older ages in Figure 17.

Thus, at young ages the main impact of human capital is to reduce Hicks labour supply elasticities on the *extensive* margin. The arguments in Imai and Keane (2004) and Keane (2015) indicate that elasticities at young ages should be dampened by human capital, but why does the effect occur almost entirely on the extensive margin? We conjecture this is because of the discreteness of hours options in our model, a feature introduced to account for the observed

[25] The effect of including (or ignoring) human capital in the life-cycle model consists of: (i) the impact of human capital on *estimates* of preference parameters, and (ii) the impact of human capital on the *behaviour* of the model *conditional* on preference parameters. We have just discussed (i), and now turn to (ii).
[26] Why does human capital cause extensive margin elasticities to fall at older ages? With human capital agents work more at young ages and accumulate more assets. Because they work less at young ages, workers in the "low HC" and "no HC" scenarios need to work more at older ages (despite falling wages) to finance retirement.

bunching of hours.[27] Intuitively, the lost human capital from not working at all at young ages is substantial, but whether one works 30 vs. 40 hours per week is presumably much less important.

Next we consider the Frisch elasticity on the extensive margin. As we see in Figure 19, in both the baseline model and the "low HC" model the Frisch elasticity is essentially zero until age 55, at which point it starts to increase substantially. But in the "no HC" scenario the Frisch is fairly sizable at young ages (<25). These results are qualitatively similar to those for the Hicks extensive margin elasticity, but quantitatively human capital is much less influential here.

Finally, consider the Frisch elasticity on the intensive margin. In Figure 20, we see that reducing the importance of human capital substantially increases the Frisch intensive margin elasticity at young ages. Elasticities for the three models get closer over time, but (with minor exceptions) remain larger in the "low HC" and "no HC" models until about age 60. Thus, the main impact of human capital is to reduce Frisch elasticities on the intensive margin, more so at younger ages, with only very modest impacts on the extensive margin.

8. Conclusion

Using data on males from the CPS, HRS and CEX, we have estimated a labour supply model that includes both human capital and an active extensive margin, along with several other key features of the economic environment, such as progressive taxation, savings, bequests, and social security benefits. We conclude by summarizing some of our main results:

Our model implies compensated (Hicks) elasticities of total lifetime hours with respect to a (hypothetical) flat-rate tax imposed on labour earnings of 1.01, 0.65 and 0.74 for dropout, high school and college workers, respectively. Averaging by population shares we obtain an aggregate Hicks elasticity of 0.75. The aggregate Marshallian elasticity is 0.23.

The progressive tax structure amplifies elasticities considerably. For instance, aside from imposing a flat-tax on earnings, we also implement a more realistic experiment of proportionally scaling up all rates in the existing progressive tax structure. Responses to this experiment imply an aggregate Hicks elasticity of 0.91, and an aggregate Marshallian elasticity of 0.42.

In a further experiment, we increase taxes in a way that increases progressivity of the tax structure. We pivot the tax schedule so the rate at $100k of taxable income increases from 26% to 41%. The labor supply response to this policy change implies an aggregate Hicks elasticity of

[27] Recall we assume that annual hours are chosen from the set $h_t \in [0, 500, 1000, 1500, 2000, 2500]$.

1.15 and a Marshallian elasticity of 0.57. These results highlight the point emphasized by Keane and Rogerson (2012) that labour supply elasticities are not fixed parameters – rather, they vary with a number of factors, including the tax structure itself.

As we noted in the introduction, models with endogenous human capital, such as Imai and Keane (2004), predict labour supply elasticities should grow with age. In contrast, models with an extensive margin (and exogenous human capital), such as Prescott et al (2009) and Rogerson and Wallenius (2009), predict a U-shape for elasticities. Our model predicts a more complex pattern that combines features of both models, and that also differs by education:

For high school dropouts, for whom human capital returns to work experience are small, the Hicks elasticity follows a U-shape pattern over the life-cycle (see Figure 13). This looks very much like the patterns predicted by extensive margin models. In contrast, for college and high school types, the Hicks elasticity is only a bit larger for young workers than for prime-aged males; but it increases dramatically from age 55 onward, just as in Imai and Keane (2004).

We also ask how much of observed labour supply responses occur on the intensive vs. extensive margins. Our model implies that for dropouts of all ages, and all men over 55, the extensive margin is very important (see Figure 15). But for high school and college men below age 55, most labour supply response occurs on the intensive margin. Thus, both margins are important, but the extensive margin is relatively more important for less skilled workers. In fact, except for old workers (65+), extensive margin elasticities fall substantially with education (i.e., initial human capital), while intensive margin elasticities increase very modestly (if at all).

Finally, we examine how endogenous human capital affects labour supply elasticities on the extensive vs. intensive margins. Our analysis leads to three key findings: 1) the main effect of ignoring human capital in *estimation* is downward bias in *intensive* margin elasticities. In terms of *behaviour* of the model (for given preferences), the main impact of endogenous human capital is to: 2) substantially reduce Hicks *extensive* margin elasticities at young ages, while increasing them at old ages, and 3) reduce Frisch *intensive* margin elasticities, more so at young ages.

Conversely, we find that ignoring human capital does not much affect: 1) estimates of *extensive* margin elasticities, 2) values of Hicks *intensive* margin elasticities, or 3) values of Frisch *extensive* margin elasticities. These findings help clarify many results in the male labour supply literature. For instance, work that ignores endogenous human capital generally finds small intensive margin elasticities and large extensive margin elasticities (see Keane (2011)). Our results show this pattern does not necessarily imply that all the “action” in male labour supply is on the extensive margin. In fact, this pattern is perfectly consistent with a world where human capital is important and intensive margin elasticities are actually large.

In this paper we have focussed on our model’s implications for labour supply elasticities, but it could also be used to address a number of other important issues:

An obvious avenue for further research is to explore the model’s implications for changes in Social Security rules. In a recent paper, Wallenius (2013) calibrates a representative agent equilibrium model with intensive and extensive margins, endogenous human capital, and Social Security. She finds that differences in Social Security systems can explain a large part of the differences in labour supply at older ages between the US and continental European countries.[28] It would be interesting to use our model to see how her results are affected by allowing for worker heterogeneity, as well as factors like progressive taxation, private pensions and Medicare.

Another important issue is how progressive taxation affects income distribution. Guvenen et al (2012) look at the effects of progressive taxation on inequality in a model with endogenous human capital but no extensive margin. They find that a more progressive tax structure reduces inequality by causing high skill workers to spend less time investing in human capital. Our model includes a non-convexity in the mapping from hours to earnings, which helps generate an active extensive margin. This non-convexity makes workers less flexible in the choice of hours conditional on work, which may lessen the effect of progressive taxes on wage inequality.

Other areas where the implications of the model could be more fully explored are the effects of private pensions, health care costs, Medicare benefits and unemployment benefits on labor supply, retirement and asset accumulation. It would also be worthwhile to more carefully examine the impact of our assumption of discrete hours levels on the behavior of the model. Finally, for policy evaluation, it would obviously be useful to embed our model in an equilibrium setting, as in van der Klaauw and Wolpin (2008). But that represents a formidable challenge.

Michael P. Keane
University of Oxford and University of New South Wales

[28] In related work, Erosa et al (2014) look at effects of progressive taxes, Social Security and disability in a model with an extensive margin but no human capital. They also find that differences in Social Security rules are an important source of lower hours in Europe than the US. They find that progressivity of taxes plays a smaller but still important role.

Nada Wasi

Survey Research Center, University of Michigan

References

Aaberge, R., Dagsvik, J. and S. Strøm S. (1995). ‘Labor supply responses and welfare effects of tax reforms’, *The Scandinavian Journal of Economics*, vol. 97(4), pp. 635-659.

Altonji, J. (1986). ‘Intertemporal substitution in labor supply: evidence from micro data’, *Journal of Political Economy*, vol. 94(3 Part 2), pp. S176–215.

Arias, E. (2002). ‘United States Life Tables 2000’, *National Vital Statistics Reports*, vol. 51(3), National Center for Health Statistics, Hyattsville Maryland.

Blau, D. M. (2008). ‘Retirement and consumption in a life cycle model’, *Journal of Labor Economics*, vol. 26(1), pp. 35-71.

Blomquist, S. (1983). ‘The effect of income taxation on the labor supply of married men in Sweden’, *Journal of Public Economics*, vol. 22(2), pp. 169–97.

Blomquist, S. (1985). ‘Labour supply in a two-period Model: The Effect of a Nonlinear Progressive Income Tax’, *Review of Economic Studies*, vol. 52(3), pp. 515–24.

Blundell, R. and Walker, I. (1986). ‘A life cycle consistent empirical model of family labour supply using cross section data’, *Review of Economic Studies*, vol. 53(4), pp. 539-558.

Browning, M., Deaton, A. and Irish, M. (1985). ‘A profitable approach to labor supply and commodity demands over the life-cycle’, *Econometrica* vol. 53(2), pp. 503–44.

Brown, J., Liebman, J. and Pollet, J. (2002). ‘Estimating life tables that reflect socioeconomic differences in mortality’, in (M. Feldstein and J. Liebman, eds.), *Distributional Aspects of*

Social Security and Social Security Reform, pp. 447-58, Chicago: University of Chicago Press.

Chetty, R. (2012). ‘Bounds on elasticities with optimization frictions: a synthesis of micro and macro evidence on labor supply’, *Econometrica*, vol. 80(3), pp. 969-1018.

Dagsvik, J. and Strøm, S. (1992). ‘Labor supply with non-convex budget sets, hours restrictions and non-pecuniary job attributes,’ Discussion Paper 76, Central Bureau of Statistics, Oslo, Norway.

Dickens, W. and Lundberg, S. (1985). ‘Hours restrictions and labor supply’, Working Paper 1638, National Bureau of Economic Research.

Eckstein, Z. and Wolpin, K. (1989). ‘Dynamic labor force participation of married women and endogenous wage growth’, *Review of Economic Studies*, vol. 56(3), pp. 375-390.

Erosa, A., Fuster, L. and Kambourov, G. (2014). ‘Towards a micro-founded theory of aggregate labor supply’, Working Paper 516, Dept. of Economics, Univ. of Toronto.

Erosa, A., Fuster, L. and Kambourov, G. (2012) ‘Labor supply and government programs: a cross-country analysis’, *Journal of Monetary Economics*, vol. 59(1), pp. 84–107.

French, E. (2005). ‘The effects of health, wealth and wages on labour supply and retirement Behaviour’, *Review Economic Studies*, vol. 72(2), pp. 395–427.

French, E. and Jones, J. B. (2011). The effects of health insurance and self-insurance on retirement behavior’, *Econometrica*, vol. 79(3), pp. 693-732.

Guvenen, F., Kuruscu, B. and Ozkan, S. (2014). ‘Taxation of human capital and wage inequality: a cross-country analysis’, *Review of Economic Studies*, vol. 81(2), pp. 818– 850.

Hausman, J. (1981). ‘Labor Supply’, in (H. Aaron and J. Pechman, eds.), *How Taxes Affect Economic Behavior*, pp. 27-71, Washington, D.C.: Brookings Institution Press.

Hausman, J. (1985). ‘Taxes and Labor Supply’, in (A. Auerbach and M. Feldstein, eds.), *Handbook of Public Economics,Volume* 1, pp. 213-63, New York and Oxford: North Holland.

Heckman, J. and MaCurdy, T. (1982). ‘Corrigendum on a life cycle model of female labour supply’, *Review of Economic Studies*, vol. 49(4), pp. 659-660.

Imai, S. and Keane, M. (2004). ‘Intertemporal labor supply and human capital accumulation’, *International Economic Review*, vol. (45)2, pp. 601-642.

Keane, M. (2011). ‘Labor supply and taxes’, *Journal of Economic Literature*, vol. 49(4), pp. 961-1075.

Keane, M. (2015). ‘Life-cycle labour supply with human capital: econometric and behavioural implications’, *The Economic Journal*, this issue.

Keane, M. and Wolpin, K. (1997). ‘The career decisions of young men’, *Journal of Political Economy*, vol. 105(3), pp. 473-522.

Keane M. and Rogerson, R. (2012). ‘Micro and macro labor supply elasticities: a reassessment of conventional wisdom’, *Journal of Economic Literature*, vol. 50(2), pp. 464–76.

Kimmel J. and Kniesner, T. (1998). ‘New evidence on labor supply: employment vs. hours elasticities by sex and marital status’, *Journal of Monetary Economics*, vol. 42(2), pp. 289–301.

MaCurdy, T. (1981). ‘An empirical model of labor supply in life-cycle setting’, *Journal of Political Economy*, vol. 89(6), pp. 1059-1085.

MaCurdy, T. (1992). ‘Work disincentive effects of taxes: a reexamination of some evidence’, *American Economic Review*, vol. 82(2), pp. 243–49.

Moffitt, R. (1984). 'The estimation of a joint wage-hours labor supply model', *Journal of Labor Economics*, vol. 2(4), pp. 550-566.

Prescott, E., Rogerson, R. and Wallenius, J. (2009). 'Lifetime aggregate labor supply with endogenous workweek length', *Review of Economic Dynamics*, vol. 12(1), pp. 23–36.

Rogerson, R. (1988). 'Indivisible labor, lotteries and equilibrium', *Journal of Monetary Economics*, vol. 21(1), pp. 3–16.

Rogerson R. and Wallenius, J. (2009). 'Micro and macro elasticities in a life cycle model with Taxes', *Journal of Economic Theory*, vol. 144(6), pp. 2277–92.

Van der Klaauw, W. and Wolpin, K. (2008). 'Social security and the retirement and savings behavior of low income households', *Journal of Econometrics*, vol. 145(1-2), pp. 21-42.

Van Soest, A., Woittiez, I. and Kapteyn, A. (1990). 'Labor supply, income taxes, and hours restrictions in the Netherlands', *Journal of Human Resources*, vol. 25(3), pp. 517–58.

Wallenius J. (2011). 'Human capital and the intertemporal elasticity of substitution of labor: How large is the bias?', *Review of Economic Dynamics*, vol. 14(4), pp. 577–91.

Wallenius, J. (2013). 'Social security and cross-country differences in hours: a general equilibrium analysis', *Journal of Economic Dynamics and Control*, vol. 37(12), pp. 2466–2482.

Ziliak, J. and Kniesner, T. (1999). 'Estimating life-cycle labor supply tax effects', *Journal of Political Economy*, vol. 107(2), pp. 326–59.

Table 1

Parameter Estimates

Panel A. *Utility function*

$$u(c_t, h_t) = \frac{c_t^{a_1}}{a_1} - b_n \frac{h_t^{a_2}}{a_2}$$ where b_n is taste for work of type *n*, *n* = 1,2

	Dropout	HS	college
a_1	0.2529	0.2623	0.2519
a_2	1.5049	1.5201	1.4969
b_1	0.000220	0.000203	0.000215
b_2	0.000130	0.000109	0.000143
Fraction Type 1	0.547	0.508	0.397

Panel B. *Bequest function*

$$B(A_{t+1}) = 3\log(\mathrm{A}_{t+1} + \phi) - 1 - 3\log(\phi) \ \text{ if } \mathrm{A}_{t+1} > 0, \quad \left(\frac{\mathrm{A}_{t+1} + \phi}{\phi}\right)^3 \ \text{ otherwise}$$

	dropout	HS	college
ϕ	30000	30000	30000
δ	35.29	32.66	25.76

Panel C. *Human Capital and Wage Process*

$$k_{t+1} = g(k_t, h_t, age_{t+1})\varepsilon_{t+1}$$

$$\ln g(k_t, h_t, t) = \lambda_0 + \lambda_1 lnk_t + \lambda_2 \max(h_t - \bar{h}, 0) + \lambda_3 \max((h_t - \bar{h})^2, 0) + \lambda_4(t - t_0(e)) + \lambda_5(t - t_0(e))^2$$

where $t_0(e)$= 16,18 and 22 for dropout, HS and college, and hours are divided by 100.

	dropout	HS	College
λ_0 :	0.1677	0.1792	0.2387
λ_1 :	0.9192	0.9199	0.9065
λ_2 :	0.0039453	0.0047937	0.005947
λ_3 :	-0.0000907	-0.0000896	-0.0000933
λ_4 :	0.0001260	0.0001251	0.0001350

λ_5 :	-0.0000050	-0.0000050	-0.0000038
$\bar{h}$:	49.84	50.15	50.45
σ	0.118	0.091	0.130

Panel D. *Fixed Costs of Work*

$$E_t = w_t * \max\{h_t - fc_h, 0\} - fc_m \cdot I(h_t > 0)$$

	dropout	HS	college
fc_h	107.70	115.51	127.05
fc_m	723.10	803.66	819.94

Panel E. *Probability of Receiving Job Offer*

$$pjob_t^* = m^1 + m^{21}(t - t') * I(t \leq t') + m^{22}(t - 30)I(t < 30) + m^{23}(t - 30)I(t > 30)$$

$$+m^{24}(t - 50\,)I(t > 50) + m^{25}(t - 59\,)I(t > 59)$$

$$+m^{30}(1 - p_{t-1}) + m^{31}(1 - p_{t-1})I(t > 30)$$

$$+m^{32}(1 - p_{t-1})(t - 40)I(t > 40) + m^{33}(1 - p_{t-1})(t - 59\,)I(t > 59)$$

where *t’* = 23 for dropout and HS, and 26 for college

	Dropout	HS	college
m^1	2.480	2.762	2.964
m^{21}	0.0536	0.0650	0.0203
m^{22}	-0.0221	-0.0198	-0.0288
m^{23}	-0.0230	-0.0209	-0.0286
m^{24}	-0.0082	-0.0291	-0.0500
m^{25}	-0.0868	-0.0700	-0.0460
m^{30}	-2.877	-2.500	-2.652

m^{31}	0.000	0.000	0.758
m^{32}	-0.0275	-0.0450	-0.0070
m^{33}	0.0009	0.0000	-0.0150

Panel F. *Social Security Earnings Test Parameter*

$$SSinc_t = SSinc^* - \theta \cdot BW(w_t h_t, t, SSinc^*, t_B)$$

	Dropout	HS	College
θ	0.403	0.053	0.171

Panel G. *Unemployment Benefit*

	Dropout	HS	College
B	2513	1898	2412

Table 2

Labour Supply Elasticities at Various Ages

Panel A: ***Short-Run Tax Effects in the Year the Tax Change is Implemented***

	Dropout			**High School**			**College**		
Age	Marshall	Hicks	Frisch	Marshall	Hicks	Frisch	Marshall	Hicks	Frisch
$\boldsymbol{t_0(e)}$	2.30	**4.80**	**2.17**	0.09	0.15	0.16	0.05	**0.24**	**0.08**
20	0.31	**0.77**	**0.50**	0.28	**1.07**	**0.93**			
25	0.25	0.96	1.10	0.23	0.88	0.90	0.37	**1.56**	**1.09**
30	0.25	0.96	1.21	0.11	0.87	0.88	0.08	**0.67**	**0.37**
35	0.25	0.83	1.06	0.09	**0.52**	**0.46**	0.09	**0.38**	**0.29**
40	0.35	0.89	1.02	0.13	**0.54**	**0.50**	0.13	**0.41**	**0.37**
45	0.47	0.92	1.07	0.18	**0.51**	**0.49**	0.12	**0.42**	**0.37**

50	0.65	1.02	1.04	0.31	0.68	0.71	0.23	**0.53**	**0.50**
55	1.13	1.39	1.28	0.39	0.63	0.63	0.34	**0.64**	**0.59**
60	1.32	1.53	1.25	0.97	1.21	0.94	0.80	**1.14**	**0.96**
Average	0.63	1.31		0.24	0.69		0.22	0.65	

Panel B: ***Effects of Tax Changes on Labour Supply over the Remainder of the Working Life***

	Dropout		**High School**		**College**	
Age	Marshall	Hicks	Marshall	Hicks	Marshall	Hicks
$\boldsymbol{t_0(e)}$	**0.33**	**1.01**	**0.19**	**0.65**	**0.22**	**0.74**
20	0.26	0.93	0.18	0.66		
25	0.27	0.95	0.17	0.65	0.20	0.75
30	0.31	1.00	0.19	0.66	0.20	0.72
35	0.39	1.04	0.23	0.69	0.23	0.77
40	0.52	1.15	0.29	0.76	0.29	0.85
45	0.68	1.29	0.39	0.85	0.37	0.95
50	0.96	1.46	0.58	1.03	0.53	1.10
55	1.24	1.62	0.88	1.27	0.82	1.35
60	1.52	1.80	1.39	1.65	1.39	1.86

Note: All elasticities are for a tax change that takes effect unexpectedly at the indicated age. In the case of the Marshall and Hicks, the tax change is permanent (and expected to be permanent). $t_0(e)$ indicates the first year of the working life, which is 16, 18 or 22 for dropouts, high school workers or college workers, respectively. In panel A the figures in bold indicate cases where the Hicks elasticity exceeds the Frisch in the short-run (i.e., the short-run effect of a compensated permanent tax change exceeds that of a transitory tax change). In Panel B, the Figures in bold indicate lifetime effects of changes in the tax regime. **Figure 1**

Average Hours by Age and Education

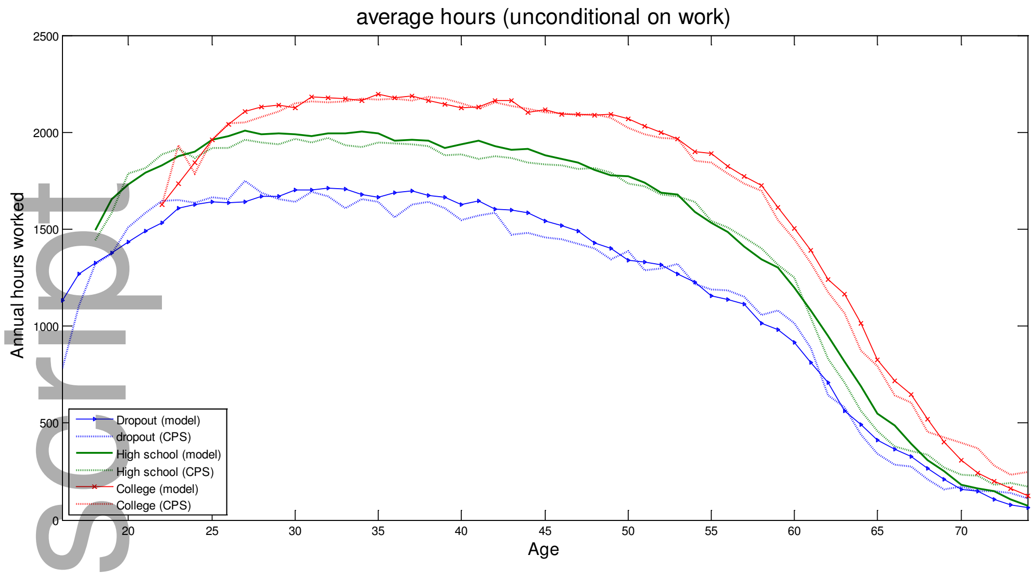


Figure 2

Employment by Age and Education

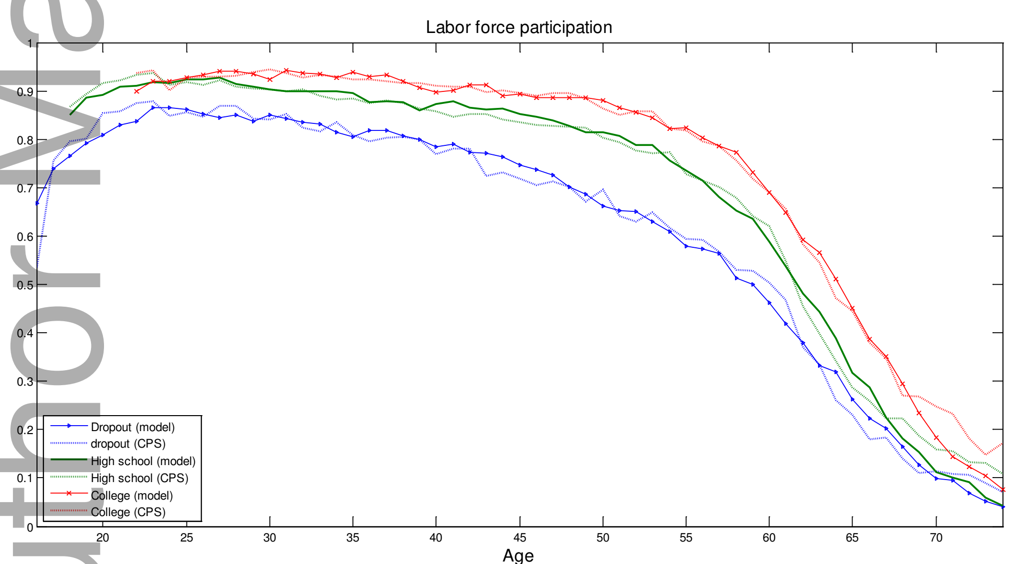


Figure 3

***Hours Conditional on Employment* (*by Age and Education*)**

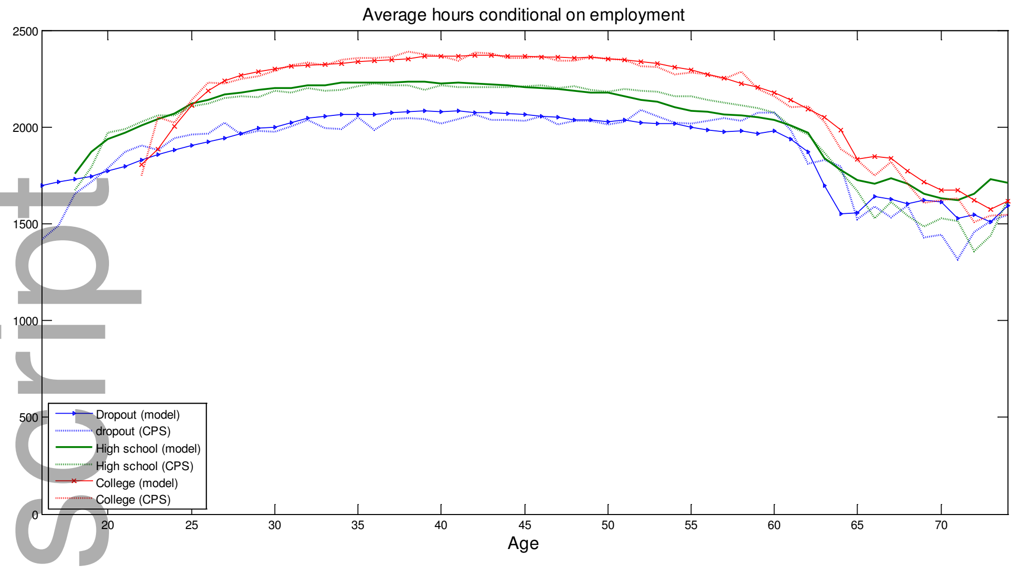


Figure 4

Full-time Hourly Wage Rate by Age and Education (Median)

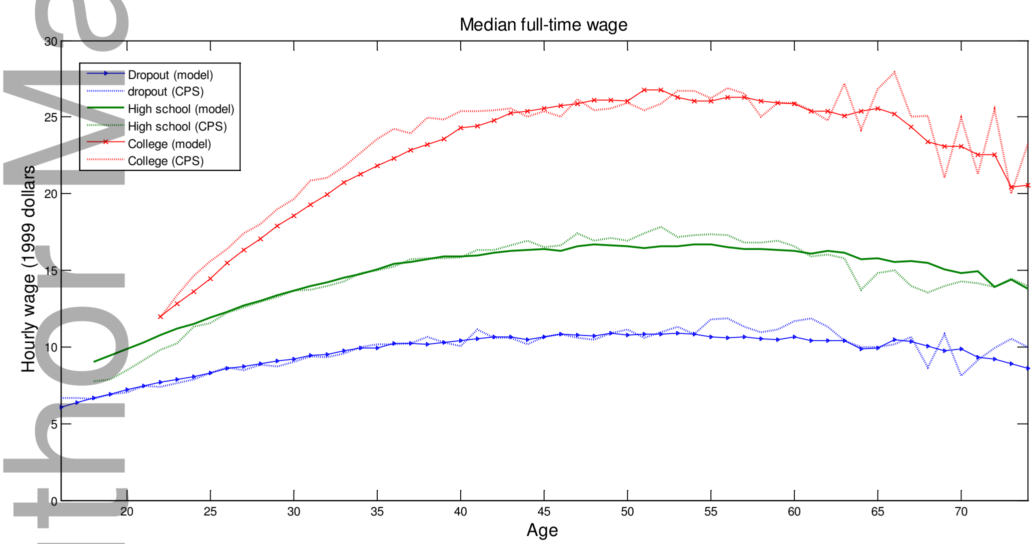


Figure 5

Mean Consumption by Age and Education

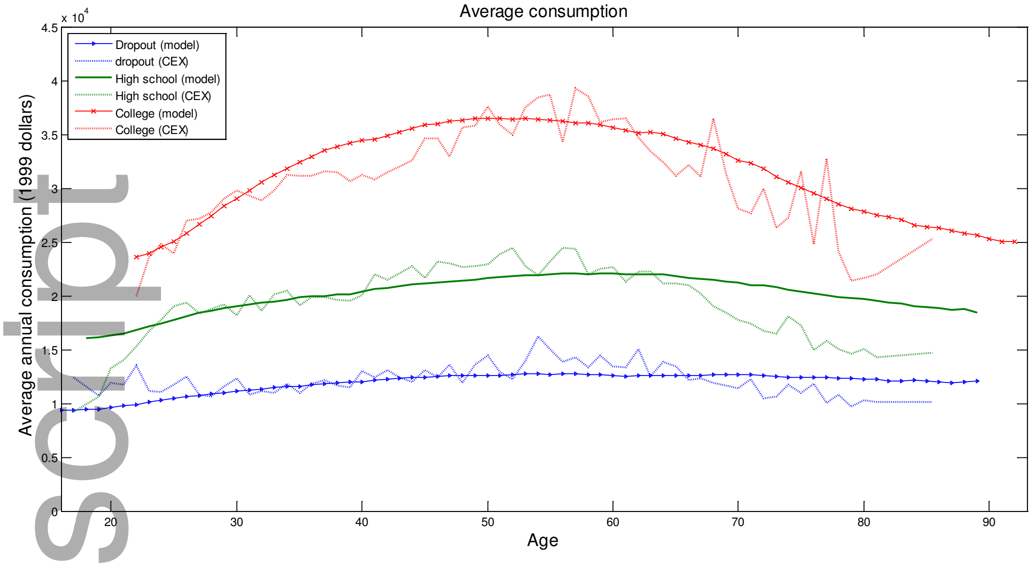


Figure 6

Social Security Claiming Age (by Education)

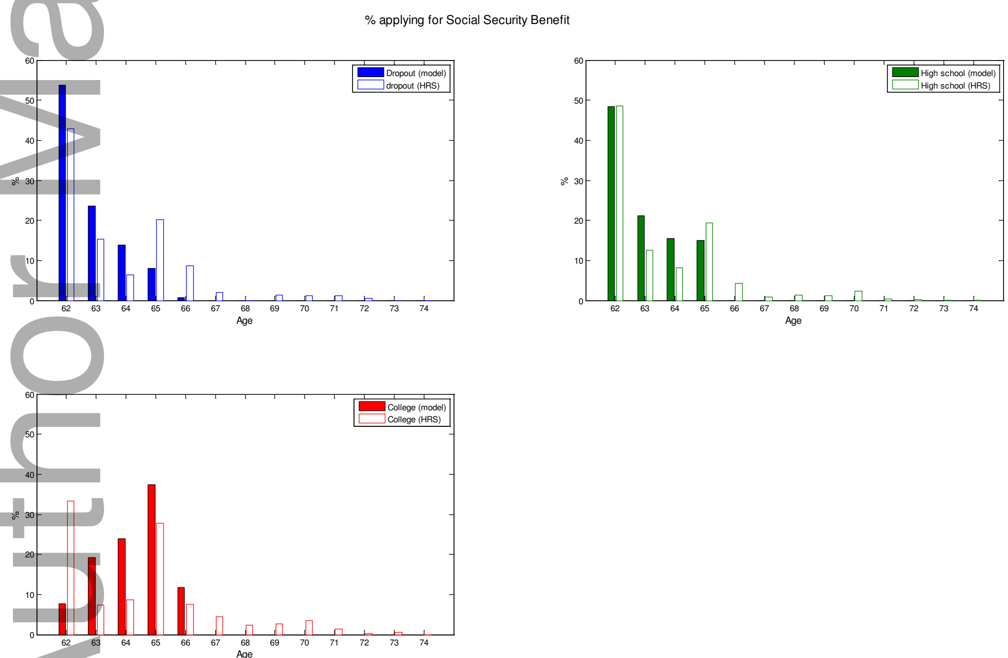


Figure 7

Standard Deviation of Hours (by Age and Education)

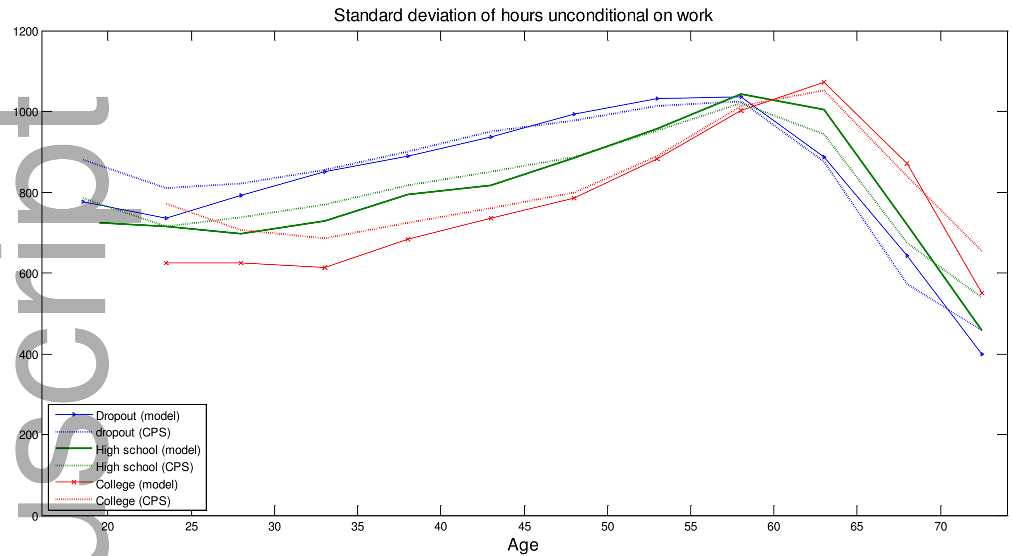


Figure 8

Standard Deviation of Log Hourly Wage (by Age and Education)

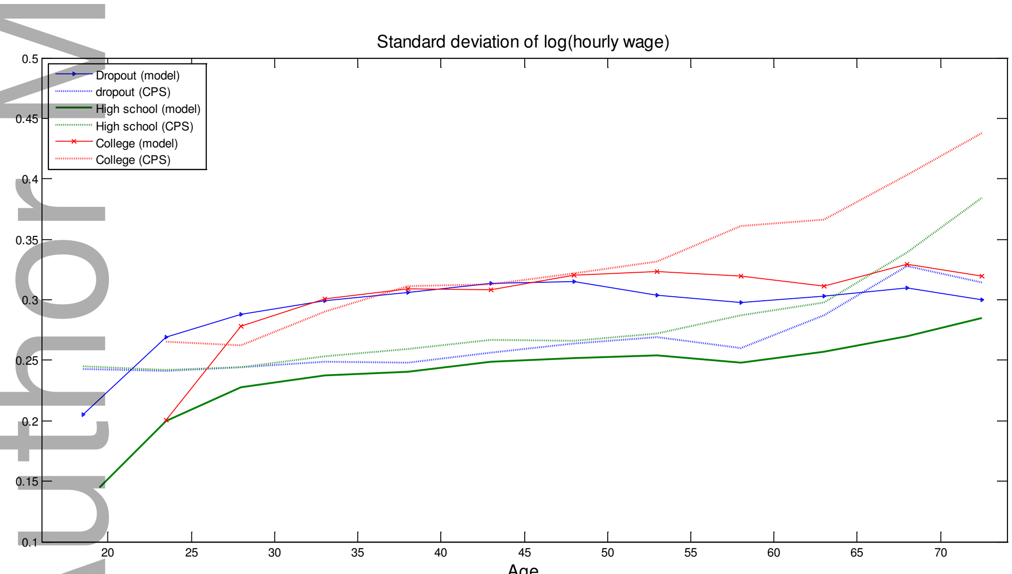


Figure 9

***Standard Deviation of Consumption* (*by Age and Education*)**

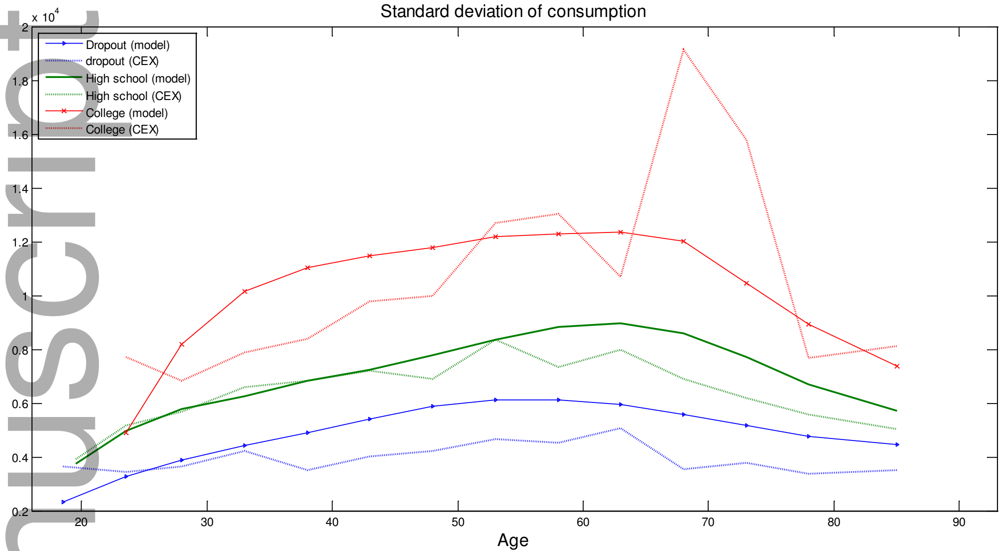


Figure 10

***Transition Rate from Work to Work* (*by Age and Education*)**

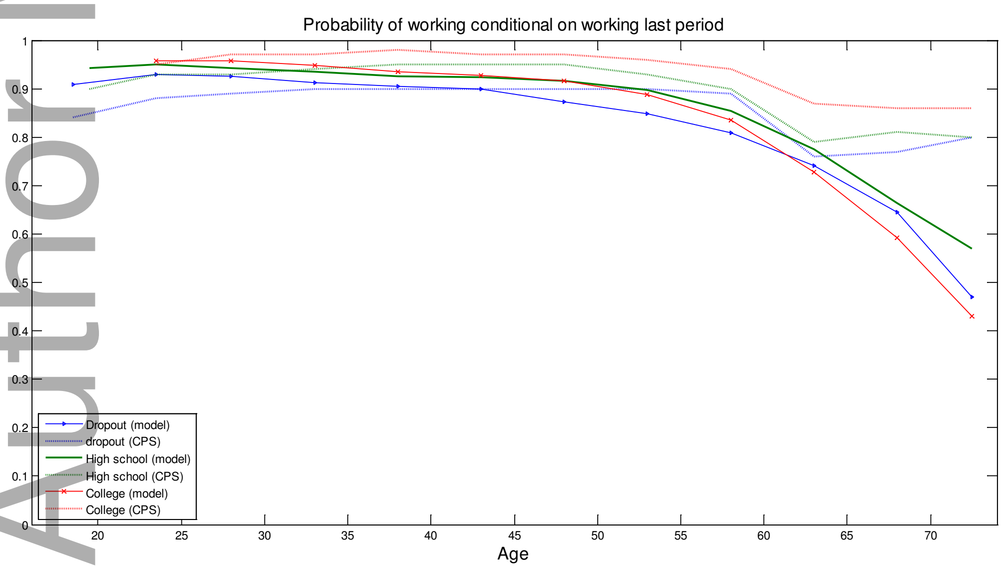


Figure 11

***Transition Rate from Not-Working to Work* (*by Age and Education*)**

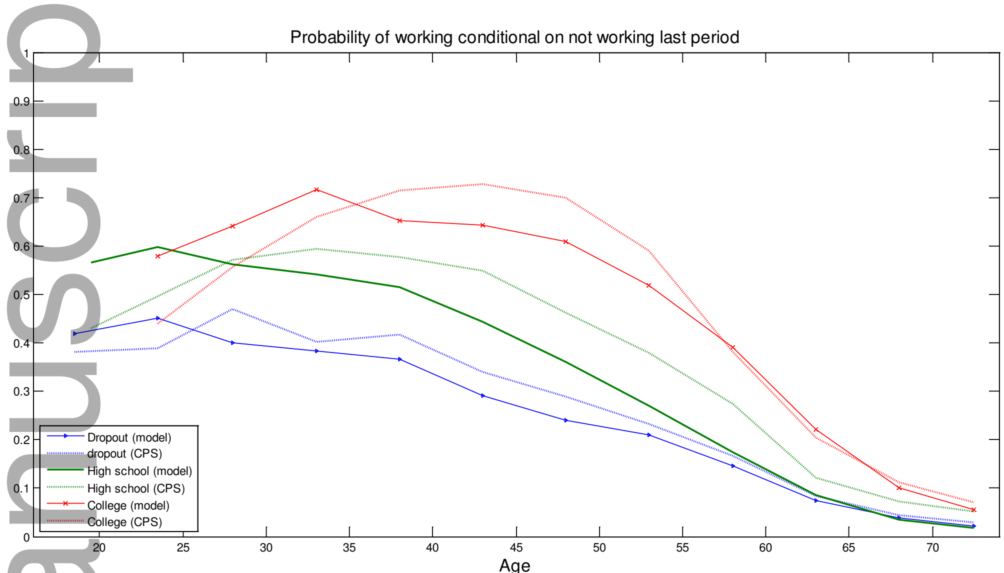


Figure 12

Frisch Elasticities by Age and Education

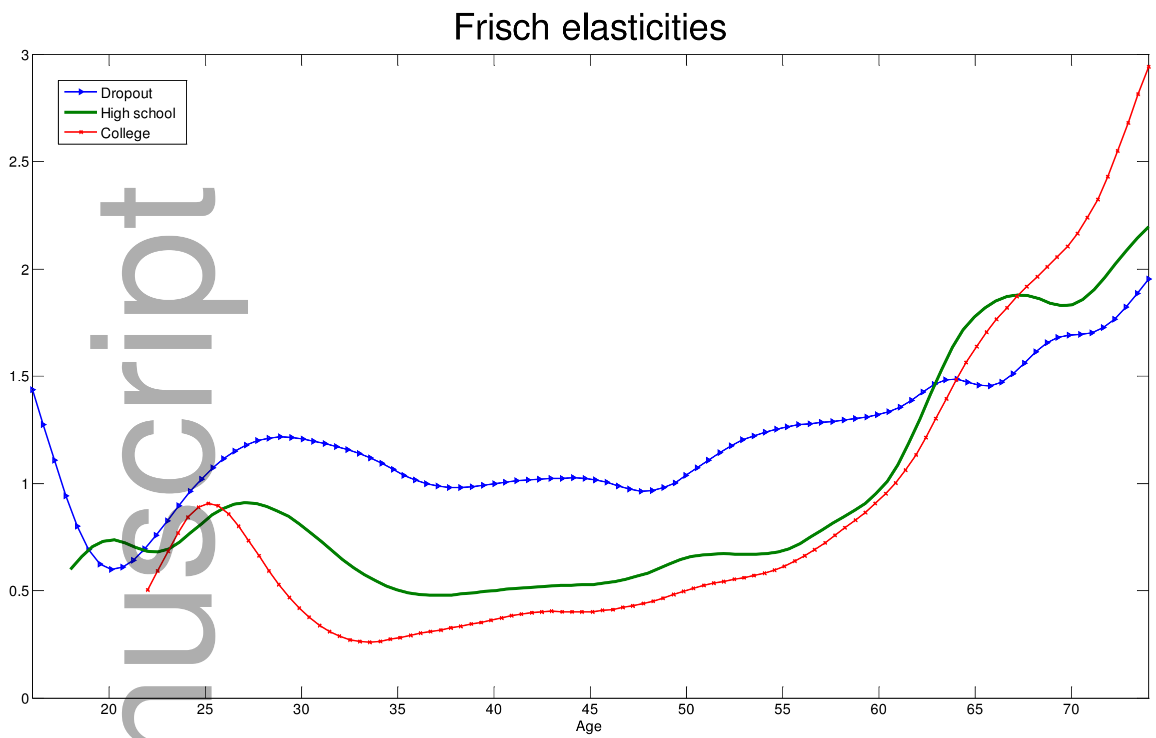


Figure 13

Hicks Elasticities by Age and Education

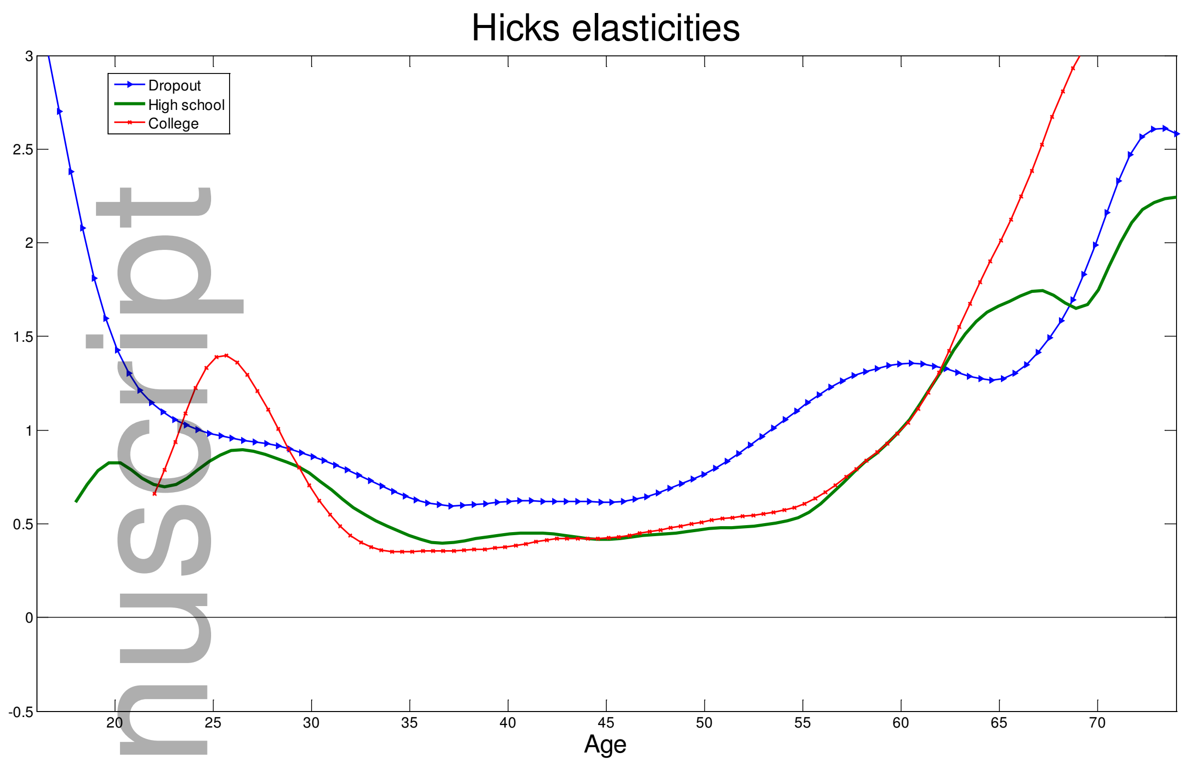

Hicks elasticities
Dropout
High school
College
Age
3
2.5
2
1.5
1
0.5
0
-0.5
20
25
30
35
40
45
50
55
60
65
70

Figure 14

Marshallian Elasticities by Age and Education

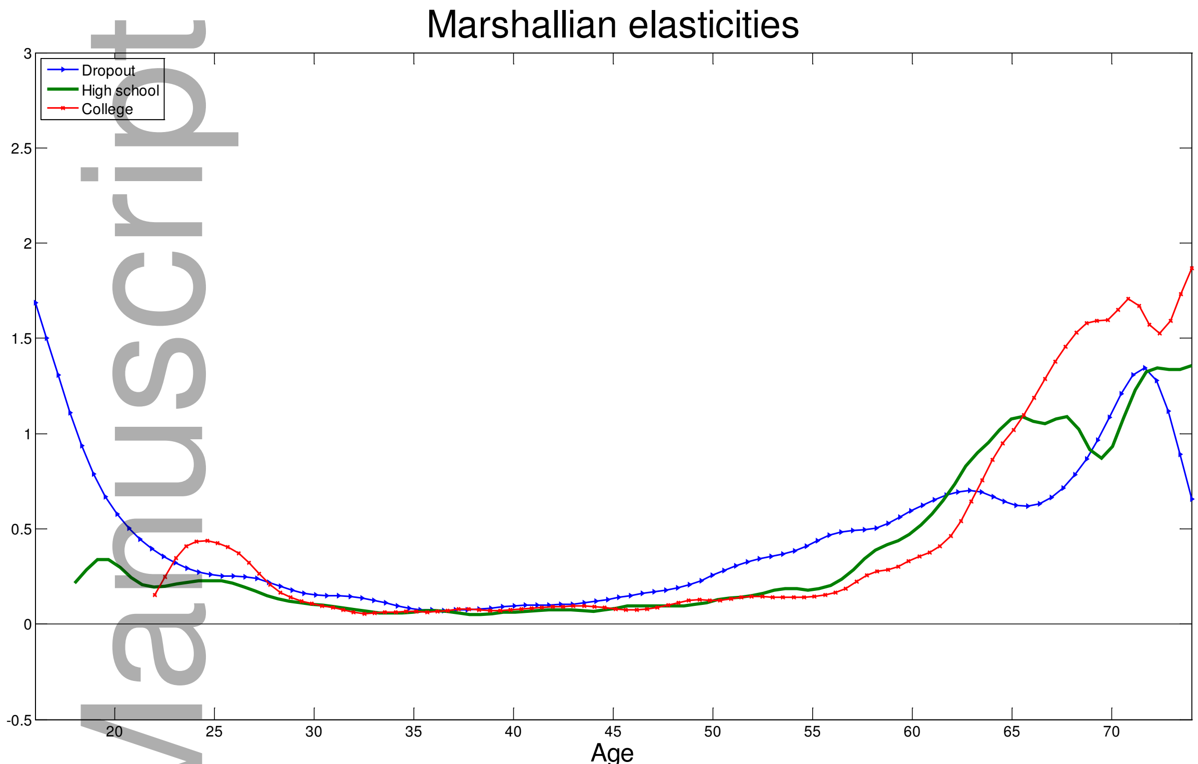


Figure 15

Intensive vs. Extensive Margin Elasticities by Age and Education

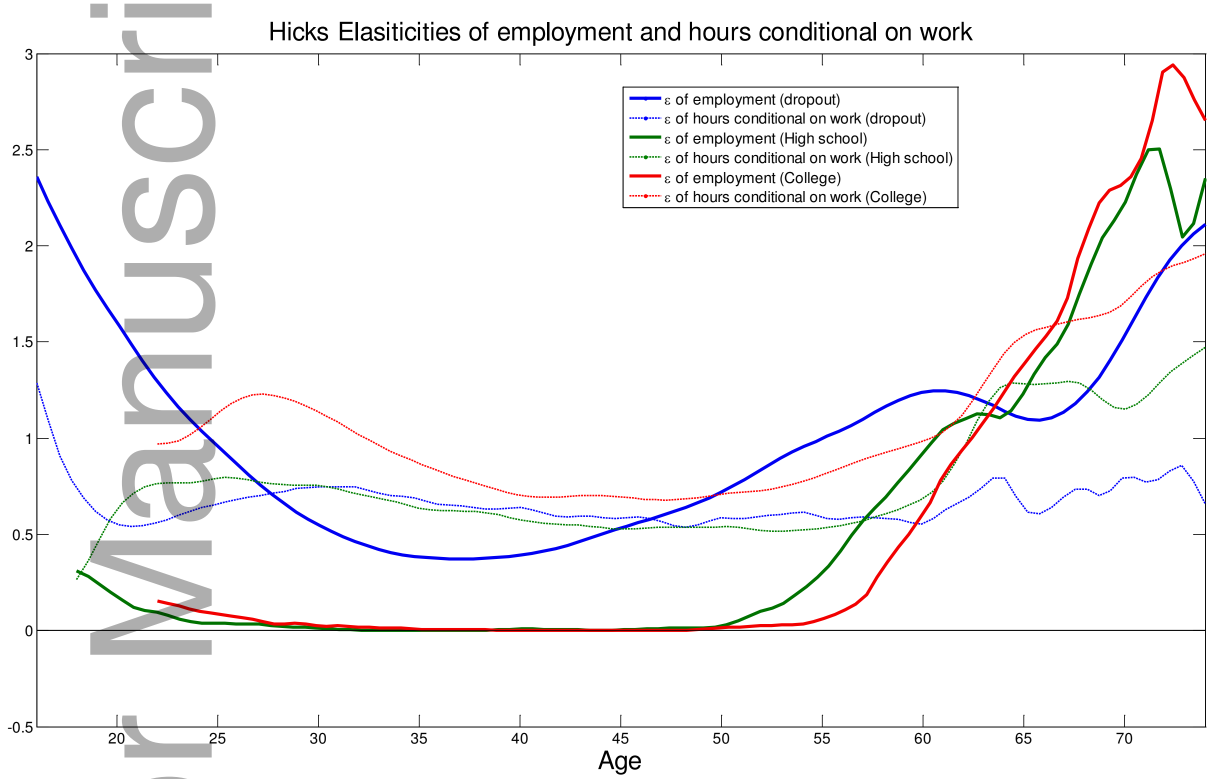


Figure 16

Effect of Reducing or Eliminating Human Capital Effects on Labour Supply

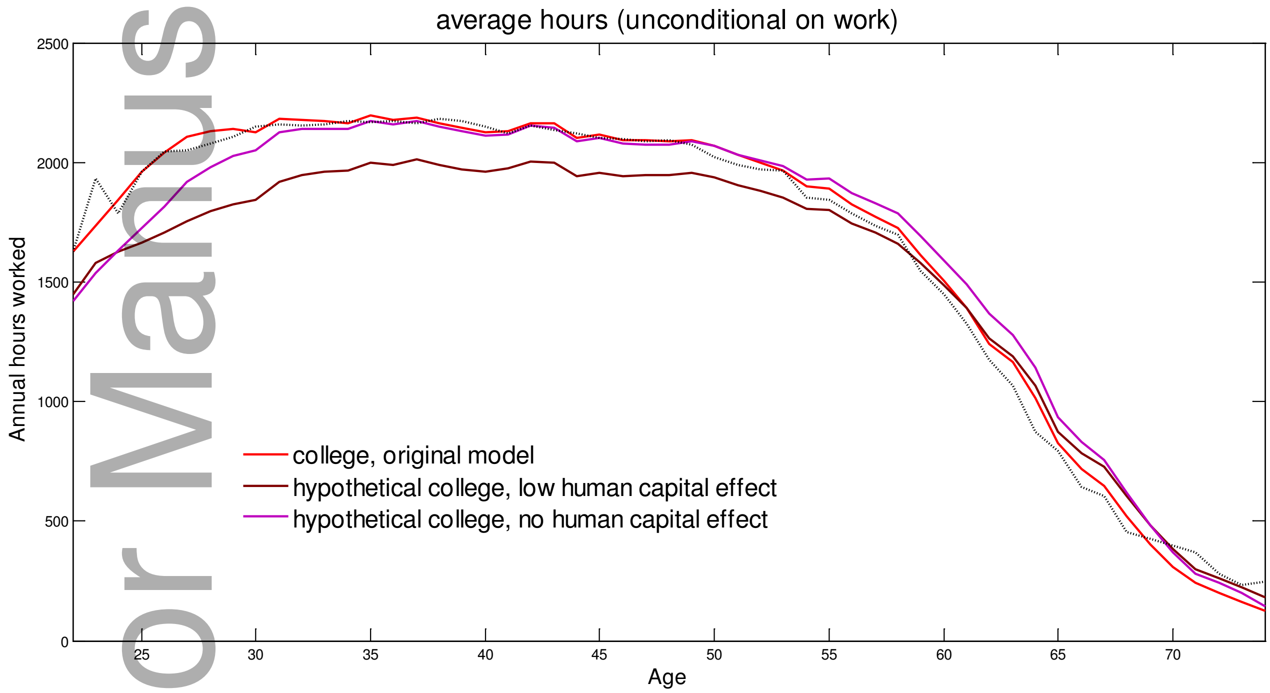


Figure 17

Effect of Human Capital – Hicks Elasticity – Extensive Margin

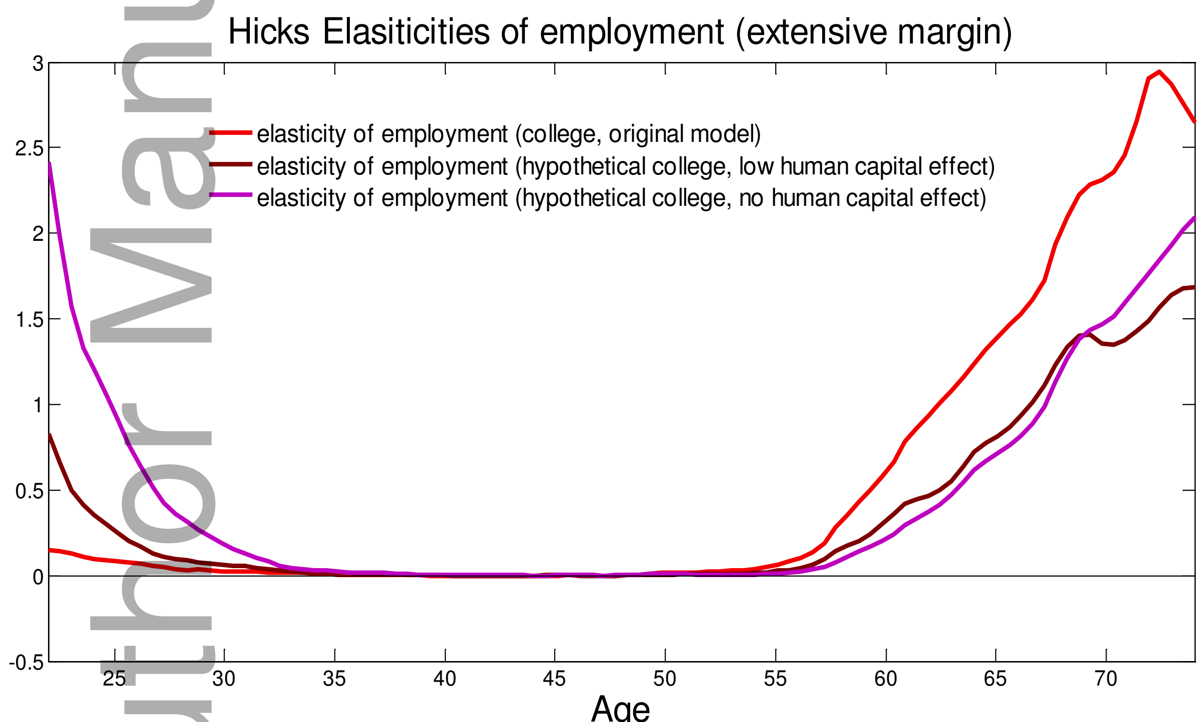


Figure 18

Effect of Human Capital – Hicks Elasticity – Intensive Margin

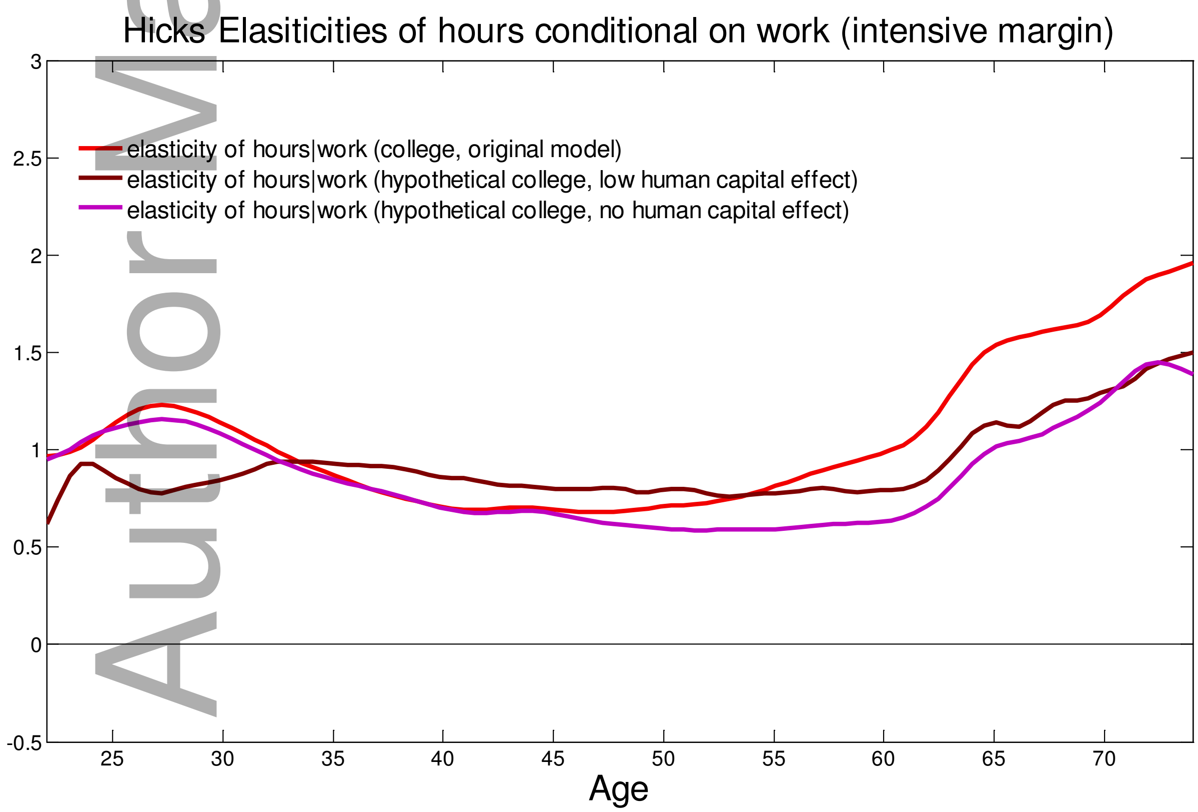


Figure 19

Effect of Human Capital – Frisch Elasticity – Extensive Margin

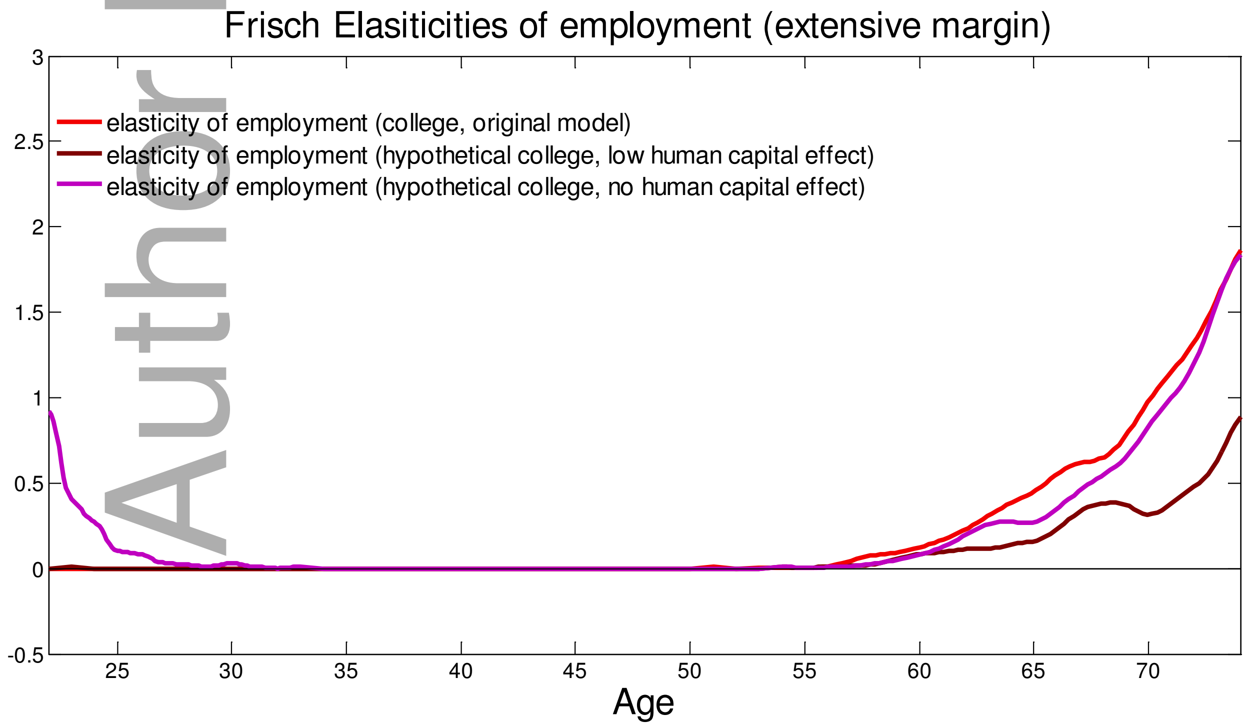
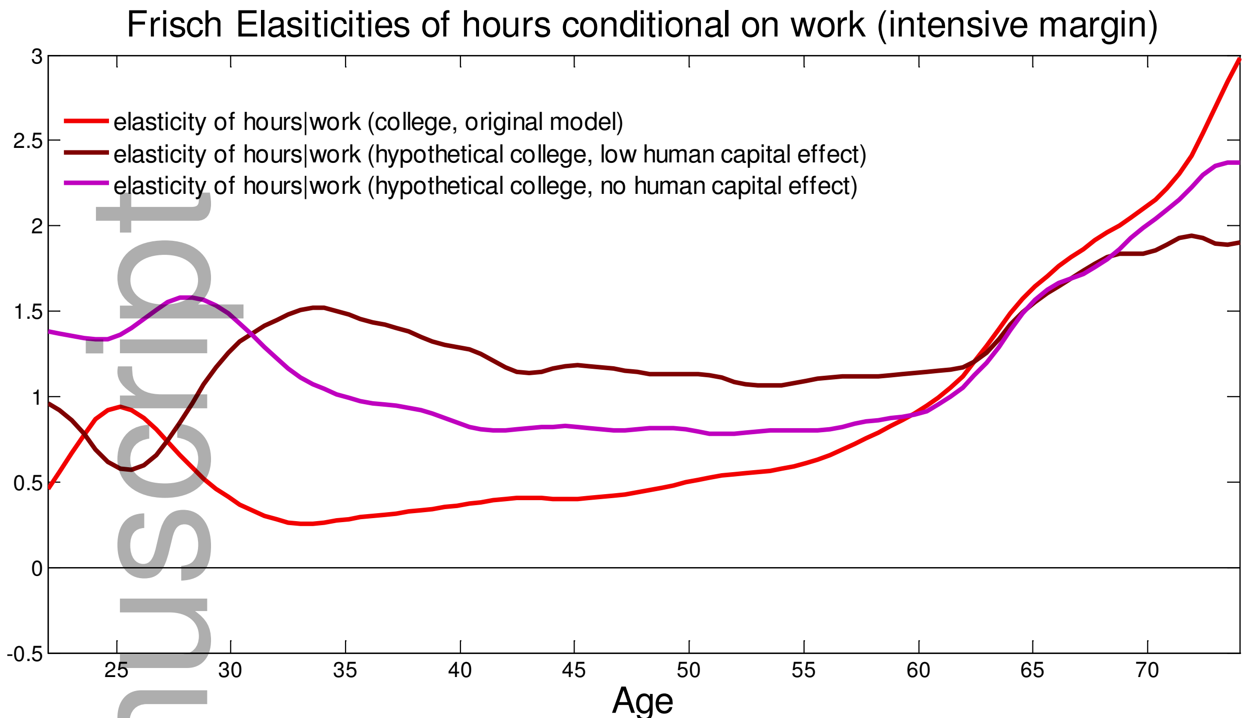


Figure 20

Effect of Human Capital – Frisch Elasticity – Intensive Margin


Frisch Elasiticities of hours conditional on work (intensive margin)
elasticity of hours|work (college, original model)
elasticity of hours|work (hypothetical college, low human capital effect)
elasticity of hours|work (hypothetical college, no human capital effect)
Age
3
2.5
2
1.5
1
0.5
0
-0.5
25
30
35
40
45
50
55
60
65
70