Reconciling Micro and Macro Labor Supply Elasticities: A Structural Perspective

by Michael Keane and Richard Rogerson

James J. Heckman

THE UNIVERSITY OF CHICAGO

Econ 350, Winter 2023

1. INTRODUCTION

- This review deals with an issue that is extremely important for a wide range of economic issues—the magnitude of the elasticity of aggregate labor supply with respect to transitory and permanent changes in wages.

- This issue is highly controversial: There is a long-standing controversy driven by the fact that labor economists typically estimate relatively small labor supply elasticities from micro data, whereas macroeconomists who use representative agent models to study aggregate outcomes typically employ parameterizations that imply large aggregate labor supply elasticities.

- A key point we wish to stress is that, in general, labor supply elasticities are neither a single number nor a primitive feature of preferences.

- Rather, labor supply responses (individual or aggregate) to a particular change in the economic environment will typically depend on features of technology and market structure, as well as preferences.

- And they will typically be heterogeneous, differing by worker characteristics such as age, gender, and skill level. In a dynamic setting, labor supply responses will generally change over time, as long- and short-run effects will differ.

- A key implication is that it is important to adopt a framework in which the choice problems of individuals are explicitly formulated.

2. BACKGROUND AND OVERVIEW

2.1. A Benchmark Model

- We begin with a benchmark life-cycle model that serves to clarify the macro-micro labor supply controversy.

- Each period, a T-period lived individual is born with preferences

$$\sum_{a=0}^{T} \beta^{a} [\frac{1}{1-\frac{1}{\eta}} c_{a}^{1-\frac{1}{\eta}} - \frac{\alpha}{1+\frac{1}{\gamma}} h_{a}^{1+\frac{1}{\gamma}}]$$

  where c_a and h_b are consumption and hours worked at age α, respectively.

- There are four preference parameters: β, η, α, and γ.

- We strive to use these parameters consistently throughout the article.

- The individual has one unit of time each period and faces an exogenous productivity sequence, denoted by e_a, so that working h_a units of time at age α yields e_{ah_a} units of labor services.

- There is a constant returns to scale aggregate production function $F(K_t, H_t)$where K_t and H_t are aggregate capital and units of labor services, respectively.

- In steady state, these satisfy

$$H = \sum_{a=0}^{T} e_a h_a, K = \sum_{a=0}^{T} k_a$$

where h_a and k_a are the steady-state life-cycle profiles for hours worked and capital holdings, respectively.

- Output can be used as consumption or investment, and capital depreciates at rate δ.

- We consider the following tax and transfer system: Labor earnings are taxed at the constant rate τ, and the resulting revenues fund a lump-sum transfer.

- To avoid issues of intergenerational redistribution, one assumes that the lump-sum transfer received by any generation is equal in present value to their tax payments.

- With infinitely lived agents, the steady-state interest rate is unaffected by this policy and equals $\frac{1}{\beta} - 1$.
- This need not hold in an overlapping generations economy.
- But because our interest is in the effects of taxes controlling for changes in other factors, such as interest rates, we assume that the steady-state interest rate is not affected by τ and equals $\frac{1}{\beta} - 1$.
- We note that there is always a government debt policy that would support this interest rate as a steady-state equilibrium.
- Constant returns to scale of F then imply that the wage per unit of labor services, denoted by w, is also independent of τ.

- An individual thus solves the following problem in steady-state equilibrium:

$$\max_{c_a, h_a} \sum_{a=1}^{T} \beta^a \left[\frac{1}{1-\frac{1}{\eta}} c_a^{1-\frac{1}{\eta}} - \frac{\alpha}{1+\frac{1}{\gamma}} h_a^{1+\frac{1}{\gamma}} \right],$$

$$\text{s.t.} \sum_{a=1}^{T} \beta^a c_a = \sum_{a=1}^{T} (1-\tau)\beta^a e_a h_a w + T.$$

- Letting λ denote the Lagrange multiplier on the budget equation, we have the following first-order conditions:

$$c_a^{-\frac{1}{\eta}} = \lambda, \qquad (1)$$

$$\alpha h_a^{\frac{1}{\gamma}} = (1-\tau)\lambda e_a w. \qquad (2)$$

- Equation 1 implies that c_a is constant over the life cycle.

- Taking logs of Equation 2 gives a simple version of the equation used by MaCurdy (1981) and others in their estimation exercises using micro data:

$$\log h_a = b + \gamma \log e_a, \qquad (3)$$

- where $b = \gamma[\log \lambda + \log w + \log(1 - \tau) - \log \alpha]$is constant for an individual over his or her life cycle in steady state.

- Because changes in $\log e_a$ are equivalent to changes in log wages for individuals over the life cycle, this equation provides a strategy for uncovering the preference parameter γ using individual panel data.

- As described below, one can also uncover the value of η.

- Estimates of these preference parameters from micro data also allow one to infer aggregate effects of changes in τ.

- Equation 3 is a useful starting point but is not sufficient.

- The reason is that if we are comparing h_a across steady-state equilibria that correspond to different values of τ, then the value of λ will also differ.

- Hence, to determine the change in h_a, we need to also derive an expression for the change in λ .

- To do this, note from Equation 2 that given the optimal value of h_0, the rest of the profile satisfies

$$h_a = \left[\frac{e_a}{e_0}\right]^{\gamma} h_0. \tag{4}$$

- Total labor income is therefore proportional to h_0.
- Because the present value of the transfer received by each individual is equal to the present value of his or her own tax payments, in steady-state equilibrium, we have

$$\sum_{a=1}^{T} \beta^a c_a = \sum_{a=1}^{T} \beta^a e_a h_a w.$$

- As ca is constant over the life cycle, its value is proportional to h_0 and w.
- Write this as $c_a = \bar{c} w h_0$. Equation 1 then implies

$$\log \lambda = -\frac{1}{\eta} [\log h_0 + \log w + \log \bar{c}]. \quad (5)$$

- Using Equation 4, we have

$$\log \lambda = -\frac{1}{\eta}\left[\gamma \log \frac{e_o}{e_a} + \log h_a + \log w + \log \bar{c}\right]. \quad (6)$$

- Given that w is independent of τ, Equation 3 implies

$$\log h_a = -\frac{\gamma}{\eta}\left[\gamma \log \frac{e_o}{e_a} + \log h_a + \log \bar{c} + \log w\right] + \gamma \log w + \gamma \log(1-\tau) - \gamma \log \alpha + \gamma \log e_a. \quad (7)$$

- Rearranging gives

$$\log h_a = \frac{\gamma}{\eta+\gamma}\left[(\eta-1)\log w - \eta \log \alpha - \log \bar{c} - \gamma \log e_0\right] + \frac{\gamma\eta}{\eta+\gamma}\log(1-\tau) + \gamma \log e_a. \quad (8)$$

- Here, the coefficient on $\log e_a$ is the Frisch elasticity.
- This is the effect of life-cycle variation in wages.
- The coefficient on $\log(1-\tau)$ is the Hicks elasticity.
- A key distinction between the two is that the Frisch elasticity holds the marginal utility of consumption constant, whereas the Hicks does not.
- The Hicks elasticity is smaller than the Frisch, with equality as $\eta \to \infty$ (i.e., when utility is linear in consumption and there are no income effects).

- Equation 8 implies that a change in τ causes hours to change proportionally at all ages.

- Because H is simply the sum of h_a, it follows that

$$\log H = B + \frac{\gamma\eta}{\eta + \gamma} \log(1 - \tau), \tag{9}$$

where B is a constant.

- Macroeconomists often impose $\eta = 1$ so preferences are consistent with balanced growth.

- Then the coefficient on $\log(1 - \tau)$ is purely a function of γ.

2.2. Micro Evidence Based on the Benchmark Model

- We consider three of the most influential papers, MaCurdy (1981), Browning et al. (1985), and Altonji (1986), each of which estimates the intertemporal elasticity of substitution, or Frisch elasticity.

- Details of their approaches differ, but all involve regressing changes in hours on changes in wages.

- For example, MaCurdy (1981) uses the basic model described above extended to allow for heterogeneity and uncertainty to derive the change in hours equation:

$$\Delta \log h_{it} = \gamma \Delta \log w_{it}(1 - \tau_{it}) - \gamma \log \beta(1 + r_t) + \alpha\gamma\Delta X_{it} + \gamma\xi_{it} + \gamma\Delta\varepsilon_{it}. \tag{10}$$

- The parameters α, β, and γ are as above, and the tax rate (τ_{it}) is allowed to vary across time and individuals.

- The X_{it} are controls for exogenous shifts in tastes for work, the ε_{it} represent unobserved taste shocks, and ξ_{it} represents the surprise part of the change in the marginal utility of wealth (or of consumption) from t - 1 to t[5]

- The literature has focused on three issues:
  - First, the ξ_{it} will be correlated with wage changes to the extent that wage changes are not fully anticipated at t - 1.
  - Second, tastes for work may be correlated with wages (e.g., those with a higher taste for work may also work harder or acquire more skills, but also lower the after-tax wage by pushing one into a higher tax bracket).
  - Third, the wage is presumably measured with considerable error.

- MaCurdy (1981) also shows that the results of estimating Equation 10 allow one to infer responses to permanent wage changes.

- Estimation of Equation 10 uncovers all parameters of the hours equation in levels,

$$\log h_{it} = \gamma \log(w_{it}(1-\tau_{it})) + \gamma \log \lambda_{i0} - \gamma \log \rho (1+r_t)^t + \alpha \gamma X_{it} + \gamma \varepsilon_{it}, \qquad (11)$$

  except for $\gamma \log(\lambda_{i0})$, which is the individual specific constant (or fixed effect) in the levels equation (λ_{i0} is the marginal utility of wealth at t = 0).

- Thus, he backs out the value of $\gamma \log(\lambda_{i0})$ in a second stage after estimating Equation 10 in the first stage.

- He can then in principle regress them on the whole set of life-cycle wages.

- His estimates imply that a 10% (fully anticipated) increase in wages at all ages increases labor supply by only 0.8%—a very small effect.

2.3. Macroeconomic Models

- Although the view that labor supply elasticities are small is clearly the majority position among microeconomists, this view is less well accepted among macroeconomists.
- Beginning with Lucas&Rapping (1969), many macroeconomists have argued that relatively large Frisch elasticities are required to account for labor market fluctuations over the business cycle.
- Prescott (2004) shows that a relatively large labor supply elasticity is also required to rationalize trend changes in hours of work among G-7 economies since 1970.
- In fact, in the infinitely lived stand-in household models that remain the normin much of the macro literature, it is standard to assume that the period utility function is log linear in consumption and leisure.
- If one-third of available time is spent in market work, this implies a Frisch elasticity of 2.0.

2.4. Overview of the Review

- One could challenge the claim that the micro literature offers a clear consensus on labor supply elasticities.
- And as with any empirical work, one could criticize the studies that find small labor supply elasticities on their own terms.
- That is, one could accept the basic empirical framework (e.g., Equation 10) but question the implementation.
- Specifically, one could question the instruments for wages, the controls for tastes for work, the functional forms for the labor supply function, measurement of wages, taxes, and so on.

- We instead focus on three other issues.
- The first, pursued in Section 3, questions fundamental assumptions of the empirical framework of Equation 10.
- The second approach, pursued in Section 4, questions whether standard micro data estimates are relevant for determining aggregate labor supply responses.
- The third (and related) approach, pursued in Section 5, highlights that most micro empirical work finding small elasticities focuses on adjustment along the intensive margin

3. MICRO EVIDENCE BASED ON EXTENSIONS OF THE BASIC MODEL

3.1. Human Capital Accumulation

- The classic MaCurdy (1981) life-cycle model assumes that wages evolve exogenously, precluding the possibility that workers acquire human capital via learning-by-doing or on-the-job investment.
- Heckman (1976) studies a model with on-the-job investment in which workers are paid only for the time they spend on productive work (not the time they spend learning).
- Shaw (1989) includes learning-by-doing in a life-cycle model.
- These models share two key properties: (a) The observed wage is less than the true price of time for young workers, and (b) the observed wage grows more quickly than the price of time over the life cycle.

- Notably, neither Heckman (1976) nor Shaw (1989) directly assesses the implication of these properties for estimates of preference parameters and labor supply responses.

- This issue is addressed by Imai&Keane (2004), who argue that abstracting from human capital accumulation would downwardly bias estimates of g.

- To illustrate the logic, assume wages evolve according to

$$w_{t+1} = \left(1 + \kappa \sum_{j=1}^{t-1} h_{t-j}\right) w_1, \qquad (12)$$

where k > 0, and w1 is the individual's wage when first entering the labor market.

- A unit increase in ht raises wt by kw1 in all future periods.

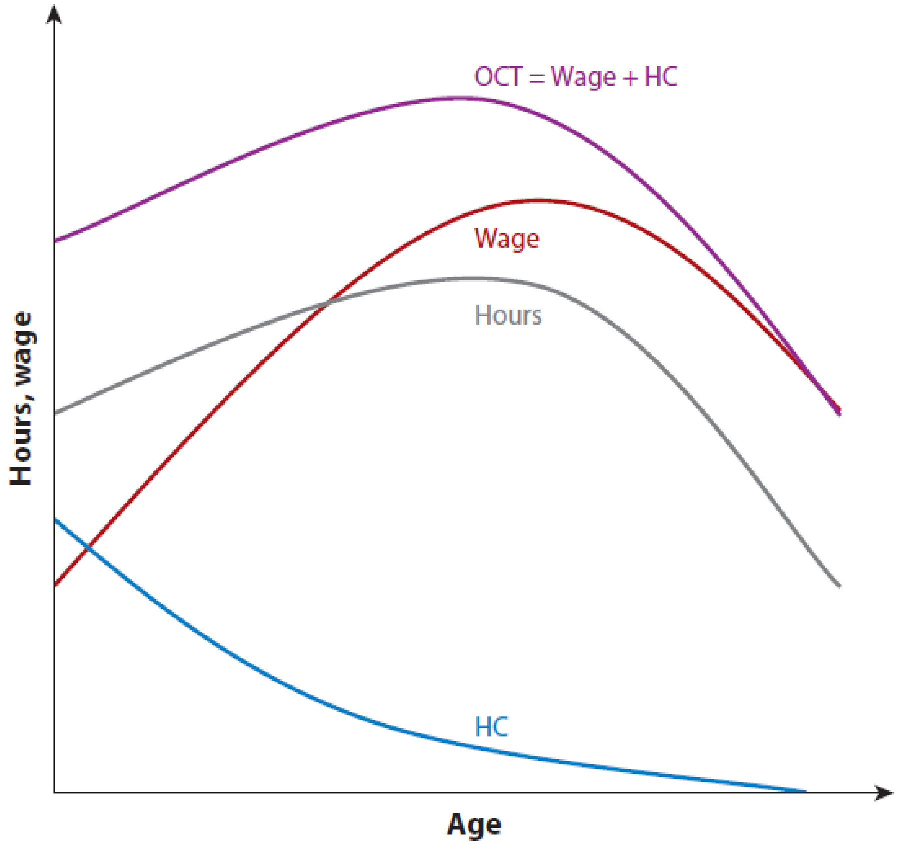
- In this model, the return to an hour of work, which Imai&Keane call the opportunity cost of time (OCT), consists of the current after-tax wage plus the expected present value of increased (after-tax) earnings in all future periods.

- Imai&Keane refer to this second component as the human capital term.
- The optimality condition for an interior solution equates the marginal rate of substitution between consumption and leisure to the OCT.
- Assuming the utility function from the benchmark model of Section 2, this gives

$$\frac{\alpha h_t^{\frac{1}{\gamma}}}{c_t^{-\frac{1}{\eta}}} = w_t(1-\tau_t) + E_t \sum_{j=0}^{T-t} \frac{\kappa w_1 h_{t+1+j}(1-\tau_{t+1+j})}{(1+r)^{1+j}}. \tag{13}$$

- A model without human capital equates the marginal rate of substitution to the after-tax wage itself.
- The human capital term creates a wedge between the OCT and the after-tax wage.
- Importantly, this wedge declines with age owing to the shrinking time horizon for recouping returns to human capital investment.

Figure1



Notes: Hours, wages, and price of time over the life cycle. Human capital (HC) denotes the return to an hour of work experience, in terms of increased present value of future wages. The opportunity cost of time (OCT) is Wageþ HC. Figure reproduced from Keane & Rogerson (2012).

- Figure 1 displays (stylized) life-cycle profiles for male wages and earnings, in addition to the OCT and human capital curves described above.

- The wage rate exhibits the familiar hump shape found in many studies (i.e., wages grow rapidly early in the life cycle, peak in the forties, and then
- decline).

- Annual work hours also have a hump shape but with much less curvature (see, e.g., the descriptive regressions in Pencavel 1986).

- Graphically, the OCT curve is the vertical sum of the wage and human capital curves.

- Because the human capital curve declines with age (owing to shrinking in the remaining horizon and possibly decreasing returns to accumulating human capital), the OCT curve is much flatter than the wage curve.

- The intuition for why ignoring human capital downwardly biases estimates of γ is now straightforward.

- Viewed through the lens of a MaCurdy (1981)–type model with exogenous wages, the relatively slow growth in hours relative to wages over the first half of the life cycle can only be rationalized if workers are very unwilling to substitute labor intertemporally, implying a small value for γ.

- In contrast, in Imai&Keane (2004), it is the slope of the hours curve relative to the OCT (rather than the wage) that matters for estimating g, thus implying a much larger estimate of g. Indeed, Imai & Keane estimate that $\gamma = 3.8$.

- Importantly, and in contrast to MaCurdy (1981), human capital breaks the direct link between γ and the Frisch elasticity.

- Another key prediction of the human capital model is that labor supply elasticities with respect to (anticipated) transitory wage changes increase steadily with age.

3.2. Borrowing Constraints

- In a model with credit constraints, reallocation of hours across time may require reallocating consumption across time, and the willingness to substitute labor intertemporally may be limited by the willingness to reallocate consumption.

- Technically, the Frisch elasticity, defined as the change in hours in response to a change in the wage, holding the marginal utility of consumption fixed, no longer exists; any reallocation of hours to the current period and away from other periods will reduce the marginal utility of consumption in the current period while increasing it in other periods.

- Nevertheless, the more general concept of an intertemporal elasticity of substitution in labor supply still exists.

- Domeij&Floden (2006) argue that credit constraints may explain why researchers obtain low estimates of the intertemporal elasticity of substitution when estimating equations like Equation 10.

- Domeij&Floden argue that credit constraints are important in the US economy and that many households hold little wealth.

- Domeij&Floden (2006) assume the same period utility function as in the benchmark model of Section 2, but the flow budget equation is now

$$A_{it} = (1+r)[A_{it} + w_{it}h_{it} - c_{it}], \quad A_{it} \geq 0. \tag{14}$$

- The stochastic process for wages is

$$\log w_{it} = \psi_t + z_{it}, \text{ where } z_{it} = \rho z_{it-1} + \varepsilon_{it}. \tag{15}$$

- MaCurdy's (1981) instrumental variable (IV) procedure to estimate γ in Equation 10 does not require one to specify a particular wage process.

- However, once we introduce extensions such as human capital or credit constraints, it becomes necessary to specify the complete model, including the wage process.

- Let ϕ_{it} denote the marginal utility of borrowing for person i at time t.

- Of course, ϕ_{it} is zero when optimal assets are positive, but it is positive if the optimal asset level is negative (i.e., the nonnegativity constraint binds).

- The marginal utility of consumption evolves according to

$$\Delta \log \lambda_{it} = \log \beta(1 + r_t) - \frac{\phi_{it-1}}{\lambda_{it-1}} + \xi_{it}, \tag{16}$$

and Equation 10 becomes

$$\Delta \log h_{it} = \gamma \Delta w_{it}(1 - \tau_{it}) - \gamma \frac{\phi_{it-1}}{\lambda_{it-1}} - \gamma \log \beta(1 + r_t) + \alpha\gamma \Delta X_{it} + \gamma \Delta \varepsilon_{it}. \tag{17}$$

- The term $\phi_{it-1}/\lambda_{it-1}$ can be interpreted as an omitted variable in the conventional IV estimation method.

- Higher expected wage growth from $t-1$ to t tends to increase the marginal utility of borrowing at time $t-1$.

- That is, ceteris paribus, a steeper future wage profile increases one's desire to borrow against future income to finance current consumption.

- Thus, ϕ_{it-1} is positively correlated with expected wage growth.

- Higher expected wage growth from $t-1$ to t also increases the worker's perceived wealth, and this reduces the marginal utility of consumption at time $t-1$.

- Thus, the entire term $\phi_{it-1}/\lambda_{it-1}$ is positively correlated with expected wage growth.

- Moreover, as is evident from Equation 17, the term $\phi_{it-1}/\lambda_{it-1}$ has a negative effect on hours growth.
- Intuitively, when people are liquidity constrained (i.e., $\phi_{it-1} > 0$), they tend to work more than they would if they could borrow against future income.
- Thus, $\phi_{it-1}/\lambda_{it-1}$ is positively correlated with expected wage growth and negatively correlated with hours growth.
- Hence, its omission will lead to downward bias in estimates of γ.

3.3. Optimization Frictions

- Chetty (2012) argues that abstracting from fixed costs of adjusting labor supply may also downwardly bias labor supply elasticities.
- Chetty attempts to bound the magnitude of the bias in elasticity estimates that might be attributed to ignoring fixed costs.
- Chetty (2012) argues that elasticity estimates are likely to be biased downward, perhaps substantially.
- This result stems from an asymmetry in how adjustment costs affect behavior when elasticities are high versus low.
- If the elasticity is large, then the objective function is fairly flat in the vicinity of optimal hours, so a sizeable departure from optimal hours causes only a small welfare loss.

- To proceed, assume a simple quasi-linear utility function,

$$U_i = wh_i(1-\tau_i) - \frac{\alpha}{1+\frac{1}{\gamma}} h_i^{1+\frac{1}{\gamma}}. \tag{18}$$

- As there are no income effects, the Marshall, Hicks, and Frisch elasticities are equivalent.

- Optimal hours are

$$h_t^* = \left[\frac{(1-\tau)w}{\alpha}\right]^{\gamma}, \tag{19}$$

- and utility evaluated at the optimum is

$$U(h_t^*|\tau_t) = \frac{1}{1+\gamma}\left[\frac{1}{\alpha}\right]^{\gamma}\left[(1-\tau_t)w\right]^{1+\gamma}. \tag{20}$$

- Consider a change in $(1-\tau)$.

- The impact on utility can be decomposed into the direct effect of the change, holding h fixed, plus the effect induced by the behavioral response of changing h :

$$U(h_{t+1}^*|\tau_{t+1}) - U(h_t^*|\tau_t) = \left[U(h_t^*|\tau_{t+1}) - U(h_t^*|\tau_t)\right] + \left[U(h_{t+1}^*|\tau_{t+1}) - U(h_t^*|\tau_{t+1})\right]. \qquad (21)$$

- From Equation 18, the first term on the right-hand side is just $wh_t^*\Delta(1-\tau)$, the change in c holding h fixed.

- The second term, the hours adjustment term, is a second-order effect that can be ignored in the case of small tax changes.

- From Equation 20, we have that $\frac{dU(h_t^*|\tau_t)}{d(1-\tau_t)} = wh_t^*$.

- So $\frac{d^2U(h_t^*|\tau_t)}{d(1-\tau_t)^2} = \gamma wh_t^*/(1-\tau_t)$

- Thus, using a Taylor series approximation, we have that, to second order,

$$U(h^*_{t+1}|\tau_{t+1}) - U(h^*_t|\tau_t) = wh^*_t \Delta(1-\tau_t) + \frac{1}{2}\gamma \frac{wh^*_t}{(1-\tau_t)} \Delta(1-\tau)^2. \qquad (22)$$

- Now assume a worker will not adjust hours if the utility gain is less than a fraction δ of consumption:

$$U(h^*_{t+1}|\tau_{t+1}) - U(h^*_t|\tau_{t+1}) = \frac{1}{2}|U''(h^*_t)|(h^*_{t+1} - h^*_t)^2 < \delta wh^*_t(1-\tau_t). \qquad (23)$$

- With quasi-linear utility, one obtains $|U''(h^*_t)| = \alpha\left(\frac{1}{\gamma}\right) h^{*\frac{1}{\gamma}-1}_t$.

- Assuming that hours were at their optimal level at t, we obtain a bound on the maximum percentage deviation of hours at $t+1$ from their optimal level:

$$\frac{h^*_{t+1} - h^*_t}{h^*_t} < [2\gamma\delta]^{1/2}. \qquad (24)$$

3.4. Summary

- Our presentation considers each of these extensions to the benchmark framework in isolation to emphasize the distinctive economic forces in each case.
- It is of interest for future work to consider these extensions jointly, both to evaluate how they interact and to assess their relative importance.
- Integrating human capital accumulation, credit constraints, and precautionary savings motives into the benchmark life-cycle model seems both natural and straightforward.
- In some cases, the effects we have studied may partially offset each other; for example, if human capital accumulation concerns lead younger workers to work more hours, then the impact of credit constraints may be less relevant.
- Based on existing work, it is our view that the inclusion of human capital accumulation as a way to account for life-cycle changes in wages is of paramount importance for the analysis of labor supply.

4. AGGREGATE LABOR SUPPLY IN MODELS WITH EXTENSIVE MARGIN ADJUSTMENT

4.1. Indivisible Labor Models

- The starting points for our discussion are the indivisible labor papers by Hansen (1985) and Rogerson (1988), who study homogeneous agent models in which all adjustments at the individual level were assumed to occur at the extensive margin (i.e., the intensive margin was fixed by assumption).

- Specifically, individuals had preferences given by

$$\sum_{t=0}^{\infty} \beta^t \left[u(c_t) + v(1 - h_t)\right],$$

- but the choice of h_t was restricted to zero or $\hat{h}$. A key result was that, assuming a set of markets sufficiently rich to decentralize optimal allocations, aggregate allocations in this economy were identical to those that would emerge from an economy with a representative household that made all labor supply adjustment at the intensive margin but had preferences given

$$\sum_{t=0}^{\infty} \beta^t \left[u(c_t) - \alpha h_t\right],$$

- where α is a constant.

- One issue with this result is that it assumes identical households.
- Both Cho (1995) and Mulligan (2001) demonstrate theoretically that the implication of an infinite Frisch elasticity for aggregate labor supply is not robust to including heterogeneity.
- More generally, the Frisch elasticity for aggregate labor supply would depend on the nature and extent of heterogeneity.Akey issue was to assess the implications of empirically relevant sources of heterogeneity.
- A quantitative analysis of this issue was undertaken by Chang &Kim (2006).
- They consider an aggregate model in which labor supply is indivisible but also assume that individuals are subject to idiosyncratic shocks and face incomplete markets for credit and insurance.

- Households consist of a male and a female, with household preferences given by

$$\sum_{t=0}^{\infty} \beta^t \left[2\log(.5c_t) - \alpha_m \frac{h_{mt}^{1+\frac{1}{\gamma}}}{1+\frac{1}{\gamma}} - \alpha_f \frac{h_{ft}^{1+\frac{1}{\gamma}}}{1+\frac{1}{\gamma}} \right],$$

- where c_t is household consumption, and h_{mt} and h_{ft} are hours worked by the male and female household member, respectively.

- Each individual can only supply zero or $\hat{h}$ units of labor in any period.

- Individual productivity, denoted by x_t, is stochastic and follows the stochastic process

$$\log x_{jt+1} = \rho_j \log x_{jt} + \varepsilon_{jt+1},\ j = m, f. \tag{25}$$

- The process is the same for all individuals of a given gender, and innovations are independently and identically distributed across individuals.

- A worker of productivity x_t has labor earnings $w_t x_t \hat{h}$ if working, where w_t is the wage per efficiency unit of labor.

- Chang&Kim proceed to study the properties of individual and aggregate labor supply in their calibrated model.

- First, they consider a sample of 50,000 households in the steady state, simulate their histories for 120 quarters, and then aggregate the observations to annual frequencies.

- In the spirit of Altonji (1986), they run a panel regression of the following form using individuals with positive hours in each year:

$$\log h_{it} = \gamma(\log w_{it} - \log c_{it}) + \varepsilon_{it}. \qquad (26)$$

- They obtain estimates of γ equal to 0.41 and 0.78 for males and females, respectively.

- The key finding is that standard labor supply regressions on individual data generated by the model yield relatively small estimates of the labor supply elasticity parameter for men, although a moderate estimate for women.

- Second, Chang&Kim subject the economy to an AR(1) aggregate technology shock, simulate the economy for 30,000 quarters, compute aggregates, and run the regression in Equation 26 using aggregate time-series data. The resulting estimate for γ is now 1.08.

- Third, they consider a stand-in household model with preferences of the form

$$\sum_{t=0}^{\infty} \beta^t \left[\log(c_t) - \tilde{\alpha} \frac{h_t^{1+\frac{1}{\tilde{\gamma}}}}{1+\frac{1}{\tilde{\gamma}}} \right],$$

- where h_t is now allowed to take on any value in the interval [0, 1].

- Assuming the same process for aggregate technology shocks, Chang & Kim find that a $\tilde{\gamma}$ of approximately 2 generates fluctuations in aggregate hours that are the same as in the heterogeneous agent economy.

- That is, using a stand-in household model to mimic the business cycle statistics for the heterogeneous agent economy requires a value of $\tilde{\gamma}$ that is roughly five times as large as the estimate based on individual data for male workers.

4.2. Models with Intensive and Extensive Margin Adjustment

- How important is the extreme assumption that all adjustments take place along the extensive margin?

- To answer this question, we next explore the aggregate properties of models that feature adjustments along both the intensive and extensive margins.

- We begin by describing the analysis in Rogerson&Wallenius (2009), which generalizes the model in Prescott et al. (2009).

- Consider an individual with the length of life normalized to one and preferences

$$\int_0^1 \left[u(c(a)) - v(h(a))\right] da,$$

- where $c(a)$ is consumption at age a, and $h(a)$ is time devoted to market work at age a.

- Individual productivity varies over the life cycle and is denoted by $e(a)$.

- Following Prescott et al. (2009), the key feature of the model is a nonconvexity in the mapping from time devoted to work to the resulting labor services:

- When a worker of age a devotes h units of time to market work, it generates labor services of $\max\{h-\bar{h},0\}\, e(a)$.

- With $\bar{h}>0$, the model can generate "retirement" as an endogenous outcome, in the sense of a worker who switches from fulltime work to no work despite continuous changes in fundamentals.

- Let w be the constant wage rate per unit of labor services.

- Assuming complete credit markets and a zero interest rate, one finds that the present value budget equation for each individual is

$$\int_0^1 c(a)da = w\int_0^1 \max\left\{h(a)-\bar{h},0\right\}e(a)da. \tag{27}$$

- Three key results emerge. First, if one uses the micro data from the model to run a regression of the form

$$\log(h(a)) = b_0 + \tilde{\gamma}\log\left(w^h(a)\right) + \varepsilon(a), \qquad (28)$$

- the estimated value of $\tilde{\gamma}$ is only about half as large as the true underlying value of γ.

- Second, the response of aggregate hours to the permanent tax and transfer policy change is to first order independent of the value of γ.

- Third, although γ has virtually no effect on the change in aggregate hours, it determines how the change in aggregate hours is decomposed into changes in working life versus changes in hours worked while employed.

- To summarize, in this life cycle economy with operative intensive and extensive margins, labor supply elasticities estimated on micro panel data using workers with positive hours are not particularly relevant in predicting the aggregate effects of permanent changes in taxes. Moreover, the aggregate elasticity is large.

4.3. Summary

5. ADJUSTMENT ON THE EXTENSIVE MARGIN: EVIDENCE FROM MICRO DATA

5.1. Early Work on Structural Models of Participation

- To study female labor supply, in which nonparticipation is prevalent, Heckman & MaCurdy (1980, 1982) modify the utility function in MaCurdy (1981) to

$$U_{it}(c_{it}, h_{it}) = \nu_{it}\eta^{-1}c_{it}^{1-\frac{1}{\eta}} + \alpha_{it}\gamma^{-1}(H_{\max} - h_{it})^{1-\frac{1}{\gamma}}.$$

- Although this generates a reservation wage for participation (the marginal disutility of work is not zero at full leisure), optimal hours are a continuous function of wages, implying that we should observe some women who work very low hours if the wage distribution is continuous.

- In fact, few women are observed to work small positive hours. To match this pattern, Cogan (1981) introduces fixed costs of work into a static labor supply model, generating what he calls a "reservation hours" level.

- Specifically, consider the quasi-linear utility function

$$u(c,h) = c + \alpha\frac{(\overline{H} - h)^{1-\frac{1}{\gamma}}}{1-\frac{1}{\gamma}}.$$

- With w, Y, and m denoting the wage rate, nonlabor income, and the fixed (monetary) costs of working, respectively, utility as a function of hours worked is

$$U(h) = (wh + Y - m) + \alpha \frac{\left(\overline{H} - h\right)^{1-\frac{1}{\gamma}}}{1 - \frac{1}{\gamma}}. \qquad (29)$$

- Optimal hours conditional on working are

$$h^* = \overline{H} - \left(\frac{w}{\alpha}\right)^{-\gamma}. \qquad (30)$$

- Working is optimal if $U(h^*) > U(0)$, which reduces to

$$h^* = \overline{H} - \left(\frac{w}{\alpha}\right)^{-\gamma} > \frac{m}{w} + \frac{1}{w} \frac{\alpha}{1 - \frac{1}{\eta}} \left[\overline{H}^{1-\frac{1}{\gamma}} - \left(\frac{w}{\alpha}\right)^{\gamma-1}\right] = h_{\mathrm{R}} > 0. \qquad (31)$$

- Equation 31 implies that a person works only when optimal hours exceed the reservation hours level h_R.

5.2. Life-Cycle Models with a Participation Margin

- Kimmel & Kniesner (1998) extend the basic MaCurdy (1981) framework to include fixed costs.

- Rather than structurally estimating the model's primitives, they estimate a life-cycle labor supply equation analogous to Equation 11 jointly with a participation decision rule and an offer wage function:

$$\log h_{it} = f_{hi} + \gamma_I \log w_{it} + \alpha_h Z_{it} + \varepsilon_{hit}, \tag{32}$$

$$P(h_{it} > 0) = F\big(f_{pi} + \tilde{\gamma}_P \log w_{it} + \alpha_p Z_{it}\big), \tag{33}$$

- where f_{hi} captures the marginal utility of wealth, along with any fixed effects in tastes for work; f_{pi} captures these and any individual heterogeneity in the fixed costs of work; and F is a cumulative distribution function.

- Kimmel & Kniesner assume that F is normal, giving a probit model.

- There are now two elasticity concepts of interest: γ_I is the conventional Frisch elasticity of labor supply conditional on employment (i.e., the elasticity on the intensive margin), and γ_P, defined by,

$$\gamma_P = \frac{\partial \log P(h_{it} > 0)}{\partial \log w_{it}} = \tilde{\gamma}_P \frac{F'(\cdot)}{F(\cdot)}, \tag{34}$$

- is a Frisch participation elasticity.

- Kimmel&Kniesner (1998) estimate this model using data on 2,428 women from the Survey of Income Program Participation (SIPP).

- They find that $\gamma_I = 0.66$ and $\gamma_P = 2.39$. Let $H = P\hat{h}$ be average hours in the population, where $\hat{h}$ is the average hours of the employed and P is the percentage employed.

- Then

$$\frac{\partial \log H}{\partial \log w} = \frac{\partial \log P}{\partial \log w} + \frac{\partial \log \hat{h}}{\partial \log w} = 0.66 + 2.39 = 3.05. \tag{35}$$

- They also obtain results for men and find that $\gamma_I = 0.39$ and $\gamma_P = 0.86$ so that $\gamma_I + \gamma_P = 1.25$.

- In summary, the participation elasticity is much larger than the hours elasticity for both women and men, and the overall elasticity is quite a bit larger for women than for men.

- These results strongly suggest that failure to account for participation decisions may lead one to substantially underestimate the overall responsiveness of labor supply to wage changes.

- Kimmel & Kniesner (1998) avoid full solution of agents' dynamic optimization problem by relying on a participation condition (Equation 33) derived from the first-order condition for hours evaluated at $h = 0$.

- Eckstein & Wolpin (1989) were the first to adopt a fully structural approach to estimating female labor supply. Their model includes work decisions on the extensive margin and human capital accumulation through work experience.

- The most comprehensive modeling effort to date is by Keane&Wolpin (2007, 2010). They extend earlier work to include marriage, fertility, school attendance, part-time work, and welfare participation as choices.

- Simulations of their model imply a “long-run” labor supply elasticity in response to permanent wage changes of approximately 2.8.

- This “long-run” elasticity has a very different interpretation from elasticities reported in the more conventional labor supply literature.

- First, it measures how a person born into a higher wage (or lower tax) regime would be affected once he or she reaches adulthood.

- Second, aside from labor supply, the simulation allows for adjustments along several other dimensions.

- For example, if a wage increase causes a woman to work more in the current period, she will not only have more human capital in the next period, but her expected number of children is also reduced, thereby further enhancing labor supply in the next period, and so on.

- Finally, Blundell et al. (2013) develop a dynamic model of female labor supply that incorporates asset and human capital accumulation and that endogenizes education.
- There are fewer structural estimation papers for males that incorporate the extensive margin as it has generally been viewed as a less important factor for men, given their high participation rate.
- However, research suggests that the extensive margin is important for males who are young, near retirement, or members of minority groups.
- In a series of papers, Keane& Wolpin (1997, 2000, 2001) study the career decisions of young men. Their models allow for work decisions on the extensive margin, along with schooling and occupation choices, all of which influence the evolution of human capital.

5.3. Summary

- The literature on estimating extensive margin elasticities in dynamic structural models is relatively young.

- However, based on the existing studies, there appears to be a very consistent pattern of high estimated labor supply elasticities for women at the extensive margin, as well as for males who have relatively low participation rates (i.e., the young, the old, and minorities).

6. CONCLUSION

- Based on the survey of the micro labor supply literature by Blundell&MaCurdy (1999), it is fair to say that the consensus view among labor economists was (and still is) that labor supply elasticities are small.
- In contrast, macroeconomists generally work with equilibrium models in which Hicks (or compensated) and Frisch (or intertemporal) labor supply elasticities are quite large (i.e., in the range of 1–2).
- In this review, we describe a relatively new literature that seeks to reconcile these conflicting micro and macro views on labor supply.

- This literature can be viewed as consisting of two branches.

- The first focuses on the micro perspective.

- In the basic life-cycle labor supply model of MaCurdy (1981), the only source of dynamics is borrowing/saving.

- Several authors have considered extensions of this model to include other potentially important sources of dynamics, such as human capital and credit constraints.

- Some work has also allowed for corner solutions in labor supply. This work has shown that if the true model (or data-generating process) contains such mechanisms, but the data are viewed through the lens of the basic model, then estimates of labor supply elasticities will tend to seriously understate their true values.

- The second branch focuses on the macro perspective.
- This literature emphasizes issues of aggregation in the presence of the extensive margin and worker heterogeneity.
- This literature has shown that small (intensive margin) elasticities at the individual level are consistent with large elasticities at the aggregate level.
- These two literatures share one key point in common.
- In the basic life-cycle model of MaCurdy (1981), there is a direct link between individual-level preference parameters and labor supply elasticities at the aggregate level.
- All the extensions to the basic model that we describe above break that direct link.