Journal of Economic Literature 2011, 49:4, 961–1075
http:www.aeaweb.org/articles.php?doi=10.1257/jel.49.4.961

Labor Supply and Taxes: A Survey

Michael P. Keane*

I survey the male and female labor supply literatures, focusing on implications for effects of wages and taxes. For males, I describe and contrast results from three basic types of model: static models (especially those that account for nonlinear taxes), life-cycle models with savings, and life-cycle models with both savings and human capital. For women, more important distinctions are whether models include fixed costs of work, and whether they treat demographics like fertility and marriage (and human capital) as exogenous or endogenous. The literature is characterized by considerable controversy over the responsiveness of labor supply to changes in wages and taxes. At least for males, it is fair to say that most economists believe labor supply elasticities are small. But a sizable minority of studies that I examine obtain large values. Hence, there is no clear consensus on this point. In fact, a simple average of Hicks elasticities across all the studies I examine is 0.31. Several simulation studies have shown that such a value is large enough to generate large efficiency costs of income taxation. For males, I conclude that two factors drive many of the differences in results across studies. One factor is use of direct versus ratio wage measures, with studies that use the former tending to find larger elasticities. Another factor is the failure of most studies to account for human capital returns to work experience. I argue that this may lead to downward bias in elasticity estimates. In a model that includes human capital, I show how even modest elasticities—as conventionally measured—can be consistent with large efficiency costs of taxation. For women, in contrast, it is fair to say that most studies find large labor supply elasticities, especially on the participation margin. In particular, I find that estimates of "long-run" labor supply elasticities—by which I mean estimates that allow for dynamic effects of wages on fertility, marriage, education and work experience—are generally quite large. (*JEL* D91, J13, J16, J22, J31, H24)

1. *Introduction*

The literature on labor supply is one of the most extensive in economics. There are many reasons why the topic is of such great interest. One key reason is that understanding the responsiveness of labor supply to after-tax wages and transfers is crucial

* University of New South Wales and Arizona State University. This research has been supported by Australian Research Council grant FF0561843 and by the AFTS Secretariat of the Australian Treasury. But the views expressed are entirely my own. An early version was presented at Australia's Future Tax and Transfer Policy Conference (June 18, 2009). I thank Roger Gordon and two anonymous referees for suggestions that have greatly improved the exposition of the survey.

for the effective design of tax/transfer systems, and for assessing the efficiency costs of labor income taxation. Thus, it is not surprising that much of the literature on labor supply focuses on how taxes on labor earnings affect peoples' decisions about hours of work. Given the importance of this issue, I will survey the labor supply literature with a particular focus on what it implies about the elasticity of labor supply with respect to wages and taxes.[1]

Unfortunately, there is no clear consensus on this issue. Indeed, the labor supply literature is characterized by a number of sharp controversies, many of which revolve around the magnitudes of labor supply elasticities, and the methods used to estimate them. *At least for men*, it is fair to say that the majority of studies find rather small elasticities with respect to after-tax wage rates. This, in turn, implies that efficiency costs of distortionary income taxation are small. But, as we'll see, a sizable minority of studies makes a strong case for larger elasticities. I will admit up front that my sympathies are with the minority group. My own judgment is that many prior studies have obtained male labor supply elasticities that are likely biased toward zero. I will explain why I think this is so, while at the same time attempting to present as balanced a view of the literature as possible.

For women, in contrast, most studies have found rather large labor supply elasticities, especially on the participation margin. This is particularly true of studies that allow for long-run dynamic effects of wages on fertility, marriage, education, and work experience.

To begin, section 2 gives a short summary of the optimal tax literature to motivate why labor supply elasticities are so important. Next, section 3 describes standard models of labor supply. I first discuss static models and then dynamic (or life-cycle) models. I also derive the three main elasticity concepts (Marshall, Hicks, and Frisch). Section 4 discusses, in a general way, some of the main econometric problems that arise in attempting to estimate labor supply models. This list is not meant to be exhaustive, but rather to highlight the difficultly of the problem. Additional issues are brought out in later sections.

Section 5 serves as a roadmap that links the theoretical sections 2–4 with the specific empirical papers that will be discussed in sections 6–7. Given the size of the empirical literature, it is helpful to begin with a fairly comprehensive roadmap that outlines how different papers seek to address the econometric problems noted in section 4—as well as various other problems that are also considered. As will become clear, no paper can claim to do more than address a subset of these problems. Thus, an understanding of biases that are likely to result from ignoring certain problems is crucial for interpreting the literature.

Section 6 surveys empirical results on male labor supply. It is divided into three parts. Section 6.1 covers "static" models that consider only choice of work hours but take assets and human capital as given. Section 6.2 covers "life-cycle" models that incorporate saving. Section 6.3 covers life-cycle models that also account for human capital (and other sources of dynamics). Next, section 7 surveys results on female labor supply. Following Jacob Mincer (1962), I feel it is harder to ignore life-cycle issues when considering female labor supply, so this section is much more focused on life-cycle models. Finally, section 8 concludes.

[1] Labor supply elasticities also play a key role in business cycle models, where they govern the extent to which fluctuations in real wages over the cycle can explain movements in hours and employment. Understanding labor supply behavior is important in the design of public welfare programs, where the goal is typically to provide income support in such a way as to minimize work disincentives. One could list other important applications.

2. *Labor Supply and Optimal Taxation*

One of the main reasons for interest in labor supply elasticities is that they play a key role in the design of tax systems. So, by way of motivation, it is important to understand why. Therefore, I'll start with a brief and informal summary of the "optimal taxation" literature, pioneered by James A. Mirrlees (1971). The optimal tax literature starts with two key problems: (1) government needs revenue to pay for public goods (e.g., education, health care, defense forces), income support for the poor, and other desirable programs and (2) the use of labor income taxation to raise revenue causes people to work less. This leads to a decline in overall economic output (generating an efficiency loss).

There is clearly a trade-off between the desirable government services that income taxation can fund and the undesirable negative impact of taxation on labor supply. Mirrlees (1971) developed mathematical models of this trade-off, and used them to derive optimal levels of taxation and government spending. His basic conclusion was that the efficiency costs of taxation are greater to the extent that the "shrinking pie" problem (2) is more severe. The more elastic is labor supply to after-tax wages, the lower is the optimal tax rate.

To give a concrete example, consider a progressive income tax, and suppose we want to choose the optimal tax rate for the top income bracket. To simplify, assume the top bracket is sufficiently high that government (or society) places no value whatsoever on an extra dollar of income for those people in it. The government's only goal is to raise as much revenue from them as possible. In this case, Emmanuel Saez, Joel Slemrod, and Seth H. Giertz (2009) give the following simple formula for the revenue maximizing top bracket tax rate, which I denote τ:

$$\tau = \frac{1}{1 + a \cdot e}. \tag{1}$$

Here e is the labor supply elasticity (i.e., the percent increase in labor supply that accompanies a 1 percent increase in the after-tax wage rate $w(1-\tau)$, where w is the pretax wage),[2] and a is the "Pareto parameter," an (inverse) measure of income dispersion within the top bracket.

Specifically, a is defined as $a = z_m/(z_m - z)$, where z is the income level where the top bracket starts, and z_m is the average income of people in the top bracket. For example, if the top bracket starts at $500,000, and average income in that bracket is $1,000,000, then $a = 1/(1–0.5) = 2$. In contrast, if average income in the top bracket is $2,000,000 (implying more dispersion/inequality), we would have $a = 2/(2–0.5) = 1.33$. As $z_m \to \infty$, meaning inequality becomes extreme, a approaches its lower bound of 1. From (1), we see that the optimal top bracket rate *increases* if there is more inequality in the top bracket (i.e., a smaller value of a).

Quite a few papers have estimated the Pareto parameter for many different countries using different points for the top bracket cutoff. There is a great deal of consistency in the estimates: it is generally found that a is stable and in the vicinity of 1.5 to 2 for income levels where the top bracket rate would typically apply. For example, Saez (2001) looks at U.S. tax return data from 1992–93 and finds a is stable at about 2 for income levels above $150,000, while Anthony B. Atkinson, Thomas Piketty, and Saez (2011) find a is around 1.5 in recent years. For the United Kingdom, Michael Brewer, Saez, and Andrew Shepard (2010) report a value of 1.67.

Now, consider what equation (1) implies for the optimal top bracket tax rate given

[2] At this point I abstract from the fact that there are multiple definitions of the labor supply elasticity, depending on what is held fixed as wages vary. This is discussed in the next section. Later we'll see that, in a static model, it is the Hicks elasticity concept that is relevant for determining the optimal top bracket rate.

TABLE 1
OPTIMAL TOP BRACKET TAX RATES FOR DIFFERENT LABOR SUPPLY ELASTICITIES

Labor supply elasticity (e)	Optimal top-bracket tax rate (τ)		
	$a = 1.50$	$a = 1.67$	$a = 2.0$
2.0	25%	23%	20%
1.0	40%	37%	33%
0.67	50%	47%	43%
0.5	57%	54%	50%
0.3	69%	67%	63%
0.2	77%	75%	71%
0.1	87%	86%	83%
0.0	100%	100%	100%

Note: These rates assume the government places essentially no value on giving extra income to the top earners.

different values of the labor supply elasticity e and the Pareto parameter a.

Table 1 reveals quite strikingly how sensitive the optimal top bracket tax rate is to the labor supply elasticity. Take the $a = 1.5$ case. If the elasticity is only 0.2, then the optimal top rate is a very high 77 percent. But if the elasticity is 2.0, the optimal top rate is only 25 percent.

Recall that table 1 assumes the government places no value on additional income for people in the top bracket: its only goal is revenue extraction. Nevertheless, the optimal top rate is generally below 100 percent—otherwise there would be no incentive to earn income above level z. The one exception, as we see in table 1, is if labor supply is totally inelastic ($e = 0$). It is also worth noting that, if government does place some value on marginal income for high earners, optimal top rates would be lower than those presented in table 1 (see below).

For a flat rate tax system (i.e., a system with no brackets, and a single flat rate tax on all income starting at \$0), we would have $z = 0$ and $a = z_m/z_m = 1$. Then, if the government's goal is purely revenue maximization, equation (1) reduces to simply

$$\tau = 1/(1 + e). \tag{2}$$

It is also simple to derive (2) directly. Let h denote hours of work, and assume that $\ln(h) = e \cdot \ln(w(1 - \tau))$, so e is the labor supply elasticity. Then $h = [w(1 - \tau)]^e$. Let R denote tax revenue. We have $R = (wh)\tau = w[w(1 - \tau)]^e \cdot \tau$. It is instructive to look at the derivative of R with respect to τ, which is $dR/d\tau = w[w(1 - \tau)]^e - ew^2[w(1 - \tau)]^{e-1} \cdot \tau$. The first term, which is *positive*, is the mechanical effect of the tax increase, holding labor supply fixed. The second term, which is *negative*, is the behavioral effect: the loss in revenue due to reduced labor supply. Setting this derivative equal to zero and solving for the revenue maximizing τ gives equation (2).

Using equation (2), table 2, column 2 gives optimal (revenue maximizing) flat-tax rates for different values of the labor supply

TABLE 2
REVENUE MAXIMIZING FLAT TAX RATES GIVEN DIFFERENT LABOR SUPPLY ELASTICITIES

Elasticity (e)	Optimal tax rate (τ)	
	$g = 0$	$g = 0.5$
2.0	33%	20%
1.0	50%	33%
0.67	60%	43%
0.5	67%	50%
0.3	77%	63%
0.2	83%	71%
0.1	91%	83%
0.0	100%	100%

elasticity e.[3] As in table 1, the optimal rate falls sharply as the elasticity increases. For instance, if $e = 0.5$, the revenue maximizing flat rate is a very high 67 percent. But if $e = 2.0$, the revenue-maximizing rate is only 33 percent.

Note that, because a is smaller in table 2 than in table 1 (i.e., $a = 1.0$), the tax rates in table 2 are generally higher. This may seem surprising, as we are now considering a flat rate tax, as opposed to a top bracket tax. Recall, however, that models in the optimal tax literature assume taxes are used largely to finance inequality-reducing transfers. Under the flat rate scheme, low to middle income tax payers pay high rates but also receive large transfers.

Now, suppose the government does not only seek to maximize revenue. Instead, it places a value of g cents on a marginal dollar of private after-tax income. Then Brewer, Saez, and Shepard (2010) show that (1) becomes $\tau = (1 - g)/(1 - g + a \cdot e)$.[4] Given $a = 1$ and $g = 0.5$, we obtain the figures in table 2, column 3. Not surprisingly, optimal tax rates fall. But more interesting is that rates become even more sensitive to the labor supply elasticity.

Both tables 1 and 2 illustrate the key role of labor supply elasticities in determining optimal tax rates. With this background in mind, I turn to a review of the empirical evidence on the magnitudes of labor supply elasticities. But before beginning it is worth giving a brief summary of the discussion. It is fair to say that, regardless of which of the various definitions of the labor supply elasticity you use, the majority of the economics profession—whether accurately or not—believes

[3] I again abstract from the fact that there are multiple labor supply elasticity concepts. As we'll see later, for a flat rate tax the relevant concept depends on what is done with tax revenue. In a static model where the revenue is returned to the population via lump-sum transfers, it is again the Hicks elasticity concept that is most relevant. But if the revenue is used to finance public goods that do not influence labor supply, it is the Marshallian elasticity concept that is relevant.

[4] This formula assumes for simplicity that all government revenue is used for redistribution (i.e., there is no minimum tax needed to provide essential services). Note that the same adjustment would apply to the top rate formula if the government (or society) places a value of g on a marginal dollar of income for a person in the top bracket. Then. if society has egalitarian preferences, we have $1 > g \geq 0$.

labor supply elasticities are fairly small (i.e., well below 0.50).

This majority view is summed up nicely in a recent survey by Saez, Slemrod, and Giertz (2009), who state: ". . . optimal progressivity of the tax-transfer system, as well as the optimal size of the public sector, depend (inversely) on the . . . elasticity of labor supply. . . . With some exceptions, the profession has settled on a value for this elasticity close to zero . . . In models with only a labor–leisure choice, this implies that the efficiency cost of taxing labor income . . . is bound to be low as well." This view implies, for instance, that the optimal top-bracket tax rate is toward the high end of the figures given in table 1.

In sections 6–7 of this survey I will argue there are important reasons to question whether this majority view is really an accurate reflection of the empirical evidence. As a result, the efficiency cost of labor income taxation may be higher than is conventionally supposed. But first, I turn to an exposition of basic labor supply models.

3. *Basic Models of Labor Supply*

Before discussing the empirical literature on labor supply, it is necessary to lay out the theoretical framework on which it is based. Labor supply models can be broadly classified into two main types, static and dynamic. There are many variations within each type, but for our purposes this simple division will prove useful.

3.1 *The Basic Static Labor Supply Model*

In the basic static model, utility in period t depends positively on consumption C_t and leisure.[5] Alternatively, it is often convenient to write that utility depends negatively on hours of work, h_t. Consumption is given by the static budget constraint $C_t = w_t(1 - \tau)h_t + N_t$, where w_t is the (pretax) wage rate, τ is the tax rate on earnings, and N_t is nonlabor income.[6]

In order to give a concrete exposition of the model it is useful to choose a particular utility function. Furthermore, to facilitate comparison of static versus dynamic models, it will be useful to exposit each using the *same* utility function. The following utility function is very commonly used in the literature on life-cycle models, primarily because it is very convenient:

$$(3) \qquad U_t = \frac{C_t^{1+\eta}}{1+\eta} - \beta_t \frac{h_t^{1+\gamma}}{1+\gamma}, \qquad \eta \leq 0, \quad \gamma \geq 0.$$

Here, utility has a CRRA form in consumption, and the disutility of labor is convex in hours of work. The parameter β_t captures tastes for leisure, which may change over time. The reason (3) is so convenient is that, as we will see below, the single parameter γ governs the strength of substitution effects, while η governs the strength of income effects.

In contrast to dynamic models, for static models it is hard to point to a utility function that is so generally used as (3). This is because in static models it is often more convenient to specify a labor supply function *directly*. Indeed, it is a bit awkward to exposit the static model using equation (3), but I think this approach is justified by the comparability so achieved.

The static model can be viewed as a special case of a dynamic model where all intertemporal linkages have been shut down.

[5] The definition of a "period" in labor supply models is somewhat arbitrary. In empirical work it is often chosen to be a year, although shorter periods are sometimes examined.

[6] So as to focus on earnings taxes, I ignore taxes on nonlabor income. N_t might be interpreted as after-tax nonlabor income, or as a tax-free transfer. The key thing is that we want to consider changes in τ holding N fixed.

First, workers do not borrow or save, so current consumption is equal to current after-tax income.[7] Second, human capital accumulation is ignored. That is, workers decide how much labor to supply today based only on today's wage rate, not considering how today's decisions may affect future wages (e.g., via accumulation of work experience). Third, history dependence in preferences is ignored. Fourth, fertility, marriage and other demographics are taken as given. Alternatively, the static model can be viewed as a special case of a dynamic model where such intertemporal linkages do exist, but where workers are myopic and ignore them when deciding on current labor supply.[8]

To solve the static model for optimal hours of work, we use the budget constraint to substitute for C_t in equation (3), and take the derivative with respect to hours, obtaining

$$(4)\quad \frac{dU_t}{dh_t} = [w_t(1-\tau)h_t + N_t]^{\eta} w_t(1+\tau) - \beta_t h_t^{\gamma} = 0.$$

This can be reorganized into the familiar marginal rate of substitution (MRS) condition:

$$(5)\quad \text{MRS} = \frac{MUL(h)}{MUC(h)} = \frac{\beta_t h_t^{\gamma}}{[w_t(1-\tau)h_t + N_t]^{\eta}} = w_t(1-\tau).$$

Equation (5), one of the most basic in economics, says to choose hours so as to equate the marginal rate of substitution between consumption and leisure to the after-tax wage, $w_t(1-\tau)$. Of course, the MRS is the ratio of the marginal utility of leisure, $\beta_t h_t^{\gamma}$ (the negative of the marginal disutility of hours), to the marginal utility of consumption, $[w_t(1-\tau)h_t + N_t]^{\eta}$.

Unfortunately, equation (5) does not give a closed form solution for hours (this is presumably why it has not been a popular choice in the static literature). But by implicitly differentiating (5) we obtain that the "Marshallian" labor supply elasticity (also known as the "uncompensated" or "total" elasticity), which holds N_t fixed. It is given by

$$(6)\quad e \equiv \left.\frac{\partial \ln h_{it}}{\partial \ln w_{it}}\right|_{N_{it}} = \frac{1+\eta \cdot S}{\gamma - \eta \cdot S},$$

$$\text{where}\quad S \equiv \frac{w_t h_t(1-\tau)}{w_t h_t(1-\tau) + N_t}.$$

Here S is the share of earned income in total income. If nonlabor income is a small share of total income, then, to a good approximation, the Marshallian elasticity is simply $(1+\eta)/(\gamma-\eta)$.

Next we use the Slutsky equation to decompose the Marshallian elasticity into separate substitution and income effects. Recall that the Slutsky equation is

$$(7)\quad \frac{\partial h}{\partial w} = \left.\frac{\partial h}{\partial w}\right|_{u} + h\frac{\partial h}{\partial N}.$$

It is convenient to write the Slutsky equation in elasticity form, so the Marshallian elasticity appears on the left-hand side. So we manipulate (7) to obtain

$$(8)\quad \frac{w}{h}\frac{\partial h}{\partial w} = \left.\frac{w}{h}\frac{\partial h}{\partial w}\right|_{u} + \frac{wh}{N}\left[\frac{N}{h}\frac{\partial h}{\partial N}\right].$$

The first term on the right is the "Hicks" or "compensated" labor supply elasticity. The second term is the income effect, which equals the income elasticity times $S/(1-S)$.

[7] One consequence of this assumption is that the static model has no explanation for the evolution of asset income. Nonlabor income can only be sensibly interpreted as exogenous transfers.

[8] Given myopia, the worker has no reason to save for the future, so all of current income is consumed and a static budget constraint holds (even if saving is technically possible).

Again applying implicit differentiation to (5), we obtain that the income elasticity is

$$(9)\quad e_I \equiv \left.\frac{\partial \ln h_{it}}{\partial \ln N_{it}}\right|_{w_{it}} = \frac{\eta}{\gamma - \eta \cdot S}(1 - S)$$

and hence the income effect is

$$(10)\quad ie \equiv \frac{w_t h_t (1-\tau)}{N_t} e_I = \frac{S}{1-S}\frac{\eta}{\gamma - \eta \cdot S}(1-S) = \frac{\eta \cdot S}{\gamma - \eta \cdot S} < 0.$$

The income effect must be negative, as $\eta < 0$ and $\gamma > 0$ (conditions required for diminishing marginal utility of consumption and leisure). If nonlabor income is a small share of total income, so $S \approx 1$, then, to a good approximation, the income effect is simply $\eta/(\gamma - \eta)$.[9]

It is intuitive that the magnitude of the negative income effect is increasing in the magnitude of the parameter η. If η is a larger negative number, then the marginal utility of consumption diminishes more quickly as consumption increases. Thus, the tendency to reduce labor supply in response to an increase in nonlabor income is greater.

Finally, using (6), (8), and (10), the Hicks elasticity is simply given by

$$(11)\quad e_H \equiv \left.\frac{\partial \ln h_{it}}{\partial \ln w_{it}}\right|_U = e - ie = \frac{1}{\gamma - \eta \cdot S} > 0.$$

The Hicks elasticity must be positive, as $\eta < 0$ and $\gamma > 0$. If nonlabor income is a small share of total income, the Hicks elasticity is approximately $1/(\gamma - \eta)$. Also note that, because $\eta < 0$, the Hicks elasticity in (11) must be greater than the Marshallian elasticity in (6). The two approach each other as $\eta \to 0$, in which case there are no income effects.[10]

In the special case of $N_t = 0$, we can use (5) to obtain the labor supply equation:

$$(12)\quad \ln h_t = \frac{1+\eta}{\gamma - \eta}\ln[w_t(1-\tau)] - \frac{1}{\gamma - \eta}\ln \beta_t .$$

From (12) we can see directly that the Marshallian labor supply elasticity is given by

$$(13)\quad e = \frac{\partial \ln h_t}{\partial \ln w_t(1-\tau)} = \frac{1 + \eta}{\gamma - \eta}.$$

As $\eta < 0$ and $\gamma > 0$ the denominator in (13) must be positive. But aside from this, economic theory tells us little. It is possible for the numerator to be negative if $\eta < -1$. Then an increase in the wage reduces hours of work. Several of the studies I review below do find this. But most find $1 + \eta$ is small and positive, so the Marshallian elasticity is also small and positive.

It is instructive to note the income effect *ie* in (8) can also be written as $w(\partial h/\partial N)$ or $\partial(wh)/\partial N$. For this reason, John H. Pencavel (1986) called it the "marginal propensity to earn."

That is, the income effect can be interpreted as the effect of an increase in nonlabor income on labor income (i.e., given an extra dollar of nonlabor income, how much

[9] If $S = 1$, then the income elasticity e_I is of course equal to zero. (If $N_t = 0$, then a one percent increase in zero is still zero). But the income effect remains well defined, as the result in (10) can also be obtained directly by implicit differentiation of (5), rather than by using (8)–(9).

[10] Much of the literature on optimal taxation assumes away income effects to simplify the analysis (see Peter A. Diamond 1998). But my review of the empirical literature suggests this assumption is questionable.

does a worker reduce his/her earnings?). As Pencavel notes, if both leisure and the composite consumption good (C_t) are normal goods, then *ie* must be between 0 and -1. Indeed, we see from (10) that *ie* approaches –1 as $\eta < 0$ goes to negative infinity. But Pencavel (1986) argues that values of *ie* near –1 are implausible. Introspection suggests people would not react to an increase in nonlabor income by reducing hours so sharply that total consumption does not increase.[11]

Knowledge of both the income and substitution effects of an after-tax wage change is important for understanding the impact of changes in tax and transfer policy. For example, suppose we have a flat rate tax system that is used to finance lump sum transfers to all members of the population (i.e., a negative income tax (NIT) scheme). Further suppose that we decide to increase the flat rate tax rate and increase the grant level. This policy discourages work in two ways. The tax increase itself reduces the reward from work, but the lump sum payments also discourage work via the income effect. To a first order approximation (ignoring heterogeneity in wages/earnings in the population) the Hicks elasticity is the correct concept to use in evaluating the labor supply effects of such a policy change.[12]

[11] Note that if $\eta < -1$, income effects dominate substitution effects, and the Marshallian elasticity turns negative. Substituting $\eta = -1$ into (6) and assuming $S \approx 1$, we see the income effect, or "marginal propensity to earn," cannot be less than $-1/(1+\gamma)$ if the Marshallian elasticity is to be positive. Note that a smaller γ implies a stronger substitution effect.

[12] Aside from population heterogeneity, another approximation here is that I ignore the distinction between "Slutsky compensation"—i.e., giving people enough of a transfer that the original consumption bundle is still feasible—and "Hicks compensation"—i.e., giving people enough of a transfer that the original utility level is maintained. For small policy changes, the two concepts are approximately equivalent, and even for large policy changes our elasticity estimates are probably not precise enough to draw a meaningful distinction between the two. Hence, I will generally ignore the distinction between Hicks and Slutsky compensation in this article.

In contrast, suppose the revenue from the increased income tax is used to finance public goods that are additively separable in (3). Then the negative effect on labor supply is less, as the income effect from transferring the tax revenue back to the population is avoided. The Marshallian elasticity is the correct concept in this case.

Another key point is that in a progressive tax system it can be shown (in the static model) that effects of changes in tax rates beyond the first bracket depend primarily on the Hicks elasticity. Hence, efficiency costs of progressive taxation are largely a function of the Hicks elasticity as well. I discuss this key point in much more detail in section 6.1.

3.2 *The Basic Dynamic Model with Savings*

The pioneering work by Thomas MaCurdy (1981, 1983) and James J. Heckman and MaCurdy (1980, 1982) introduced dynamics into empirical labor supply models by allowing for borrowing and lending across periods. This model is commonly known as the "life-cycle" model of labor supply. They considered a multiperiod model, but in order to emphasize the key points it is useful to first consider a model with two periods in the working life. Initially, I also assume workers have perfect foresight (about future wages, taxes, tastes, and nonlabor income). As before, I will assume that the per period utility function is given by equation (3).

The key change in the dynamic model is that the first period budget constraint is now $C_1 = w_1(1-\tau_1)h_1 + N_1 + b$, where b is the net borrowing in period 1, while $C_2 = w_2(1-\tau_2)h_2 + N_2 - b(1+r)$, where $b(1+r)$ is the net repayment of the loan in period 2. Here τ_1 and τ_2 are tax rates on earnings in periods 1 and 2, and r is the interest rate.[13] Of course, b can be negative (i.e., the person saves in period 1). In the life-cycle

[13] Given the focus on wage taxation, I ignore taxation of asset income. One may interpret r as an after-tax rate.

model, in contrast to the static model, there is a clear distinction between *exogenous* nonlabor income $\{N_1, N_2\}$ and asset income.[14]

In a dynamic model, a person chooses a life-cycle labor-supply/consumption plan that maximizes the present value of lifetime utility, given by

$$(14) \qquad V = U_1 + \rho U_2,$$

where ρ is the discount factor. Substituting C_1 and C_2 into (3) and then (14), we obtain

$$(15) \quad V = \frac{[w_1 h_1(1-\tau_1) + N_1 + b]^{1+\eta}}{1+\eta} - \beta_1 \frac{h_1^{1+\gamma}}{1+\gamma} + \rho\left\{\frac{[w_2 h_2(1-\tau_2) + N_2 - b(1+r)]^{1+\eta}}{1+\eta} - \beta_2 \frac{h_2^{1+\gamma}}{1+\gamma}\right\}.$$

In the standard life-cycle model, there is no human capital accumulation via returns to work experience. Thus, a worker treats the wage path $\{w_1, w_2\}$ as exogenously given (that is, it is unaffected by the worker's labor supply decisions).

In the life-cycle model, a new labor supply elasticity concept is introduced. This is the response of a worker to a *temporary* change in the after-tax wage rate. This could be induced by a temporary tax cut in period 1 that is rescinded in period 2. As the worker can now save, the response may be to work more in period 1, save part of the extra earnings, and work less in period 2. The strength of this reaction (i.e., shifting labor toward periods of high wages) is given by the "intertemporal elasticity of substitution," also known as the "Frisch" elasticity.

The first order conditions for the worker's optimization problem are simply

$$(16) \quad \frac{\partial V}{\partial h_1} = [w_1 h_1(1-\tau_1) + N_1 + b]^{\eta} \times w_1(1-\tau_1) - \beta_1 h_1^{\gamma} = 0$$

$$(17) \quad \frac{\partial V}{\partial h_2} = [w_2 h_2(1-\tau_2) + N_2 - b(1+r)]^{\eta} \times w_2(1-\tau_2) - \beta_2 h_2^{\gamma} = 0$$

$$(18) \quad \frac{\partial V}{\partial b} = [w_1 h_1(1-\tau_1) + N_1 + b]^{\eta} - \rho[w_2 h_2(1-\tau_2) + N_2 - b(1+r)]^{\eta} \times (1+r) = 0.$$

Equation (18) can be written as $[C_1]^{\eta}/[C_2]^{\eta} = \rho(1+r)$, which is the classic intertemporal optimality condition that requires one to set the borrowing level b so as to equate the ratio of the marginal utilities of consumption in the two periods to $\rho(1+r)$.[15]

Utilizing the intertemporal condition, we divide (17) by (16) and take logs to obtain

$$(19) \quad \ln\left(\frac{h_2}{h_1}\right) = \frac{1}{\gamma}\left\{\ln\frac{w_2}{w_1} + \ln\frac{(1-\tau_2)}{(1-\tau_1)} - \ln\rho(1+r) - \ln\frac{\beta_2}{\beta_1}\right\}.$$

[14] It is rather standard in life-cycle models to ignore exogenous nonlabor income $\{N_1, N_2\}$, thus assuming that all nonlabor income is asset income. But as we will see, adding exogenous nonlabor income does not complicate the analysis of MaCurdy-type life-cycle models in any significant way.

[15] An important special case is when $\rho = 1/(1+r)$, so people discount the future using the real interest rate. In that case $\rho(1+r) = 1$ so that $[C_1]^{\eta}/[C_2]^{\eta} = 1$ and hence $C_1 = C_2$. That is, we have complete consumption smoothing.

From (19) we obtain:

$$(20) \qquad \frac{\partial \ln(h_2/h_1)}{\partial \ln(w_2/w_1)} = \frac{1}{\gamma}.$$

Thus, the Frisch elasticity of substitution, the rate at which a worker shifts hours of work from period 1 to period 2 as the relative wage increases in period 2, is simply $1/\gamma$. The elasticity with respect to a change in the tax ratio $(1-\tau_2)/(1-\tau_1)$ is identical.

There is an important relation between the Frisch, Hicks, and Marshallian elasticities:

$$(21) \quad \frac{1}{\gamma} > \frac{1}{\gamma - \eta \cdot S} > \frac{1+\eta \cdot S}{\gamma - \eta \cdot S}$$

$$\Rightarrow \frac{1}{\gamma} > \frac{1}{\gamma - \eta} > \frac{1+\eta}{\gamma - \eta} \quad \text{if } S = 1.$$

That is, the Frisch elasticity is larger than the Hicks, which is larger than the Marshallian. This follows directly from $\eta < 0$ (i.e., diminishing marginal utility of consumption). This ranking of elasticities implies that if we can obtain an estimate of the Frisch it provides an upper bound on how large the Hicks and Marshallian elasticities might be.

It is straightforward to extend the life-cycle model to the case of multiple periods and uncertainty. First, note that equations like (16)–(17) must hold in any period, so we have

$$(22) \quad [w_t h_t(1-\tau_t) + N_t + b_t]^{\eta} w_t(1-\tau_t)$$

$$= \beta_t h_t^{\gamma} \qquad \Rightarrow \qquad \frac{\beta_t h_t^{\gamma}}{C_t^{\eta}} = w_t(1-\tau_t).$$

Note that (22) is almost identical to the MRS condition (5) in the static model. The only difference is that now consumption includes the borrowing/lending amount b_t that is allocated to period t. (This observation plays a key role in subsequent developments in section 6.2.)

Under uncertainty the intertemporal condition (18) only holds in expectation. Following MaCurdy (1981), we write

$$(23) \quad C_t^{\eta} = E_t \rho(1 + r_{t+1}) C_{t+1}^{\eta}$$

$$\Rightarrow \quad \rho(1 + r_{t+1}) C_{t+1}^{\eta} = C_t^{\eta}(1 + \xi_{t+1}),$$

where ξ_{t+1} is a mean zero forecast error that is uncorrelated with information known to the agent at time t. If we assume these forecast errors are "small," we obtain an approximate expression for the evolution of the marginal utility of consumption over the life cycle:

$$(24) \quad \Delta \ln C_t^{\eta} = -\ln \rho(1 + r_{t+1}) + \xi_t.$$

Taking logs and differencing (22), and using (24) to substitute out for $\Delta \ln C_{it}^{\eta}$, we obtain

$$(25) \quad \Delta \ln h_t = \frac{1}{\gamma} \Delta \ln w_t + \frac{1}{\gamma} \Delta \ln(1-\tau_t) - \frac{1}{\gamma} \ln \rho(1+r_t) - \frac{1}{\gamma} \Delta \ln \beta_t + \frac{1}{\gamma} \xi_t.$$

Under conditions I discuss below, (25) can be used to estimate the Frisch elasticity $(1/\gamma)$.

With these concepts in hand, we are in a position to talk about estimation of labor supply elasticities. In section 4, I discuss the key econometric issues that arise in estimation.

4. *Econometric Issues in Estimating Labor Supply Elasticities*

Broadly speaking, there are two main approaches to estimating labor supply elasticities in the literature. One starts by specifying a utility function, which is then fit to data on hours, wages, and nonlabor income. The alternative is to specify a labor supply function directly. The functional form is typically chosen so that estimation involves regressing hours on wages and nonlabor income, and so that convenient expressions for labor supply elasticities are obtained. I'll begin by discussing a regression approach.

Many functional forms for labor supply regressions have been tried, but there is no consensus which one is "right." To fix ideas, consider a logarithmic specification of the form

$$(26) \qquad \ln h_{it} = \beta + e \ln w_{it}(1-\tau_t) + \beta_I N_{it} + \varepsilon_{it},$$

where I now include person subscripts i to indicate that we have data on a sample of people. Thus h_{it} is hours of work for person i in period t, and so on. Crucial is the addition of the stochastic term ε_{it}, which enables the model to explain heterogeneity in behavior. If (26) is to be interpreted as a labor *supply* relationship, then the ε_{it} must be interpreted as arising from supply shocks (i.e., shocks to person i's tastes for work at time t), perhaps augmented to include optimization error and/or measurement error, but not demand shocks.[16]

It is important that equation (26) controls for nonlabor income, N_{it}. As a result, the coefficient on the log after-tax wage rate (e) is the effect of a wage change holding nonlabor income fixed, which is directly interpretable as the Marshallian elasticity. The coefficient on the nonlabor income variable ($\beta_I = \partial h_{it}/\partial N_{it}$) can be multiplied by the after-tax wage rate to obtain the income effect $ie = w_{it}(1-\tau)\beta_I$. Then, of course, the Hicks elasticity can be backed out using the Slutsky equation as $e_H = e - w_{it}(1-\tau)\beta_I$.[17]

[16] In the utility function approach, the ε_{it} would typically have been obtained by assuming there exists a part of the taste for leisure term β_t in an equation like (3) that is unobserved by the econometrician.

[17] It is important to note that (26) is not consistent with the utility function in (3) because, as we saw in (10), equation (3) gives rise to an income effect that takes on a rather different form.

I will use the static labor supply model in (26) as a guide to discuss the econometric problems that arise in attempting to estimate labor supply elasticities. There are a multitude of problems, but I highlight seven of the most important. These include: (i) endogeneity of wages and nonlabor income arising from correlation with tastes for work, (ii) endogeneity arising from simultaneity, (iii) the treatment of taxes, (iv) measurement error in wages and nonlabor income, and (v) the problem that wages are not observed for nonworkers, and, more generally, the treatment of the participation margin. Additional problems arise from dynamics, including (vi) the treatment of nonlabor income and savings, and (vii) other sources of dynamics like human capital accumulation, fertility, etc.

4.1 *Endogeneity of Wages and Nonlabor Income Arising from Correlation with Tasks for Work*

Perhaps the most obvious problem is endogeneity of wages and nonlabor income arising from possible correlation with tastes for work. For example, people who choose to work long hours (because they have a low taste for leisure) may also work harder and be more productive when they do work. Then ε_{it} would be positively correlated with the wage. Furthermore, those who are relatively hard working might also tend to save more, leading to relatively high asset income, creating a positive correlation between ε_{it} and nonlabor income.

These problems are not merely academic. Pencavel (1986, p. 23) reports a simple ordinary least squares (OLS) regression of annual male hours of work on wage rates, various types of nonlabor income, and a long list of demographic controls (e.g., education, age, marital status, children, race, health, region) using data from the 1980 U.S. census. He finds that the coefficient on asset income is actually *positive*, implying that \$10,000 in additional nonlabor income would *increase* annual hours by 46 hours. This contradicts the assumption that income effects should be negative.[18] He also finds that the coefficient on the wage rate is negative, implying that a dollar per hour wage increase would reduce annual hours by 14. As noted earlier, a negative Marshallian elasticity is theoretically possible, but only due to a large negative income effect. So, prima facie, the sign pattern found here seems to completely contradict economic theory. But it is quite likely the result of endogeneity and/or other econometric problems.

One way to deal with endogeneity arising from correlated tastes is to adopt a fixed effects specification, where the error term is decomposed as

$$\varepsilon_{it} = \mu_i + \eta_{it}. \tag{27}$$

Here μ_i is an individual fixed effect, which captures person i's (time invariant) taste for work, while η_{it} is an idiosyncratic taste shock (e.g., person i may have been sick in a particular period). In the fixed effects approach, it is assumed that the fixed effect μ_i may be correlated with wages and nonlabor income, but that the idiosyncratic shocks η_{it} are not. Methods such as first differencing or demeaning the data can be used to eliminate μ_i from the error term. The η_{it} terms that remain are then assumed exogenous.[19] In addition, labor supply studies typically also include various observable control variables (known as "taste shifters") that might shift tastes for work, such as age, number and ages of children, marital status, etc..

A second approach is instrumental variables (IV). Here one seeks instruments that are correlated with wages and nonlabor income, but uncorrelated with tastes for work (ε_{it}). For example, changes in the world price of iron ore, bauxite, or coal would shift wage rates in Australia, but are presumably uncorrelated with tastes for work. Thus, mineral prices would be sensible instruments for wage rates. In most contexts, however, validity of instruments is controversial. We'll see examples of this when discussing particular papers below.

4.2 *Endogeneity Arising from Simultaneity*

The second problem in estimating an equation like (26), emphasized by Pencavel (1986, p. 59), is whether we are estimating a labor supply curve or demand curve, or just some combination of the two. As an extreme example, say there is a large (uncompensated) tax cut, but labor supply appears to be little affected. This may be because the Marshallian elasticity is close to zero. But it may also be that labor demand is very inelastic (in the short run). Then, any increase in labor supply is quickly choked off by falling wages.

More generally, for any study, the key question we need to address is *why* wages and nonlabor income vary across people/over time. For clarity, I will focus on the problem of wages (assuming nonlabor income is exogenous). A common (although not universal) perspective is that wages are a payment for skill. Each person has a skill level S_{it}, and the equilibrium of the economy determines a rental price on skill (p_t). Thus, the wage rate is given by

$$w_{it} = p_t\, S(X_{it}). \tag{28}$$

[18] A positive income effect for hours, implying a negative income effect for leisure, means leisure is not a normal good (i.e., people do not demand more leisure as they become wealthier). While not theoretically impossible this seems highly unintuitive, and it undermines the standard labor supply model.

[19] A limitation of fixed effects is that η_{it} must be "strictly exogenous," not merely exogenous. That is, η_{it} must be uncorrelated with all leads and lags of wages and nonlabor income, not just contemporary values. This is a very strong assumption. It implies, e.g., that an adverse health shock that lowers tastes for work today cannot affect wages in subsequent periods. Yet, one can easily imagine it would (e.g., if working less in the current period causes human capital to depreciate). Michael P. Keane and David E. Runkle (1992) provide an extensive discussion of this issue.

Here $S(X_{it})$ is a function that maps a set of individual characteristics X_{it} into the skill level S_{it}. X_{it} would include the person's skill endowment, along with education, experience, etc.

Now we modify (26) to include a set of observables Z_{it} that shift tastes for work

$$(29) \quad \ln h_{it} = \beta + e \ln w_{it}(1 - \tau_t) + \beta_I N_{it} + \beta_T Z_{it} + \varepsilon_{it}.$$

One approach to identification of the supply curve in (29) is that there exist some variables in X_{it} that can be plausibly excluded from Z_{it}. Unfortunately, such variables are hard to find.

For example, as we'll see, quite a few authors assume that age and education enter X_{it} but not Z_{it}. Yet it is perfectly plausible that age and education are related to tastes for work, so they belong in Z_{it} as well. Indeed, the profession has had difficulty agreeing on any particular variable or set of variables that could be included in X_{it} and excluded from Z_{it}.

Another approach becomes apparent if we assume that $X_{it} = Z_{it}$. Further assume we can write that $p_t = p(F_t)$, where F_t is a set of exogenous factors that shift the equilibrium skill rental price. Then substitute (28) into (29) to obtain the "reduced form" equation

$$\begin{aligned}(30) \quad \ln h_{it} &= \beta \\ &\quad + e\{\ln p(F_t) + \ln S(Z_{it}) + \ln(1 - \tau_t)\} \\ &\quad + \beta_I N_{it} + \beta_T Z_{it} + \varepsilon_{it} \\ &= \beta + e\{\ln p(F_t) + \ln(1 - \tau_t)\} \\ &\quad + \beta_I N_{it} + \beta_T^* Z_{it} + \varepsilon_{it}.\end{aligned}$$

Here the term $\beta_T^* Z_{it} = e \ln S(Z_{it}) + \beta_T Z_{it}$ subsumes all of the common skill and taste shift variables. As we see from (30), one way to identify the Marshallian elasticity e in the supply equation is to exploit exogenous variation in the skill rental price p_t (induced by elements of F_t that are excluded from Z_{it}) or in the tax rate rate τ_t. What are plausible elements of F_t?

As I alluded to under problem (1), (real) prices of raw materials like oil, iron ore, or coal could serve as "demand side instruments" that shift the skill rental price but are plausibly unrelated to tastes for work. Also, while tax rates τ_t enter (30) directly, they are generally also elements of F_t, as they may shift equilibrium rental prices (see the example at the start of this section). (An exception is if a tax change only affects a very small group of people.)

However, the use of tax rates as instruments is not so straightforward. As I discuss next under problem (3), the marginal tax rates that people actually face may be endogenous given progressive taxation. But the generic tax *rules* that people face are often plausibly exogenous. Thus, one might consider estimating an equation like (29) using raw material prices and/or tax rules as instruments for after-tax wages.[20,21]

4.3 *The Treatment of Taxes*

The third main problem involved in estimation of (26) is that real world tax schedules are not typically the sort of flat rate schedules that I assumed in the theoretical discussion of section 3. The typical "progressive" rate schedule in OECD countries involves transfers to low income individuals, a rate at which these transfers are taxed away as

[20] These issues may apply to nonlabor income as well. Again, a possible approach is to instrument for nonlabor income using the rules that determine transfer benefits. This approach is taken in Raquel Bernal and Keane (2011).

[21] Another possible source of endogeneity is aggregate taste for work shocks. If such aggregate shocks exist, then ε_{it} will not be mean zero in the population. Such shocks will alter equilibrium skill rental prices (by shifting the labor supply curve). This issue is discussed in sections 6.2.5–6.2.6.

income increases, and a set of income brackets, with progressively higher tax rates at higher income levels. We can summarize this by saying the tax rate τ_i that a person faces, as well as their nonlabor income N_{it}, are functions of their wage rate and hours of work. I will denote these functions as $\tau_i(w_{it}, h_{it})$ and $N_{it}(w_{it}, h_{it})$. Then (26) becomes

$$(31)\quad \ln h_{it} = \beta_0 + \beta_w \ln w_{it}(1 - \tau_i(w_{it}, h_{it})) + \beta_I N_{it}(w_{it}, h_{it}) + \varepsilon_{it}.$$

This creates a blatant endogeneity problem, as the after-tax wage rate and nonlabor income depend directly on hours, the dependent variable. For example, for a given pretax wage and nonlabor income, a person with a low taste for leisure (i.e., a high ε_{it}) will work more hours. With progressive taxes, this may drive the person into a higher bracket and/or a lower level of transfers. Hence, progressive taxes generate a negative correlation between the error term ε_{it} and both the after-tax wage and nonlabor income. OLS assumptions are violated, and OLS estimates of (31) are rendered meaningless. (It is worth emphasizing that this endogeneity problem arises by construction even if the pretax wage rate w_{it} is actually exogenous.)

Another problem created by progressive taxation is that tax rates do not usually vary smoothly with income. Rather, they tend to take discrete jumps at certain income levels. An example is given in figure 1, which shows the budget constraint created by a tax system with two brackets. In bracket 1, the tax rate is τ_1, while in bracket 2 it jumps to τ_2. The person in the graph moves into the upper bracket if he/she works more than H_2 hours, at which point his/her income $(wH_2 + N)$ exceeds the cutoff for bracket 2. Then, the slope of the budget constraint drops from $w(1 - \tau_1)$ to $w(1 - \tau_2)$, creating a "kink" where the constraint does not have a well-defined slope. The theory discussed in section 3 is based on the idea that hours are determined by setting the MRS between consumption and leisure equal to the after-tax wage, which is the slope of the budget line. This approach breaks down if there are kinks.

Several methods for dealing with the problems created by piece-wise linear budget constraints (like that in figure 1) have been proposed in the literature, and this has become quite a controversial area. I discuss these methods in detail in section 6.1.

4.4 *Measurement Error in Wages and Nonlabor Income*

A fourth major problem in estimation of labor supply elasticities is measurement error in wages and nonlabor income. There is broad consensus that wages are measured with considerable error in available micro datasets. As is well known, classical measurement error will cause OLS estimates of the coefficient on the wage variable to be biased toward zero, leading to underestimates of labor supply elasticities.

But the measurement error in wages may not be classical. In many studies wage rates are constructed by taking the *ratio* of annual earnings to annual hours. If hours are measured with error this leads to "denominator bias." That is, measurement error induces a negative correlation between measured hours and the ratio wage measure, biasing the wage coefficient in a negative direction. This may in part account for the negative wage coefficient found by Pencavel (1986, p. 23). Many of the studies I describe below use ratio wage measures, including a disproportionate number of studies that obtain small labor supply elasticities.

One way to deal with measurement error is to instrument for after-tax wages. Notice that in discussing estimation of (26) and (29) I already indicated that IV procedures may be necessary to deal with endogeneity problems. In many studies, the use of IV serves

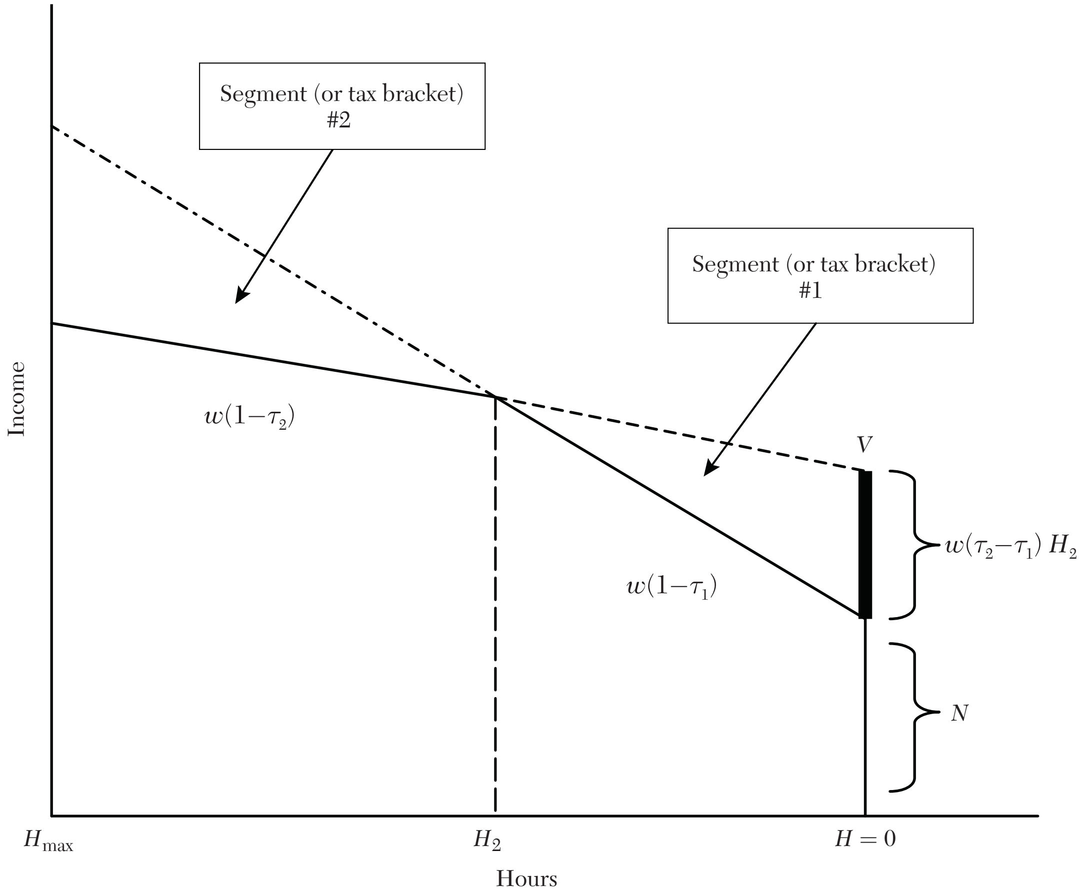

Figure 1. The Piecewise Linear Budget Constraint Created by Progressive Taxation

the dual role of dealing with endogeneity and measurement error. In contrast, in the fully structural approach (sections 6.3 and 7.4) it is necessary to model the measurement error process.

It is likely that error in measuring nonlabor income is more severe than for wages. As we'll see, some methods for modeling labor supply in the presence of taxes require modeling details of workers' budget constraints. Yet knowing the actual constraint a worker faces given modern tax systems is difficult. One key problem is that taxes apply to taxable income, and the typical system offers an array of deductions. In most datasets it is very difficult to know which deductions a worker takes, so deductions are often imputed. Other problems are accounting for all sources of nonlabor income, and measuring fixed costs of work.

4.5 *The Problem that Wages Are Not Observed for Nonworkers*

The fifth main problem involved in estimation of (26) is that wages are not observed for people who choose not to work. This leads to the well-known problem of “selection bias” if we attempt to estimate (26) using data on workers alone. Assume that, ceteris paribus, the probability of working is increasing in the wage. Then, people we see working despite relatively low wages will be those with relatively high tastes for work. This induces a negative correlation between w_{it} and ε_{it} among the subpopulation of workers, even if w_{it} is exogenous in the population as a whole.

Pioneering work by Heckman (1974) began a large literature on methods to deal with the selection problem. Unfortunately, there is no solution that does not involve making strong assumptions about how people select into employment. This means that empirical results based on these methods are necessarily subject to some controversy.[22]

In the literature on male labor supply, it is common to ignore selection on the grounds that the large majority of adult non-retired men do participate in the labor market, so selection can be safely ignored. Whether selection is really innocuous is unclear, but this view is adopted in almost all papers on males that I review. In contrast, dealing with selection by modeling participation decisions is central to the modern literature on female labor supply.

4.6 *The Treatment of Nonlabor Income and Savings*

I now turn to issues that become apparent when one views labor supply in a dynamic context. The sixth main problem in estimation of (26) is interpretation of the nonlabor income variable. In the static model, current nonlabor income is treated as a measure of wealth. However, much of nonlabor income is asset income, and asset levels are, of course, driven by life-cycle consumption and savings patterns. Specifically, we expect assets to follow an inverted U-shaped path over the life cycle: low when people are young, have low incomes, and borrow to finance consumption, high in the middle of the life cycle as they build up assets for retirement, and then declining in retirement. This means a person’s asset level at a particular *point in time* is not a good indicator of their lifetime wealth.

For example, a 40 year old with a high level of skills who has gone rather heavily into debt to set up a household and pay for his children’s education may in reality be wealthier (in a life-cycle sense) than a 60-year who has positive savings but at a level that is inadequate to fund retirement. The income effect creates a greater inducement to supply labor for the latter than the former, despite the fact that latter person has a higher level of current assets.

One approach to this problem, pursued by MaCurdy (1983) and Richard Blundell and Ian Walker (1986), is to estimate versions of (26) where the nonlabor income variable is redefined in a manner that is consistent with life-cycle asset allocation. An alternative, due to MaCurdy (1981), is to estimate a labor supply function like (25) that is explicitly derived from a life-cycle model. I discuss these approaches in section 6.2 and 7.2. A third approach is full structural estimation of the life-cycle model, which I discuss in section 6.3 and 7.4.

4.7 *Other Sources of Dynamics*

Several other issues arise from viewing labor supply in a life-cycle context. One is the issue of human capital. If work experience builds human capital, then the current

[22] Recent years have seen the development of semiparametric selection methods that make weak assumptions about the functional form of the selection rule. But these methods typically require strong assumptions on the “excluded instruments” (i.e., variables that enter the work decision rule but do not affect wages directly). Aside from the exclusion itself, which is often controversial, common requirements are availability of an instrument that can drive the employment probability close to one, or the assumption that the instrument can be manipulated to keep the selection probability fixed as determinants of wages vary (an index assumption).

labor supply decision affects future wages. As I discuss in section 6.3, this has fundamental implications for the estimation of labor supply elasticities. Other sources of dynamics are state dependence in tastes for work, and, of particular importance for women, joint decision making about labor supply, fertility, and marriage. This is the focus of section 7.4.

5. *A Roadmap to the Empirical Literature*

There have been several major surveys of the labor supply literature, including Mark R. Killingsworth (1983), Jerry A. Hausman (1985b), Pencavel (1986), Killingsworth and Heckman (1986), Blundell and MaCurdy (1999), and Costas Meghir and David Phillips (2010). They typically sort results by demographic group and/or econometric models employed. I follow the same approach: the literature on men is covered in section 6, while that on women is covered in section 7. Different empirical approaches are then covered in subsections.

5.1 *The Male Labor Supply Literature*

As best as I can determine, the male labor supply literature using micro data begins with Marvin Kosters (1969), but he fails to address the key issues I raised in section 4. Subsequent papers by Orley Ashenfelter and Heckman (1973) and Michael J. Boskin (1973) used IV techniques to deal with measurement error. They found very small (even negative) labor supply elasticities.

There followed a series of papers that attempt to deal with progressive taxation and piecewise-linear budget constraints. I discuss these in section 6.1.1. Following early work by Robert E. Hall (1973), more sophisticated methods were developed by T. J. Wales and A. D. Woodland (1979) and Hausman (1981). These papers, and similar papers that followed, generally found large Hicks elasticities for males. As Hausman (1981) pointed out, this implied large efficiency losses from progressive taxation. And for a (brief) time, a consensus emerged that the small labor supply elasticities of earlier studies were an artifact of failure to account for taxes.

The NIT experiments in the United States spurred the development, by Gary Burtless and Hausman (1978), of a method to deal with piecewise-linear *nonconvex* budget constraints. This work was also novel in that it relied for identification on experimentally generated variation in budget constraints (albeit with flaws in the experimental design). I discuss this work in section 6.1.2. The consensus from work on the NIT experiments was that labor supply elasticities for low-income workers were quite modest.

Section 6.1.3 provides a brief summary of the literature before 1990. At that time the piecewise-linear approach had become dominant. But that changed as a result of the so-called "Hausman–MaCurdy controversy," which I describe in detail in section 6.1.4. The controversy began when MaCurdy, David Green, and Harry J. Paarsch (1990) argued that Hausman's (1981) approach to dealing with progressive taxation imposed a large Hicks elasticity a priori. They proposed a different method, based on a smooth approximation to the piecewise-linear constraint.[23] Using similar data, they found much smaller elasticities than Hausman (1981).

However, Eklöf and Hans Sacklén (2000) provide convincing evidence that differences in results between MaCurdy, Green, and Paarsch (1990) and Hausman (1981) (and a number of other studies) are not mostly due to differences in econometric method. Rather, whether a study finds relatively large labor supply elasticities is better predicted by whether it deals with denominator bias in wage measures. As dealing with denominator bias is presumably a good thing, I take this

[23] Alternatively, Soren Blomquist, Matias Eklöf, and Whitney Newey (2001) model labor supply nonparametrically (section 6.1.4.2).

as supporting a view that elasticities are relatively large.

Finally, in section 6.1.5, I describe a paper by Arthur van Soest, Isolde Woittiez, and Arie Kapteyn (1990) that attempts to explain why we rarely observe people working a small number of hours. They introduce a demand side constraint on possible hours choices. Other approaches to this problem are to introduce fixed costs of work or to assume hourly wages are lower for part-time work. There is more focus on this issue in the literature on females (see below).

In section 6.2, I consider the basic life-cycle model with saving. Two unsatisfactory features of the static model are (i) current labor supply only depends on the current wage, not how it compares to wages in other periods, and (ii) current nonlabor income may be a poor proxy for lifetime wealth. Even if the current wage is high and nonlabor income is low, there may be little incentive to work if one expects wages to be much higher in the future. But this is only true if one can borrow/lend across periods. Thus, in the male labor supply literature, the first and most common extension of the static model is to allow for borrowing/saving.

In section 6.2.1, I describe two simple approaches to estimating the life-cycle model with saving developed in MaCurdy (1983). Both use consumption to proxy lifetime wealth. The first, which I call the "MRS method," involves estimating MRS conditions like (22). The second is known as the "life-cycle consistent" method. This involves simply replacing nonlabor income with consumption minus current after-tax earnings in a conventional static labor supply model. MaCurdy (1983) found very large labor supply elasticities using both methods. But subsequent work by Joseph G. Altonji (1986) and Blundell and Walker (1986) failed to confirm this. MaCurdy conceded his results may have come from using a small nonrepresentative sample.

Section 6.2.2. Blomquist (1985) noted that the two methods developed in MaCurdy (1983) need to be modified to deal with progressive taxation. James P. Ziliak and Thomas J. Kniesner (1999) implement this change, and find large efficiency costs of progressive taxation.

In section 6.2.3, I discuss a paper by Ziliak and Kniesner (2005) that allows for nonseparability of leisure and consumption in the life-cycle model. I describe how this can have important implications for the behavior of the model. They again find large efficiency costs of progressive taxation (although here they do not adopt the Blomquist 1985 approach).

In section 6.2.4, I describe a third approach to estimating the life-cycle model, due to MaCurdy (1981). This approach involves estimating hours growth equations like (25). It only uncovers a subset of utility function parameters, and it is used specifically to estimate the Frisch elasticity. Using this approach, MaCurdy (1981) obtained a very small but also very imprecise estimate of the Frisch elasticity.

There followed a large literature seeking to improve upon MaCurdy's (1981) method. The focus was on better instruments for wage growth, ways to reduce measurement error, or ways to measure expected wage growth. I discuss the key papers in sections 6.2.5 to 6.2.7. Most continue to obtain small (imprecise) estimates of the Frisch elasticity, although Joshua D. Angrist (1991) and Luigi Pistaferri (2003) obtain precise estimates as large as 0.63 to 0.70. Notably, it is common in this literature to ignore taxes. In section 6.2.8, I discuss Daniel Aaronson and Eric French (2009), who build progressive taxation and tied wage–hours offers into the model.

In section 6.3, I consider models that incorporate both saving and human capital in the life-cycle framework. Key papers here are Heckman (1976), Kathryn L. Shaw

(1989), Keane and Kenneth I. Wolpin (2001), and Susumu Imai and Keane (2004). Imai and Keane argue that the failure to account for human capital probably led prior work to severely understate the intertemporal elasticity of substitution. I discuss implications of their model for elasticities with respect to permanent tax changes as well. Finally, section 6.4 summarizes the male labor supply literature.

5.2 *The Female Labor Supply Literature*

In section 7, I turn to the literature on female labor supply. Mincer (1962) argued that in modeling female labor supply it is crucial to consider the allocation of labor supply, and in particular labor force participation, over the life cycle. I agree with this view, so I start in section 7.1 with the attempt by Heckman and MaCurdy (1980, 1982) to extend the basic life-cycle model of MaCurdy (1981) to incorporate participation decisions.

Section 7.1.1 describes work that extends the life-cycle model to incorporate fixed costs of work. The key paper here is Jean Kimmel and Kniesner (1998). Section 7.1.2 describes work by Sumru Altug and Robert A. Miller (1998) that extends the model to also include human capital. When they are successful in obtaining estimates, papers in section 7.1 obtain Frisch elasticities of 2.35 to 3.05, with most of the action on the participation margin. These are much higher than values obtained for men. Note that none of the papers in 7.1 account for taxes.

In section 7.2, I turn to the "life-cycle consistent" approach. The key paper here is Blundell, Alan Duncan, and Meghir (1998). The main motivation of their paper is to use a series of British tax law changes to help identify the model. They estimate a Hicks elasticity for employed women of 0.20. The participation elasticity is not estimated.

Section 7.3 describes work by Robert Moffitt (1984) that estimates a life-cycle model with both fertility and human capital. Instead of doing fully structural estimation, he estimates what I call an "approximate reduced form" of a very complicated structural model (which, given 1984 computer technology, would have been infeasible to fully solve and estimate).

In section 7.4, I describe papers that implement full-solution structural estimation of female labor supply models. This starts, in section 7.4.1, with the paper by Zvi Eckstein and Wolpin (1989) in which participation is the only decision and work experience builds human capital. I then show, in sections 7.4.2 and 7.4.3, how the literature sought to endogenize fertility, marriage, schooling, and welfare participation, culminating in the papers by Keane and Wolpin (2007, 2010) where all these variables are treated as choices. The consistent finding of these structural papers is that labor supply elasticities are quite large for women. A limitation of all the papers in sections 7.3–7.4 is that they do not explicitly account for taxes.

Then, in section 7.5, I discuss the nonstructural approach of Nada Eissa (1995, 1996a), who analyzed the effects of the major U.S. tax reforms of 1981 and 1986 on labor supply of married women. Finally, section 7.6 summarizes the female labor supply literature.

6. *A Survey of the Male Labor Supply Literature*

In this section, I focus on labor supply of men, and consider results from static models, life-cycle models with savings, and life-cycle models with both savings and human capital. I will discuss the literature on women in section 7.

As should be clear from section 4, there are many econometric problems to confront when estimating labor supply elasticities. And, as we will see below, there are many alternative approaches to dealing with these problems. Unfortunately, no consensus has emerged on a "correct" approach. Indeed, the controversy between advocates of alternative approaches has often been rather intense, as will at times become clear in what follows.

6.1 *A Summary of Results from Static Labor Supply Models*

Pencavel (1986) notes that the first labor supply function estimation using individual (as opposed to aggregate) level data was by Kosters (1969). He considered employed married men, aged 50–64, from the 1960 U.S. Census. Estimating an equation with log hours as the dependent variable and logs of wages and nonlabor income as independent variables (along with various controls for tastes), he obtained a Marshallian elasticity of –0.09 (i.e., backward bending labor supply) and a small (negative) income effect (–0.14). However, this early study ignored endogeneity, taxes, and essentially all the key problems listed in section 4.

A number of subsequent studies tried to instrument for the wage and/or nonlabor income to deal with measurement error. But these studies generally continued to obtain small negative Marshallian elasticities for married men. For instance, Ashenfelter and Heckman (1973) instrument for nonlabor income and obtain $e = -0.15$, $ie = -0.27$ and a Hicks elasticity of 0.12. Boskin (1973) instruments for the wage and obtains a Marshallian elasticity of –0.07, an income effect of –0.17 and a Hicks elasticity of 0.10. These studies ignored taxes.

6.1.1 *Attempts to Deal with Progressive Taxation (Piecewise-Linear Budget Constraints)*

6.1.1.1 *Studies Based on U.S. Data*

Hall (1973) developed a method to deal with the piecewise linear budget constraints created by progressive taxation. The idea, illustrated in figure 1, is to model each person *as if* they choose labor supply subject to a *hypothetical* linear budget constraint created by taking the segment on which they are observed to locate, and extending it from $h = 0$ to $h = H_{\max}$. In figure 1, these extensions of segments 1 and 2 are indicated by the dotted lines. For example, the hypothetical budget constraint for a person on segment 2 is characterized by the slope $w(1 - \tau_2)$ and the "virtual" nonlabor income level $V_2 = N + w(\tau_2 - \tau_1)H_2$. As noted by Hall (1973), if preferences are strictly convex (as implied by diminishing marginal returns to consumption and leisure) a person facing such a hypothetical budget constraint would make the same choice as when facing the actual budget constraint.[24]

Pencavel (1986) excluded Hall's results from his extensive survey because "many different estimates are presented and I gave up the attempt to summarize them adequately." But Hall's figures 3.5 and 3.6 seem to provide a concise summary of the results. His sample consisted of all men and women from the 1967 U.S. Survey of Economic Opportunity (SEO), an augmented version of the CPS that included better wage and hours measures and an over-sample of the low income population. Hall's figures 3.5 and 3.6 present labor supply curves averaged across various demographic groups. Figure 3.6 shows backward bending labor supply above an after-tax wage rate of about \$2.00 per hour. But figure 3.5 shows a Hicks elasticity at 2,000 hours of approximately 0.45.[25] Thus, my interpretation is that Hall's results imply backward bending labor supply but a strong income effect and a large Hicks elasticity.

[24] It is common in applying this method to instrument for wages and nonlabor income to deal with measurement error. Hall (1973) does this as well.

[25] The graph of the compensated labor supply function that Hall (1973) presents in figure 3.5 is rather flat over a very wide range. This is not true of the uncompensated graph.

It should be noted that Hall's (1973) approach does not deal with the endogeneity of after-tax wages and nonlabor income created by the choice of segment. If tastes for work are stochastic, as in (26), then the segment where one chooses to locate is determined not only by ones wage rate and nonlabor income, but also by the taste shock ε_{it}. By taking the segment where a person chooses to locate as *given*, we are in effect truncating the range of the taste shock—e.g., people who locate on a high hours segment will tend to be those with high tastes for work. As I noted earlier, this induces a negative correlation between the after-tax wage and tastes for work, which tends to bias labor supply elasticities in a negative direction.

The papers by Burtless and Hausman (1978) and Wales and Woodland (1979) were the first to model choice of segment. Thus, in estimating labor supply elasticities, they account for the correlation between taste shocks and after-tax wages. Other well-known examples of this "structural approach," in which one models in detail how people make labor supply decisions subject to a nonlinear tax schedule, include Hausman (1980, 1981), Blomquist (1983) and Moffitt (1983). The basic idea of the structural approach is illustrated by the following simple example: Assume a two-segment budget line as in figure 1. Further assume that person i locates on budget segment 2, with slope $w_i(1-\tau_2)$ and virtual income V_i. The person has to work at least H_2 hours to be on this segment. Now, to keep things simple (and highlight the key idea) I will assume that we know a person's segment without error, but that hours are measured with error. This could occur, for example, if we had access to tax records to determine a person's earnings and tax bracket, but had to rely on survey data to measure hours. Assume the person's choice of hours is determined by the equation

$$(32) \quad \ln h_i = \beta_0 + e \ln w_i(1-\tau_i) + \beta_I V_i + \varepsilon_i.$$

Here, in contrast to equation (26), labor supply is expressed as a function of V_i, the level of virtual income for the segment on which person i locates. Observed hours are given by

$$(33) \quad \ln h_i^o = \ln h_i + v_i = \ln h_i^* + \varepsilon_i + \nu_i,$$

where ν_i is measurement error and $\ln h_i^*$ is the predicted level of log hours from (32). Then the likelihood contribution for person i is given by

$$(34) \quad P(\ln h_i^o, \ln h_i > \ln H_2) = P(\varepsilon_i + v_i, \varepsilon_i + \ln h_i^* > \ln H_2) = f(\varepsilon_i + v_i \mid \varepsilon_i > \ln H_2 - \ln h_i^*) \times P(\varepsilon_i > \ln H_2 - \ln h_i^*).$$

Note that the density of the error in the hours equation is now conditioned on the event $\varepsilon_i > \ln H_2 - \ln h_i^*$, accounting for the fact that those who locate on segment 2 have relatively high taste shocks. Conversely, the analogous term to (34) for people who locate on segment 1 includes the truncation $\varepsilon_i < \ln H_2 - \ln h_i^*$. Thus, building the likelihood based on terms like (34) will give an estimator that is consistent in the presence of endogenous segment location.

For males, the first study to model the full complexity of the budget constraint created by *progressive* taxation, and model men as choosing labor supply subject to this constraint, was Wales and Woodland (1979). They assume wages and nonlabor income are measured without error, and preferences are homogeneous in the population. Then, the econometrician can determine a worker's optimal hours level (and true budget segment), conditional on the model parameters. Deviations between observed and predicted hours are explained by measurement error. The Wales and Woodland estimates, obtained using married men from the Panel Study of Income Dynamics (PSID), are quite different from the earlier literature. They obtain a Marshallian elasticity of 0.14 (finally positive!), a large income effect of –0.70, and a Hicks elasticity of 0.84.

Hausman (1981) extended the Wales and Woodland (1979) approach to include taste heterogeneity. Specifically, he let β_I in (32) be random in the population (while interpreting ε as measurement error). With heterogeneity in β_I, the worker's choice of segment is no longer deterministic. Conditional on the latent β_I, the likelihood contribution for a worker is the density of the measurement error that reconciles observed and predicted hours. However, to form the likelihood, one must now integrate out the latent β_I from this density. This means integrating over all possible segments and kink points, weighting each by the probability β_I is in the appropriate range such that a worker would choose it. This obviously makes estimation much more difficult than when the true segment is known (as in Wales and Woodland 1979 or my simple example in (32)–(34)). For estimation, Hausman (1981) also used married men in the PSID. He obtained a Marshallian elasticity of close to 0 and an income effect of –0.77.

An important point, stressed by Hausman (1981), is that, even with a small (or zero) Marshallian elasticity, large Hicks elasticities of the type estimated by Hall (1973), Wales and Woodland (1979), and Hausman (1981) imply large negative labor supply effects of progressive taxation (as well as large efficiency losses). To understand why, consider again figure 1. A person in bracket #1 has an after tax wage rate of $w(1-\tau_1)$ and nonlabor income of N. If this person increases his/her hours above level H_2, so that he/she earns enough to be in bracket #2, then not only does his/her marginal wage fall to $w(1-\tau_2)$, but, in addition, the level of "virtual" nonlabor income that is relevant for his/her decision making increases to $V = N + wH_2(\tau_2 - \tau_1)$. Thus, even if the Marshallian elasticity is close to zero, a large income effect (or, equivalently, a large Hicks elasticity) can have a strong negative effect on labor supply by discouraging workers from increasing hours above H_2.

Furthermore, following MaCurdy (1992), one can show, to a good approximation, that the Hicks elasticity determines the labor supply response of tax payers already in the higher brackets. Suppose the tax rate on segment #2 is increased from τ_2 to $(\tau_2 + \Delta)$. This causes the after-tax wage to fall by Δw and virtual nonlabor income to increase by $\Delta w H_2$. To keep things simple, assume a simple linear labor supply function (as in Hausman 1981):

$$(35)\quad h = \beta + \beta_w w(1-\tau_2) + \beta_I V_2 + \varepsilon,$$

where τ_2 and V_2 are the tax rate and virtual income on segment #2, respectively. Plugging in the new values for the tax rate and virtual income we get

$$(35')\quad h' = \beta + \beta_w w(1-\tau_2-\Delta) + \beta_I(V_2 + \Delta w\, H_2) + \varepsilon.$$

Thus, we have that $h' - h = -\beta_w w\Delta + \beta_I w\Delta H_2 = -(\Delta w)(\beta_w - H_2\beta_I)$. That is, the change in the after-tax wage $(-\Delta w)$ is multiplied by $(\beta_w - H_2\beta_I)$. If we multiply this quantity by w/H_2 we get precisely the Hicks elasticity from equation (8), for the linear model (35), evaluated at hours level H_2.[26]

[26] In the linear model, $\beta_w = \partial h/\partial w$ is the uncompensated wage effect and $h \cdot \beta_I = h \cdot \partial h/\partial N$ is the income effect.

Thus, we see that, to a good approximation, the Hicks elasticity determines the labor supply response of taxpayers in the higher brackets.[27] Hence, a large Hicks elasticity implies large efficiency loses from progressive taxation.[28] Given their findings of substantial Hicks elasticities, Hall and Hausman became strong advocates for a flat rate tax.

Pencavel's (1986) classic survey of male labor supply emphasized that the income effect, or "marginal propensity to earn," could, in the static model, also be calculated from consumption data, by looking at how consumption/earnings respond to changes in nonlabor income. In fact, Angus Deaton (1982) did this—using the U.K. Family Expenditure Survey (FES) of 1973, he obtained an estimate of *ie* near zero (i.e., *wh* is hardly affected by an increase in nonlabor income). Based on this result, Pencavel (1986) concludes estimates of the income effect that differ much from zero are suspect. He goes on to discount results of several studies that obtain large income effects, such as Wales and Woodland (1979) and Hausman (1981).

In my view, this conclusion goes too far. The Deaton (1982) result is hard to interpret as a causal effect of nonlabor income on consumption, as nonlabor income is likely to be endogenous in a consumption equation. And in a life-cycle model, high nonlabor income may simply indicate high permanent income, causing it to be highly positively correlated with consumption. Furthermore, there is substantial evidence that people mostly save the proceeds from temporary tax rebates.[29] As indicated earlier, introspection may suggest that very large effects of N on wh (that is, values of *ie* very near –1) are implausible, but I would not conclude based on Deaton (1982) that only effects near zero are plausible.

6.1.1.2 *Studies Based on Non-U.S. Data*

Until now, I have discussed only labor supply studies based on U.S. data. As Pencavel (1986) notes, the British literature took a somewhat different tack for two reasons. First, it has always focused on effects of taxation, so wages and nonlabor income are always treated as after-tax. Second, it is largely based on the FES, which contains both hours and consumption data. Thus, it has usually estimated labor supply and consumption functions jointly.[30] The eight British studies Pencavel cites all find small negative Marshallian elasticities (with a mean of –0.16), income effects in the –0.04 to –0.50 range (with a mean of –0.29), and Hicks elasticities ranging from 0.30 to slightly less than 0 (with a mean of 0.13). Note, however, that these papers do not adopt the piecewise-linear approach.

The influential paper by Blomquist (1983) used the piecewise-linear method to study labor supply in Sweden in 1973. The country had a highly progressive tax structure at that time. Blomquist studied married men from the Level of Living Survey who were of prime working age (i.e., 25–55 years old). His estimates implied a Marshallian elasticity

[27] A limitation of this analysis is that it does not consider the behavior of taxpayers who are initially located at budget constraint kink points prior to the tax change. These people are unaffected by small tax changes.

[28] In fact, Hausman (1981) found that the efficiency loss from progressive taxation was 22 percent of tax revenues. He found that a shift to a flat rate tax would reduce this to only 7 percent.

[29] Note that a one-for-one increase in consumption in response to an increase in nonlabor income, if interpreted causally, is wildly at variance with the life-cycle model—i.e., only *unanticipated* changes in nonlabor income should alter consumption at all. And even an unanticipated change would be smoothed out over the whole life cycle, and therefore would have little effect in any one period. Only an unanticipated change in nonlabor income that is also expected to be highly persistent should have much impact on current consumption.

[30] Of course, given a utility function defined over both leisure (or hours) and consumption, as in (3), along with a budget constraint, one can derive both labor supply and consumption functions.

of 0.08 and an income effect of $ie = -0.03$ at mean values in the data. The implied Hicks elasticity is 0.11.

Blomquist (1983) stressed the key point that, in nonlinear budget constraint models, labor supply elasticities cannot (in general) tell us how people will respond to changes in the constraint. Hence, he used his model to simulate the consequence of Sweden switching from the highly progressive tax regime in place in 1973 to a flat tax, a lump sum tax, and a no tax regime. Under the (existing) progressive income tax, the model predicts average annual hours of work of 2,143 hours (close to the sample average). It predicts that complete elimination of taxes would increase annual hours of work from 2,143 to 2,443, a 14 percent increase. Blomquist also calculates that a 34 percent flat rate tax would raise the same revenue as the progressive tax. Given a flat rate tax, average annual hours would be 2,297 hours, a 7.2 percent increase.[31]

Comparing the proportional and no tax worlds, Blomquist finds a 34 percent tax increase (wage reduction) leads to a 6 percent reduction in hours. The implied Marshallian elasticity is roughly $6/34 = 0.18$. This is quite a bit larger than the Marshallian elasticity of 0.08 implied by the estimates at the mean values of after-tax wages and hours in the data.[32] This illustrates how elasticities calculated assuming linear budget constraints can be quite misleading in a piecewise-linear context. It may also indicate that mean values of elasticities can be quite misleading with regard to population responses in models with heterogeneous workers.[33]

The compensating variation (CV) is the lump sum payment needed to make a person in a progressive or flat-tax world equally well off as a person in a no-tax world. For the flat rate tax it is 16,417 SEK while for the progressive tax it is 18,059 SEK. This compares to 16,103 SEK in revenue per person (under either tax). One way to measure deadweight loss from a tax is the CV as a percent of revenue. This gives $(18{,}059 - 16{,}103)/16103 = 12$ percent for the progressive tax and 2 percent for the flat tax. Thus, the implied efficiency losses for the progressive tax system are rather large. This is despite the quite modest estimates of the Marshallian and Hicks elasticities at the mean of the data (0.08 and 0.11 respectively).

Blomquist and U. Hansson-Brusewitz (1990) performed a similar analysis on Swedish data from 1980 on 602 married men aged 25–55. An innovation is use of an hours equation with a quadratic in wages. This provides a significantly better fit than a linear specification, but has little impact on the main results. The authors obtain modest positive Marshallian elasticities (0.12 to 0.13 in their preferred models) and income effects of only about –0.005.

A novel feature of Blomquist and Hansson-Brusewitz (1990) is that they plot both the "structural" labor supply equation that would obtain if people maximized utility subject to a linear budget constraint, and whose parameters can be used to infer the underlying utility function, and the "mongrel" or "reduced form" equation that gives desired hours as a function of wages, nonlabor income and the *existing tax structure*.

[31] It is important to note that this is a partial equilibrium analysis. As both of these experiments lead to substantial increases in labor supply, they would presumably also lead to a reduction of wages in equilibrium.

[32] Of course, for such a large change, the direction in which we do the calculation matters. Going from the proportional tax world to the no tax world, hours increase 6.4 percent while wages increase 52 percent, so the implied elasticity is $6.4/52 = 0.12$. This is still 50 percent greater than Blomquist's 0.08 estimate at mean values.

[33] It is also interesting to compare a no tax world to lump sum tax world. Blomquist simulates that a 16,103 SEK lump sum tax would increase hours from 2,443 to 2,506, or 2.6 percent. His nonlabor income coefficient of –0.0042 (per thousand) implies an increase in hours of $(0.0042)(16{,}103) = 68$ hours, which is quite close.

This reduced form equation varies as the tax system varies. Strikingly, although the "structural" labor supply curve is linear with a positive Marshallian elasticity throughout, the reduced form supply curve is backward bending for wage rates above 26 SEK per hour. This compares to an average gross wage rate of 41.75 SEK and an average after-tax rate of only 14.83 SEK. Thus, a reduced form analysis that fails to account for progressive taxation could easily conclude labor supply is backward bending when this is only a feature induced by the tax system, not by underlying preferences.

Finally, when Blomquist and Hansson-Brusewitz (1990) simulate the consequence of shifting to a flat rate tax (which must be 37 percent to generate equivalent revenue) they find that the efficiency loss from taxation falls from 16 percent to 5 percent of revenue collected, while annual hours of work increase from 2099 to 2238 (or 6.7 percent). They also simulate a cut in the national tax rate in the top several brackets by 5 percentage points, from a range of 44 percent to 58 percent to a range of 39 percent to 53 percent. They simulate that this would increase labor supply by 0.4 percent while actually increasing tax revenue by 0.6 percent. This implies that the upper bracket tax rates in Sweden in 1980 actually exceeded the revenue maximizing rates (see equation (1)).[34]

6.1.2 *The NIT Experiments (Nonconvex Budget Constraints)*

A good deal of work on labor supply was stimulated by the NIT experiments conducted in several U.S. cities beginning in 1968. The NIT experiments were intended to have treatment and control groups. Members of the treatment groups received a grant level G that was taxed away, at a fairly high rate, as they earned income. Thus, G serves as the guaranteed minimum income for a person with no earnings or nonlabor income. At a certain income level a person reaches the "break-even point" where G is totally taxed away. Then, they revert to the normal income tax rate, with is typically less than the benefit tax rate. This creates a *nonconvex* budget constraint, since tax rates *fall* as income rises.

Figure 2 illustrates the shape of a typical nonconvex budget constraint created by the NIT or other types of welfare programs. The budget constraint connects points a, b, c, and e. The figure has been drawn so a person who works zero hours receives G. If they begin to work their income drops (from a to b), due to fixed costs of working (FC). I have drawn an example where, as the person works more hours, the grant G is taxed away at a 100 percent rate as earnings increase. This is represented by the flat dotted line from point b to point c. The tax rate in the NIT program was only 40 percent or 60 percent, but it has not been uncommon for other types of welfare programs to have rates up to 100 percent. A good example is the Aid to Families with Dependent Children (AFDC) program in the United States. Finally, point c is the breakeven point. Above that the person is off the program and faces the regular income tax schedule.

Unfortunately, people in the NIT experiments were not actually assigned randomly to the "treatment" and "control" groups, and there is a substantial literature on why this was the case. Nevertheless, the NIT experiments generated useful variation in budget constraints across workers that can be used to help estimate labor supply elasticities.

A very well known analysis of the NIT experiments was by Burtless and Hausman (1978). The approach is similar to the Wales and Woodland (1979) and Hausman (1981)

[34] Note that Sweden had an array of payroll, value added and local taxes that brought overall rates to well above the 58 percent top bracket national rate. In 1980, the upper limit for the sum of national and local rates was set at 85 percent.

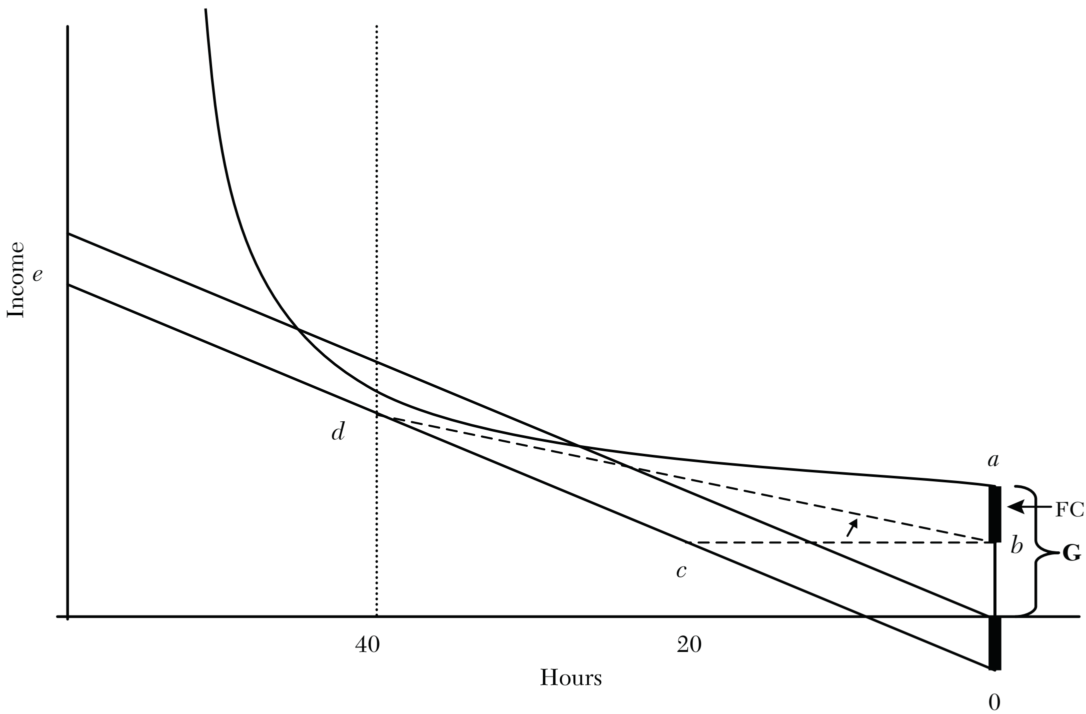

Figure 2. The Nonconvex Budget Constraint Created by NIT or AFDC type Programs

Note: The budget constraint created by the program goes through points *a*, *b*, *c*, *e*. It is generated by the program grant level (G), the fixed cost of working (FC) and the program tax rate, which render the constraint nonconvex. The straight line through the origin is the after-tax-wage line that would be the budget constraint in the hypothetical situation of a flat rate tax. The dotted line from *b* to *d* shows the shift in the budget constraint when the program tax rate on earnings is reduced to 50 percent.

studies mentioned earlier. That is, the authors model how men choose labor supply subject to the complex nonlinear budget constraint created by the NIT, including the choice of which segment to locate on. But, while the previously mentioned studies dealt with the convex budget constraints created by progressive taxation, the Burtless–Hausman study was the first to deal with the *nonconvex* budget constraint created by a typical transfer program.

It is important to note that Hall's (1973) simplifying idea (i.e., that a person's hours choice would be unchanged if he/she had faced a hypothetical linear budget constraint through the observed hours point) does not work in this case. Hall's idea, which allows one to work with linear labor supply functions provided one appropriately defines "virtual" nonlabor income, requires a *convex* budget constraint.[35] But as Burtless and Hausman (1978) discuss, given a *nonconvex* budget constraint, one must specify the

[35] As an obvious example, the person whose indifference curve is drawn in figure 2 would choose point *a*, but he would make a different choice if he faced a linear budget constraint through point *a*.

utility function in order to model labor supply decisions. Given the utility function, one can assess the maximized utility level on each segment and kink point of the constraint, and determine the utility maximizing kink or segment (and utility maximizing hours on that segment).

Still, Burtless and Hausman (1978) argued that, for consistency with prior literature, it is more intuitive to specify a familiar hours equation and work back (using Roy's identity) to the implied utility function.[36] Burtless and Hausman choose to use a double log specification:

$$(36)\quad \ln h_i = \beta + e \ln w_i(1 - \tau_i(w_i, h_i)) + e_I N_i(w_i, h_i) + \tilde{\varepsilon}_i.$$

Here e and e_I *would be* the Marshallian and the income elasticities in the hypothetical case of a person facing a linear budget constraint. Equation (36) implies the indirect utility function:

$$v(w, N) = \exp(\beta_i) \frac{w^{1+e}}{1+e} + \frac{N^{1+e_I}}{1+e_I}.$$

Burtless and Hausman introduce a stochastic element by letting $\exp(\beta_i) = X_i \beta_T + \varepsilon_i$ where X_i and ε_i are observed and unobserved taste shifters, respectively, and letting $e_I \sim TN(\mu, \sigma_{e_I}^2)$ with a truncation from above at zero. This restricts the *hypothetical* income elasticity, in the linear budget constraint case, to be negative. It is worth emphasizing, however, that given the nonconvex budget set, the estimates of e and e_I will not tell us anything about how a person would respond to particular changes in the tax structure. In a model of this type, that would require simulating the person's optimal behavior under the new regime.[37]

The implications of this point are far reaching and worth emphasizing. In particular, given nonconvexities and piecewise linear budget constraints, utility function parameters are no longer tightly linked with any particular elasticity concept. Thus, labor supply may appear to be "elastic" or "inelastic," depending on the type of budget constraint shift one considers.

This point is illustrated in figure 2. The budget constraint goes through a, b, c, e, and I have drawn the indifference curve so utility is maximized at a, where $h = 0$. Now consider the response of a person with such preferences to various changes in the budget constraint:

First, consider the program's tax rate on earnings (or "benefit reduction rate"). The dotted line from b to d represents how the budget constraint shifts if the tax rate on earnings is reduced from 100 percent to 50 percent. As we see, this has no effect whatsoever on hours of work.

In contrast, a small increase in the worker's market wage rate would cause him/her to jump from 0 to 40 hours of work per week (by slightly raising point d). This is true whether the program tax rate is 100 percent or 50 percent. Similarly, reductions in the grant level or in the fixed costs of working would have large effects.

Thus, given data that contained wide historical variation in program tax rates, a researcher studying a program like that depicted in figure 2 might well conclude labor supply is inelastic, so it would be very difficult to induce members of the target

[36] Today, I suspect many economists would be more comfortable specifying utility functions.

[37] This point was emphasized by all the authors who pioneered this literature. For instance, Blomquist (1983) states: "A change in the gross wage rate, nonlabor income, or parameters of the tax system changes the whole form of the budget set . . . the elasticities presented above should therefore *not* be used to calculate [their] effects . . . " (emphasis added) (p. 186).

population to work. Historically, this is roughly what happened with the AFDC program in the United States. Years of tinkering with the AFDC tax rate in attempts to create work incentives had little effect, leading to a conventional wisdom that labor supply was "inelastic" for single mothers.

Thus, most of the economics profession was taken by surprise when changes in policy in the mid-1990s, including wage subsidies (EITC) and child care subsidies (CCDF), as well as a strong macroeconomy that raised wage rates, led in a short period of time to dramatic labor supply increases for this group (see Hanming Fang and Keane 2004 for a detailed discussion). Notably, however, work by Keane and Moffitt (1998) and Keane (1995), who modeled the budget constraint created by AFDC (along with other programs and fixed costs of work) in great detail, predicted that, while large AFDC tax rate reductions would have little effect, labor supply of single mothers would be quite sensitive to wage subsidies, EITC and fixed cost of work subsidies (or work bonuses). This illustrates the value of a structural approach.[38]

Still, the labor supply literature has had a strong tendency to report parameters like e in (36) as "the" Marshallian elasticity obtained by the study in question. I will generally follow this ingrained tradition, but the reader should always keep this caveat in mind: when one sees a typical labor survey that contains a list of Marshallian and Hicks elasticities, one should recall that in many cases these are statements about the shape of workers' utility functions, not about how they would respond to particular tax changes.

That being said, note that Burtless and Hausman obtained a "Marshallian elasticity" of $e \approx 0$ and an elasticity of hours with respect to nonlabor income of $e_I = -0.048$. As we see from (8), to obtain the income effect from the income elasticity we need to multiply by wh/N. Given the population under study, reasonable values (on a weekly basis) appear to be roughly $w = \$3.00$, $h = 35$, $N = \$70$ so that $wh/N = 105/70 = 1.5$, giving a typical value of $ie \approx -0.072$.[39] The overall conclusion was that the income guarantee in the NIT experiments led to only modest reductions in labor supply (i.e., an hours reduction of about 7.5 percent).[40]

Pencavel (1986) summarizes results of eight other studies of the NIT experiments. Again, the estimates of the Marshallian elasticity are all small, but the mean is positive (0.03). Income effects range from about 0.02 to –0.29 (mean −0.10). Hicks elasticity estimates are bunched fairly tightly around the mean of 0.13.

6.1.3 *A Brief Summary of the Literature up to 1990*

Here it is useful to summarize the state of the literature up to 1990. I discussed four papers that used sophisticated econometric methods to model labor supply with progressive taxes. These studies tended to find larger

[38] As noted by Hausman (1980), "Structural econometric models which make labor force participation a function of . . . wages, income transfer levels and the tax system can attempt to answer questions such as the effect of lowering the marginal tax rates on labor force participation. The more traditional reduced form models which do not explicitly parameterize the tax system will be unable to answer such questions" (p. 161).

[39] Burtless and Hausman (1978) do not go into much detail about characteristics of the sample. I choose $h = 35$ because they indicate this was the mean of hours, and I choose $N = \$70$ because their examples imply that G was approximately \$3,500 per year. $w = \$3.00$ seems plausible given the time and sample, which was very low income. Alternatively, they evaluate wh/N at the first kink point in the budget constraint for control subjects (see the first row of their table 2). This gives a higher ie of $[(1.67) \times (43.16)/(27.8)](-0.048) = (2.6)(-0.048) = -0.125$.

[40] Burtless and Hausman (1978) have been criticized because they let the income elasticity e_I be heterogeneous in the population, and a large fraction of the estimates were bunched near zero. See MaCurdy, Green, and Paarsch (1990). The implication is that much of the mass would have been on positive values for the income elasticity if this had been allowed. Even so, it seems the main conclusion of small income effects would not be altered.

labor supply elasticities than in earlier work using "simpler" methods. They also found large efficiency losses from taxation. Hausman (1981), Blomquist (1983), and Blomquist and Hansson-Brusewitz (1990) estimated efficiency losses from progressive taxation of 22 percent, 12 percent, and 16 percent of revenue, respectively, and that these losses could be greatly reduced by shifting to a flat tax; to 7 percent, 2 percent, and 5 percent, respectively. The fourth paper, Wales and Woodland (1979), did not calculate tax effects, but they presumably would have found similar results, given their large Hicks elasticity estimate (0.84). Hausman (1985b) argued this body of work provided a strong case for a flat rate tax. In contrast, the NIT experiments produced small estimates of labor supply elasticities for low-wage workers.

6.1.4 *The "Hausman–MaCurdy Controversy"*

In a very influential paper, MaCurdy, Green, and Paarsch (1990) challenged the conclusions about progressive taxation described above. In fact, they argued that the whole approach to estimating piecewise-linear budget constraint models represented by Hausman (1981) was fundamentally flawed because the method was biased toward finding large Hicks elasticities. This became known as the "Hausman–MaCurdy controversy."

To understand the issue, consider a linear specification as in (35). For a person on segment #1 in figure 1, the labor supply equation is

$$\text{(37a)}\quad h = \beta + \beta_w\, w(1-\tau_1) + \beta_I N + \varepsilon,$$

while, for a person located on segment #2, the labor supply equation is

$$\text{(37b)}\quad h = \beta + \beta_w\, w(1-\tau_2) + \beta_I[N + w(\tau_2-\tau_1)H_2] + \varepsilon.$$

In (37b), I have substituted $V_2 = N + w(\tau_2 - \tau_1)H_2$.

Now, the taste shock ε has to be above a certain threshold (such that desired hours are at least H_2) for the person to locate on segment #2. And ε has to be below some threshold in order for the person to choose to locate on segment #1.[41] Crucially, there is an *intermediate range* of ε such that a person chooses to locate precisely at the kink point H_2. This occurs if

$$\text{(38a)}\quad \beta + \beta_w w(1-\tau_2) + \beta_I[N + w(\tau_2-\tau_1)H_2] + \varepsilon < H_2$$

$$\text{(38b)}\quad \beta + \beta_w w(1-\tau_1) + \beta_I N + \varepsilon > H_2.$$

Equation (38a) says, given a hypothetical budget line that extends segment #2 down to $h = 0$, the person would choose hours *less* than H_2. Equation (38b) says, given a hypothetical budget line extending segment #1 up to $h = H_{\max}$, the person would choose hours *greater* than H_2. For the actual two-segment constraint, the best choice is to locate precisely at the kink point H_2.

Now, rearranging (38) to express it as a range on ε, we obtain

$$\text{(38')}\quad \begin{aligned} \varepsilon &< H_2 - \beta - \beta_w\, w(1-\tau_2) - \beta_I[N + w(\tau_2-\tau_1)H_2] \equiv U(\varepsilon) \\ \varepsilon &> H_2 - \beta - \beta_w\, w(1-\tau_1) - \beta_I N \equiv L(\varepsilon). \end{aligned}$$

[41] Of course, this dependence of the range of the errors on the observed segment is precisely why the errors do not satisfy standard OLS assumptions in models with progressive taxation.

Here I use $U(\varepsilon)$ and $L(\varepsilon)$ to denote the upper and lower bounds on ε such that a person wants to locate at the kink point. Obviously we must have $U(\varepsilon) > L(\varepsilon)$ in order for the probability of locating at the kink point to be positive. Indeed, the opposite case of $U(\varepsilon) > L(\varepsilon)$ would imply the logical impossibility that the probability is negative, implying an internal inconsistency within the model. The condition that $U(\varepsilon) > L(\varepsilon)$ can be written as

$$-\beta_w w(1-\tau_2) - \beta_I[N + w(\tau_2 - \tau_1)H_2] > -\beta_w w(1-\tau_1) - \beta_I N,$$

which can be further simplified to

$$\beta_w[w(1-\tau_1) - w(1-\tau_2)] - \beta_I w(\tau_2 - \tau_1)H_2 > 0 \tag{39}$$

or simply $\beta_w - \beta_I H_2 > 0$, which we can put in elasticity terms to obtain

$$(w/H_2)[\beta_w - \beta_I H_2] > 0. \tag{40}$$

The left hand side of (40) is simply the definition of the Hicks elasticity from equation (8), for the linear model (35), evaluated at hours level H_2. Thus, MaCurdy, Green, and Paarsch (1990) argued that the Hausman approach to piece-wise linear tax models requires the Hicks elasticity to be positive (at all kink points) to avoid generating negative probabilities.[42]

Notice that, if $\beta_I > 0$ (i.e., the income effect has the "wrong" sign, implying leisure is not a normal good), then (40) will have to turn negative for large enough values of H_2. Thus, for all practical purposes, if confronted with a tax system with kinks at high levels of income, the Hausman approach requires that $\beta_I < 0$.[43] Indeed, Burtless and Hausman (1978), Hausman (1981), and Blomquist (1983) all restrict $\beta_I < 0$ in estimation.[44]

To gain intuition for why (40) is necessary to induce people to locate at kink, suppose $\beta_I > 0$. Then, for a person located at H_2, the increase in virtual nonlabor income that occurs if he/she increases hours above H_2 is actually an inducement to increase hours, not a deterrent. Thus, the only thing to keep the person from increasing hours beyond H_2 is if the Marshallian elasticity is large enough to outweigh the perversely signed income effect (as the wage drops if the person moves above H_2). But this means by definition the Hicks elasticity is positive.

To proceed, MaCurdy, Green, and Paarsch (1990)—referring to surveys by Pencavel (1986) and Hausman (1985b)—noted how papers that used "simple" empirical methods tended to obtain small Hicks elasticities, including even perverse negative values. In contrast, the papers that used the piecewise-linear budget constraint approach tended to get large Hicks elasticities. MaCurdy, Green, and Paarsch (1990) argued that the difference in results did *not* arise because the piecewise-linear budget constraint models did a better job of incorporating taxes. Instead, they argued the difference

[42] Reversing this logic, we can infer the magnitude of the Hicks elasticity from how many people locate at kinks (see Saez 2010). But Raj Chetty et al. (2009) show that, with a small amount of friction in hours adjustments or a small amount of measurement error, large Hicks elasticities can be consistent with a small amount of bunching.

[43] Equation (39) says the uncompensated wage effect (β_w), times the drop in the wage in going from segment #1 to segment #2, must exceed the income effect (β_I) times the increase in virtual nonlabor income. Normally, we would expect $\beta_I < 0$, so the second term in (39) is positive. Then (39) simply constraints how negative β_w, the sign of which is theoretically ambiguous, can be (i.e., the Marshallian elasticity can't be too negative). But, if β_I has the "wrong" sign (i.e., $\beta_I > 0$), then the second term is negative and increasing in H_2. In that case, it becomes very difficult to satisfy (39) for large values of H_2 unless the Marshallian elasticity is a very large positive.

[44] All these papers assume the income effect is randomly distributed in the population with a truncation at zero.

arose simply because the piecewise-linear approach imposed the restriction in (40) that the Hicks elasticity be positive.[45] This criticism was highly influential, leading many to discount the large Hicks elasticities obtained using the piecewise linear methods, and contributing to the consensus that the Hicks elasticity is small.

MaCurdy, Green, and Paarsch (1990) proposed an alternative idea of approximating a piecewise linear convex budget constraint by a smooth (i.e., kink free and differentiable) polynomial function. Suppose that the tax function is a differentiable function of earnings, which I'll denote by $\tau(w_t h_t)$. Then, for example, equations (3) and (5) become

$$(3') \quad U_t = \frac{[w_t h_t + N_t - \tau(w_t h_t)]^{1+\eta}}{1+\eta} - \beta_t \frac{h_t^{1+\gamma}}{1+\gamma} \quad \eta \leq 0, \ \gamma \geq 0$$

$$(5') \quad MRS = \frac{MUL(h)}{MUC(h)} = \frac{\beta_t h_t^{\gamma}}{[w_t h_t + N_t - \tau(w_t h_t)]^{\eta}} = w_t(1 - \tau'(w_t h_t)).$$

Comparing (5) and (5′), we see that the constant tax rate τ in (5) is replaced by $\tau'(w_t h_t)$, the derivative of the tax function evaluated at earnings level $w_t h_t$ (i.e., the tax on a marginal dollar of earnings). $\tau(w_t h_t)$ can be chosen to provide a good approximation to the actual tax system.

Now, while it is undeniable that the piecewise-linear budget constraint approach must constrain the Hicks elasticity to be positive to generate a sensible econometric model (with probabilities guaranteed to be positive), it is not obvious that this can explain the difference in results between piecewise-linear budget constraint studies and those that use simpler linear regression methods. I say this for two reasons: First, a number of studies that use a piecewise-linear budget constraint approach do nevertheless find Hicks elasticities that are close to zero. Conversely, some papers using simpler methods to handle taxes find large Hicks elasticities.

For example, consider what happened when MaCurdy, Green, and Paarsch (1990) applied the same approach as Hausman (1981) to a sample of 1,017 prime age men from the 1975 PSID. Like Hausman, they assume a linear hours equation as in (35) with a random coefficient on nonlabor income. Strikingly, MaCurdy, Green, and Paarsch obtained a wage coefficient of essentially zero and a (mean) income coefficient of –0.0071 (see their table 2, first column). The latter implies an income effect of roughly $w \cdot (\partial h/\partial N) = (4.4)(-0.0071) = -0.031$ and hence a Hicks elasticity of roughly 0.031 at the mean of the data. Thus we have an example where the piece-wise linear approach does yield a very small Hicks elasticity.

There have been other applications where the piecewise-linear approach yielded small Hicks elasticities. A good example is Robert K. Triest (1990) who applies methods very similar to Hausman (1981) to study 978 married men aged 25–55 in the 1983 PSID. He obtains an income elasticity of essentially

[45] To quote MaCurdy, Green, and Paarsch (1990): "As documented in the surveys of Pencavel (1986) and Hausman (1985b), empirical studies . . . based on econometric approaches incorporating piecewise-linear constraints produce . . . estimates of compensated substitution responses that have the sign predicted by economic models of consumer choice, which is in contrast to much of the other empirical work on labor supply. This finding of greater consistency with economic theory has been interpreted . . . as evidence confirming the merits of accounting for taxes using the piecewise-linear approach. Contrary to this interpretation, this paper shows that the divergence in the estimates . . . follows directly from features of the econometric models that implicitly restrict parameters . . . The simple estimation approaches impose no restrictions, but maximum likelihood techniques incorporating piecewise-linear budget constraints require . . . Slutsky condition to hold at various points in estimation" (pp. 416–17).

zero and Marshallian and Hicks elasticities of roughly 0.05. And recall that the Blomquist (1983) study that I discussed earlier obtained a Hicks elasticity of roughly 0.11 and an income effect of –0.03, which can hardly be called large.

Next consider papers that use "simple" methods but still obtain large Hicks elasticities. A prime example is the classic paper by Hall (1973). Recall that he linearized the budget constraint around the observed wage/hours combination, but did not model the choice of segment. But, like Hausman (1981), he obtained a large Hicks elasticity (0.45).

As for the "simple" approach of assuming a smooth approximation to the kinked budget constraint, as in (3′)–(5′), MaCurdy, Green, and Paarsch (1990) note that this approach also constrains the Hicks elasticity, except now the constraint is a bit weaker: instead of requiring it to be positive, it requires that it can't be "too negative." But the situation is not fundamentally different. As the smooth approximation to the budget constraint is made more accurate, the bound on the Hicks elasticity gets tighter, converging to a lower bound of zero as the approximation approaches the true constraint. When MaCurdy, Green, and Paarsch (1990) apply this approach, they conclude (see page 458): "there is no perceptible difference in the estimates obtained assuming differentiable and piecewise-linear tax functions."

Thus, it is not clear that use of piecewise linear budget constraint methods versus simpler methods can explain the large divergence in results across studies. It is particularly puzzling that Wales and Woodland (1979), Hausman (1981), MaCurdy, Green, and Paarsch (1990), and Triest (1990) all applied the piecewise linear approach to data on married men in the PSID, using data from nearby (sometimes identical) waves, and yet the former two studies obtained very large Hicks elasticities and income effects while the latter two studies obtained negligible values for each. Indeed, the latter two papers explicitly make note of the fact that this contrast is puzzling.

6.1.4.1 *An Attempt to Resolve the Controversy—Eklöf and Sacklén (2000)*

The excellent replication study by Eklöf and Sacklén (2000) sheds a great deal of light on the reasons for the divergence in results between Hausman (1981) and MaCurdy, Green, and Paarsch (1990). Both papers study married men aged 25 to 55 in the 1976 wave of the PSID. The MaCurdy, Green, and Paarsch sample size is a bit smaller (1018 versus 1084), because they apply slightly more stringent selection criteria,[46] but Eklöf and Sacklén (2000) show this is not a main reason for differences in results. Rather, the difference appears to arise because the two studies adopt very different definitions of the wage and nonlabor income variables.

A key point about the PSID is that it contains questions both about the interview week (e.g., What is your current wage rate?) and about the prior year (e.g., What were your annual earnings and annual hours during the past year?). Hausman (1981) uses the current wage question as his measure of the wage rate, while MaCurdy, Green, and Paarsch (1990) use the ratio of annual earnings to annual hours. Both of these wage measures have problems:

Hausman's current wage measure is missing for 87 workers and for 4 workers who were not employed in the survey week, and it is top coded at $9.99 per hour for 149 workers. Hausman imputes these missing wage observations for $240/1084 = 22$ percent of the sample using a regression method. In

[46] The main difference is Hausman (1981) requires that workers not be self-employed at the 1976 interview, while MaCurdy, Green, and Paarsch (1990) requires they not be self-employed in *both* 1975 and 1976. This costs 55 people.

addition, even an accurately measured current wage is presumably a noisy measure of the wage rate that is relevant for the whole prior year.

MaCurdy, Green, and Paarsch's ratio wage measure suffers from the denominator bias problem discussed in section 4: Say observed hours equal $h^* = h + \varepsilon$, where h is true hours and ε is measurement error, and we construct the wage as $w^* = E^*/(h + \varepsilon)$, where E^* is measured earnings. Then the measurement error in hours tends to induce negative covariance between h^* and w^*. This denominator bias has the potential to drive the wage coefficient negative.

In addition, Hausman (1981) and MaCurdy, Green, and Paarsch (1990) take very different approaches to measuring nonlabor income. Hausman simply imputes an 8 percent return to equity in owner occupied housing (the only financial asset measured in the PSID). In contrast to this narrow measure, MaCurdy, Green, and Paarsch (1990) construct a very broad measure by taking total household income minus labor earnings of the husband. The broad measure has the problem that it includes the wife's income, which may be endogenous (i.e., jointly determined with husband earnings). In contrast, Hausman's narrow measure simply leaves out many types of nonlabor income. The sample mean of MaCurdy, Green, and Paarsch's nonlabor income measure is three times greater than Hausman's. Neither includes imputed service flows from consumer durables.

Finally, Hausman (1981) and MaCurdy, Green, and Paarsch (1990) use different hours measures. MaCurdy, Green, and Paarsch (1990) use a direct question about hours of work in 1975. Hausman (1981) uses questions about usual hours per week and number of weeks worked in 1975. The mean of MaCurdy, Green, and Paarsch's hours measure is 2,236 while that of Hausman's hours measure is 2,123.

Using the same data as MaCurdy, Green, and Paarsch, Eklöf and Sacklén (2000) are able to replicate their results (for the piecewise linear approach) almost exactly. That is, the wage coefficient bumps up against the nonnegativity constraint and has to be pegged at zero. And the mass of the random nonlabor income coefficient piles up near zero. Then Eklöf and Sacklén report results of an experiment where, either one by one or in combination, they shift to Hausman's wage measure, nonlabor income measure, sample selection criteria, and/or hours measure.

A subset of the results is reproduced in table 3. The first row presents the replication of MaCurdy, Green, and Paarsch (1990). The only difference is a slight change in the computation procedure that leads to a small increase in the estimated income effect (from about –0.037 to –0.068).[47] The second row shows the effect of adopting Hausman's sample selection criteria. This leads to a doubling of the income effect to –0.136. But the wage coefficient remains pegged at zero.

In the third row, the authors switch to Hausman's narrower definition of nonlabor income. This has a dramatic effect on the results, as the income effect jumps to –0.488. This is actually quite disconcerting. Given that each paper's definition of nonlabor income is quite debatable, and that, as noted in section 4, it is not at all obvious how one should define nonlabor income in a static model (as nonlabor income evolves over the life cycle based on savings decisions), it is unfortunate that results are so sensitive to how it is defined.[48]

[47] MaCurdy, Green, and Paarsch (1990) reported it was necessary to constrain the variance of the random income effect to obtain sensible estimates, but Eklöf and Sacklén (2000) did not have this problem in the replication.

[48] I was puzzled why the authors maintained a peg of the wage coefficient at zero in this model. With an income effect of –0.488, the Marshallian elasticity can go well negative while maintaining a positive Hicks elasticity.

TABLE 3
Eklöf and Sacklén (2000) Analysis of Hausman versus MaCurdy–Green–Paarsch (M–G–P)

Wage measure	Nonlabor income measure	Sample selection criteria	Hours measure	Coefficient on: Wage	Coefficient on: Nonlabor income	Marshall elasticity	Income effect	Hicks elasticity
M–G–P	M–G–P	M–G–P	M–G–P	0.0	−0.011	0.000	−0.068	0.068
M–G–P	M–G–P	Hausman	M–G–P	0.0	−0.022	0.000	−0.136	0.136
M–G–P	Hausman	M–G–P	M–G–P	0.0	−0.079	0.000	−0.488	0.488
Hausman	M–G–P	M–G–P	M–G–P	10.3	−0.004	0.030	−0.025	0.055
Hausman	Hausman	M–G–P	M–G–P	26.5	n/a	0.078	n/a	n/a
Hausman	Hausman	Hausman	M–G–P	26.9	−0.036	0.078	−0.222	0.300
Hausman	Hausman	Hausman	Hausman	16.4	−0.036	0.048	−0.222	0.270
Hausman's reported results				0.2	−0.120	0.000	−0.740	0.740

Notes: For the sake of comparability all elasticities and income effects are calculated using the mean wage of $6.18 and the mean hours 2,123 from Hausman (1981). In the authors' attempt to replicate Hausman's data set, the corresponding figures are 6.21 and 2,148. The mean values of both hours and wages are a bit higher in the MaCurdy et al. data set, but this makes little difference for the calculations. For the random nonlabor income coefficient, the table reports the median.

The fourth row gives results using Hausman's wage measure. Strikingly, the wage coefficient is now positive, implying a small but positive Marshallian elasticity (0.03). But the income effect remains very small (–0.025), implying a Hicks elasticity of only 0.055.

The fifth row shows the effect of simultaneously adopting Hausman's wage and nonlabor income measures. This causes the Marshallian elasticity to jump further to 0.078, but unfortunately the authors do not report the income coefficient for this case. The sixth row shows the effect of simultaneously adopting Hausman's wage and nonlabor income measures, and his sample selection criteria. The Marshallian elasticity remains at 0.078 and the income effect is –0.222, giving a fairly large Hicks elasticity of 0.300.

Finally, the sixth row also adopts Hausman's hours measure. Having adopted all of his variable definitions and sample screens, row 6 is in fact the authors attempt to replicate Hausman (1981). The results have a similar flavor to Hausman's: the Marshallian elasticity is modest (0.048) but the income effect is –0.220, giving a fairly large Hicks elasticity of 0.270.

Based on these results, the authors conclude it is not the piecewise linear budget constraint approach per se that explains why Hausman (1981) obtained a much larger Hicks elasticity than other authors. Instead, Eklöf and Sacklén (2000) argue that the key differences were Hausman's use of a direct wage measure and his narrow definition of nonlabor income. In particular, the evidence suggests that measuring the wage as an annual earnings divided by annual hours does lead to denominator bias that tends to drive the wage coefficient negative.

As further evidence of this assertion, they point to the special issue on labor supply in the 1990 *Journal of Human Resources*. In three studies that use a ratio wage measure (Triest 1990, MaCurdy, Green, and Paarsch 1990, Ugo Colombino and Daniela Del Boca 1990), the Hicks elasticity is either negative or runs up against a nonnegativity constraint. But in three studies that use a direct wage measure (Blomquist and Hansson-Brusewitz 1990, van Soest, Woittiez, and Kapteyn 1990 and their own version of MaCurdy, Green, and Paarsch (1990)), the Hicks elasticity is positive.[49]

The last two rows of table 3 compare Eklöf and Sacklén's replication of Hausman (1981) with Hausman's reported results. Clearly, there are large differences. While Hausman obtained a Marshallian elasticity close to zero, the authors obtain 0.048. And while Hausman obtained a large income effect of –0.740, Eklöf and Sacklén obtain a perhaps more plausible value of –0.222.[50] What accounts for these differences? The authors note that they could not match Hausman's sample exactly. They also note that the likelihood is quite flat near the optimum. A wide range of different values for the mean and variance of the random coefficient on nonlabor income produce similar likelihood values. Thus, they speculate that fairly minor changes in the dataset could have produced a fairly large change in the estimates.

Recall that Hausman (1981) calculated that tax progressivity led to an efficiency loss of 22 percent of revenues. This value is driven largely by his large estimate of the Hicks elasticity. As Eklöf and Sacklén (2000) obtain a mean Hicks elasticity about a third of Hausman's, one is tempted to conclude the implied efficiency loss is about a third as large as well. But these models assume a distribution of income effects, and Eklöf and Sacklén (2000) obtain both a lower mean and a higher variance. Thus, it is not at all clear what simulation of their model would imply about efficiency losses. It is unfortunate such a simulation is not available.

6.1.4.2 *A Nonparametric Approach—Blomquist, Eklöf, and Newey (2001)*

One possible reaction to the "Hausman–MaCurdy" controversy is to adopt a less structured approach to modeling the effect of progressive taxation. This is done in Blomquist, Eklöf, and Newey (2001). Their idea is to run a nonparametric regression of hours on a set of variables that completely characterize the budget constraint. In the case of a linear constraint, this is just a nonparametric regression of hours on the wage and nonlabor income (meaning, in practice, using polynomials in these variables as regressors). But with a piecewise linear constraint the regressors are the intercepts and slopes of each segment.[51]

They apply this idea to data on married men (aged 20–60) from the Swedish Level of Living Surveys of 1973, 1980, and 1990. The progressivity of the tax system was greatly reduced from 1980 to 1990, so this study implicitly uses this variation in tax rates to identify labor supply elasticities. But a difficulty arises in applying the nonparametric approach because the Swedish tax system

49 Blomquist (1996) conducted a Monte Carlo study where he compared the performance of the piecewise linear ML approach with the simpler Hall (1973) approach of linearization combined with IV. For the case where both earnings and hours are measured with error and a ratio wage measure is used, he finds both ML and IV give wage coefficients severely biased toward zero. But, if a noisy direct wage measure is used, both estimators perform fairly well. But the problems created by error in measuring *taxable* nonlabor income can be much more severe. It can create severe bias for both ML and IV, and the direction of the bias depends on the nature of the error (i.e., classic measurement error, imputing deductions, etc.).

50 Recall that Pencavel (1986) argued the income effect estimated by Hausman (1981) was implausibly large.

51 Given that one regresses hours on the "the whole budget set," the problem of endogenous choice of segment is circumvented. However, as the budget set facing an individual depends on their own (pretax) wage and nonlabor income, endogeneity of (pretax) wages and nonlabor income is still an issue.

had many segments (e.g., twenty-seven in 1980). Thus, the tax system in each of the three years is approximated by a three segment constraint. Intercepts and slopes of these three segments are then used as regressors in the nonparametric regression.

The authors report a Marshallian elasticity at the mean of the data of 0.075 and an income elasticity of –0.038. Based on their tables 1 and 2, it appears a "typical" person in 1980 faced the linearized constraint $I = 59{,}110 + (0.25)(wh)$ with $w = 54$ SEK and $h = 2{,}080$. This implies an income effect of $ie = (0.48)(-0.038) = -0.019$ and thus a Hicks elasticity of 0.094. Estimation of a more conventional parametric model gives a Marshallian elasticity of 0.137 and a Hicks elasticity of 0.144 (evaluated at the same point).

Blomquist, Eklöf, and Newey (2001) simulate effects of different aspects of the reform, but consider just the drop in marginal rates. From 1980 to 1991, the marginal rate for the "typical" person fell from 75 percent to 69 percent, while virtual income fell only 1,000. Using the nonparametric model, the authors simulate that the reduction in progressivity from 1980 to 1991 increased (mean) hours from 2,082 to 2,157 (or 3.9 percent), while reducing revenue by 9 percent.[52] To get a rough idea of the implied efficiency gain from the reform, imagine reversing it. If there were no behavioral response (i.e., labor supply stays fixed at the 1991 level), tax revenue would be roughly 4 percent higher in 1980 than under the baseline. Thus, revenue would increase by roughly 14 percent in the absence of a behavioral response. Instead, the increase is only 10 percent due to the labor supply reduction. So the implied efficiency loss from reverting to the 1980 system is substantial.

6.1.5 *A Look at the Distribution of Hours of Work*

The paper by van Soest, Woittiez, and Kapteyn (1990) uses the Dutch Organization of Strategic Labor Market Research 1985 survey. It contains a direct question about wages on a weekly or monthly basis. Using a piecewise linear approach, the authors obtain a Marshallian elasticity of 0.19 and an income effect of –0.09; so they have no problem with the nonnegativity constraint on the Hicks elasticity (0.28).[53] In my view, the more important aspect of this paper is that, as far as I can discern, it was the first (for males) to use simulated data from the model to examine model fit.[54] A rather striking failure of the labor supply literature (which it shares with many other literatures in economics) is the lack of effort to examine model fit. What the authors find, perhaps not surprisingly, is that a simple linear labor supply function like (35), combined with a piecewise linear budget constraint, does a very poor job of fitting the observed distribution of hours. In particular, it is completely unable to generate the substantial bunching of male hours at forty hours per week (see their figure 1).

The authors attempt to rectify this problem by introducing a demand side constraint on possible hours choices. Each worker is assumed to draw a set of hours points at which he may locate, and a probability of each point is estimated. Of course, offers of forty hours are estimated to be much more likely than offers of lower hours levels. So this model does fit the spike in hours at

[52] The parametric model predicts a larger increase in hours, from 2,084 to 2,215, or 6.3 percent.

[53] Based on our discussion of Eklöf and Sacklén (2000), this is as we would expect given that they use a direct wage measure. Van Soest, Woittiez, and Kapteyn (1990) do not give information on the construction of their nonlabor income variable, but in private correspondence the authors told me that they used a fairly narrow measure that consists only of child benefits (that do not depend on income) and capital income (which few households have).

[54] For females, see the discussion of Cogan (1981) in section 7.1.1.

forty (as well as the distribution over other points) quite well. What seems unsatisfactory is that the model contains no rationale for why offers of lower levels of hours are uncommon. One explanation would be start up costs at work, so that productivity rises with hours but starts to decline somewhere after forty. An alternative supply side story for why low levels of hours are uncommon is fixed costs of work (see John F. Cogan 1981 in section 7).

6.1.6 *Summary of the Static Literature on Male Labor Supply*

To summarize, the literature on static models has not produced a clear consensus on male labor supply elasticities. Some claim that piecewise linear budget constraint methods produce large Hicks elasticities while "simpler" IV methods produce small elasticities. But Eklöf and Sacklén (2000) find that differences in results between several such studies are better explained by different definitions of the wage rate and nonlabor income. The use of annual ratio wage measures, as opposed to hourly or weekly measures, tends to generate smaller elasticities, presumably due to denominator bias. And more narrow definitions of nonlabor income tend to generate larger income effects. Given the problem of denominator bias, it seems fairly clear that the use of ratio wage measures should be avoided in favor of hourly measures.[55] But the best way to measure nonlabor income is not at all clear.

In general, nonlabor income may include many components, such as interest income from assets, the service flow from durables, government transfer payments, transfers from relatives, and, in a household context, spouse's income (or some share thereof). Determining the "right" measure of nonlabor income in a static labor supply model is difficult in part because the static model does not provide a framework to even think about asset income. Indeed, in a static model assets should not even exist, as there is no motive for saving. This leads us naturally to an examination of life-cycle labor supply models with savings.

6.2 *The Life-Cycle Labor Supply Model with Savings*

In dynamic models, workers make labor supply decisions jointly with decisions about consumption/savings, and the evolution of nonlabor income becomes part of the model. But, as I will discuss below, estimation of dynamic models is difficult. Thus, some authors have sought to develop alternative approaches that maintain the simplicity of static models while producing estimates that are still consistent with life-cycle behavior.

6.2.1 *Simple Methods for Estimating the Life-Cycle Labor Supply Model—MaCurdy (1983)*

In an important paper, MaCurdy (1983) developed a scheme for estimating the parameters of a life-cycle labor supply model using techniques no more complicated than instrumental variables. To see how his method works, we return to the life-cycle model of section 3.2 and rewrite equation (22) as

$$(41)\qquad \frac{\beta_t h_t^{\gamma}}{[w_t(1-\tau_t)h_t + N_t + b_t]^{\eta}} = \frac{\beta_t h_t^{\gamma}}{[C_t]^{\eta}} = w_t(1-\tau_t).$$

It is important to note that, while (41) assumes a flat rate tax, an optimality condition analogous to (41) will also hold in a world with progressive taxation. Then, τ_t is the marginal tax rate the person faces at time

[55] This is not to say that an hourly wage measure is ideal. Its drawback is that we are typically modeling labor supply over a longer period, such as a year. Indeed, this is presumably the reason that many studies chose to use annual wage measures (to match the time period of the wage with that of the observed labor supply behavior).

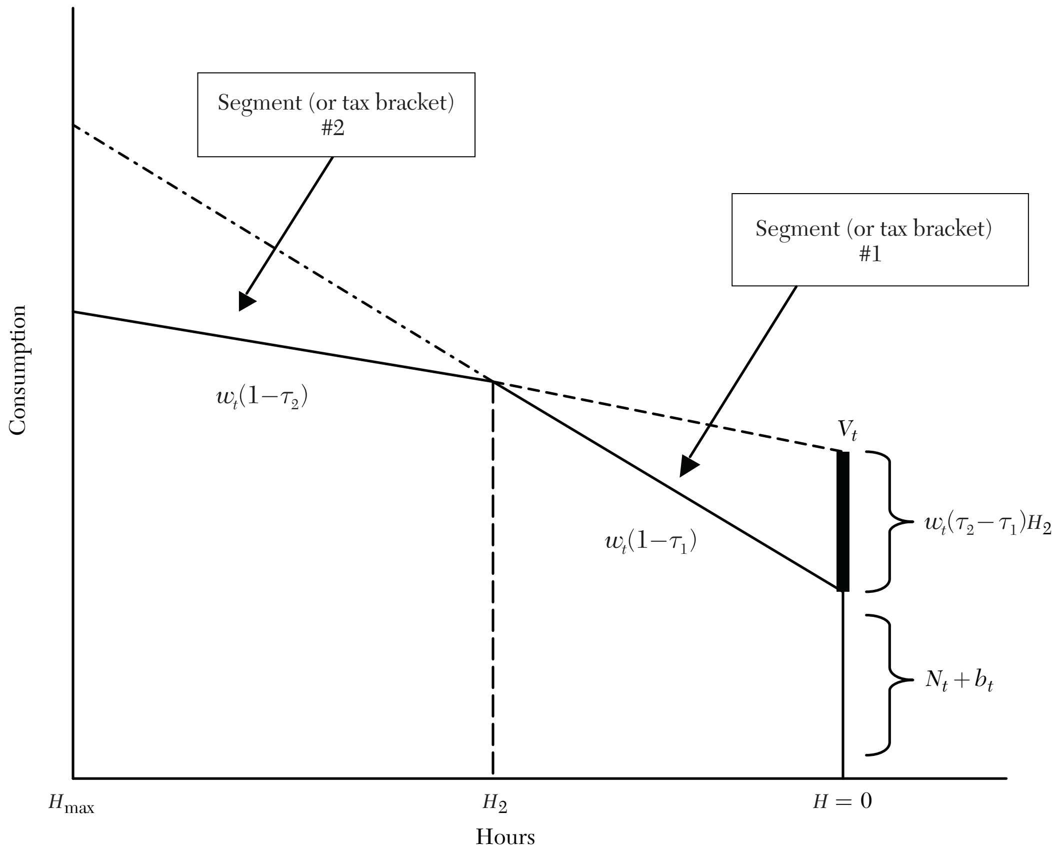

Figure 3: The Budget Constraint Created by Progressive Taxation in the Presence of Saving

t, for the tax bracket in which he/she sits at that time (as in equation (5′)).[56] Also, the equation for consumption must be modified. This is illustrated in figure 3 for a system with two brackets. Consider a person who chooses to locate on segment 2, so $h_t > H_2$, where H_2 is the hours level that renders the person's earnings high enough that he/she enters tax bracket #2. The consumption level for this person is

$$\text{(42a)} \qquad C_t = w_t(1 - \tau_2)h_t + V_t,$$

where "virtual" nonlabor income V_t is given by

$$\text{(42b)} \qquad V_t = w_t(\tau_2 - \tau_1)H_2 + N_t + b_t.$$

MaCurdy (1983) noted these points, and also that the optimality condition (41) contains only variables dated at time t. Hence, despite

[56] Condition (41) would fail to hold for a person who locates at a kink point. Thus, MaCurdy assumes the tax system is approximated by a smooth function, ruling out kink points.

the fact that we have a dynamic model with saving, the preferences parameters γ and η can be estimated from a *single period of data*, provided we utilize not only data on hours and wages, but also on consumption. MaCurdy proposed two methods for doing this, and both play a key role in the subsequent literature:

Method 1: Estimate (41) using two stage least squares. As MaCurdy notes, (41) must hold at the optimal hours choice, regardless of whether there is a flat-rate or progressive tax (provided the person is not at a kink point). To put (41) in a form that can be estimated, we need to introduce a source of stochastic variation in hours and consumption choices. Let the parameter β_t, which shifts the MRS between consumption and leisure, be given by

$$\beta_{it} = \exp(X_{it}\,\alpha - \varepsilon_{it}). \tag{43}$$

Here X_{it} represents *observed* characteristics of person i that shift tastes for consumption versus leisure, and ε_{it} represents unobserved taste shifters. Now, taking logs of (41), and putting i subscripts on all variables to indicate person specific values, we get

$$\ln w_{it}(1-\tau_{it}) = \gamma \ln h_{it} - \eta \ln C_{it} + X_{it}\,\alpha - \varepsilon_{it}. \tag{44}$$

Note that (44) is not a typical labor supply equation (with hours as the dependent variable). Rather, it is simply a relationship among three endogenous variables—the after-tax wage, hours and consumption—that must hold if person i is making work/consumption choices as suggested by economic theory. All three variables are endogenous as they are correlated with the taste shocks ε_{it}. This occurs for reasons we have already discussed.[57] As a result, it is just a matter of convenience which endogenous variable we call the "dependent" variable.

Given the endogeneity of hours and consumption, we should estimate (44) using IV. Instruments must be correlated with wages, hours, and consumption but not with unobserved tastes for work ε_{it}. Naturally, the choice of instruments tends to be controversial in any such approach. Estimation of (44) gives values of the structural parameters of preferences γ and η, from which we can construct the Marshall, Hicks, and Frisch elasticities, as in equation (21).

Method 2: Estimate a labor supply function consistent with the life-cycle framework. The idea here is to extend the Hall (1973) approach to the dynamic case simply by redefining virtual nonlabor income for period t to include b_t. The approach is illustrated in figure 3 for the case of a two bracket tax system. Note that, if a person locates on segment #1, then his/her after-tax wage is $w_t(1-\tau_1)$ and virtual nonlabor income is $V_t = N_t + b_t$. If a person locates on segment #2, the after-tax wage is $w_t(1-\tau_2)$ and virtual nonlabor income is $V_t = w_t(\tau_2 - \tau_1)H_2 + N_t + b_t$. Notice that, regardless of segment, virtual nonlabor income is given by

$$V_t = C_t - w_t(1 - \tau_t)h_t. \tag{45}$$

where τ_t denotes the tax rate for the segment on which the person locates at time t. Thus, MaCurdy suggests estimating labor supply equations of the form

$$h_{it} = h(w_{it}(1 - \tau_{it}), V_{it}, X_{it}). \tag{46}$$

[57] Obviously, hours are endogenous as a person with a low value of ε_i tends to work more hours. The after-tax wage is endogenous because a person who is hard working will: (i) tend to have a high pretax wage because he/she puts in greater effort, and (ii) tend to face a higher tax rate because he/she works enough hours to be pushed into a high bracket. And consumption is endogenous because it is a function of the endogenous w and h.

To implement this procedure, one must pick a particular functional form for the labor supply function in (46). For example, one might chose the linear specification in (35) or the double log specification in (36). Also, as both the after-tax wage rate and virtual nonlabor income are endogenous, we must instrument for them, analogous to the approach in method #1.

MaCurdy (1983) implements methods #1 and #2 on a sample of 121 married men in the control group of the Denver Income Maintenance Experiment (DIME), a negative income tax experiment, in 1972–75. To implement method #1—equation (44)—MaCurdy includes as observed taste shifters (X_{it}) number of children and race. His main instruments are quadratics in age and education, and interactions between the two. This makes sense given the strong correlation between education and lifetime earnings, and the fact that both wages and hours follow hump shapes over the life cycle. The interactions capture that the shapes of these humps differ by education level. But use of these instruments requires the strong assumption that age and education are uncorrelated with tastes for work.

MaCurdy estimates that $\gamma = 0.16$ and $\eta = -0.66$. To compare these values to prior literature, he calculates what they would imply about labor supply elasticities given a linear budget constraint. It turns out the estimates imply highly elastic labor supply. Using our formulas for a person with no nonlabor income (see equation (21)) the implied Marshallian elasticity is $(1+\eta)/(\gamma-\eta) = 0.42$, the Hicks is $1/(\gamma-\eta) = 1.22$, the income effect is $\eta/(\gamma-\eta) = -0.80$, and the Frisch is $1/(0.16) = 6.25$. At the mean of the data, MaCurdy calculates a Marshallian elasticity of 0.70, a Hicks elasticity of 1.47, and an income effect of $w\partial h/\partial N = -0.77$.

Turning to method #2, MaCurdy considers both linear and double log specifications, using the same controls and instruments as in method #1. For the double log, he obtains

$$\ln h_{it} = \underset{(0.53)}{0.69} \ln w_{it}(1-\tau_{it}) - \underset{(0.0010)}{0.0016}\, V_{it} + X_{it}\,\alpha + \varepsilon_{it},$$

while for the linear specification he obtains

$$h_{it} = \underset{(13.8)}{19.4}\, w_{it}(1-\tau_{it}) - \underset{(0.07)}{0.16}\, V_{it} + X_{it}\,\alpha + \varepsilon_{it},$$

where the figures in parentheses are standard errors. The log specification gives a Marshallian elasticity of 0.69, almost identical to that obtained via method #1 at the mean of the data.

MaCurdy (1983) also evaluates the other elasticities at the mean of the data. This requires knowing that the mean of $V_{it} = C_{it} - w_{it}(1-\tau_{it})h_{it} = \$133$ per month, the mean of the after-tax wage is \$2.75 per hour and the mean of hours is 170 per month. In the double log model, the income effect is then $(wh)(1/h)\partial h/\partial V = (468)(-0.0016) = -0.75$. This is again almost identical to the value obtained using method #1. The Hicks elasticity is thus $0.69 + 0.74 = 1.44$.

For the linear specification, the Marshallian elasticity is $(2.75/170)(19.4) = 0.31$, the income effect is $(2.75)(-0.16) = -0.44$ and the Hicks elasticity is 0.75. Thus, the linear model produces more modest elasticities. Nevertheless, as MaCurdy notes, all three approaches (method #1 and method #2 with a double log or linear model) give elasticities that are quite large relative to most of the prior literature. He notes, this may indicate that prior estimates were misleading because: "Existing studies of male labor supply rarely treat measures of wages and income as endogenous variables . . . Many of these studies ignore taxes or fail to account properly for the endogeneity of marginal tax rates, and none of them recognizes that a household may save or dissave during a period." But he also notes that other factors, like invalid instruments or the small and unrepresentative nature of the DIME sample, may have led to upward biased elasticity estimates. The estimates are also rather imprecise (see above).

Altonji (1986) noted that one could rewrite (44) as

$$(47)\quad \ln h_{it} = \frac{1}{\gamma}\ln w_{it}(1-\tau_{it}) + \frac{\eta}{\gamma}\ln C_{it} - X_{it}\frac{\alpha}{\gamma} + \frac{\varepsilon_{it}}{\gamma}.$$

By estimating (47) by instrumental variables, we uncover the Frisch elasticity $(1/\gamma)$ directly. Recall the Frisch elasticity is defined as the effect of a change in the wage holding (marginal utility of) lifetime wealth fixed. In (47), consumption serves as a summary statistic for lifetime wealth. If the wage changes but consumption stays fixed, it means perceived wealth stayed fixed. This means either (i) that the person expected the wage change, so it does not affect his/her perception of lifetime wealth, or (ii) the person expects the wage change to be very short lived, so that it has a negligible effect on lifetime wealth. Estimation of (47) also enables us to back out η as the ratio of the consumption coefficient to the wage coefficient.

Altonji (1986) estimates (47) using data on married men, aged 25–60, from the 1968–81 waves of the PSID. Two key differences with MaCurdy (1983) are that he uses pretax wages, and the PSID measure of consumption includes only food. Altonji also uses a more extensive set of observed taste shifters in X (i.e., besides children and race he includes age, health, region, and year dummies). Recall that one must instrument for consumption and wages both because they are measured with error and because they are presumably correlated with the unobserved tastes (ε_{it}). A novel feature of Altonji's paper is that he uses a ratio wage measure (annual earnings over hours) as the independent variable in (47), and then uses a direct question about the hourly wage as an instrument. As long as the measurement error in these two measures is uncorrelated (as seems plausible), the latter is a valid instrument.[58] As an additional instrument Altonji uses a measure of the "permanent wage," constructed by regressing the observed wage on individual fixed effects, education, a quadratic in age, an interaction between age and education, year dummies, health, and region.

Altonji (1986) estimates that $(1/\gamma) = 0.172$ (standard error 0.119) and that $(\eta/\gamma) = -0.534$ (standard error 0.386). The implied values of γ and η are 5.81 and –3.10. These imply Frisch, Hicks, and Marshall elasticities of 0.17, 0.11, and –0.24, respectively, and an income effect of –0.35. This compares to values obtained by MaCurdy (1983) of 6.25, 1.22, 0.42, and –0.80. Reminiscent of the "Hausman–MaCurdy" debate discussed earlier, we have a situation where authors obtain very different estimates for reasons that are not evident. Does MaCurdy get much higher elasticities because he accounts for taxes and/or has a more complete measure of consumption? Or because he uses different instruments? Or are his results unreliable due to the small and unrepresentative nature of the DIME sample? Does rearranging (44) to obtain (47) matter? Unfortunately, there is no replication study that attempts to reconcile the Altonji (1986) and MaCurdy (1983) results, so we don't know the answer to these questions.

A closely related paper is Blundell and Walker (1986). They develop a simple scheme for estimating life-cycle models

[58] To be in the sample, a person must have both wage measures. There are 4,367 men who satisfy this criterion. Note that this tilts the composition of the sample toward hourly workers.

based on the idea of "two-stage budgeting." In the first stage, the worker/consumer decides how to allocate his/her "full income" across all periods of life. Full income is defined as the after-tax wage rate times the total hours in a period, plus any exogenous nonlabor income, plus net dissaving. Within each period, full income is allocated between consumption and leisure. Thus we have the within period budget constraint:

$$(48)\quad F_t = w_t(1-\tau_t)T + N_t + b_t$$
$$= w_t(1-\tau_t)(T-h_t) + C_t,$$

where F_t is full income, T is total time in a period and $T - h_t$ is leisure.[59] One then estimates a labor supply function that conditions on the full income allocated to period t:

$$(49)\quad h_{it} = h(w_{it}(1-\tau_{it}), F_{it}, X_{it}).$$

Note that this method is in fact identical to MaCurdy's method #2. This is because one can always define a segment of a budget constraint in terms of the after-tax wage and *either* full income or virtual nonlabor income ($V_t = C_t - w_t(1-\tau_t)h_t$). Thus, (46) and (49) are alternative expressions for the same labor supply function.[60]

Blundell and Walker (1986, p. 545) argue there is no need to instrument for F_{it}, even if it is a choice variable, as it is plausible that taste shifters that affect allocation of resources over the life cycle are independent of those that affect choices within a period. But this argument seems strained. For instance, one would plan to allocate more resources to periods when tastes for consumption and/or leisure are likely to be high than toward other periods.

In order to derive a labor supply function, Blundell and Walker (1986) consider a case where the indirect utility function has the Gorman polar form

$$(50)\quad U_t = G\left[\frac{F_t - a(w_t(1-\tau_t))}{b(w_t(1-\tau_t))}\right].$$

Actually, as we will see below, Blundell and Walker consider a more complex model of joint labor supply of couples, where the price of consumption goods varies over time in addition to the wage. But I will abstract from those complications for now. Obviously, we have that[61]

$$(51)\quad \frac{\partial U_t}{\partial w_t(1-\tau_t)} = h_t \cdot \frac{\partial U_t}{\partial F_t}.$$

Applying (51) to (50), we can obtain the labor supply equation

$$(52)\quad h_t = \frac{G'(\cdot)\left\{\dfrac{-b'(\cdot)}{b^2(\cdot)}[F_t - a(\cdot)] - \dfrac{a'(\cdot)}{b(\cdot)}\right\}}{G'(\cdot)\dfrac{1}{b(\cdot)}}$$
$$= -a'(w_t(1-\tau_t)) - \frac{b'(w_t(1-\tau_t))}{b(w_t(1-\tau_t))} \times [F_t - a(w_t(1-\tau_t))].$$

[59] Given progressive taxes, N_t could be defined to include the virtual nonlabor income for the linearized budget constraint, just as before.

[60] The full income allocated to period t plays the same role as consumption in MaCurdy's method #1 or virtual income in method #2. If the wage increases but full income allocated to the period is held fixed, it means that the wage increase did not make the person feel wealthier (i.e., it did not relax his/her lifetime budget constraint).

[61] This equality holds because, if the after-tax wage increases by one unit, the person has h_t extra units of income to spend on consumption. But if full income increases by one unit, the person has only one extra unit of income to spend on consumption. Thus the derivative on the left of (51) must be h_t times greater than that on the right. That h_t is held fixed when the wage increases is a consequence of the envelope theorem. That is, for very small changes in the wage, a consumer can't do better than to spend all the extra income on consumption. Any utility gain that might be achieved by reallocating full income between consumption and leisure is trivially small.

As Blundell and Walker (1986) note, the researcher has great flexibility in choosing the $a(\cdot)$ and $b(\cdot)$ functions, so labor supply can be allowed to depend on the wage and full income in complex ways. This has a downside in terms of interpretability, in that, in contrast to the MaCurdy (1983) specification (equation (3)), elasticities must be simulated. As best as I can ascertain, the specification that Blundell and Walker (1986) estimate for men has the form

$$(53) \quad h_t = (T_m - \gamma_m) - \frac{\beta_m}{w_t^m} \times [F_t - \gamma_m w_t^m - \gamma_f w_t^f - \gamma_c p_t^c - 2\gamma_{fc}(w_t^f p_t^c)^{1/2}].$$

Here w_t^m is the after tax wage for the husband, w_t^f is the after tax wage for the wife, p_t^f is the price of consumption goods and F_t is full income as given by (48).

Blundell and Walker (1986) estimate this model on a sample of families from the 1980 FES. Here I focus on the male labor supply results. Some limitations of the analysis should be noted. First, as I stated earlier, the authors do not instrument for full income or wages. Second, the analysis is limited to families where the household head is a manual worker, shop assistant, or clerical worker, giving 1,378 households (with a female participation rate of 64 percent). The authors do not indicate why they choose to restrict the sample in this way. Third, the consumption measure was limited, including food, clothing, services, and energy but excluding housing, transport, alcohol, and other important categories.

Averaging over the sample, the authors simulate a Frisch elasticity for men of only 0.026 and a Hicks elasticity of 0.024. Similar small values are obtained for all demographic subgroups examined. The authors report an elasticity of male hours with respect to full income of –0.287. Based on figures reported in the paper, I calculate that full income is £267 per week on average, and that male after-tax earnings is $(2.08)(39.8) = £82.78$ per week on average. These figures imply an income effect of $(wh/F)[(F/h)(\partial h/\partial F)] = -0.089$[62] and a Marshallian elasticity of –0.065. It is notable that the authors use a ratio wage measure (i.e., usual earnings over usual hours) to construct wage rates. As discussed earlier, this may lead to downward bias in elasticity estimates, particularly if no instrument is used to correct for measurement error. This may account in part their low elasticity estimates.

6.2.2 *Progressive Taxation in the Life-Cycle Model—Ziliak and Kniesner (1999)*

Blomquist (1985) noted that the "life-cycle" consistent approach has a problem in contexts with progressive taxation. The basic idea is that an increase in labor supply in period t, holding consumption fixed, will cause a person to have more assets at the end of period t. This, in turn, leads to higher asset income in period $t + 1$. And this in turn may increase the person's tax bracket at $t + 1$. Hence, we no longer achieve the simplification that, *conditional* on full income allocated to time t, we can model a person's time t decisions *as if* he/she were choosing labor supply subject to a one-period budget constraint.

Ziliak and Kniesner (1999) attempt to deal with the problem of progressive taxation within the two-stage budgeting framework. Following Blomquist (1985), they note that the two-stage approach can be salvaged by writing labor supply in period t as conditional on assets at both the start and end of the period. The basic idea is that, by holding end of period assets fixed, one shuts down

[62] Consistent with this calculation, when the authors simulate a £50 reduction in nonlabor income it leads to an average 2.5 hour increase in weekly hours for males, implying a derivative of about $-2.5/50 = -0.05$. Multiplying this by the mean male after-tax male wage rate of 2.08 gives an income effect of about –0.10.

any channel by which increased labor supply at t might affect the budget constraint at $t+1$. Thus, they estimate a linear labor supply equation of the form

$$(54)\quad h_{it} = \beta + \beta_w w_{it}(1-\tau_{it}) + \beta_{A1}A^*_{i,t-1} + \beta_{A2}A_{i,t} + X_{it}\alpha + \mu_i + \varepsilon_{it}.$$

In this equation, A^*_{t-1} is "virtual wealth," which plays a role analogous to virtual nonlabor income in static piecewise-linear budget constraint models (see figure 1). It is defined as

$$(55)\quad A^*_{t-1} = A_{t-1} + \frac{(\tau_{it} - \tau^A_{it})I_{it}}{r_t},$$

where τ^A_{it} is the average tax rate paid by person i in period t on their income I_{it}, and r_t is the risk free rate of interest. Note that A^* can be interpreted as a virtual asset level.[63]

Ziliak and Kniesner (1999) estimate (54) using data on 532 married men from the PSID. They were 22–51 years old in 1978 and worked in every year from 1978 to 1987. The asset measure is home equity plus the capitalized value of rent, interest, and dividend income. In constructing the wage measure, Ziliak and Kniesner seek to avoid denominator bias by using the hourly wage rate for hourly workers. For workers paid weekly, they divide weekly earnings by 40 hours rather than actually observed hours (and so on for workers paid over other time periods). This procedure avoids denominator bias at the cost of introducing a different type of measurement error. In forming taxable income and marginal tax rates, the authors use information in the PSID to estimate standard and itemized deductions.[64] Observed taste shifters included in X_{it} are age, health, and number of children.

Equation (54) contains three endogenous variables: the after-tax wage, end of period assets, and start of period virtual assets. All three variables may be correlated with the individual fixed effect μ_i (i.e., a person with high tastes for work will tend to have both a high wage and high asset levels). Thus, as a first step toward estimating (54), the authors take first differences to eliminate the individual fixed effect μ_i

$$(56)\quad \Delta h_{it} = \beta_w \Delta w_{it}(1-\tau_{it}) + \beta_{A1}\Delta A^*_{i,t-1} + \beta_{A2}\Delta A_{it} + \Delta X_{it}\alpha + \Delta\varepsilon_{it}.$$

Now, $w_{it}(1-\tau_{it})$ and A_{it} are presumably correlated with ε_{it}, as a high taste for work at t will tend to both (i) shift a person into a higher tax bracket and (ii) lead to higher assets at the end of the period. Furthermore, start of period virtual wealth A^*_{t-1} is also correlated with ε_{it}. If a high ε_{it} tends to shift a worker into a higher tax bracket at time t, then it affects A^*_{t-1} as shown in (55). Thus valid instruments for estimation of (54) must be uncorrelated with both ε_{it} and ε_{it-1}. Ziliak and Kniesner (1999) argue one must lag the wage and asset variables by two periods (i.e., $w_{it-2}(1-\tau_{it-2})$, A_{it-2} and A^*_{it-3}) to obtain instruments uncorrelated with ε_{it-1}. They also use a quadratic in age, age/education interactions, and home ownership as additional instruments.

The main estimation results imply a Marshallian elasticity evaluated at the mean of the data of $(w/h)\partial h/\partial w = (10.19/2179) \times (24.66) = 0.1153$, and a very small income

[63] The quantity $(\tau - \tau^A)I/r$ is the hypothetical amount of extra assets needed to get the person up to the virtual nonlabor income level associated with their budget segment.

[64] The authors use Internal Revenue Service data to calculate the average level of itemized deductions for a person's income level. Starting in 1984, the PSID asks whether or not a person itemized, and nonitemizers are assigned the standard deduction. Prior to 1984, the authors assign either the standard or itemized deduction, whichever is larger.

effect of $w_t \partial h_t / \partial A_{t-1} = (10.19)(-0.00162) = -0.0165$. Thus, the Hicks elasticity is 0.1318. In a second stage, which I describe in detail below, they estimate the Frisch elasticity as 0.163.[65]

Ziliak and Kniesner (1999) go on to use their model to simulate the impact of various tax reform experiments. The average marginal tax rate in their data was 29 percent. One experiment simulates an across the board 10 percent rate cut by the United States in 1987. The authors simulate that this would increase average annual hours by thirteen hours (0.6 percent). This small effect is not surprising given the Marshallian elasticity of 0.12. They also simulate the 1986 tax reform act (TRA86) that substantially reduced progressivity. As we have seen, it is the Hicks elasticity, which they estimate to be 0.13, that is relevant for determining the efficiency effects of changing progressivity. They simulate only a 2 percent hours increase, but a substantial efficiency gain.

Recall that Blomquist (1983) and Blomquist and Hansson-Brusewitz (1990) found Hicks elasticities of only 0.11 and 0.13 respectively, but still found large efficiency gains from switching to a flat-rate tax. So all three of these papers are similar in finding that a modest Hicks elasticity can imply substantial efficiency gains from reducing progressivity.

6.2.3 *Nonseparability between Consumption and Leisure—Ziliak and Kniesner (2005)*

Ziliak and Kniesner (2005) address the issue of nonseparablility between leisure and consumption, which until now I have largely ignored. To further explore the implications of nonseparability, suppose we modify the utility function in equation (3) to read

$$(57)\quad U_t = G\left[\frac{C_t^{1+\eta}}{1+\eta} - \beta_t \frac{h_t^{1+\gamma}}{1+\gamma}\right]$$
$$\eta \leq 0,\ \gamma \geq 0$$

Assume that $G[\cdot]$ is a concave function, such as $G[X] = \log(X)$ or $G[X] = (1+\sigma)^{-1}X^{1+\sigma}$ for $\sigma \leq 0$. Notice that now the marginal utility of consumption is given by

$$(58)\quad \frac{\partial U_t}{\partial C_t} = G_t'(X_t)\cdot C_t^{\eta} \equiv \lambda_t,$$
$$\text{where}\quad X_t \equiv \frac{C_t^{1+\eta}}{1+\eta} - \beta_t \frac{h_t^{1+\gamma}}{1+\gamma}.$$

Thus, unlike in (3), the marginal utility of consumption is not a function of consumption alone. It also depends on X_t which is a composite of consumption and hours of work. Note that, for a given level of consumption, X_t is decreasing in h_t. And G is concave, so $G_t'(X_t)$ is increasing in h_t. Thus, ceteris paribus, the consumer would like to allocate *more* consumption to periods when hours are high. That is, hours and consumption are compliments.

The consumer still seeks to satisfy an intertemporal optimality condition like (23) but, with the new expression for marginal utility of consumption in (58), we revise (23) to obtain

$$(59)\quad G_t'(X_{it})[C_{it}]^{\eta}$$
$$= E_t\,\rho(1+r)\{G_t'(X_{i,t+1})[C_{i,t+1}]^{\eta}\}\quad \eta \leq 0.$$

Note that now, even if $\rho(1+r) = 1$, the consumer will not seek to equalize consumption across periods. As indicated above, with G concave the consumer will seek to make consumption higher when hours are high. But of course, in the life-cycle model, hours are

[65] The authors show that if they use a ratio wage measure (annual earnings over annual hours) and apply exactly the same estimation procedure, they obtain a Marshallian elasticity of –0.083 and a Hicks elasticity of –0.072. This highlights the potentially severe bias created by use of ratio wage measures that I discussed earlier.

high when wages are high. So the consumer will seek to make consumption high when earnings are high. Thus, if G is sufficiently concave, the life-cycle model can generate consumption and earnings paths that look very much like liquidity constrained behavior!

Now consider how the MRS condition (22) is altered by the utility function in (57)

$$(60)\qquad \frac{\partial V}{\partial h_t} = \{G'(X_t)C_t^{\eta}\}w_t(1-\tau_t) - \{G'(X_t)\beta_t h_t^{\gamma}\} = 0$$

$$\Rightarrow C_t^{\eta} w_t(1-\tau_t) = \beta_t h_t^{\gamma}.$$

That is, it doesn't change at all. The factor $G_t'(X_t)$ appears in both the marginal utility of consumption and the marginal utility of leisure, and so it cancels out. This point was made by MaCurdy (1983): the G function does not affect within period decisions about work and consumption, so estimation of the MRS condition tells us nothing about the form of G.

Thus MaCurdy (1983), in his method #1, proposed to estimate the form of G in a second stage. The first stage (discussed earlier) uses the MRS condition to obtain estimates of the parameters of the X_t function (i.e., γ and η in our example (58)). One can then use these estimates, along with hours and consumption data, to construct estimates of the X_t. One then treats these estimates of the X_t as data. In the second stage, one uses data on X_t, C_t, and r_t from multiple periods to estimate the unknown parameters of (59). These include the discount rate ρ and the parameters of G. For instance, if we assume $G[X] = (1+\sigma)^{-1}X^{1+\sigma}$ the only parameter of G is σ. In his study, MaCurdy (1983) estimated $\sigma = -0.14$ but with a standard error of 0.23. Thus, he couldn't reject the simple linear G case ($\sigma = 0$).

We now seek more insight into the impact of G on the Frisch elasticity. Substituting the first order condition for consumption (58), which states that $\lambda_t = G_t'(X_t)\cdot C_t^{\eta}$, into the first order condition for hours in (60), we obtain $\lambda_t w_t(1-\tau_t) = G'(X_t)\beta_t h_t^{\gamma}$. Taking logs, we have

$$(61)\quad \ln h_t = \frac{1}{\gamma}\Big\{\ln w_t + \ln(1-\tau_t) + \ln\lambda_t - \ln G'(X_t) - \ln\beta_t\Big\}.$$

So the Frisch elasticity—the effect of a wage change holding marginal utility of consumption λ_t fixed—is no longer simply $(1/\gamma)$, because in general a change in w_t will affect $G'(X_t)$.

To explore further how a concave G affects willingness to substitute labor across periods, let's assume $G[X] = (1+\sigma)^{-1}X^{1+\sigma}$ so that (61) becomes

$$(61')\quad \ln h_t = \frac{1}{\gamma}\{\ln w_t + \ln(1-\tau_t) + \ln\lambda_t - \sigma\ln X_t - \ln\beta_t\}.$$

Clearly, the Frisch elasticity (i.e., the elasticity of hours with respect to the wage holding λ_t fixed) is not simply $(1/\gamma)$ in this case because we have to consider the $\ln X_t$ term, and X_t contains C_t and h_t. The exception of course is if $\sigma = 0$ so the $\ln X_t$ term drops out.

To determine what (61′) implies about the Frisch elasticity, we can use the within period MRS condition in (60) to substitute out for consumption in X_t, obtaining an expression for X_t solely in terms of hours. Then (61′) becomes an implicit equation that relates hours and the wage, holding λ_t fixed. Implicit differentiation of this equation gives

$$(62)\quad e_F = \left.\frac{\partial \ln h_t}{\partial \ln w_t}\right|_{\lambda_t\, fixed} = \frac{1}{\gamma}\left\{\frac{X_t + \left(\frac{\sigma}{\eta}\right)C_t^{1+\eta}}{X_t + \left(\frac{\sigma}{\eta}\right)C_t^{1+\eta} - \sigma\left(\frac{\beta}{\gamma}\right)h_t^{1+\gamma}}\right\}.$$

TABLE 4
How Frisch Elasticity Varies with Willingness to Substitute Utility Over Time

σ	Frisch elasticity	Changes in hours		Changes in consumption		Changes in utility	
		Hours(1)	Hours(2)	C(1)	C(2)	G(X(1))	G(X(2))
0.0	2.00	+1.03%	−0.96%	+0.97%	+0.97%	−0.05%	+1.44%
−0.5	1.40	+0.82%	−0.58%	+1.18%	+0.58%	+0.27%	+0.87%
−1.0	1.25	+0.76%	−0.48%	+1.24%	+0.48%	+0.38%	+0.72%
−2.0	1.14	+0.73%	−0.41%	+1.27%	+0.42%	+0.41%	+0.62%
−5.0	1.06	+0.70%	−0.36%	+1.30%	+0.36%	+0.45%	+0.54%
−10.0	1.03	+0.69%	−0.34%	+1.31%	+0.34%	+0.46%	+0.51%
−40.0	1.01	+0.68%	−0.33%	+1.32%	+0.33%	+0.48%	+0.49%

If $\sigma = 0$, then (62) reduces to just $(1/\gamma)$ as we would expect. However, as $\sigma \to -\infty$ the fraction in brackets becomes less than one (as the term $-\sigma(\beta/\gamma)h_t^{1+\gamma}$ in the denominator is positive). So complimentarity between work and consumption reduces the Frisch elasticity below $(1/\gamma)$.

Numerical simulations of a simple two-period model based on (57)–(61′) reveal a lot about how σ influences behavior, and give a clear intuition for why the Frisch elasticity falls as $\sigma \to -\infty$. I start from a base case where $w_t = h_t = 100$ in both periods, and $\tau = 40$ percent. I set $\rho(1 + r) = 1$ so consumption is 6,000 in both periods. In a two period model where each period corresponds to roughly twenty years of a working life, a plausible value for $1 + r$ is about $(1 + 0.03)^{20} \approx 1.806$, so $\rho = (1 + r)^{-1} \approx 0.554$. The utility function parameters are set to $\gamma = 0.5$ and $\eta = -0.5$.

Then, from (21), the Marshallian elasticity, which indicates how hours respond to a permanent (i.e., two-period) wage change, is 0.5. The Hicks elasticity, where this permanent wage change is compensated by a lump sum transfer, is 1.0. From (60) these elasticities are invariant to σ. In contrast, the Frisch elasticity is $(1/\gamma) = 2.0$ if $\sigma = 0$, but it varies with σ.

In table 4, I simulate a 1 percent after-tax wage increase at $t = 1$ (from 60 to 60.6) induced by cutting the tax rate from 0.40 to 0.394. Results are shown for σ ranging from 0 to –40. When $\sigma = 0$ the worker increases hours at $t = 1$ by 1.03 percent and reduces hours at $t = 2$ by 0.96 percent. Thus, labor supply increases by 2 percent at $t = 1$ relative to $t = 2$, as expected given a Frisch elasticity of 2. Note also that the consumer continues to smooth his/her consumption over time: consumption increases by 0.97 percent in both periods.[66] The consequence is that utility actually falls slightly in period 1 while rising by 1.44 percent in period 2.

As σ decreases, the consumer is less willing to sacrifice utility at $t = 1$ to achieve higher utility at $t = 2$. Given $\sigma = -40$, the consumer is very unwilling to substitute utility across periods. Note how he/she allocates consumption/hours so utility increases by 0.48 percent at $t = 1$ and 0.49 percent at $t = 2$. To achieve this balance, the worker shifts consumption into period 1 to compensate for having to

[66] Earnings increase by about 2 percent in period 1 and drop by about 1 percent in period 2. This causes the present value of lifetime earnings (and hence of consumption) to increase by $[1.02 + (0.554)(0.99)]/1.554 \approx 1.0097$ or 0.97 percent.

work more hours at $t = 1$ (i.e., consumption increases by 1.32 percent at $t = 1$ versus only 0.33 percent at $t = 2$). The worker also shifts less labor supply toward $t = 1$ than in the $\sigma = 0$ case. Now hours only increase 0.68 percent at $t = 1$ and only fall 0.33 percent at $t = 2$. This implies a Frisch elasticity of 1.01. As $\sigma \to -\infty$ we get a Leontieff utility function where the consumer only cares about maximizing the minimum utility in any period, and the Frisch elasticity approaches 1.0. This is exactly the Hicks elasticity for a permanent (two-period) wage change.[67]

To summarize: In the linear case $(G(X_t) = X_t)$, and with X_t additive in consumption and hours, there is a separation of the labor supply and consumption problems. The worker shifts labor supply toward high wage periods while using savings to smooth consumption across periods. This means sacrificing utility in the high-wage periods.

But if G is concave, the worker/consumer tries to equalize utility across periods by shifting consumption into high-wage/high-hours periods. Thus, the consumer's willingness to substitute consumption intertemporally puts a damper on his/her willingness to substitute labor. This tends to reduce intertemporal substitution in labor supply. As a result, the Frisch elasticity is less than $(1/\gamma)$ and in the limit (as $\sigma \to -\infty$) it equals the Hicks elasticity.

It is worth recalling that the within period MRS condition (41) still holds regardless of the form of G. So one might estimate equations like (47) to uncover $(1/\gamma)$, and fail to realize it is not the Frisch elasticity. One can still use (47) to obtain the Hicks and Marshallian however.

Returning to the empirical literature, Ziliak and Kniesner (2005) adopt MaCurdy's method #1, and allow for nonseparablity between leisure and consumption, both via the G function and by including an interaction between leisure and consumption in the within period utility function X_t. Specifically, they adopt a translog within-period utility function

$$(63)\quad U_t = G\Big[\alpha_1 \ln(\overline{L} - h_t) + \alpha_2 \ln C_t - \alpha_3 \ln(\overline{L} - h_t)\ln C_t - \alpha_4 \times [\ln(\overline{L} - h_t)]^2 - [\ln C_t]^2\Big],$$

with $G[X] = (1 + \sigma)^{-1}X^{1+\sigma}$. If $\alpha_3 > 0$, hours and consumption are compliments in X_t. This appears to be the only paper allowing for within period nonseparability and taxes in a dynamic model.

As in most U.S.-based work that uses the within period MRS condition, Ziliak and Kniesner (2005) use the PSID, which only contains food consumption.[68] However, they try to improve on this using a method proposed by Blundell, Pistaferri, and Ian Preston (2002) to impute nondurable consumption. Using the Consumer Expenditure Survey, which has much more complete consumption information, they develop an

[67] Notice, that if we take the limit of (62) as $\sigma \to -\infty$, we get that

$$e_F \to \frac{1}{\gamma}\left\{\frac{\left(\frac{\sigma}{\eta}\right)C_t^{1+\eta}}{\left(\frac{\sigma}{\eta}\right)C_t^{1+\eta} - \sigma\left(\frac{\beta}{\gamma}\right)h_t^{1+\gamma}}\right\} = \frac{1}{\gamma}\left\{\frac{C_t^{1+\eta}}{C_t^{1+\eta} - \eta\left(\frac{\beta}{\gamma}\right)h_t^{1+\gamma}}\right\}.$$

For the parameter values in the simulation, the term in curly brackets is equal to 1/2. But if $\eta = 0$ (i.e., no income effects) the term in brackets is exactly equal to 1. So, if utility is linear in consumption, σ has no impact on the Frisch elasticity. This is because the consumer is willing to equalize utility across periods via consumption shifting alone, leaving him/her free to substitute hours of work toward high wage periods as much as if $\sigma = 0$. Ironically, a high elasticity of substitution in consumption ($\eta \approx 0$) combined with curvature in G makes the consumer behave in a way that looks like liquidity-constrained behavior (consumption closely tracks income). Finally, when $\eta = 0$, the Frisch $(1/\gamma)$ and Hicks $1/(\gamma - \eta)$ elasticities are exactly equal (regardless of the value of σ).

[68] Recall that the Denver data used by MaCurdy (1983) had a very comprehensive consumption measure.

equation to predict nondurable consumption based on food consumption, food prices and demographics. They also try a method proposed by Jonathan Skinner (1987) to predict total consumption using food consumption, house value and rent.

Ziliak and Kniesner use the 1980–99 waves of the PSID. One advantage over prior work is the long sample period, which encompasses five significant tax law changes. This provides variation in the budget constraint to help identify utility function parameters.[69] The sample includes 3,402 male household heads who were at least 25 in 1980, no older than 60 in 1999 and who are observed for at least five years. They use the hourly wage rate question for workers paid by the hour, and, in an effort to reduce denominator bias, for salaried workers they use the same procedure of hours bracketing as in their 1999 paper discussed earlier.

A key challenge in incorporating taxes is to estimate taxable income. The authors assume all married men filed joint returns while unmarried men filed as heads of households (the latter is the more likely source of error). The income of working wives is included when calculating adjusted gross income (AGI). Deductions are estimated using Internal Revenue Service estimates of the average levels of itemized deductions by AGI. From 1984 onward, the PSID reports whether a person itemized or took the standard deduction. Following MaCurdy, Green, and Paarsch (1990), the authors use a smooth approximation to the piecewise linear tax schedule.

The parameters α_1 and α_2 in (63) are allowed to depend on children, race, and age of youngest child to capture how these variables may shift tastes for work and consumption. Besides these, the instruments used to estimate the MRS equation, which should be correlated with after-tax wages and consumption but uncorrelated with unobserved tastes for work, are age, education, health, home ownership, and industry, occupation and region dummies.

The authors estimate $\alpha_3 > 0$, implying hours and consumption are compliments in the within period utility function (i.e., if hours are higher the marginal utility of consumption is higher). Ziliak and Kniesner let σ vary with age and obtain $\sigma = 0.844 - 0.039 \cdot Age$. So σ is about zero at age 20 and falls to –1.5 at age 60.[70] But the age effect is imprecisely estimated.

Given the translog within period utility function in (63) there is no closed form for the Marshallian and Hicks elasticities. At the mean of the data, the authors calculate a Marshallian elasticity of –0.468 (standard error 0.098) and a Hicks elasticity of 0.328 (standard error 0.064). This implies a very large income effect (–0.796) that is comparable to values obtained by Hausman (1981) and Wales and Woodland (1979). In the second stage, incorporating information from the intertemporal condition (59), they obtain a Frisch elasticity of 0.535. If they restrict $\alpha_3 = 0$ they obtain Marshallian, Hicks, and Frisch elasticities of –0.157, 0.652, and 1.004 respectively. Thus, ignoring the complimentarity between work hours and consumption

69 The Economic Recovery Tax Act of 1981, the Tax Reform Act of 1986, the Omnibus Reconciliation Tax Acts of 1990 and 1993, and the Taxpayer Relief Act of 1997. Notably, however, while these tax changes are in the data, the authors do not fully exploit it by using aspects of the tax rules as instruments (see also section 6.2.8).

70 It is difficult to conceptualize what it means for σ to vary with age, given that σ governs a person's willingness to substitute utility over time. Does a 20 year old with $\sigma \approx 0$ solve his/her lifetime planning problem as if he/she is very willing to substitute utility intertemporally and then engage in replanning each year as his/her σ drops? Or does a person plan his/her life knowing that σ will fall over time? If so, exactly how does one do that? Does a person take preferences of his/her future selves into account? Apparently, one can circumvent such questions in estimating intertemporal conditions like (59). But they would have to be confronted to obtain a full solution of a person's lifetime optimization problem. I discuss models that involve full solutions in section 6.3.

TABLE 5
LABOR SUPPLY ELASTICITIES BASED ON ALTERNATIVE CONSUMPTION MEASURES

Consumption measure	Marshall	Hicks	Income effect	Frisch
Blundell et al (2001)	−0.468	0.328	−0.796	0.535
Skinner (1987)	−0.313	0.220	−0.533	0.246
PSID unadjusted	−0.442	0.094	−0.536	0.148

appears to cause upward bias in all three labor supply elasticities.[71]

Interestingly, Ziliak and Kniesner (2005) examine how their results are affected by the use of different consumption measures. The comparison is shown in table 5. Clearly, Hicks and Frisch elasticity estimates are *very* sensitive to the consumption measure used (while the Marshallian is relatively insensitive). In section 6.1, I noted how Eklöf and Sacklén (2000) found that elasticity estimates in static models are quite sensitive to the wage and nonlabor income measures. Here we see evidence the same is true of the consumption measure used in life-cycle models (for estimation methods based on the MRS condition).

Finally, the authors use the estimated Hicks elasticity ($e_H = 0.328$) to calculate the marginal efficiency cost of tax increases that raise all tax rates proportionately. This turns out to be 16 percent of the revenue raised. However, if they do the same calculation using the Hicks elasticity obtained using the PSID unadjusted food consumption measure ($e_H = 0.094$), the efficiency loss is only 5 percent of the revenue raised.

A limitation of this study, as well as Ziliak and Kniesner (1999), is the measurement of taxable income. Recall (section 6.1.4) how Blomquist (1996) found that imputing tax deductions, and the resultant measurement error, may have large effects on estimates.

6.2.4 *Methods for Estimating the Frisch Elasticity—MaCurdy (1981)*

Starting with MaCurdy (1981), a number of studies have attempted to use equations similar to (25) from section 3.2 to estimate the Frisch elasticity directly. In order to put (25) in a form amenable to estimation, we use (43) to substitute for the taste shifter β_{it} and obtain

$$(64)\quad \Delta \ln h_{it} = \frac{1}{\gamma}\Delta \ln w_{it}(1-\tau_{it}) - \frac{1}{\gamma}\ln \rho(1+r_t) - \frac{\alpha}{\gamma}\Delta X_{it} + \frac{1}{\gamma}\xi_{it} + \frac{1}{\gamma}\Delta\varepsilon_{it}.$$

In contrast to MaCurdy's methods #1 and #2, estimation of (54) does not require consumption data. However, it does require hours and wage data for at least two periods, and it will only deliver an estimate of γ and not of η.

A key point is that the error term in (64) consists of two components. The first is the surprise change in the marginal utility of consumption, multiplied by the factor $(1/\gamma)$,

[71] I believe the intuition for this result is as follows: If work and consumption are compliments within a period, then a wage increase affects hours through three channels. There are the usual substitution and income effects. But in addition, a wage increase will, ceteris paribus, increase consumption. This reduces the marginal utility of leisure at the initial hours level, giving an additional reason for hours to increase. As a result, a smaller substitution effect is required to explain any given level of responsiveness of hours to wages.

which I denote by $\zeta_{it} = (1/\gamma)\xi_{it}$. This arises in part due to surprise wage growth. Surprise wage growth makes a person wealthier, creating a negative income effect. So long as wage growth contains a surprise component we have $\text{Cov}(\xi_{it}, \Delta w_{it}) < 0$, and wage growth is endogenous in (64). Of course, errors in forecasting wage growth arise due to new information revealed between $t-1$ and t (e.g., unexpected recession, illness, plant closure). Thus, valid instruments for estimation of (64) should have the property that they were in the agent's information set at time $t - 1$, so they could have been used to forecast wage growth.[72]

The second component of the error term arises from the change in tastes for work from $t - 1$ to t, which I denote by $\Delta\varepsilon^*_{it} = (1/\gamma)\Delta\varepsilon_{it}$. I have already discussed at length in section 4 the multiple reasons why we would expect tastes for work to be correlated with both gross and after-tax wages. Thus, valid instruments for estimation of (64) should also have the property that they are uncorrelated with changes in tastes for work.[73]

MaCurdy (1981) estimates (64) using annual data on 513 married men observed from 1967–76 in the PSID. To be included in the sample, they must have been 25 to 46 years old in 1967 and married to the same spouse during the sample period. MaCurdy uses a complete set of time dummies to pick up the log interest rate terms in (64), rather than using a particular interest rate variable. No observed taste shifter variables X are included.[74] It is notable that MaCurdy does not adjust wages for taxes.[75] Hence, the specification is simply a regression of annual log hours changes on log wage changes, along with a set of time dummies.

MaCurdy presents his analysis in a setting where workers have perfect foresight about future wages. But, as he notes, his results extend to the uncertainty case provided he uses as instruments for wages variables that were known to a worker at time t or before, so that the worker could have used these variables to forecast wage growth from time $t - 1$ to t. If workers forecast wage growth rationally, such instruments are uncorrelated with the forecast error ξ_{it}. The instruments MaCurdy uses to predict wage growth are by now familiar: quadratics in age and education as well as age/education interactions, parental education and year dummies.

To gain intuition for how this procedure identifies the Frisch elasticity, it is useful to consider a two-stage least squares (2SLS) approach. In stage one, regress wage growth on functions of age and education to obtain predicted wage growth over the life-cycle. In stage two, regress hours growth on predicted wage growth. The coefficient on predicted wage growth captures how hours respond to *predictable* variation of wages over the life-cycle—i.e., the extent to which people substitute their time intertemporally and allocate more work hours to periods when wages are relatively high. This is exactly the Frisch elasticity concept.

Using this approach, MaCurdy (1981) obtained a Frisch elasticity of only 0.15, with a standard error of 0.98. It is striking

72 This idea of using variables that economic agents use to make forecasts as instruments in dynamic models originated in work by Bennett T. McCallum (1976) and Thomas J. Sargent (1978).

73 In section 4, I noted that many authors adopt a view that tastes for work can be decomposed into a permanent component that may be correlated with wages, and idiosyncratic shocks that are exogenous. In that case, the first differencing in (64) eliminates the permanent component. It may be more plausible that certain instruments are uncorrelated with *changes* in tastes for work than with tastes in levels. Education is a good example.

74 Taste shifters often included in X in other studies include age, number and ages of children, marital status, etc.

75 It may be argued that taxes will largely drop out of (64) if the marginal tax rate a person faces does not change too much from year to year. Altonji (1986) makes this argument explicitly.

to compare this to the Frisch elasticity of 6.25 obtained by MaCurdy (1983) using the DIME data, where he adopted the "MRS approach" of using consumption to proxy the marginal utility of wealth (see section 6.2.1). But, as Altonji (1986) and Blundell and Walker (1986) obtained estimates of 0.172 and 0.026, respectively, using similar consumption-based approaches, the high Frisch elasticity in MaCurdy (1983) starts to look like an extreme outlier. Furthermore, as the Frisch is in theory an upper bound on the Hicks and Marshallian, this would lead one to conclude labor supply elasticities are small for men in general.

But before reaching this conclusion, it is important to keep two points in mind: One is the large standard error (0.98) on MaCurdy's estimate. This suggests instruments such as age and education do a very poor job of predicting wage *changes*. Second, taking the change in wages as in (64) exacerbates measurement error problems, which may bias the wage coefficient toward zero. These issues are related, as one needs good predictors of true wage changes to correct the measurement error problem. I'll explore various attempts to improve upon MaCurdy (1981) in the next few sections (sections 6.2.5–6.2.8).

6.2.5 *Attempts to Deal with Measurement Error and Weak Instruments*

Altonji (1986) tried to address the measurement error and weak instrument problems by using a better instrument for wage changes. As I noted earlier, he uses two wage measures from the PSID, one serving as the wage measure in the labor supply equation, the other serving as an instrument. Using a PSID sample similar to MaCurdy (1981) and using similar predictors of wage changes (quadratic in age and education, etc.), he gets an R-squared in the first stage prediction equation of only 0.008. Then, in estimating (64) he gets a Frisch elasticity of 0.31, with a large standard error of 0.65. But when Altonji uses his alternative wage change measure as an additional instrument, he gets a much better R-squared of 0.031 in the first stage equation.[76] Then, in estimating the labor supply equation, he gets a Frisch elasticity of 0.043, with a standard error of only 0.079. Thus, we seem to have a rather tight estimate of a small Frisch elasticity. The problem here, of course, is that the alternative wage change measure is only a valid instrument under the strong assumption that workers do have perfect foresight about wage changes. Otherwise, any wage change measure will be correlated with ξ_{it}.[77]

Angrist (1991) proposes to deal with the measurement error problem using grouped data estimation. That is, he works with the equation

$$(65)\quad \overline{\ln h_{it}} - \overline{\ln h_{i,t-1}} = \frac{1}{\gamma}\left\{\overline{\ln w_{it}(1-\tau_t)} - \overline{\ln w_{i,t-1}(1-\tau_{t-1})}\right\} + f(t) + \overline{\zeta_{it}} + \frac{1}{\gamma}\overline{(\varepsilon_{it} - \varepsilon_{i,t-1})} + e_t,$$

where $\overline{X}_{it}$ denotes the sample mean of X_{it} over all people i observed in year t. The idea is that, while the *individual* log hours and log wages are measured with error, this will largely cancel out when we average over people.[78] The additional error term e_t arises from error in calculating true means of log wages and hours using a finite sample.

[76] This may still seem small, but is actually not bad given the large sample size of roughly 4,000 observations, as indicated by the highly significant F statistic of 129.

[77] Given this problem, Altonji tried using the lagged wage change as an instrument, as the lag would have been known at time t. But it is a poor predictor, and the standard error jumps to 0.45.

[78] Of course, this requires that the measurement error in log wages (and in log hours) is additive.

Notice that I have substituted a function of time $f(t)$ for the interest rate variable in (64). MaCurdy (1981) and Altonji (1986) both used a complete set of year dummies to pick up interest rates. But that will not work here because a complete set of year dummies would enable one to fit changes in average hours perfectly and the Frisch elasticity $(1/\gamma)$ would not be identified. Identification requires that $f(t)$ be specified as a low order polynomial in time.

Now, it is important to consider whether estimation of (65) actually uncovers labor supply parameters, or some mongrel of supply and demand factors. For estimation of (65) to identify the Frisch elasticity, we require that the variation in average wages be induced by *anticipated* shifts in labor demand (e.g., anticipated productivity growth).

But the error term in (65) includes the mean of the aggregate surprise variable $\overline{\zeta}_{it}$. This aggregate surprise term would arise from *unexpected aggregate productivity shocks*. Of course, these aggregate shocks would alter the average wage, and induce income effects on aggregate labor supply. Thus, for estimation of (65) to identify the Frisch elasticity, $f(t)$ must capture such unexpected aggregate shocks, so that $\overline{\zeta}_{it}$ drops out.

The error term in (65) also includes $\overline{(\varepsilon_{it} - \varepsilon_{i,t-1})}$, the average change in tastes for work. If there are aggregate taste shocks, we would expect a negative correlation between $\overline{(\varepsilon_{it} - \varepsilon_{i,t-1})}$ and the change in average wages (as increased labor supply would drive down wages in equilibrium). So we must assume aggregate taste shocks are captured by the time polynomial $f(t)$ as well. In summary, for (65) to represent a supply equation, it is necessary that $f(t)$ adequately captures aggregate taste and productivity shocks. [An alternative is to use demand side instruments that predict productivity growth, but Angrist does not pursue this route.]

With these caveats in mind, let's consider Angrist's results. He uses PSID data from 1969–79, and takes a sample of 1,437 male household heads aged 21 to 64 with positive hours and earnings in each year. He constructs average log hours and log wages for the sample members in each year, and uses these to estimate (65). If the trend term $f(t)$ is left out, the estimate of the Frisch elasticity is –0.132 (s.e. = 0.042), which violates economic theory. This clearly occurs because in 1969–79 there was a secular downward trend in average male hours and a secular upward trend in wages. When a linear trend is included, it picks this up and Angrist obtains a Frisch elasticity of 0.556 (s.e. = 0.124).[79] Using a quadratic trend, he obtains 0.634 (s.e. = 0.205). Specification tests do not reject the linear trend model, although the test may have little power given the small sample size.

Regardless, these estimates provide some evidence for higher values of the Frisch elasticity than results from most of the prior literature would suggest. However, it is unclear if Angrist obtains the higher value because of a superior method of handling measurement error or because the results are contaminated by aggregate shocks (not captured by $f(t)$) that induce a positive correlation between wages and hours.

A related paper is by Martin Browning, Deaton, and Margaret Irish (1985), who show how one can estimate the Frisch elasticity using repeated cross section data instead of true panel data. First, they show how to derive a version of the Frisch labor supply function that has the hours change in levels (not logs) as the dependent variable, giving an equation of the form

[79] It is interesting that when Angrist simply estimates (65) on the micro data, using a linear trend, he obtains a Frisch elasticity of –0.267 (s.e. = 0.008). But when he estimates the hours equation in levels, he obtains –0.063 (s.e. = 0.005). This illustrates how first differencing exacerbates the downward bias in the wage coefficient.

$$(66)\quad h_{it} - h_{i,t-1} = \beta\{\ln w_{it} - \ln w_{i,t-1}\} - \beta \ln \rho(1 + r_{it}) - \alpha\{X_{it} - X_{i,t-1}\} + \zeta_{it} + (\varepsilon_{it} - \varepsilon_{i,t-1}).$$

To estimate (66), the authors use data on married men from the FES for the seven years from 1970–76. The FES does not track individuals through time. Rather it takes a random sample of the population in each year. Thus, it is not possible to take first differences like $h_{it} - h_{i,t-1}$ for individual people i. Instead Browning, Deaton, and Irish construct eight cohorts from the data: men who were 18–23 in 1970, men who were 24–28 in 1970, up to men who were 54–58 in 1970. (Note: members of the first cohort are 24–29 in 1976 when the data ends, while members of the last cohort are 60–64. Thus, the data cover all ages from 18 to 64.) The authors then take cohort specific means of each variable in (66) for each year of data. This gives

$$(67)\quad h_{ct} - h_{c,t-1} = \beta\{\ln w_{ct} - \ln w_{c,t-1}\} - \beta \ln \rho(1 + r_t) - \alpha\{X_{ct} - X_{c,t-1}\} + \zeta_{ct} + (\varepsilon_{ct} - \varepsilon_{c,t-1}).$$

Here, for instance $\ln w_{ct}$ is the mean of the log wage for people in cohort $c, c = 1,\ldots,8$ in year t, $t = 1970,\ldots,76$. Notice that ζ_{ct} arises from the mean of the surprise shock to the marginal utility of wealth for members of cohort c in year t. This may differ among cohorts because different cohorts are affected differently by aggregate shocks in period t. For example, an unexpected recession in year t may lead to larger unexpected wage reductions for younger workers.

Similarly, ε_{ct} is the mean taste shock for cohort c in year t. As I noted earlier, writing the labor supply equation in terms of aggregate or cohort means highlights the potential existence of aggregate shocks. If aggregate taste shocks exist, they would alter equilibrium wages, and (67) would no longer represent a labor supply relationship. To deal with this problem we require "demand side" instruments that generate exogenous variation in wages. And, given the existence of aggregate surprise changes to lifetime wealth (captured by the ζ_{ct}), it is necessary that any instruments we use to predict wage growth be known at time $t - 1$.

Browning, Deaton, and Irish (1985) use time dummies to pick up the aggregate shocks, an approach that is feasible for them (in contrast to Angrist) because they observe multiple cohorts at each point in time. However, this approach assumes aggregate shocks affect all cohorts in the same way, a point raised by Altug and Miller (1990) that I'll return to in section 6.2.6.

To estimate (67), Browning, Deaton, and Irish (1985) use as instruments a quadratic in age along with lagged wages. They include number of children as observed taste shifters in X_{ct}. The wage measure is "normal" weekly salary divided by normal weekly hours, and taxes are not accounted for. The main results, which they report in their table 4, row 4.6, indicate that the Frisch elasticity is very small. The estimate of β in (67) is 0.13, with a standard error of 0.27. Given this functional form, the Frisch elasticity is roughly β/h which is $3.77/43 = 0.09$ at the mean of the data, implying very little intertemporal substitution in labor supply. Indeed, only the year dummies (and, marginally, children) are significant in the equation.

Based on this result, the authors argue "there is a marked synchronization over the life-cycle between hours worked and . . . wage rates . . ." but "the characteristic hump-shaped patterns of . . . hours . . . though explicable in terms of life-cycle wage variation . . . can be explained as well as or better . . . as the response of credit-constrained consumers to the variation in needs accompanying the birth, growth and departure of children." This quote illustrates one of two possible reactions to finding the Frisch elasticity is very small:

One could maintain that the life-cycle model is valid but that preferences are such that people are very unwilling to substitute hours intertemporally (i.e., $\gamma >> 0$). Then, as the Frisch elasticity is an upper bound on the Hicks and Marshallian, one must conclude they are small as well. Alternatively, one could conclude, as do Browning, Deaton, and Irish (1985), that consumers are credit constrained. Then the life-cycle model is invalid, and the static model is in fact appropriate. Thus, the Frisch elasticity is meaningless, and estimates that it is small tell us nothing about possible values of the Hicks or Marshallian. Of course, if we abandon the life-cycle model we need an alternative explanation for variation in assets over the life-cycle.[80]

6.2.6 *The Problem of Aggregate Shocks—Altug and Miller (1990)*

Now let's further consider the issue of aggregate shocks. It is important to note that the presence of aggregate surprise variables like $\overline{\zeta}_{it}$ or ζ_{ct} is not only an issue in studies like Angrist (1991) and Browning, Deaton, and Irish (1985) that work with sample or cohort means. Taking means just makes the issue more salient. The same issue is implicitly present in studies like MaCurdy (1981) and Altonji (1986) that use micro panel data to estimate versions of (64). The potential problems created by aggregate shocks were stressed by Altug and Miller (1990). In particular, they argue that use of time dummies to "soak up" the mean of the aggregate shock in each period may not solve these problems.

Specifically, let $(\zeta_{it} - \overline{\zeta}_{it})$ be the idiosyncratic surprise for household i at time t. Having included time dummies D_t in equation (64), it takes the form

$$\text{(68)}\quad \Delta \ln h_{it} = \frac{1}{\gamma}\Delta \ln w_{it}(1-\tau_t) + D_t - \frac{\alpha}{\gamma}\Delta X_{it} + (\zeta_{it} - \overline{\zeta}_{it}) + \frac{1}{\gamma}(\varepsilon_{it} - \varepsilon_{i,t-1}),$$

where now the error term includes only the idiosyncratic surprise terms $(\zeta_{it} - \overline{\zeta}_{it})$ along with unobserved taste shifters.[81] The key point in Altug and Miller (1990) is that, while $(\zeta_{it} - \overline{\zeta}_{it})$ is mean zero by construction, it may be systematically related to instruments like age and education that are typically used to predict wage growth in this literature.

For example, suppose the sample period contains an adverse productivity shock in year t, but that low education workers are more adversely affected. Thus, low education workers have relatively large values for ζ_{it} (as a *positive* ζ_{it} represents a surprise *negative* shock to lifetime wealth). Letting S_i denote education, we have $\mathrm{Cov}[S_i, (\zeta_{it} - \overline{\zeta}_{it})] < 0$. Now, this would not invalidate education as an instrument if the sample contained some other years where shocks to lifetime wealth tended to favor low education workers. However, the key point is that the sample must consist of a fairly large number of years before we could be confident that such

[80] There is of course a huge parallel literature testing the life-cycle model by looking for evidence of liquidity constraints that prevent people from using assets to smooth consumption over the life-cycle (see. e.g., Keane and Runkle 1992). It is beyond the scope of this survey to discuss that literature, except to mention that whether liquidity constraints are important determinants of savings behavior remains controversial.

[81] Note that for simplicity this discussion ignores the possible existence of aggregate taste shocks. In that event, (68) would be modified to also subtract off the time specific means of those taste shocks.

favorable and unfavorable shocks roughly canceled out.

Altug and Miller (1990) do not seek to "solve" this problem. Instead, they are explicit about the assumptions needed to make it vanish. Specifically, they assume that workers have complete insurance against idiosyncratic shocks, so the $(\zeta_{it} - \overline{\zeta}_{it})$ terms vanish. Of course, all models are abstractions, so we should not dismiss a model just because of an implausible assumption like complete insurance. And, as Altug and Miller argue, with the existence of unemployment insurance, family transfers, etc., the existence of complete insurance, while obviously false, might not be such a bad approximation. The real questions are (1) what does the assumption buy you, and (2) does it severely bias our estimates of parameters of interest?

Now, consider how Altug and Miller (1990) exploit the complete insurance assumption. First, return to equation (23) and make the person subscripts explicit:

$$(23') \quad [C_{it}]^{\eta} \;=\; E_t\, \rho(1 + r_{t+1})[C_{i,t+1}]^{\eta} \quad \eta \leq 0.$$

Recall that $[C_{it}]^{\eta}$ is the marginal utility of consumption at time t, and (23′) describes how it evolves over the life-cycle, given the worker makes optimal consumption/savings decisions. Note that the marginal utility of consumption in period t is equivalent to the "marginal utility of wealth" at time t, which I'll denote λ_{it}. This is the increment in lifetime utility that the consumer can achieve if we give him/her an extra unit of assets (or wealth) at the start of period t (more formally, the Lagrange multiplier on the budget constraint). The equivalence $\lambda_{it} = [C_{it}]^{\eta}$ arises because, for a small increment in wealth at time t, the consumer can't do significantly better than to simply spend it all at once.[82]

Now, to assume away idiosyncratic risk, Altug and Miller (1990) assume that

$$(69) \qquad \lambda_{it} \;=\; \mu_i\, \lambda_t.$$

A person i with a low μ_i has a relatively low marginal utility of wealth, meaning he/she is relatively rich. The person's position in the wealth distribution stays constant over time. Aside from interest rates and discounting, the only source of over-time variation in the marginal utility of wealth is aggregate shocks that cause movements in λ_t.

Given this assumption, Altug and Miller can rewrite (23′) as

$$\begin{aligned}(70) \quad \lambda_{it} \;&=\; E_t\, \rho(1 + r_{t+1})\lambda_{i,t+1}\\ \Rightarrow \quad \mu_i \lambda_{\mathrm{t}} \;&=\; \rho\, \mu_i E_t(1 + r_{t+1})\lambda_{t+1}\\ \Rightarrow \quad \lambda_t \;&=\; \rho E_t(1 + r_{t+1})\lambda_{t+1}.\end{aligned}$$

Then (23) and (24) become

$$\begin{aligned}(71) \quad \rho(1 + r_{t+1})\lambda_{t+1} \;&=\; \lambda_t(1 + \xi_{t+1})\\ \Rightarrow \quad \Delta \ln \lambda_t \;&\approx\; -\ln \rho(1 + r_t) + \xi_t.\end{aligned}$$

Notice also that, using (71), equation (64) becomes

$$\begin{aligned}(72) \quad \Delta \ln h_{it} \;=\; &\frac{1}{\gamma}\Delta \ln w_{it}(1 - \tau_{it})\\ &- \frac{\alpha}{\gamma}\Delta X_{it} - \frac{1}{\gamma}[\ln\rho(1 + r_t)]\\ &+ \frac{1}{\gamma}\,\xi_t + \frac{1}{\gamma}(\varepsilon_{it} - \varepsilon_{i,t-1}).\end{aligned}$$

Compared to (64), this equation has the almost imperceptible difference that the surprise term $\zeta_t = \xi_t/\gamma$ no longer has an i subscript, so it really is just an aggregate shock,

[82] This is a simple application of the "envelope theorem." If we give the consumer a small increment of assets at the start of period t, any incremental gain in lifetime utility he/she might achieve by optimally allocating tiny increases in consumption over all remaining periods of the life, so as to satisfy (23), would be trivially small (i.e., second order).

and it can be appropriately captured with time dummies.

But Altug and Miller (1990) do not make this point simply as a critique of other work (or at least its interpretation). They note that if we adopt the assumption (69), and utilize the fact that $\eta \ln C_{it} = \ln \lambda_{it} = \ln \lambda_{it} + \ln \mu_i$, then we can first difference (47) to obtain

$$(73)\quad \Delta \ln h_{it} = \frac{1}{\gamma}\Delta \ln w_{it}(1-\tau_{it}) - \frac{\alpha}{\gamma}\Delta X_{it} - \frac{1}{\gamma}[\ln \lambda_t - \ln \lambda_{t-1}] + \frac{1}{\gamma}\Delta \varepsilon_{it}.$$

Using (73), we see that we can actually estimate the changes in the $\ln \lambda_t$ from time dummy coefficients. That in turn means that, given data on interest rates r_t, we can also estimate the asset pricing equation (71), with the only unknown parameter being ρ.

So the main point of Altug and Miller (1990) is to use data on hours, consumption, wages, and rates of return to *jointly* estimate (i) a within period optimality condition like (47), a first difference hours equation like (73), and an asset pricing equation like (71), using the cross equation restrictions (e.g., γ appears in multiple places) to get a more efficient estimate of the Frisch elasticity.[83] They estimate their model on a sample of continuously married men from the PSID for 1967–80. The men had to be no older than 46 in 1967.

A complication is that Altug and Miller (1990) do not use the simple utility function (3) that was used by MaCurdy (1981) and Altonji (1986). They use a more complex function where wives' leisure is nonseparable from consumption and male leisure. In that case, the analogue to (47) is the husband's demand for leisure equation

$$(74)\quad \ln l_{it}^{h} = \frac{1}{\tilde{\gamma}}\ln w_{it} + \frac{\eta}{\tilde{\gamma}}\ln C_{it} - \frac{\alpha}{\tilde{\gamma}}X_{it} + \frac{\pi}{\tilde{\gamma}}\ln l_{it}^{s} + \frac{\varepsilon_{it}}{\tilde{\gamma}},$$

where now $(1/\tilde{\gamma})$ is the intertemporal elasticity of substitution in *leisure* and l_{it}^{s} is the wives' leisure. In estimating (74) jointly with the rest of the system, the authors obtain a Frisch elasticity of leisure with respect to the wage of 0.037, with a standard error of 0.013. This precise estimate contrasts with an estimate of 0.018 with a standard error of 0.087 that they obtain when they do not include the first difference hours equation and the asset equation in the system. Thus we see that their approach does lead to a substantial efficiency gain.

Given that leisure is normalized to a fraction of total time, we have that the Frisch elasticity of male labor supply with respect to the wage implied by the authors' estimate is

$$\frac{\partial \ln h}{\partial \ln w} = \frac{w}{h}\frac{\partial h}{\partial w} = \frac{w}{h}\frac{\partial(1-l)}{\partial w} = -\frac{w}{h}\frac{\partial l}{\partial w} = -\frac{w}{h}\frac{l}{w}\left[\frac{w}{l}\frac{\partial l}{\partial w}\right] = -\frac{l}{h}(-0.037) \approx \frac{8760}{2300}(0.037) = 0.14.$$

Thus, despite the different methodology, the estimate is similar to the rather small values obtained by MaCurdy (1981) and Altonji (1986).

I conclude by pointing out some limitations of the study: It uses a ratio wage measure, it does not incorporate taxes, and it relies on the poor PSID consumption measure. Finally,

[83] For good measure, they throw in a wage equation as well. There are no cross equation restrictions between this and the other three equations, but allowing for the error covariance increases efficiency.

an odd aspect of Altug and Miller's (1990) results is that the coefficient on consumption in (74) is 0.003, which implies $\eta = 0.08$. This violates the theoretical restriction $\eta < 0$ (diminishing marginal utility of consumption). However, the coefficient is so imprecisely estimated that one can't reject that utility is linear in consumption ($\eta = 0$). But log utility ($\eta = -1$) is rejected. So all that can be discerned is that $-1 < \eta < 0$, almost the whole plausible range for η.[84]

6.2.7 *A New Approach: Measuring Expectations—Pistaferri (2003)*

A novel twist in the literature is the paper by Pistaferri (2003). He estimates hours change regressions (i.e., equation (64)), as in the MaCurdy (1981) and Altonji (1986) papers discussed earlier. But Pistaferri adopts a very different approach. The earlier papers treated the *expected* change in wages from time $t-1$ to t as unobserved, and they used instruments dated at time $t-1$ to construct predicted wage growth in the first stage of a 2SLS procedure. This approach relies on the assumption that the econometrician knows quite a bit about how workers make forecasts. Specifically, he/she must pick instruments that (i) are uncorrelated with the workers' forecast errors, and (ii) are good predictors of the wage growth workers actually expect. But as we saw, these papers suffered from the problem that coming up with good predictors for wage growth is difficult (i.e., first stage R^2s are low). Perhaps workers can predict their own wage growth better than we can. Furthermore, as we don't actually know how workers forecast wages, we can't be sure that all variables dated at time $t-1$ are in fact used to make forecasts, so we can't be sure they are valid instruments. Pistaferri's (2003) innovation is to use actual data on expectations to construct measures of workers' anticipated and unanticipated wage growth.

The data that Pistaferri uses is the Bank of Italy Survey of Households' Income and Wealth (SHIW) from 1989, 1991, and 1993. The survey is conducted every two years, and a fraction of subjects are reinterviewed (creating a panel component). The survey contains questions about expected *earnings* growth, not *wage* growth. I'll discuss the problems this creates below, but first consider how we could use *wage* expectations data if we had it.

Recall the hours growth equation (64) contained actual wage growth as a regressor. Surprise wage growth was relegated to a part of the residual denoted ξ_{it}, representing how the surprise altered the marginal utility of consumption λ_{it}. The presence of ξ_{it} in the residual meant the instruments used to predict wage growth had to be correlated with expected wage growth and uncorrelated with *un*-expected wage growth. Specifically, recall that

$$(75)\quad \zeta_{it} \equiv \frac{1}{\gamma}\,\xi_{it}$$
$$\equiv \frac{1}{\gamma}\{\ln \lambda_{it} - E_{t-1} \ln \lambda_{i,t-1}\}$$
$$= \frac{1}{\gamma}\,\frac{d\ln \lambda_{it}}{d\psi_{it}}\{\Delta \ln w_{it} - E_{t-1}\Delta \ln w_{it}\},$$

where I have defined $\psi_{it} \equiv \{\Delta \ln w_{it} - E_{t-1}\Delta \ln w_{it}\}$, the unexpected wage change from $t-1$ to t.

The first two equalities in (75) are only definitions. But the third states an assumption that all surprise changes in the marginal

[84] The greater imprecision of the η estimate compared to prior studies may stem from attempting to estimate the extent of nonseparability between female nonmarket time and both consumption and male labor supply, which adds female nonmarket time as an additional regression in the labor supply equations. Altug and Miller (1990) reject the joint hypothesis that female nonmarket time is separable from both consumption and male labor supply, but their estimates are too imprecise to determine which quantity it is nonseparable with.

utility of consumption are due to surprise wage changes. The term $d \ln \lambda_{it}/d\,\psi_{it}$ captures how wage surprises affect the marginal utility of consumption. An assumption that only wage surprises move λ_{it} is fairly strong. It rules out, e.g., unexpected transfers of assets. But it is important for Pistaferri's approach, as I describe below.

Now, if expected wage growth could actually be measured, we could rewrite (64) as

$$(76)\quad \Delta \ln h_{it} = \frac{1}{\gamma}(\Delta \ln w_{it}) - \frac{\alpha}{\gamma}\Delta X_{it} - \frac{1}{\gamma}\ln \rho(1+r_t) + \left[\frac{1}{\gamma}\frac{d \ln \lambda_{it}}{d\,\psi_{it}}\right] \times \{\Delta \ln w_{it} - E_{t-1}\Delta \ln w_{it}\} + \frac{\Delta\varepsilon_{it}}{\gamma}.$$

Furthermore, if we decompose the first term on the right hand side of (76)—actual wage growth—into parts that were anticipated versus unanticipated at time $t-1$, we obtain

$$(77)\quad \Delta \ln h_{it} = \frac{1}{\gamma}(E_{t-1}\Delta \ln w_{it}) - \frac{\alpha}{\gamma}\Delta X_{it} - \frac{1}{\gamma}\ln \rho(1+r_t) + \left[\frac{1}{\gamma} + \frac{1}{\gamma}\frac{d \ln \lambda_{it}}{d\,\psi_{it}}\right]\psi_{it} + \frac{1}{\gamma}\Delta\,\varepsilon_{it}.$$

Equation (77) captures how anticipated wage changes $E_{t-1}\Delta \ln w_{it}$ have only a Frisch substitution effect $(1/\gamma) > 0$ on hours. But unanticipated wage changes $\psi_{it} = \{\Delta \ln w_{it} - E_{t-1}\Delta \ln w_{it}\}$ have both a substitution effect $(1/\gamma)$ and an income effect $(1/\gamma)(d \ln \lambda_{it}/d\,\psi_{it}) < 0$. Thus, the sign of the effect of unanticipated wage changes is theoretically ambiguous.

A number of authors, including Blundell and MaCurdy (1999), have argued that tax reforms are generally unexpected and, to a good approximation, assumed to be permanent by workers.[85] If we grant this, then, as Blundell and MaCurdy (1999) state, the coefficient on unanticipated wage changes, $(1/\gamma) + (1/\gamma)(d \ln \lambda_{it}/d\,\psi_{it})$, is what we should be concerned with for evaluating labor supply effects of tax reforms.[86] But many subtle issues are involved here:

The coefficient $(d \ln \lambda_{it}/d\,\psi_{it})$ depends on many things, including: (i) how do consumers forecast future wages (i.e., to what extent do they expect surprise changes to be permanent or transitory?), (ii) how do consumers forecast future taxes (i.e., to what extent do they expect tax rule changes to be permanent or transitory?), and (iii) do the answers to questions (i) and (ii) depend on the source of the wage or tax change? (E.g., If a wage change occurs due to an unexpected change in tax law is it expected to be more or less persistent than if it occurs due to an unexpected promotion or layoff?) Many more questions of this type could be asked.

The first fundamental issue that one must deal with is how workers map unanticipated wage changes into expectations of future wages. To do this one must specify a model of the wage process, and make an assumption about how consumers forecast future wages.

[85] Note that the two assumptions are really two sides of the same coin: If one always thinks the current tax regime is permanent, then one will always be surprised by changes.

[86] See Blundell and MaCurdy (1999) page 1603: "As most tax and benefit reforms are probably best described as once-and-for-all unanticipated shifts in net-of-tax real wages today and in the future, the most appropriate elasticity for describing responses to this kind of shift is $\alpha_I + \gamma_0$," where α_I and γ_0 correspond, in their notation, to the two coefficients $(1/\gamma) + (1/\gamma)\ (d \ln \lambda_{it}/d\,\psi_{it})$ on *un*expected wage changes in equation (77).

Pistaterri (2003) assumes that log wages follow a random walk process with drift

$$(78) \qquad \ln w_{it} \;=\; \ln w_{i,t-1} + X_{i,t-1}\,\theta + \psi_{it}$$

$$E_{t-1}\,\psi_{it} \;=\; 0,$$

where ψ_{it} is unexpected wage growth. Pistaferri (2003) further assumes that workers know (78) is the wage process, and use (78) to forecast future wages. Ironically, having data on wage expectations does not obviate the need to make assumptions about how expectations are formed![87] The key behavioral assumption implied by (78) is that workers view all wage innovations as permanent: an unexpected wage change ψ_{it} shifts a worker's expectation of all his/her future wages by exactly ψ_{it}.

Next, similar to MaCurdy (1981), Pistaferri (2003) must make an assumption about how current and expected future wages, as well as current assets, map into the marginal utility of consumption. Specifically, he assumes that

$$(79) \quad \ln \lambda_{it} \;=\; \Gamma_{at}\,A_{it} + \Gamma_{0t}\ln w_{it} + \sum_{\tau=t+1}^{T}\Gamma_{\tau-t,t}\,E_t\ln w_{i,\tau}.$$

There is a key difference between MaCurdy (1981) and Pistaferri (2003) however, in that MaCurdy approximates the marginal utility of wealth in a model with perfect foresight. This is a function of the whole life-cycle wage path and initial assets, and it only varies over time according to the deterministic relationship $\lambda_{it} = \rho(1+r_t)\lambda_{i,t+1}$ (equation (18)). Thus MaCurdy is trying to estimate a single $\ln \lambda_{i0}$, while Pistaferri is trying to estimate the time varying $\ln\lambda_{it}$.[88]

A key thing an approximation to $\ln \lambda_{it}$ ought to capture is that, as the time horizon grows shorter, a temporary wage increase should have a larger effect on the marginal utility of consumption (i.e., In the terminal period, a wage increase is used entirely to increase current consumption, while in an earlier period it would be spread over all future periods).[89] This is why the $\{\Gamma_{kt}\}$ terms in (79) are allowed to vary over time. Each term has both a subscript $k = 0,\ldots,T-t$ that indicates the effect of the expected wage at time k on perceived wealth at time t, and a time subscript t that allows these effects to change over time. Of course, if one allowed the $\{\Gamma_{kt}\}$ terms to vary in an unconstrained way over k and t there would be severe proliferation of parameters. So Pistaferri constrains them to vary linearly.[90]

[87] This point has been made in a different context (forecasting future prices of durable goods) by Tulin Erdem et al. (2005). They show that when enough periods are available one can use the expectations data to estimate the expectations formation process, but one still has to impose some a priori structure on the process. Pistaferri cannot pursue this approach because he only has two periods of expectations data.

[88] MaCurdy (1981) backs out $\ln \lambda_{i0}$ in a second stage after estimating the differenced hours equation (64) in the first stage. This is possible because estimation of (64) uncovers all parameters of the hours equation in levels, $\ln h_{it} = (1/\gamma)w_t(1-\tau_t) + (1/\gamma)\ln\lambda_{i0} - (1/\gamma)t\cdot\ln\rho(1+r) + (\alpha/\gamma)X_{it} + (1/\gamma)\varepsilon_{it}$ except for $(1/\gamma)\ln \lambda_{i0}$, which is the individual specific constant (or "fixed effect") in the levels equation. Given these constants, MaCurdy can, in principle, regress them on the whole set of life-cycle wages. However, he only observes wages for his ten-year sample period. So he fits a life-cycle wage profile for each person using ten years of data. He then regresses the $(1/\gamma)\ln \lambda_{i0}$ on the individual specific parameters of this (assumed quadratic) profile. Using the coefficient on the wage equation intercept, MaCurdy can determine how an upward shift in the whole wage profile would affect $(1/\gamma)\ln \lambda_{i0}$, and hence labor supply. He estimates a 10 percent increase in wages at all ages would increase labor supply by only 0.8 percent. Of course, the problem with this procedure relative to MaCurdy (1983) method #1 is the need to extrapolate out of sample wage information rather than using current consumption to proxy for lifetime wealth.

[89] The same argument holds for an increase in assets. For instance, a 60 year old who wins a million dollars in the lottery should be much more likely to retire than a 30 year old.

[90] This is not indicated in the notation in his paper (see Pistaferri equation (8)), but Pistaferri has confirmed this to me in private correspondence.

From (78)–(79) we get that the surprise change in the marginal utility of consumption is related to the surprise change in the wage as follows:

$$(80)\quad \ln \lambda_{it} - E_{t-1} \ln \lambda_{it} = \Gamma_a[A_{it} - E_{t-1}A_{it}] + \Gamma_0\{\ln w_{it} - E_{t-1}\ln w_{it}\} + \sum_{\tau=t+1}^{T} \Gamma_{\tau-t}\{E_t \ln w_{i,\tau} - E_{t-1}\ln w_{i,\tau}\} = \Gamma_a[A_{it} - E_{t-1}A_{it}] + \Gamma_0\{\psi_{it}\} + \sum_{\tau=t+1}^{T}\Gamma_{\tau-t}\{\psi_{it}\} = \Gamma_a \cdot 0 + \Gamma \cdot \psi_{it},$$

where I have suppressed the time subscripts on the Γ to conserve on notation.

In (80), all the Γ are negative because a surprise increases in assets, a surprise increase in the current wage, or in any future wage, all increase the consumer's perceived wealth. This leads to higher current consumption and hence a lower marginal utility of consumption. The second line of the equation utilizes the fact that, given the random walk wage process in (78), the changes in *all* future wage expectations $E_t \ln w_{i,\tau} - E_{t-1}\ln w_{i,\tau}$ for $\tau = t+1, \ldots, T$ are equal to the current wage surprise ψ_{it}. Again, this is because that surprise is expected to persist forever. At the opposite extreme, if we had instead assumed that consumers perceive all wage surprises as purely transitory, then we would have $E_t \ln w_{i,\tau} - E_{t-1}\ln w_{i,\tau} = 0$ for all $\tau = t+1, \ldots, T$ and the third term after the equal sign would vanish. Finally, the last line of (80) invokes Pistaferri's assumption of no unexpected asset changes, and defines $\Gamma = \Gamma_0 + \Gamma_1 + \cdots + \Gamma_{T-t}$.[91]

Now, given (80), we have that $d \ln \lambda_{it}/d\,\psi_{it} = \Gamma\psi_{it}$ and hence we can rewrite (77) as

$$(81)\quad \Delta \ln h_{it} = \frac{1}{\gamma}(E_{t-1}\Delta \ln w_{it}) - \frac{\alpha}{\gamma}\Delta X_{it} - \frac{1}{\gamma}\ln \rho(1+r_t) + \left[\frac{1}{\gamma} + \frac{\Gamma}{\gamma}\right] \times \{\Delta \ln w_{it} - E_{t-1}\Delta \ln w_{it}\} + \frac{\Delta\varepsilon_{it}}{\gamma}.$$

This gives us Pistaferri's essential idea. We can use the coefficient on expected wage changes to estimate the intertemporal elasticity of substitution $(1/\gamma)$, while using the coefficient on unexpected wage changes to estimate the "total" effect of a wage change, which includes both the substitution effect and the income effect. Taking the difference between the two coefficients enables us to isolate the income effect of a permanent wage increase (Γ/γ).

Estimation of (81) has some key advantages over the conventional approach pursued in MaCurdy (1981) and Altonji (1986). First, instruments need not be uncorrelated with wage surprises, as unexpected wage changes are controlled for (rather than being relegated to the error). This also circumvents the problem, noted earlier, that it is hard to find good predictors of wage growth. Second, as Pistaferri notes, the best predictors are usually age and education, but it is a strong assumption that these are excluded instruments not appearing in the hours equation itself (i.e., that they do not shift tastes for work). Third, the problem of aggregate productivity shocks is avoided, as

[91] Notice that the wealth effect term Γ gets (mechanically) smaller as t gets larger, simply because there is less of a future horizon over which wages will increase, so fewer Γ_t terms are being added up. Counteracting that, as I argued earlier, is that the wealth effect of each individual (period specific) wage increase should grow larger as one gets closer to the end of the planning horizon.

the average forecast error no longer enters the error term.

Unfortunately, there are important gaps between this excellent idea and its actual empirical implementation. The first problem Pistaferri faces is that the Bank of Italy Survey does not really contain expectations of *wage* changes, but only of *earnings* changes. Pistaferri shows how to construct a version of (81) where expected and unexpected *earnings* replace wages, and coefficients are suitably modified. However, as Pistaferri notes, this introduces a major problem: Unobserved shifts in tastes for work will of course alter earnings (as earnings are a function of hours). And, as unobserved tastes for work ($\Delta\varepsilon_{it}$) enter the error term in the hours equation, expected and unexpected earnings changes are endogenous.

Second, expected earnings changes are presumably measured with error. As variables like hours and earnings are measured with error, it would be highly implausible to assume a more subtle concept like the expected change in earnings is not measured with error as well. Furthermore, there may be systematic errors arising from how respondents interpret the survey question, which reads "We are interested in knowing your opinion about your labor earnings or pensions 12 months from now." Is it clear whether respondents would include expected tax changes when answering such a question? And, while the expectations question asked about earnings 12 months hence, the data on wages, hours and earnings were collected in 1989, 1991, and 1993. In order to align the two-year time interval of the earnings data with the one-year forecast horizon, Pistaferri assumes a person would have projected their earnings growth rate forecast to persist for two years, an additional source of measurement error.

Both of these problems suggest that it may be necessary to instrument for expected and unexpected changes in earnings, using variables that help predict these variables but that are uncorrelated with taste shocks and measurement error. In that case, a key advantage of Pistaferri's procedure (i.e., not needing to instrument) is lost. Pistaferri (2003) does not in fact attempt to deal with these problems, and he estimates his version of (81) by least squares.

For estimation, Pistaferri uses data on male household heads aged 26 to 59 in 1989. There are 1,461 person-year observations in the unbalanced panel. As observed taste shifters he uses age, education, region, family size, whether the wife or other household members work, and number of children in various age ranges. He estimates that the Frisch elasticity is 0.704 with a standard error of 0.093 and the income effect (Γ/γ) is –0.199 (standard error 0.091). Thus, the elasticity of labor supply with respect to a surprise permanent upward shift in the wage profile is 0.51. That is, a permanent unexpected 10 percent wage increase would cause a 5 percent increase in labor supply. This is a very large uncompensated wage effect, and it implies that permanent tax changes have very large effects on labor supply. The result contrasts sharply with MaCurdy's (1981) comparable estimate of only an 0.8 percent increase. We should view both results with some caution however, given the data limitations noted above.[92]

[92] Pistaferri (2003) contains a few other elements I haven't mentioned. His version of (81) includes a measure of the perceived variance of earnings, also obtained from the survey of expectations. But he finds that variance is not significant in the hours equation. He also tests for separability between leisure and consumption but does not find strong evidence of nonseparability. Finally, Pistaferri also uses his data to estimate an hours change regression like that in MaCurdy (1981) and Altonji (1986), using a cubic in age and education and interactions between age and education as instruments. This produces a Frisch elasticity of 0.318 with a standard error of 0.319. The R^2 in the first stage regression is only 0.0025 with an F-statistic of 1.73. A key advance in econometric practice since the 1980s is the attention that is now paid to the problem of weak instruments in the first stage of 2SLS. A common rule of thumb is that the F-statistic should be at least 5 before results are to be trusted.

Another key point is that Italy had a recession in 1993. Pistaferri (2003) includes a 1993 dummy in (81) and obtains a coefficient of –0.068 (standard error of 0.023), which implies a 6.8 percent decline in hours not explained by the model. This may suggest that workers in Italy are not free to adjust hours in the short run, and that there was demand induced rationing.

6.2.8 *Progressive Taxation and Tied Wage–Hours Offers—Aaronson and French (2009)*

The papers in sections 6.2.4 to 6.2.7 all ignore taxation in estimating the Frisch elasticity. Aaronson and French (2009) account for both progressive taxation and tied wage–hours offers. In their model, the wage is given by $w = w_0 h^\theta$ and labor earnings are given by $E = wh - \tau(wh + N)$, where $\tau(\cdot)$ is a tax function. The effective wage is $dE/dh = (1 - \tau'(\cdot))[d(wh)/dh] = w(1 + \theta)(1 + \tau'(\cdot))$. Equation (22) is modified so the marginal utility of leisure is set equal to the marginal utility of consumption λ times the effective wage, as in $\beta_t h_t^\gamma = \lambda_{it} w_{it}(1 + \theta)(1 - \tau'(\cdot))$. A key point is that if we now perform the usual log and first-difference procedures, so as to obtain the analogue of (25), the term $(1 + \theta)$ drops out. Thus, at least given this functional form for the tied wage–hours relationship, we cannot identify θ in the MaCurdy (1981) framework. So ignoring tied wage-hours offers will not bias estimates of the preference parameter γ.

Nevertheless, if $\theta > 0$, the Frisch hours response to a tax change is magnified (i.e., it is greater than $(1/\gamma)$). Intuitively, if a tax cut raises hours then it raises the pretax wage, which further raises hours, etc.. Aaronson and French show the Frisch elasticity with respect to $(1 - \tau)$ is $e_F = (1/\gamma)/[1 - \theta(1/\gamma)]$. But to estimate θ we need outside information—e.g., as Aaronson and French note, Moffitt (1984) and Keane and Wolpin (2001) both estimate $\theta = 0.20$, the former using a wage on hours regression (using taste shifters like children to instrument for hours), the latter in a structural model (see Section 6.3.3). If $(1/\gamma) = 1$ and $\theta = 0.20$ we get $e_F = 1/0.80 = 1.25$. Then, if we ignore tied wage hours offers we understate the wage elasticity by 25 percent.

Aaronson and French note that the simple regression $\ln(1 - \tau) = \pi_0 - \pi_1 \ln(wh + N)$ with $\pi_1 = 0.11$ gives a fairly good fit to the U.S. progressive tax structure for the 1977–96 period. They derive that in this case, if taxes are ignored, the usual IV method of estimating (25) will lead to downward (asymptotic) bias in the estimate of $(1/\gamma)$ by a factor of $(1 - \pi_1)/[1 + \pi_1(1/\gamma)]$. Thus, for example, if $(1/\gamma) = 1$ and $\pi_1 = 0.11$ then the bias factor is 0.80. That is, our estimate of the Frisch elasticity would be biased down by 20 percent due to ignoring progressive taxation.

If we combine the problems of ignoring taxes in estimation and ignoring tied wage-hours offers in prediction, we see the downward bias in predicting the hours response may plausibly be on the order of 50 percent. This is certainly important, but on the other hand, it is not adequate to reconcile the predominantly small Frisch elasticities estimated in the literature with a view that labor supply responses are actually quite large.

Using PSID data from 1977–96 and CPS data from 1979–2003, the authors obtain a range of Frisch elasticities from 0.161 to 0.608 depending on details of the specification.[93] Interestingly, *the estimates are almost completely insensitive to whether or not they include the progressive tax term*. But, as the authors note, this does *not* show that failure

[93] The estimates are precise due to large sample sizes, despite first stage R^2s in the 0.001 to 0.002 range. In some specifications a ratio wage measure is used. In other specifications, to avoid denominator bias, the authors use earnings instead of wages as the independent variable. In this case, the earnings coefficient is $1/(1 + \gamma)$. The latter approach leads to somewhat larger estimates of the Frisch elasticity.

to account for progressive taxation is innocuous. Rather, this result obtains because the usual age and education polynomial instruments used in this literature do a very poor job of predicting the tax term. (I am puzzled why the authors did not try using tax rules as additional instruments).

6.3 *The Life-Cycle Model with Both Human Capital and Savings*

A fundamental problem with the labor supply models in sections 6.1 and 6.2 is that they treat (pretax) wages as exogenous. That is, they ignore the possibility that work experience may lead to increased wages. Existence of such experience effects has rather striking implications for all estimation methods discussed in sections 6.1 and 6.2.

To see this, let's return to the simple two period model of (15) but assume that the wage in period 2, rather than being exogenously fixed, is an increasing function of hours of work in period 1. For expositional purposes, I will assume the simple function

$$(82) \qquad w_2 = w_1(1 + \alpha h_1),$$

where α is the percentage growth in the wage per unit of work. Once we introduce human capital accumulation via work experience as in (82), equation (15) is replaced by

$$(83) \quad V = \frac{[w_1 h_1(1-\tau_1) + b]^{1+\eta}}{1+\eta} - \beta_1 \frac{h_1^{1+\gamma}}{1+\gamma} + \rho\left\{\frac{[w_1(1+\alpha h_1)h_2(1-\tau_2) - b(1+r)]^{1+\eta}}{1+\eta} - \beta_2\frac{h_2^{1+\gamma}}{1+\gamma}\right\},$$

where I have ignored nonlabor income for simplicity. The first order conditions are now

$$(84) \quad \frac{\partial V}{\partial h_1} = C_1^{\eta} w_1(1-\tau_1) - \beta_1 h_1^{\gamma} + \rho C_2^{\eta} w_1\, \alpha\, h_2(1-\tau_2) = 0$$

$$(85) \quad \frac{\partial V}{\partial h_2} = C_2^{\eta} w_1(1+\alpha\, h_1)(1-\tau_2) - \beta_2 h_2^{\gamma} = 0$$

$$(86) \quad \frac{\partial V}{\partial b} = C_1^{\eta} - \rho(1+r)C_2^{\eta} = 0.$$

Comparing (84) to (16) we see it now includes the extra term $\rho C_2^{\eta} w_1 \alpha h_2(1-\tau_2)$, the effect of an extra hour of work at $t = 1$ on the present value of earnings at $t = 2$. If we perform the usual manipulations on (84) to obtain the within period MRS condition, we now obtain

$$MRS = \frac{MU_L}{MU_C} = \frac{\beta h_1^{\gamma}}{C_1^{\eta}} = w_1(1-\tau_1) + \frac{\rho C_2^{\eta}}{C_1^{\eta}}\alpha w_1 h_2(1-\tau_2.)$$

This can be simplified by using (86) to eliminate $C_2^{\eta}/C_1^{\eta} = [\rho(1+r)]^{-1}$, giving

$$(87) \quad \frac{\beta h_1^{\gamma}}{C_1^{\eta}} = w_1(1-\tau_1) + \frac{\alpha w_1 h_2(1-\tau_2)}{1+r}.$$

It is useful to compare (87) to (22), the MRS condition for the model without human capital. Here the opportunity cost of time (OCT) is no longer simply the after-tax wage rate. Instead, it is augmented by the term $\alpha w_1 h_2(1-\tau_2)/(1+r)$, which captures the effect of an extra hour of work at $t = 1$ on the present value of earnings at $t = 2$.

The fact the MRS condition for $t = 1$ depends not just on variables dated at $t = 1$ but also on h_2 means the key idea of MaCurdy's method #1 and #2—i.e., that current consumption is a sufficient statistic for all future variables—no longer holds. Say we ignore this problem and attempt to estimate preferences by estimating an MRS condition like (22), without the human capital term. The size of the resultant bias depends on the behavior of the human capital term. I'll discuss this in detail in section 6.3.2.3 after we examine some empirical results.

6.3.1 *Early Attempts to Include Human Capital—Heckman (1976), Shaw (1989)*

I now turn to the empirical literature on male life-cycle labor supply that includes human capital accumulation. Unfortunately, there are very few papers of this type. As far as I am aware, the first paper to estimate a life-cycle model with human capital was Heckman (1976). The computing technology available at that time did not permit estimation of a model where workers decide jointly on savings and human capital investment, particularly not while also allowing for uncertainty in wages and stochastic taste shocks. Thus, Heckman's model is deterministic and only attempts to fit "typical" life cycle paths of wages and hours.

The Heckman (1976) approach is rather different from the "learning by doing" model in (82)–(83). Instead, he follows Yoram Ben-Porath (1967) and William J. Haley (1973) in using a model where a worker may choose to devote some fraction to his/her work time to investment. The worker is paid only for productive time, not time spent learning. But observed labor supply is the sum of all time at work: productive time plus investment time. Hence, the observed market wage rate in period t is $w_t = w_t^*(1 - S(t))$, where w_t^* is the worker's actual productivity and $S(t)$ is the fraction of his/her time at work that the worker spends investing in human capital.

The key similarity between Heckman's model and the learning-by-doing model is that the observed market wage rate (w_t) is not the opportunity cost of time. Instead it is w_t^*, the worker's productivity, as that is what he/she gives up per unit of time spent in leisure or learning. The cost of time exceeds the wage rate by the multiplier$1/(1 - S(t))$, which is an increasing function of the fraction of time at work that the worker spends investing in human capital. So fundamentally Heckman's model is quite similar in spirit to that in equations (82)–(83), in that human capital investment causes the opportunity cost of time to exceed the wage.

Heckman (1976) estimates an equation for $S(t)$ jointly with an equation for observed hours and wages (derived from a particular functional form mapping investment time into wages). He uses data for 23–65 year old males from the 1970 U.S. Census. As the model is deterministic, it is fit to average wages and hours (by age). Heckman's estimate of the $S(t)$ function implies that 23-year old males spend roughly 35 percent of their "work" time on human capital investment. Hence, their opportunity cost of time exceeds their observed wage rate by roughly 54 percent. The fraction of time spent investing is estimated to drop steadily, becoming near 0 percent at about age 40. Thus, his estimates imply the OCT grows only 65 percent as much as the observed wage from age 23 to age 40 (i.e., $w_{40}^*/w_{23}^* = w_{40}/[w_{23}/(1 - 0.35)] = (0.65)(w_{40}/w_{23})$).

Shaw (1989) substantially extended Heckman (1976) by estimating a model where workers make joint decisions about savings and human capital investment, incorporating uncertainty about future wages and hours. Her approach is to estimate an equation similar to equation (87), the MRS condition. However, to take the model to data one must first extend it to multiple periods and introduce uncertainty. A simple way to do this is to rewrite (87) as

$$(88)\quad \frac{\beta h_t^{\gamma}}{C_t^{\eta}} = w_t(1 - \tau_t) + E_t \sum_{\tau=0}^{T-t} \frac{(\alpha w_1) h_{t+1+\tau}(1 - \tau_{t+1+\tau})}{(1 + r)^{1+\tau}},$$

where I assume the wage equation (82) has the form $w_{t+1} = (1 + \alpha \sum_{\tau=1}^{t} h_\tau) w_1$. Then, a one-unit increase in h_t raises the wage by (αw_1) in all future periods. This raises earnings by $(\alpha w_1) h_{t+1+\tau}$ for $\tau = t+1, \ldots, T$. The second term in (88) is the expected present value of increased (after-tax) earnings in all future periods obtained by working an extra unit of time at time t.

Unfortunately (88) involves hours of work in all future periods, which is not available in any data set. So Shaw (1989) uses the following trick: First, rewrite (88) as

$$(88') \quad \frac{\beta h_t^\gamma}{C_t^\eta} = w_t(1-\tau_t) + \frac{(\alpha w_1) h_{t+1}(1-\tau_{t+1})}{(1+r)} + E_t \sum_{\tau=0}^{T-t-1} \frac{(\alpha w_1) h_{t+2+\tau}(1-\tau_{t+2+\tau})}{(1+r)^{2+\tau}}.$$

Now take (88) and date it forward one period:

$$(88'') \quad \frac{\beta h_{t+1}^\gamma}{C_{t+1}^\eta} = w_{t+1}(1-\tau_{t+1}) + E_{t+1} \sum_{\tau=0}^{T-t-1} \frac{(\alpha w_1) h_{t+2+\tau}(1-\tau_{t+2+\tau})}{(1+r)^{1+\tau}}.$$

Now, notice that the summation terms on the right-hand sides of (88′) and (88″) are identical, except for a factor of $1/(1+r)$ and the dating of the expectation. So, premultiplying (88″) by $1/(1+r)$ and taking the expectation at time t we obtain

$$(89) \quad E_t\left\{\frac{1}{1+r}\left[\frac{\beta h_{t+1}^\gamma}{C_{t+1}^\eta} - w_{t+1}(1-\tau_{t+1})\right]\right\} = E_t \sum_{\tau=0}^{T-t-1} \frac{(\alpha w_1) h_{t+2+\tau}(1-\tau_{t+2+\tau})}{(1+r)^{2+\tau}}.$$

The intuition for why these manipulations are useful is that, at time t, the worker knows (or, rather, we assume he/she knows) that at time $t+1$ he/she will choose hours and consumption to satisfy (88″). Thus, we can use the worker's own labor supply and consumption behavior at $t+1$, described by the simple expression on the left of (89), to infer what he/she believes about the complex expectation term sitting on the right.[94]

So, using (89) to substitute for the summation term in (88′), we obtain

$$(90) \quad \frac{\beta h_t^\gamma}{C_t^\eta} = w_t(1-\tau_t) + \frac{(\alpha w_1) h_{t+1}(1-\tau_{t+1})}{(1+r)} + E_t\left\{\frac{1}{1+r}\left[\frac{\beta h_{t+1}^\gamma}{C_{t+1}^\eta} - w_{t+1}(1-\tau_{t+1})\right]\right\}.$$

This equation is feasible to estimate, as it only requires data on hours at time t and $t+1$, wages at t and $t+1$, and consumption at time t and $t+1$. The final step is to replace the expectation term with its actual realization, while appending a forecast error:

$$(91) \quad \frac{\beta h_t^\gamma}{C_t^\eta} = w_t(1-\tau_t) + \frac{(\alpha w_1) h_{t+1}(1-\tau_{t+1})}{(1+r)} + \frac{1}{1+r}\left[\frac{\beta h_{t+1}^\gamma}{C_{t+1}^\eta} - w_{t+1}(1-\tau_{t+1})\right] + \xi_{t+1}.$$

Equation (91) is the basic type of equation that Shaw (1989) uses. The estimation is done in two stages. In the first stage, a wage equation is estimated to determine how wages grow with work experience (i.e., α

[94] Interestingly, this is a continuous data analogue of the procedure developed by V. Joseph Hotz and Miller (1993) to infer agents' expectations from their discrete choices in discrete choice dynamic programming models.

in equation (82)). In the second stage, the wage equation parameters are treated as known and (91) is estimated by instrumental variables.[95] Valid instruments are known by workers at time t, so they are uncorrelated with the forecast error ξ_{t+1}.

While (91) is similar to what Shaw (1989) estimates, she does not include taxes. On the other hand, she introduces a number of additional features. First, rather than (3) she uses a translog utility function as in (63), with $G(X) = X$. As a consequence the marginal utility of consumption and leisure terms in (91) become more complicated. Second, she lets the taste for work parameter β vary across workers based on schooling level. Third, in the wage equation she allows the rental rate on human capital to vary over time.

It is interesting that Shaw (1989) does not introduce stochastic variation in tastes as in the previous studies we have examined. The reason why can be seen by looking at the simple MRS condition for the model without human capital, (22), and following the steps that led to equation (64), where expectation errors and taste shocks enter as a composite additive error. Hence, one can estimate (64) by IV without having to assume any distribution for the forecast errors and taste shocks. In contrast, in (91), we see that if β is allowed to have a stochastic component it will enter the equation in a highly nonlinear way. Thus, the taste shock will not "pop out" into an additive error that can be combined with ξ_t. This makes the simple application of instrumental variables estimation infeasible.[96]

To proceed, Shaw (1989) assumes a worker's human capital, K, evolves according to

$$(92)\quad K_{i,t+1} = \alpha_1 K_{it} + \alpha_2 K_{it}^2 + \alpha_3 K_{it} h_{it} + \alpha_4 h_{it} + \alpha_5 h_{it}^2 + \tau_t + \varepsilon_{it}.$$

That is, current human capital is a quadratic function of last year's human capital and last year's hours of work. The $\{\tau_t\}$ are year specific aggregate shocks while the $\{\varepsilon_{it}\}$ are person specific idiosyncratic shocks to human capital production (i.e., illness, job separations). The wage rate is the aggregate rental price of human capital (R_t) times the stock of human capital:

$$(93)\quad w_{it} = R_t K_{it} \quad \Rightarrow \quad K_{it} = w_{it}/R_t.$$

Shaw (1989) allows rental prices to vary over time in an unconstrained way. However, as the units of human capital are arbitrary, the rental price must be normalized in one year ($R_1 = 1$). An estimable wage equation is obtained by substituting (93) into (92). The parameters to be estimated are the $\{\alpha\}$, the rental rates, $\{R_t\}_{t=2}^{T}$, and the time dummies, $\{\tau_t\}$.

Shaw (1989) estimates the wage equation using data on white males, aged 18 to 64, from the 1968 to 1981 waves of the PSID. Valid instruments should be uncorrelated with the human capital production shock ε_{it} in (92). Shaw uses a polynomial in current wages[97] and hours, along with schooling, age, local unemployment, a South dummy and year dummies.

It is worth commenting on the use of current hours h_{it} as an instrument. In general, we would expect the person specific productivity shock ε_{it} to enter the decision rule for hours. For example, if ε_{it} is high, a person realizes his/her human capital is going to rise substantially at time $t + 1$, even if he/she has low current hours of work. Thus, given

[95] Shaw (1989) also substitutes for consumption at $t + 1$ using the familiar relationship $C_t^{\eta} = E_t\, \rho(1 + r)C_{t+1}^{\eta}$.

[96] Recently, Keane (2009) has proposed a simulated maximum likelihood method for estimating models where multiple stochastic terms enter the first order conditions nonlinearly.

[97] Note that ε_{it} does not affect the wage until $t + 1$. So the current wage is a valid instrument given the timing.

diminishing returns to human capital, we would expect the person to work less at time t. Under this scenario, current hours are not a valid instrument. The key assumption that would validate using hours as an instrument is if ε_{it} is not revealed until *after* the worker decides on current hours of work.

Another important point is that, unlike conventional studies in the human capital literature, the wage equation estimated here does not include an individual effect to capture a person's unobserved skill endowment. Shaw (1989) makes the point that this is not necessary here, because the lagged level of human capital proxies for unobserved ability.

Given (92), the derivative of human capital with respect to hours of work is

$$\frac{\partial K_{i,t+1}}{\partial h_{it}} = \alpha_3 K_{it} + \alpha_4 + 2\alpha_5 h_{it}.$$

The estimates are $\alpha_3 = 0.30$, $\alpha_4 = -3.55$, and $\alpha_5 = 0.69$. To interpret these figures, let $R = 1$, and note that mean hours in the data is 2,160 while the mean wage rate is \$3.91. Then, noting that h_{it} is defined as hours divided by 1,000, we have, at the mean of the data

$$\frac{\partial K_{i,t+1}}{\partial h_{it}} = (0.30)(3.91) - 3.55 + 2(0.69)(2.16) = 0.60.$$

This implies, for example, that an extra 500 work hours at time t (an increase in h_t of 0.5) increases the wage rate at $t + 1$ by 30 cents per hour. In percentage terms, this is a 23 percent hours increase causing an 8 percent wage increase—a very strong effect of work experience on wages.

Notice that the positive estimate of α_3 implies that hours of work and human capital are compliments in the production of additional human capital. That is, wages rise more quickly with work experience for high-wage workers than low-wage workers.

The estimates also imply that human capital rental rates are quite volatile, although the year specific rental rates are quite imprecisely estimated. Interestingly, Shaw (1989) reports that the series of annual rental rates for the fourteen years of data has a correlation of –0.815 with an index for the price of fuel. This is consistent with results in Keane (1993b) showing that oil price movements in the 1970s and 1980s had very large effects on real wages in the United States.

Shaw (1989) estimates the first order condition (91) using a subset of the data (10 years), as the PSID did not collect food consumption in 1967–68 and 1975. The instruments, assumed uncorrelated with the forecast error ξ_{t+1}, include a fully interacted quadratic in the *time t* values of leisure (defined as 8,760 minus hours of work), food consumption, and the wage rate (constructed as annual earnings divided by annual hours). Also included are education, age, the local unemployment rate, a South dummy, and time dummies.

The parameter estimates are reasonable, implying the marginal utility of leisure and consumption are both positive, with diminishing marginal returns. The coefficient on the consumption/leisure interaction is negative, implying hours of work and consumption are compliments. The discount factor is estimated to be 0.958. More interesting however are the simulations of the model.

Unfortunately, first order conditions like (91) are inadequate to simulate the behavior of workers in a life-cycle model. The problem is that the first order condition, combined with the laws of motion for human capital (92) and assets ($A_{t+1} = (1 + r)(w_t h_t - C_t + A_t)$), only tell us how hours, wages, and assets move from one period to the next, *conditional on a particular starting point*. But the assumed starting point is arbitrary. The first order condition cannot be used to determine *optimal* first period choices implied by the model. For that, we need a "full solution" of a worker's dynamic optimization problem, an issue I turn to in section 6.3.2.

Note that this criticism is not particular to Shaw (1989). It also applies to all methods based on estimating first order conditions of life-cycle models that I discussed earlier (e.g., MaCurdy 1983 method #1), and to the life-cycle consistent methods (e.g., MaCurdy 1983 method #2).[98] Furthermore, this criticism of first order condition methods omits the further problem that, even to use first order conditions to simulate forward from an arbitrary starting point, one still needs to know the distribution of the stochastic terms (e.g., the forecast error ξ_{t+1} in (91)). The instrumental variables estimation techniques that are typically used to estimate first order conditions do not deliver estimates of the distributions of the stochastic terms of the model, making even this limited type of analysis infeasible.[99]

These problems are why, when authors have estimated dynamic models using first order conditions or life-cycle consistent methods, they have sometimes used the estimated preference parameters to simulate how workers would respond to tax changes under the hypothetical that they live in a static world (with a static budget constraint). An example of this is MaCurdy (1983). In some cases such simulations are informative. For instance, in the simple life-cycle model of (15), workers' response to a permanent anticipated tax change is given by the Marshallian elasticity (6). But this is no longer true in a life-cycle model with human capital: If a tax change alters labor supply at time t it will also alter the pretax wage at $t+1$. Thus, the response will change with age/time (see section 6.3.2.4).[100]

Consistent with the above discussion, Shaw (1989) conducts her simulations by choosing arbitrary $t = 1$ values for wages, hours, and assets, and setting the stochastic terms to zero. Despite these limitations, the simulations are interesting. Take a worker starting at age 18 with a wage of \$3.30 per hour and working 2,200 hours per year. The simulations imply that such a worker's wage would rise to roughly \$3.65 over the first eight years of employment (an 11 percent increase), but his/her hours are essentially flat (in fact, they decline very slightly).

This is a more extreme version of the pattern found in Heckman (1976). Even though the wage increases by 11 percent over the first eight years, the OCT does not rise at all, as the drop in the human capital return to experience is sufficient to completely outweigh it. As a result, hours do not rise. Thus, a researcher looking at these simulated data through the lens of a model that ignores human capital would conclude there is no intertemporal substitution whatsoever in labor supply, yet we know that, in the true model that generates the data, there is

98 MaCurdy (1983) himself emphasized the limitations of all these approaches. As he stated: "Implementing the above procedures yields estimates required to formulate the lifetime preference function, but . . . this . . . is not sufficient to determine how a consumer will respond to various shifts in budget or asset accumulation constraints, such as those arising from changes in wages or in tax policies. . . . To form predictions for such responses, it is necessary to introduce sufficient assumptions to provide for a complete . . . formulation of the lifetime optimization problem . . . which, in addition to a function for preferences, requires a full specification for a consumer's expectations regarding current and future opportunities . . . Given a particular formulation for the lifetime optimization problem, one . . . [can conduct] . . . simulation analysis which involves numerically solving the consumer's optimization problem for the different situations under consideration." The numerical procedure that MaCurdy describes here is what I refer to as a "full solution" of the optimization problem.

99 Keane (2009) develops an estimation method that involves estimating the distribution of stochastic terms that enter first order conditions.

100 Indeed Keane (2011) argues that, in a model with human capital, tax changes cannot be viewed as inducing exogenous changes in after-tax wages, because the worker's labor supply response to the tax change affects his/her wage path, rendering the wage change endogenous.

intertemporal substitution with respect to the OCT.[101]

6.3.2 *Full Solution Estimation with Human Capital and Assets—Imai and Keane (2004)*

Imai and Keane (2004) was the first paper to use full solution methods to estimate a life-cycle labor supply model that includes *both* human capital investment and assets, along with a continuous choice of hours. As their model is rather complex, I present a simplified version that captures the main points. Assume a worker's human capital evolves according to

$$K_{i,t+1} = (1 + \alpha\, h_{it})K_{it}. \tag{94}$$

K_{i1} is the person's skill endowment at the time of labor force entry.[102] A person's wage at t is equal to the current stock of human capital times the (constant) rental price of skill R. Human capital is subject to a transitory productivity shock. Specifically:

$$w_{it} = RK_{it}(1 + \varepsilon_{it}). \tag{95}$$

The period specific utility function is given by (3), as in MaCurdy (1981), and assets evolve according to $A_{t+1} = (1 + r)(A_t + w_t h_t(1 - \tau) - C_t)$. Given this setup, an agent's state at the start of any period t is fully characterized by the vector of state variables $\Omega_t \equiv \{K_t, A_t, \varepsilon_t, \beta_t\}$.

Imai and Keane (2004) assume that the shocks to wages (ε_t) and tastes (β_t) are independent over time. Such independence assumptions are common in the literature, because, as we'll see, they greatly reduce the computational burden of obtaining a full solution to the agents' dynamic optimization problem.[103] The next section describes the solution method. Readers not interested in the technical details can skip to 6.3.2.2.

6.3.2.1 *How to Solve the Dynamic Programming Problem—A Simple Exposition*

To describe the full solution method, I first take the value function for the simple two-period model (83) and extend it to a multi-period setting (with uncertainty):

$$\begin{aligned} V_t(K_t, A_t, \varepsilon_t, \beta_t) &= \left[\frac{C_t^{1+\eta}}{1+\eta} - \beta_t \frac{h_t^{1+\gamma}}{1+\gamma}\right] \\ &\quad + E_t\left\{\sum_{\tau=t+1}^{T} \rho^{\tau-t}\left[\frac{C_\tau^{1+\eta}}{1+\eta} - \beta_\tau \frac{h_\tau^{1+\gamma}}{1+\gamma}\right] \middle| (K_{t+1}, A_{t+1})\right\}. \end{aligned} \tag{96}$$

The value function now has a t subscript, as it is specific to time period t, as opposed to being a lifetime value function. The arguments of the value function are the complete vector of state variables. The first term on the right-hand side of (96) is current utility at time t. The second term is the *expected*

[101] Shaw (1989) admits that her model actually provides a rather poor fit to the data because hours for youth do in fact exhibit a moderate rise in the first several years after they enter the labor market. She attributes this to factors omitted from the model. Note, however, that it is the very large experience return in her model that drives this result, by causing the opportunity cost of time to greatly exceed the wage at $t = 1$.

[102] Imai and Keane (2004) actually assume a much more complex process, designed to capture patterns of complimentarity between human capital and hours of work in the human capital production function. But use of the simpler form in (94) helps to clarify the key points.

[103] Stochastic terms such as tastes for work are often assumed to consist of a part that is constant over time and a part that is stochastic. The constant part is no different from any other utility function parameter (i.e., η or γ).

present value of utility in all periods from $t+1$ until the terminal period T.

The notation $E_t\{\cdot|(K_{t+1}, A_{t+1})\}$ indicates that the expectation E_t is taken conditional on next period's state variables K_{t+1} and A_{t+1}. This is possible because the model is set up so human capital and assets evolve deterministically—i.e., given (K_t, A_t) and the current choice (C_t, h_t), the worker knows the resulting (K_{t+1}, A_{t+1}) with certainty. It is important to note that the expectation is taken assuming choices will be made optimally in all future periods. As a result, it is often called the "*Emax*" or "$Emax_t$" function for short.

Uncertainty in the model arises from only two sources: the wage shocks ε_t and taste shocks β_t.[104] Thus, the expectation $E_t\{\cdot|(K_{t+1}, A_{t+1})\}$ in (96) is taken over possible $t+1$ realizations of ε_{t+1} and β_{t+1}. Because of the independence assumption, ε_t and β_t do not help predict ε_{t+1} and β_{t+1}. Thus, they drop out of the conditioning set. This is why independence greatly reduces the computational burden of solving the agent's optimization problem.

To obtain a full solution of the agent's problem, we must (in principle) solve for the value functions in (96) at every possible state point. This is done via a "backsolving" procedure, where we start with the terminal period T. That is, we start by calculating $V_T(K_T, A_T, \varepsilon_T, \beta_T)$ for every possible state $(K_T, A_T, \varepsilon_T, \beta_T)$ at which the worker might enter period T. Note that, in the terminal period, we simply have

$$(97)\quad V_T(K_T, A_T, \varepsilon_T, \beta_T) = \max_{C_T, h_T}\left\{\frac{C_T^{1+\eta}}{1+\eta} - \beta_T\frac{h_T^{1+\gamma}}{1+\gamma}\right\}.$$

As there is no future beyond T, we have a simple static problem. Given $w_T = RK_T(1+\varepsilon_T)$ and A_T, the consumer chooses consumption and hours of work to maximize utility at time T subject to the static budget constraint $C_T = w_T h_T(1-\tau) + A_T$.[105]

The solution to this static problem for any particular state $(K_T, A_T, \varepsilon_T, \beta_T)$ is given by

$$(98)\quad \frac{\beta_T h_T^{\gamma}}{[w_T h_T(1-\tau_T) + A_T]^{\eta}} = w_T(1-\tau),$$

which can be solved for the optimal h_T via an iterative search procedure.[106] Once the optimal h_T is determined, the optimal C_T is obtained from the budget constraint, and these are both plugged into (97) to obtain $V_T(K_T, A_T, \varepsilon_T, \beta_T)$ at that particular state point.

We see immediately however that it is not feasible to solve for $V_T(K_T, A_T, \varepsilon_T, \beta_T)$ at literally *every* state point: the number of possible levels of human capital, assets, productivity shocks and tastes for work at the start of period T is extremely large, if not infinite. Keane and Wolpin (1994) developed an approach to this problem that has become quite widely used in the literature on dynamic models. The idea is to calculate

[104] Uncertainty, and hence the need to take an expectation of the time $t+1$ outcome, may arise for a number of other reasons. For instance, the rental rate on human capital may evolve stochastically, as in Shaw (1989). Or there may be a stochastic component to how interest rates or tax rates evolve. Such features may be incorporated fairly simply, but they would complicate the exposition.

[105] For expositional simplicity, I assume the end of the working life T corresponds to the end of life, and there are no bequests. Hence, the worker consumes all of his/her remaining assets at time T. In Imai and Keane (2004), the worker values carrying assets into $T+1$ as savings for retirement and for bequests. These extensions are handled by adding to (97) an additional term $f(A_{T+1})$ that represents the value of assets carried into period $T+1$.

[106] As an aside, I'd argue that the basic idea of the life-cycle model with human capital—that working hard today improves one's prospects tomorrow—is one that most people find quite intuitive. Yet one often hears academic economists argue that people can't behave *as if* they solve dynamic optimization problems because the math involved is too daunting. On the other hand, one doesn't often hear academic economists argue that people can't behave according to a static labor supply model because they can't solve an implicit equation for hours like (98). I suspect that most people would find solving an implicit equation daunting as well.

the $Emax_t$ terms like those on the right side of (96) at only a finite (and tractably small) subset of the state points. One then interpolates the $Emax_t$ values at the remaining state points.[107]

To illustrate how the procedure works, we start by using (97)–(98) to calculate $V_T(K_T, A_T, \varepsilon_T, \beta_T)$ at a set of D randomly chosen state points. Denote these solutions by $V_T(K_T^d, A_T^d, \varepsilon_T^d, \beta_T^d)$ for $d = 1, \ldots, D$. We now choose an interpolating function to approximate $Emax_T \ (K_T, \ A_T) \equiv E_{T-1}\{\cdot \mid (K_T, A_T)\}$ at points (K_T, A_T) that are not among the selected points.

Denote the interpolating function that approximates $Emax_T(K_T, A_T)$ by

$$\pi_T(K_T, A_T) \approx E_{T-1}\{V_T(K_T, A_T, \varepsilon_T, \beta_T) \mid (K_T, A_T)\}$$
$$\textit{where} \quad \frac{\partial \pi_T}{\partial K_T} > 0, \frac{\partial \pi_T}{\partial A_T} > 0.$$

We must assume that π_T is a smooth differentiable function of K_T and A_T for the next step. For expositional convenience, assume π_T is the following simple function of K_T and A_T:

$$\text{(99)} \quad \pi_T(K_T, A_T) = \pi_{T0} + \pi_{T1} \ln K_T + \pi_{T2} \ln A_T.$$

We now estimate the parameters of this function by regressing the $V_T(K_T^d, A_T^d, \varepsilon_T^d, \beta_T^d)$ on (K_T^d, A_T^d) for $d = 1, \ldots, D$.

Note that ε_T^d and β_T^d should *not* be included in the regression in (99), as the worker does not use these variables to forecast $V_T(K_T^d, A_T^d, \varepsilon_T^d, \beta_T^d)$. The regression is meant to give a prediction of $V_T(K_T^d, A_T^d, \varepsilon_T^d, \beta_T^d)$ based *only* on (K_T^d, A_T^d), which is how $Emax_T$ is defined.

Once we have fit the regression in (99), we can use it to predict or interpolate the value of $E_{T-1}\{V_T(K_T, A_T, \varepsilon_T, \beta_T) \mid (K_T, \ A_T)\}$ at any desired state point (K_T, A_T), including, in particular, values of (K_T, A_T) that were not amongst those used to fit the regression. Thus, we may proceed *as if* $E_{T-1}\{V_T(K_T, A_T, \varepsilon_T, \beta_T) \mid (K_T, A_T)\}$ is known at every possible state (K_T, A_T).

The next step of the backsolving process is to move back to period $T - 1$. To do this, we will need to be able to solve for $V_{T-1}(\cdot)$ at any particular state point $(K_{T-1}, A_{T-1}, \varepsilon_{T-1}, \beta_{T-1})$. Note that at time $T - 1$ equation (96) takes the form

$$V_{T-1}(K_{T-1}, A_{T-1}, \varepsilon_{T-1}, \beta_{T-1}) = \max_{C_{T-1}, h_{T-1}} \left\{ \left[\frac{C_{T-1}^{1+\eta}}{1+\eta} - \beta_{T-1} \frac{h_{T-1}^{1+\gamma}}{1+\gamma} \right] + \rho E_{T-1} \left\{ \left[\frac{C_T^{1+\eta}}{1+\eta} - \beta_T \frac{h_T^{1+\gamma}}{1+\gamma} \right] \middle| (K_T, A_T) \right\} \right\}.$$

But, if we substitute our approximating polynomial $\pi_T(K_T, A_T) \approx E_{T-1}\{V_T(\cdot) \mid (K_T, A_T)\}$ for the expectation term on the right, we obtain simply

$$\text{(100)} \quad V_{T-1}(K_{T-1}, A_{T-1}, \varepsilon_{T-1}, \beta_{T-1}) \approx \max_{C_{T-1}, h_{T-1}} \left\{ \left[\frac{C_{T-1}^{1+\eta}}{1+\eta} - \beta_{T-1} \frac{h_{T-1}^{1+\gamma}}{1+\gamma} \right] + \rho \pi_T(K_T, A_T) \right\}.$$

[107] Alternative solution methods are discussed in several references, including John Rust (1987), John Geweke and Keane (2001), and Victor Aguirregabiria and Pedro Mira (2010). Olympia Bover (1989) shows how, in the case of a Stone-Geary utility function (with assets as the only state variable), it is possible to obtain an exact closed form solution for the DP problem. But this functional form is very restrictive, so the model provides quite a poor fit to the data.

Using (99), and substituting in the laws of motion for assets and human capital, we obtain

$$V_{T-1} = \frac{C_{T-1}^{1+\eta}}{1+\eta} - \beta_{T-1}\frac{h_{T-1}^{1+\gamma}}{1+\gamma} + \rho\Big\{\pi_{T0} + \pi_{T1}\ln K_{T-1}(1 + \alpha\, h_{T-1}) + \pi_{T2}\ln(1+r)[w_{T-1}h_{T-1}(1-\tau) - C_{T-1} + A_{T-1}]\Big\}.$$

Notice that finding the optimal values of $C_{T\text{-}1}$ and h_{T-1} is now just like a *static* optimization problem. We have the first order conditions

$$(101a)\quad \frac{\partial V_{T-1}}{\partial h_{T-1}} = -\beta_{T-1}h_{T-1}^{\gamma} + \rho\pi_{T1}\frac{\alpha}{(1+\alpha h_{T-1})} + \rho\pi_{T2}\frac{w_{T-1}(1-\tau)}{w_{T-1}h_{T-1}(1-\tau) - C_{T-1} + A_{T-1}} = 0.$$

$$(101b)\quad \frac{\partial V_{T-1}}{\partial C_{T-1}} = C_{T-1}^{\eta} - \rho\pi_{T2}\frac{1}{w_{T-1}h_{T-1}(1-\tau) - C_{T-1} + A_{T-1}} = 0.$$

These two equations can be solved numerically to obtain the optimal (C_{T-1}, h_{T-1}). These values can be plugged into (100) to obtain $V_{T-1}(K_{T-1}, A_{T-1}, \varepsilon_{T-1}, \beta_{T-1})$. Thus, given the interpolating function $\pi_T(K_T, A_T)$, we have a simple way to solve for $V_{T-1}(K_{T-1}, A_{T-1}, \varepsilon_{T-1}, \beta_{T-1})$ at any state point $(K_{T-1}, A_{T-1}, \varepsilon_{T-1}, \beta_{T-1})$ that might arise at $T-1$.

The next step of the backsolving process is to fit an interpolating regression like (99) to obtain an approximating function $\pi_{T-1}(K_{T-1}, A_{T-1}) \approx E_{T-2}\{V_{T-1}(\cdot)\,|\,(K_{T-1}, A_{T-1})\}$ to use to approximate $Emax_{T-1}$. We have already devised a way to solve for $V_{T-1}(\cdot)$ at any particular state point $(K_{T-1}, A_{T-1}, \varepsilon_{T-1}, \beta_{T-1})$. So we randomly pick a subset of D possible state points and solve for $V_{T-1}(\cdot)$ at those points. Denote these solutions by $V_{T-1}(K_{T-1}^{d}, A_{T-1}^{d}, \varepsilon_{T-1}^{d}, \beta_{T-1}^{d})$ for $d = 1,\ldots,D$. We obtain a new interpolating function $\pi_{T-1}(K_{T-1}, A_{T-1})$ by running a regression of the $V_{T-1}(K_{T-1}^{d}, A_{T-1}^{d}, \varepsilon_{T-1}^{d}, \beta_{T-1}^{d})$ on the $(K_{T-1}^{d}, A_{T-1}^{d})$, as in (99). Using this interpolating function, we can write the (approximate) value functions at time $T-2$ as

$$(102)\quad V_{T-2}(K_{T-2}, A_{T-2}, \varepsilon_{T-2}, \beta_{T-2}) \approx \max_{C_{T-2},h_{T-2}}\left\{\left[\frac{C_{T-2}^{1+\eta}}{1+\eta} - \beta_{T-2}\frac{h_{T-2}^{1+\gamma}}{1+\gamma}\right] + \rho\,\pi_{T-1}(K_{T-1}, A_{T-1})\right\}.$$

Note this is exactly like equation (100), the (approximate) value function at time $T-1$, except here we have an approximating function $\pi_{T-1}(K_{T-1}, A_{T-1})$ with different coefficients. The first order conditions for C_{T-2} and h_{T-2} will look exactly like (101), except with different π values. So the problem at $T-2$ is essentially identical to that at $T-1$. Thus, we can keep repeating the above steps until we obtain an approximate solution for every period back to $t = 1$.

When the backsolving process is finished, the (approximate) full solution consists of the complete set of interpolating functions $\pi_t(K_t, A_t)$ for $t = 2,\ldots,T$. Using these interpolating functions we can solve simple two equation systems like (101) to find optimal choices of a worker at any point in the state space. In particular, using $\pi_2(K_2, A_2)$ we can solve for optimal labor supply and consumption in period $t = 1$, the first period of the working life. As I discussed earlier, this is what first order conditions alone do not allow one to do.

Furthermore, by drawing values for the productivity and taste shocks and repeatedly solving equations like (101) over time, one can simulate entire career paths of workers. This in turn, enables one to simulate how changes in tax rates would affect the entire life-cycle path of labor supply and consumption, as one can re-solve the model and simulate career paths under different settings for the tax parameters.

6.3.2.2 *Empirical Results—Imai and Keane (2004)*

The model in Imai and Keane (2004) is in most respects similar to that in (94)–(96). The main difference is they use a much richer specification for the human capital production function, designed to capture the empirical regularity that wages grow more quickly with work experience for higher wage workers. The parameters of the human capital production function are also allowed to differ by education level.

Imai and Keane (2004) estimate their model using white males from the National Longitudinal Survey of Youth 1979 (NLSY79). The men in their sample are aged 20 to 36 and, as the focus of their paper is solely on labor supply, those included in the sample are required to have finished school. Due to the computational burden of estimation, they randomly choose 1,000 men from the NLSY79 sample to use in estimation. People are observed for an average of 7.5 years each, and not necessarily starting from age 16.

Notably, Imai and Keane allow for measurement error in hours, earnings, and assets when constructing the likelihood of the data given their model. They use a ratio wage measure, but account for the resultant denominator bias in forming the likelihood. Given that all outcomes are assumed to be measured with error, construction of the likelihood is fairly simple. One can (1) simulate histories of hours, earnings and assets for each worker, and (2) form the likelihood of a worker's observed history as the joint density of the set of measurement errors necessary to reconcile the observed history with the simulated data.[108]

Imai and Keane estimate that $\gamma = 0.26$. In a model without human capital, this would imply a Frisch elasticity of $(1/\gamma) = 3.8$, which implies a much higher willingness to substitute labor intertemporally than in any estimation we have discussed so far (with the sole exception of MaCurdy 1983). What accounts for this wide divergence in results?

Imai and Keane argue that failure of prior studies to account for human capital led them to severely underestimate $(1/\gamma)$. The logic of their argument is described in figure 4, which gives a stylized plot of male wages and hours over the life cycle.[109] The solid line is (typical) annual hours, while the dotted line is the typical wage path. Many studies have shown that both these curves have hump shapes over the life cycle (see, e.g., the descriptive regressions in Pencavel 1986 or the descriptive statistics in Imai and Keane 2004).

Now, as we have seen, the typical study of the Frisch elasticity regresses hours growth on wage growth. To deal with endogeneity, it instruments for wages using polynomials in age and education. Predicted wages so obtained will follow a hump shape over the life-cycle as in figure 4. Thus, if we regress hours growth on *predicted* wage growth, we essentially uncover the relative slope of the hours and wage curves. The hump in wages is much steeper than the hump in hours, so the estimated elasticity of hours with respect to predicted wages is small.

The dashed line labeled "HC" in figure 4 represents the return to human capital investment—i.e., the return to an extra hour

[108] Keane and Wolpin (2001) first developed this approach to forming the likelihood in dynamic models.

[109] That is, it does not plot any particular data set, but simply illustrates the typical patterns for male wages and hours observed across a broad range of data sets.

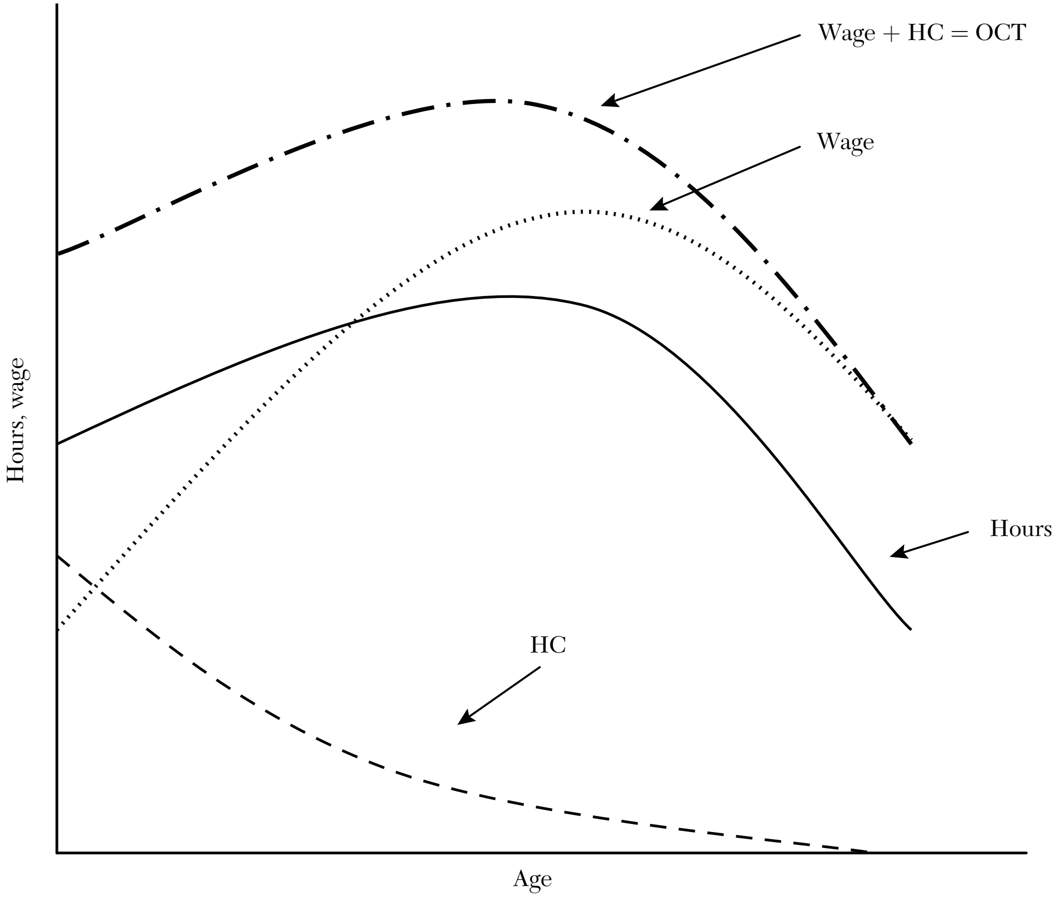

Figure 4. Hours, Wages, and Price of Time over the Life Cycle

Note: HC denotes the return to an hour of work experience, in terms of increased present value of future wages. The opportunity cost of time is Wage + HC.

of work in terms of increased future earnings—given by the second term on the right hand side of (88). The Imai and Keane (2004) estimates imply that at age 20 this human capital return is slightly larger than the wage itself, so in figure 4 the HC line is drawn as starting slightly higher than the wage line. Of course, the human capital investment return declines with age, because the worker approaches the end of the planning horizon T. By age 36, the human capital return is only 25 percent as large as the wage.

The line labeled "OCT" in figure 4 is the opportunity cost of time (OCT), which is the wage plus the human capital return. The Imai–Keane estimates imply that from age 20 to 36 the mean of the OCT increases by only 13 percent. In contrast, the mean wage rate increases by 90 percent in the actual data, and 86 percent in the simulated data. So the wage increases about 6.5 times faster than the OCT. As a result, if we use the relative slopes of the hours and OCT curves to estimate how responsive hours are to changes in the price of time, we will obtain an estimate of $(1/\gamma)$ about 6.5 times larger than if we compare the relative slopes of the hours and wage curves. This is the Imai–Keane argument for why they obtain such a large value of $(1/\gamma)$.[110]

6.3.2.3 *Assessing the Credibility of the Imai–Keane (2004) Results*

Whether the Imai and Keane estimate of $(1/\gamma)$ is credible hinges on several factors. Two in particular are: (1) Can their model replicate results from earlier studies?, and (2) is it plausible that the bias from omitting human capital is as great as Imai and Keane claim?

To address the first issue, Imai and Keane simulate data from their model, and use IV methods like those in MaCurdy (1981) and Altonji (1986) to estimate $(1/\gamma)$. They obtain estimates of 0.325 (standard error = 0.256) and 0.476 (standard error = 0.182), respectively. Thus, the model generates life-cycle histories that, viewed through the lens of models that ignore human capital, imply low Frisch elasticities like those obtained in most prior work.[111]

The authors also compare OLS regressions of hours changes on wage changes for both the NLSY79 and simulated data from their model. The estimates are –0.231 and –0.293, respectively. Thus, the model does a good job of fitting the raw correlation between hours and wage changes in the data. This shows that a negative correlation between hours and wage changes in the raw data is not inconsistent with a high willingness to substitute labor inter-temporally over the life-cycle. What reconciles these *prima facie* contradictory phenomena is the divergence between the OCT and the wage in a world with returns to work experience.

To address the second issue—Is it plausible that the bias from omitting human capital is as large as Imai and Keane claim?—I have done a simple back-of-the-envelope calculation using the two-period model of equations (82)–(87). From (84)–(86) we obtain

$$\left(\frac{h_2}{h_1}\right)^{\gamma} = \frac{\beta_1}{\beta_2}\frac{w_1(1+\alpha h_1)(1-\tau_2)}{\rho(1+r)w_1(1-\tau_1)+\rho\alpha w_1 h_2(1-\tau_2)}.$$

To obtain a more intuitive expression, I set $w_2 = w_1(1+\alpha h_1)$, assume $\tau_1 = \tau_2 = \tau$, and take logs

$$\text{(103)}\quad \ln\left(\frac{h_2}{h_1}\right) = \left(\frac{1}{\gamma}\right)\left\{\ln\left(\frac{w_2}{w_1}\right) - \ln\left(1+\frac{\alpha h_2}{1+r}\right) - \ln\rho(1+r) - \ln\left(\frac{\beta_2}{\beta_1}\right)\right\}.$$

This is the same as the first difference log wage equations often used to estimate the Frisch elasticity (see (25)), *except* for the additional term $\ln(1+\alpha h_2/(1+r))$. If $\alpha > 0$, this is positive. So the existence of learning-by-doing will, ceteris paribus, reduce the rate of hours growth over the life cycle (relative to the exogenous wage case). A model that ignores learning-by-doing will rationalize the apparently small response of hours to wage growth by understating $(1/\gamma)$.

How large is this bias likely to be? One way to look at the problem is to simplify (103) by assuming $\rho(1+r) = 1$ and $\beta_1 = \beta_2$. Then we can solve (103) for $(1/\gamma)$ to obtain

[110] French (2005), in a study of retirement behavior, also obtains a rather large value of $(1/\gamma) = 1.33$ for the intertemporal elasticity of substitution for 60 year olds in the PSID. As both Shaw (1989) and Imai and Keane (2004) note, human capital investment is not so important for people late in the life cycle. For them, the wage will be close to the OCT, and the bias that results from ignoring human capital will be much less severe.

[111] In other words, the model does not generate data that exhibit an oddly high level of positive comovement between hours and wages compared to the actual data.

$$(104)\quad \frac{1}{\gamma} = \ln\!\left(\frac{h_2}{h_1}\right) \div \left[\ln\!\left(\frac{w_2}{w_1\!\left(1+\frac{\alpha h_2}{1+r}\right)}\right)\right].$$

Equation (104) says that, if $\alpha = 0$, we could calculate $(1/\gamma)$ just by taking the ratio of hours growth to wage growth. This is analogous to the standard regression procedure for estimating the Frisch elasticity.[112] But if $\alpha \neq 0$, the human capital term $(1 + \alpha h_2/(1 + r))$, which tells us how much the OCT exceeds the wage at $t = 1$, comes into play. How large is this term?

In a two-period model each period corresponds to roughly twenty years of the working life. It is plausible (in fact, conservative) in light of existing estimates that αh_1, the percentage growth in the wage rate over the first twenty years, is on the order of 33 percent.[113] A plausible value for hours growth from age 25 to the peak at roughly age 45 is a modest value like 20 percent, which implies that αh_2 is roughly 33 percent × (1.20) = 40 percent. A reasonable value for $1/(1 + r)$ is 0.554. Thus, a plausible value for the human capital term $(1 + \alpha h_2/(1 + r))$ is about $1 + (0.40)(0.554) = 1.22$. Plugging these values into (104) we obtain $(1/\gamma) = \ln(1.20) \div \ln(1.33/1.22) = 2.1$.[114]

However, if we mistakenly ignored the human capital term, we would instead obtain $(1/\gamma) = \ln(1.20) \div \ln(1.33) = 0.6$. Thus, in this simple example, plausible (in fact conservative) values for the returns to work experience bias the estimate of $(1/\gamma)$ down by a factor of 3.5. Larger (yet still plausible) returns to experience lead to larger biases.[115]

6.3.2.4 *Implications of the Imai–Keane Model for Effects of Wage and Tax Changes*

Imai and Keane (2004) use their model to simulate responses to a 2 percent temporary and unanticipated wage increase. This generates primarily an intertemporal substitution effect, as a short-lived wage increase has a small effect on lifetime wealth (at least for relatively young workers). For a person at age 20 the increase in hours is only 0.6 percent. At first this may seem surprisingly small, given their estimate of $(1/\gamma) = 3.8$. The reconciliation lies in the fact that, according to Imai and Keane's estimates, at age 20 the wage is less than half the opportunity cost of time. As we would expect, the strength of the substitution effect rises steadily with age. At age 60, the increase in hours is nearly 4 percent, and at age 65 it is about 5.5 percent.

Unfortunately, Imai and Keane did not simulate effects of permanent tax changes, which are more relevant for tax policy. To fill this gap, Keane (2011) uses the Imai–Keane model to simulate permanent tax effects. Note that Imai–Keane estimate $\gamma = 0.26$ and $\eta = –0.74$. In a model without human

112 Of course, in this simple model we abstract from any complicating factors that would require us to use IV.

113 For instance, using the PSID, Geweke and Keane (2000) estimate that for men with a high school degree, average earnings growth from age 25 to 45 is 33 percent. For men with a college degree, they estimate 52 percent. Most of this earnings growth is, in fact, due to wage growth because the growth in hours is modest.

114 An interesting aspect of this example is that wage growth is entirely endogenous, in that it is solely due to work experience. Yet it is still possible to estimate $(1/\gamma)$ by relying on the structure of the model. Note that the OCT in period 2 is just the wage $w_2 = w_1(1 + \alpha h_1)$. The OCT at $t = 1$ is given by the wage times the human capital term, $w_1(1 + \alpha h_2/(1 + r))$. Taking the ratio of the two, we get that the growth in the OCT is $(1 + \alpha h_1)/(1 + \alpha h_2/(1 + r))$. Taking the ratio of hours growth to OCT growth delivers the correct estimate of $(1/\gamma)$.

115 For instance, a back-of-the-envelope calculation in Keane (2010) shows that, given consensus values for the return to work experience in the United States, at age 20 the OCT is roughly double the wage, while at age 40 it is only 20 percent greater. This implies the OCT grows about six times more slowly than the wage, leading to a downward bias by a factor of six in calculating $(1/\gamma)$. These figures are very similar to the Imai and Keane (2004) estimates.

capital, this implies a Marshallian elasticity of $(1+\eta)/(\gamma-\eta) \approx 0.24$ and a compensated (Hicks) elasticity of $1/(\gamma-\eta) \approx 1.0$. However, when Keane (2011) uses the Imai–Keane model to simulate effects of a permanent 5 percent tax rate increase (starting at age 20) on labor supply over the whole working life, he finds uncompensated and compensated elasticities with respect to permanent tax changes of 0.40 and 1.32, respectively.

Thus, human capital magnifies the effect of permanent tax changes beyond what we would expect from plugging the Imai–Keane preference estimates into a static model. This is because, with human capital, a tax cut has a cumulative effect: If a tax cut increases hours at t it will also raise the pretax wage at $t+1$. This causes a further increase in hours at $t+1$, etc. Over the whole life cycle, this cumulative effect is quite dramatic. For example, given a (compensated) permanent 5 percent tax increase starting at age 20, simulated reductions in hours are 3.2 percent, 3.3 percent, 4.2 percent, 5.7 percent, 8.7 percent, and 20 percent at ages 20, 30, 40, 45, 50, and 60, respectively. By age 60, the reduction in the pretax wage rate relative to the baseline case is roughly 10 percent. It is notable that, in response to a permanent tax increase, workers not only reduce labor supply, but also shift their lifetime labor supply out of older ages toward younger ages.

As we have seen, the human capital mechanism dampens the effect of transitory taxes while magnifying the effects of permanent taxes. As Keane (2011) points out, this means that—contrary to conventional wisdom—in a model with human capital, a permanent tax change may actually have a larger effect on *current* labor supply than a temporary tax change. The reason is apparent if we look at (87). A temporary $t=1$ tax increase affects only the current wage $w_1(1-\tau_1)$. But a permanent tax increase, which increases both τ_1 and τ_2, reduces the human return $\alpha w_1 h_2(1-\tau_2)/(1+r)$ as well. Thus, the permanent tax change *may* have a larger effect on the OCT at $t=1$ than a temporary tax change.

I say "may" because Keane (2011) shows the outcome is theoretically ambiguous. A permanent tax change has both (i) a larger income effect and (ii) a larger effect on returns to human capital. These forces work in opposite directions. Keane (2011) presents simulations showing that for plausible parameter values the human capital effect can dominate, so permanent tax changes have larger effects. He also shows that in the Imai–Keane model permanent tax changes have larger effects only for younger workers (those in their 20s and 30s). This is as expected, as the human capital return is most important for younger workers.

Finally, two limitations of Imai and Keane (2004) are that, like most work on dynamic male labor supply models, they ignore both progressive taxation and participation decisions. As Aaronson and French (2009) and Richard Rogerson and Johanna Wallenius (2009) point out, both these factors may have caused prior work to understate labor supply elasticities. Thus, the failure to include them may cause Imai and Keane to exaggerate the role of human capital. However, it seems unintuitive that adding these features would cause one to infer *smaller* elasticities.

6.3.3 *Education, Experience, Saving, and Participation—Keane and Wolpin (2001)*

The first paper to structurally estimate a model with human capital and savings was actually Keane and Wolpin (2001). But unlike Imai and Keane (2004), and other male labor supply papers I have discussed, they did not allow for a continuous choice of hours. Instead, the work options are discrete (full-time, part-time, or not at all). Unfortunately, they do not simulate wage elasticities, so it is hard to compare their results to those of papers I discussed previously. Indeed, their focus was not on labor supply per se, but on school attendance. Still, as we'll see, some of their estimates are directly relevant to the present discussion.

Keane and Wolpin (2001) set out a model where a person (age 16 to 65) decides every period whether to work and/or attend school full-time, part-time, or not at all. Choices are not mutually exclusive (e.g., a youth might work part-time while attending college). Somewhat unusually, the model has three decision periods per year (the two school semesters and the summer). This allows youth to work summers to finance school.

The model is fit to panel data from the NLSY79, which contains people aged 14–21 in January 1979. The sample consists of 1,051 white males who are followed from age 16 until 1992. The maximum age attained in the sample is 30. The NLSY79 collected comprehensive asset data beginning in 1985, making it possible to model savings. A key feature of the model is that it allows for liquidity constraints (i.e., an upper bound on uncollateralized borrowing that is estimated). The model fits data on assets, school, and work from age 16 to 30 quite well.

One reason the paper is of interest here is that it assumes a CRRA utility function in consumption, so it provides an estimate of the key preference parameter η in (3). It is the only paper besides Imai and Keane (2004) to do so in a context with both human capital and saving. Keane and Wolpin (2001) obtain $\eta \approx -0.50$. This compares to the $\eta = -0.74$ obtained by Imai and Keane (2004). The Imai–Keane estimate of η implies a slightly lower intertemporal elasticity of substitution in consumption than the Keane–Wolpin estimate (i.e., $(1/\eta) = 1/(-0.74) = -1.35$ versus $1/(-0.50) = -2.0$).[116] However, both estimates imply weaker income effects, and a higher willingness to substitute intertemporally, than much of the prior literature.

Keane and Wolpin (2001, p. 1078) discuss how failure to model liquidity constraints may have led to downward bias in prior estimates of η.[117] A number of other recent studies also give credibility to values of η in the –0.5 to –0.75 range. Jacob K. Goeree, Charles A. Holt, and Thomas R. Palfrey (2003) present extensive experimental evidence, as well as evidence from field auction data, in favor of $\eta \approx -0.4$ to -0.5. Patrick Bajari and Ali Hortacsu (2005) estimate $\eta \approx -0.75$ from auction data.

A second point of interest is that Keane and Wolpin allow the full and part-time wage functions to differ. They estimate that part-time wages are roughly 15 percent lower than full-time. This enables them to fit the low prevalence of part-time work (see section 6.1.5). Authors like Rogerson and Wallenius (2009) have used this result to motivate models where most of the variability in male labor supply is on the participation margin.

6.3.4 *Efficiency Costs of Taxation in a Life-Cycle Model with Human Capital*

Finally, Keane (2011) uses the Keane and Wolpin (2001) and Imai and Keane (2004) estimates of γ and η to calibrate the simple two period model of equation (83), and uses it to provide simulations of the efficiency cost of income taxation. To do this, he augments the model to include a public good P financed by taxation, as in

[116] The intertemporal elasticity of substitution in consumption is the drop in current consumption in response to an increase in the interest rate (the willingness to sacrifice current consumption for higher future consumption).

[117] Specifically, without constraints on uncollateralized borrowing, one needs a large negative η to rationalize why youth with steep age–earnings profiles do not borrow heavily in anticipation of higher earnings in later life.

$$V = \lambda f(P) + \frac{[w_1 h_1(1-\tau)+b]^{1+\eta}}{1+\eta} - \beta\frac{h_1^{1+\gamma}}{1+\gamma} + \rho\left\{\lambda f(P) + \frac{[w_2 h_2(1-\tau)-b(1+r)]^{1+\eta}}{1+\eta} - \beta\frac{h_2^{1+\gamma}}{1+\gamma}\right\},$$

where $\lambda f(P)$ indicates how people value the public good. The government provides the same level of the public good P in both periods, and the government budget constraint requires that $P + P/(1+r)$ equals the present value of tax revenues. The benevolent government sets the tax rate optimally to equate marginal utility of consumption of the public and private goods.[118]

As we have a two period model we can think of each period as roughly twenty years of a forty year working life (e.g., 25 to 44 and 45 to 64). The annual interest rate is set at 3 percent, and the (twenty-year) discount factor is set to $\rho = 1/(1+r)^{20} = 0.554$. The wage equation is similar to (82), but augmented to include a quadratic in hours and depreciation of skills. Wage equation parameters are calibrated to give 33 percent to 50 percent earnings growth from age 25 to 45, comparable to what Geweke and Keane (2000) find for men in the PSID. Results are obtained for $f(P) = \log(P)$, $f(P) = 2P^{.5}$ and $f(P) = P$, corresponding to cases where P/C declines, is stable or grows as C increases. The *qualitative* results are not very sensitive to this assumption.

[118] In the solution, workers ignore the effect of their own actions on P, as each worker makes a trivial contribution to total government revenue. Thus, workers continue to solve equations (84)–(86).

To calculate efficiency losses, Keane (2011) also solves a version of the model in which a lump sum tax is used to finance the public good. The lump sum tax is set to the level that funds the same level of the public good as in the proportional tax version of the model.

In a version of the model without human capital, it would be typical in the prior literature to estimate $(1/\gamma) \approx 0.25$. For this value, Keane (2011) finds the efficiency loss of proportional income taxation is a bit less than 6 percent of revenue raised, regardless of the form of $f(P)$ or whether η is set to –0.5 or –0.75. This is consistent with the conventional wisdom that efficiency losses from taxation are small if the substitution effect is small. In contrast, when Keane (2011) sets $(1/\gamma) \approx 2.0$, welfare losses as percent of revenue are 20 to 40 percent of revenue (depending on the degree of curvature in utility from the public good). Notably, efficiency losses are not very sensitive to η but are strongly increasing in $(1/\gamma)$.

6.4 *Summary of the Male Labor Supply Literature*

The literature on male labor supply is vast, with many contentious issues. It is thus impossible to arrive at a simple summary. One crude way to summarize the literature is to give a table that lists all the elasticity estimates from the papers I have discussed. I do this in table 6. In many ways, such a table is useless because it makes no attempt to weigh studies based on their relative merits (quality of data, soundness of approach, etc.). Table 6 in effect ignores all the important issues I discussed in sections 4–6.

On the other hand, table 6 is useful for answering the following type of question: "In the male labor supply literature, is there a clear consensus that the Hicks elasticity is small?" Recall that, in section 2, I quoted Saez, Slemrod, and Giertz (2009) as stating: "with some exceptions, the profession has

TABLE 6
Summary of Elasticity Estimates for Males

Authors of study	Year	Marshall	Hicks	Frisch
Static models				
Kosters	1969	−0.09	0.05	
Ashenfelter-Heckman	1973	−0.16	0.11	
Boskin	1973	−0.07	0.10	
Hall	1973	n/a	0.45	
Eight British studies[a]	1976–83	−0.16	0.13	
Eight NIT studies[a]	1977–84	0.03	0.13	
Burtless-Hausman	1978	0.00	0.07–0.13	
Wales-Woodland	1979	0.14	0.84	
Hausman	1981	0.00	0.74	
Blomquist	1983	0.08	0.11	
Blomquist-Hansson-Busewitz	1990	0.12	0.13	
MaCurdy-Green-Paarsch	1990	0.00	0.07	
Triest	1990	0.05	0.05	
Van Soest-Woittiez-Kapteyn	1990	0.19	0.28	
Ecklof-Sacklen	2000	0.05	0.27	
Blomquist-Ecklof-Newey	2001	0.08	0.09	
Dynamic models				
MaCurdy	1981	0.08[b]		0.15
MaCurdy	1983	0.70	1.22	6.25
Browning-Deaton-Irish	1985			0.09
Blundell-Walker	1986	−0.07	0.02	0.03
Altonji[c]	1986	−0.24	0.11	0.17
Altonji[d]	1986			0.31
Altug-Miller	1990			0.14
Angrist	1991			0.63
Ziliak-Kniesner	1999	0.12	0.13	0.16
Pistaferri	2003	0.51[b]		0.70
Imai-Keane	2004	0.40[e]	1.32[e]	0.30–2.75[f]
Ziliak-Kniesner	2005	−0.47	0.33	0.54
Aaronson-French	2009			0.16–0.61
Average		0.06	0.31	0.85

Notes: Where ranges are reported, mid-point is used to take means.
[a] = Average of the studies surveyed by Pencavel (1986).
[b] = Effect of surprise permanent wage increase.
[c] = Using MaCurdy Method #1.
[d] = Using first difference hours equation.
[e] = Approximation of responses to permanent wage increase based on model simulation.
[f] = Age range.

settled on a value for [the Hicks] elasticity close to zero."[119] But, as we see in table 6, the mean value of the Hicks elasticity across twenty-two studies reviewed here is 0.31. (Note that seven studies do not estimate this parameter).

As we have seen, a value of 0.31 for the Hicks elasticity is large enough to generate substantial efficiency costs of taxation. For instance, Ziliak and Kniesner (2005) obtain a Hicks elasticity of 0.33, and simulations of their model imply substantial efficiency costs. And Blomquist (1983) and Blomquist and Hansson-Brusewitz (1990) obtain Hicks elasticities of only 0.11 and 0.13, respectively, yet they also simulate substantial efficiency costs from progressive taxation (i.e., 12 percent and 16 percent of revenue, respectively, compared to only 2 percent or 5 percent under a flat rate tax). Similarly, Ziliak and Kniesner (1999) obtain a Hicks elasticity of 0.13, yet also simulate large efficiency costs of taxation. Based on these results, one would have to conclude that a Hicks elasticity of 0.31 is quite sufficient to generate large efficiency losses.

It is also interesting to display the estimates graphically, as I do in figure 5. Note that, of the twenty-two studies considered here, fourteen produce estimates in a tight range from 0.02 to 0.13. And eight studies produce estimates in the 0.27 to 1.32 range. As the figure makes clear, there is an odd gap between 0.13 and 0.27, with no studies falling in that range. Estimates of the Hicks elasticity seem to bifurcate into a low group versus a high group.

It would be difficult to look at figure 5 and conclude there is a broad consensus within the economics profession that the Hicks elasticity is close to zero—unless, that is, one believes all the studies bunched up in the 0.02 to 0.13 range are credible while all those in the 0.27+ range are flawed. I think such a position would be untenable, as one can also point to flaws in all the studies in the 0.02 to 0.13 range (just as in all empirical work).[120]

The notion there is consensus on a low Hicks elasticity may stem in part from a widespread perception that piecewise-linear budget constraint methods (Burtless and Hausman 1978, Wales and Woodland 1979, and Hausman 1980, 1981) have been discredited, and that *all* the high estimates come from this approach. But as I have discussed, a careful reading of literature suggests this is not the case. These methods have sometimes produced low estimates of the Hicks elasticity, while alternative methods have sometimes produced high estimates. There is no clear connection between the methods adopted and the result obtained.

Indeed, as the careful study by Eklöf and Sacklén (2000) showed, divergent results across studies may be better explained by the data used than by the particular empirical methods employed. In particular, they find that studies that use "direct wage measures" (i.e., a question about ones' wage rate per unit of time, such as hourly or weekly or monthly) tend to get higher estimates of labor supply elasticities than studies that use "ratio wage measures" (i.e., annual earnings divided by annual hours). This is presumably because the denominator bias inherent in taking the ratio biases the wage coefficient in a negative direction.

This pattern can be seen quite clearly in table 6. Specifically, of the eight studies that obtain "large" values for the Hicks elasticity (i.e., those in the 0.27+ range), six use a direct wage measure (Hall 1973, Hausman

[119] At that point, I did not note that they were specifically referring to the Hicks elasticity, as I had not yet defined the different elasticity concepts.

[120] For example, Kosters (1969) does not account for endogeneity of wages, Ashenfelter and Heckman (1973) do not account for taxes, MaCurdy, Green, and Paarsch (1990) and Triest (1990) use ratio wage measures that would lead to denominator bias, Blundell and Walker (1986) do not instrument for full income, and so on.

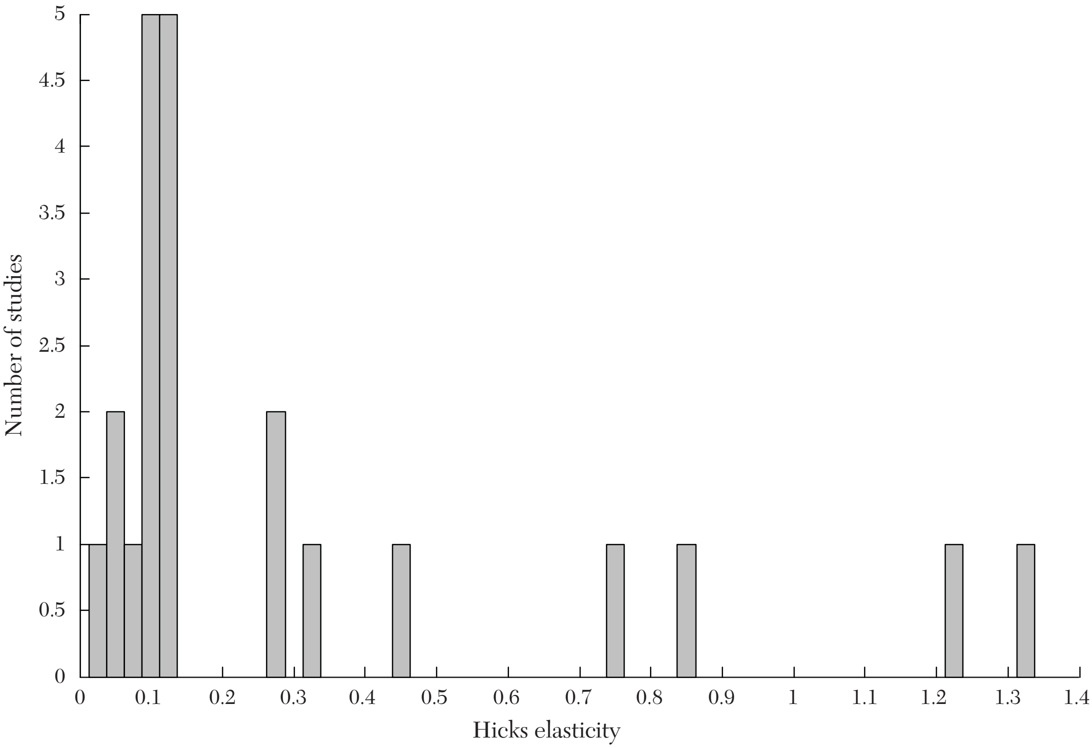

Figure 5. Distribution of Hicks Elasticity of Substitution Estimates

Note: The figure contains a frequency distribution of the twenty-two estimates of the Hicks elasticity of substitution discussed in this survey.

1981, van Soest, Woittiez, and Kapteyn 1990, MaCurdy 1983, [121] Eklöf and Sacklen 2000, Ziliak and Kneisner 2005), one works with shares to avoid ratios (Wales and Woodland 1979), and one models the measurement error process to take denominator bias into account in estimation (Imai and Keane 2004).

Thus, if we give all studies equal weight, the existing literature suggests a Hicks elasticity of 0.31. But if we were to only count the studies that use direct wage measures, we would obtain 0.43.[122] Finally, another

[121] In the Denver experiment, workers were asked a direct question about their wage rate every month. MaCurdy (1983) is a bit vague about how he constructed his wage measure, but from his description I believe he took an average of the answers to these monthly questions over twelve months to get an annual wage.

[122] Of the twenty-two studies examined, fourteen studies either used direct wage measures or made some attempt to deal with the denominator bias problem. In addition to the eight cited above, these include Burtless and Hausman (1978), Blomquist (1983), Blomquist and Hansson-Brusewitz (1990), Blomquist, Eklöf, and Newey (2001), Blundell and Walker (1986), and Ziliak and Kneisner (1999). The average Hicks elasticity among this group is 0.43.

point I have stressed is the failure of prior studies to account for human capital. The effect of human capital is to dampen the response of younger workers to transitory changes in their wage rates. This is because, for them, the wage is a relatively small part of the opportunity cost of time. I believe this has probably led to downward bias in prior estimates of labor supply elasticities. The one study that accounts for this human capital effect, Imai and Keane (2004), obtained a Hicks elasticity of 1.32.

In summary, to conclude there is consensus on a small Hicks elasticity for males, one has to put essentially all mass on the fourteen studies bunched up near zero in table 6 and figure 5. This is hard to justify, as these studies do not share any broad common feature in terms of either methodology or data construction. Indeed, the closest they come to a shared feature is that eight of the fourteen use ratio wage measures, which may well lead to downward biased estimates.

Finally, note that no existing paper deals with all of the issues I discussed in section 4. For instance, no paper has allowed for saving, progressive taxation, human capital, and participation decisions simultaneously. This would obviously be a difficult undertaking.

7. *Female Labor Supply*

Next I turn to the literature on female labor supply. The literature on women has evolved quite differently from that on men. As we saw in section 6, the literature on males has mostly ignored participation decisions because the large majority of prime-age males do work. Thus, researchers have argued (or hoped) that the selection bias induced by ignoring nonworkers would be minimal. The male literature has instead focused on the *continuous* choice of hours, and emphasized savings as the main source of dynamics.

In contrast, a large percentage of women (especially married women) do not work, so the literature has long focused on modeling the participation decision (see Heckman 1974). Nonparticipation brings to the fore: (i) the issue of fixed costs of work (see Cogan 1981) and how they are influenced by marriage and children, and (ii) the question of how tastes for work are influenced by past work decisions (see, e.g., Heckman and Robert J. Willis 1977).

Also, the prevalence of nonparticipation naturally raises the issue of depreciation of human capital. Thus, while the literature on males has mostly treated wages as exogenous, the literature on females has long focused on how work experience affects earnings (see, e.g., Yoram Weiss and Reuben Gronau 1981, Eckstein and Wolpin 1989).

Conversely, the literature on women has placed less emphasis on saving as a source of dynamics. This is no accident: as Eckstein and Wolpin (1989) note, it is very computationally difficult to model participation, human capital and saving simultaneously.[123] So the emphasis on participation and human capital has often come at the expense of not modeling savings.

At least since the pioneering paper by Mincer (1962), the literature on women has found it unsatisfactory (though often practically necessary) to treat marriage and children as exogenous to female labor supply decisions. Instead, it is natural to think of women making decisions—based on their endowments of market and nonmarket skills—about what fraction of the life-cycle to spend in school versus market work versus child rearing, as well as about

[123] Indeed, to my knowledge the only paper that attempts to do so is Keane and Wolpin (2001). That paper is on labor supply and human capital investment decisions of young men, who often have low participation rates.

the timing of marriage and fertility.[124] This life-cycle perspective is already present in Mincer (1962), Heckman and Willis (1977), and Weiss and Gronau (1981).

In Mincer (1962), variation over time in a woman's market work hours stem from her allocating work to periods when market wages are high relative to the value of home time. He hypothesized that, in a life-cycle setting, a transitory change in husband's income (which has no effect on his permanent income), should not affect a woman's labor supply decisions. But Mincer (1962) presented some informal evidence that women do work more if the husband is unemployed, which he took as evidence against a life-cycle model. Of course, alternative explanations are that leisure time of the husband and wife are nonseparable in utility, or that unemployed husbands may contribute to home production and/or child care.

7.1 *Life-Cycle Models with a Participation Margin*

The modern literature on life-cycle models of female labor supply begins with Heckman and MaCurdy (1980, 1982). I focus on the second paper, as it corrects an error in the first. The approach is similar to MaCurdy (1981), except they use the utility function

$$(105)\quad U_{it} = \alpha_{it}\eta^{-1}C_{it}^{\eta} + \beta_{it}\gamma^{-1}(H_{\max} - h_{it})^{\gamma}$$
$$\eta < 1,\ \gamma < 1.$$

Here α_{it} and β_{it} are taste shifters and leisure is given by $L_{it} = (H_{\max} - h_{it})$. The authors assume perfect foresight, so the marginal utility of consumption evolves according to $\lambda_{it} = [\rho(1+r)]^t\lambda_{i0}$. Thus, if w_{it} is the exogenously given time t wage rate, the first order condition for an interior solution for leisure is the usual MRS condition (analogous to (22))

$$(106)\quad \frac{\partial U_{it}}{\partial L_{it}} = \lambda_{it}w_{it}$$
$$\Rightarrow\ \beta_{it}L_{it}^{\gamma-1} = [\rho(1+r)]^t\lambda_{i0}w_{it}.$$

Taking logs and rearranging, this gives the Frisch demand function for leisure:

$$(107)\quad \ln L_{it} = \frac{1}{\gamma-1}\Big\{\ln w_{it} + \ln\lambda_{i0} + t\ln[\rho(1+r)] - \ln\beta_{it}\Big\}.$$

Notice that the utility function (105) admits of corner solutions, in contrast to equation (3) that MaCurdy (1981) used for males. To deal with corner solutions, Heckman and MaCurdy (1980, 1982) note that a woman will choose not to work if the marginal utility of leisure, evaluated at zero hours of work, exceeds the marginal value of working. That is, if

$$(108)\quad \left.\frac{\partial U_{it}}{\partial L_{it}}\right|_{L_{it}=H_{\max}} \geq \lambda_{it}w_{it}$$
$$\Rightarrow\ \beta_{it}H_{\max}^{\gamma-1} \geq [\rho(1+r)]^t\lambda_{i0}w_{it}.$$

Taking logs and rearranging, we can write (108) as a reservation wage condition:

$$(109)\quad h_{it} > 0 \quad \textit{iff}$$
$$\ln w_{it} > -\ln\lambda_{i0} - t\ln[\rho(1+r)] + \ln\beta_{it} - (1-\gamma)\ln H_{\max}.$$

Notice that if the woman has a lower level of lifetime wealth, and hence a higher value of

[124] In contrast, for males, marriage and children tend to be modeled as exogenous taste shifters or as variables that shift the budget constraint. I am not aware of any work on males that treats them as choice variables.

λ_{i0}, her reservation wage is correspondingly reduced.

To obtain an estimable model, Heckman and MaCurdy (1980, 1982) next assume functional forms for the taste shifter β_{it} and the wage equation as follows:

$$\text{(110a)} \quad \ln \beta_{it} = Z_{it}\,\phi + \eta_{1i} + \varepsilon_{1it}$$

$$\text{(110b)} \quad \ln w_{it} = X_{it}\,\theta + \eta_{2i} + \varepsilon_{2it},$$

where Z_{it} and X_{it} are vectors of observables that shift tastes for work and labor productivity, respectively, η_{1i} and η_{2i} are unobserved individual fixed effects, and ε_{1it} and ε_{2it} are transitory shocks to tastes and productivity. Substituting (110) into (107) and (109), we obtain reduced forms for (i) leisure conditional on participation and (ii) the participation decision rule:

$$\text{(111)} \quad \ln L_{it} = f_i + X_{it}\frac{\theta}{\gamma - 1} - Z_{it}\frac{\phi}{\gamma - 1} + \frac{\ln[\rho(1+r)]}{\gamma - 1}t + \frac{\varepsilon_{2it} - \varepsilon_{1it}}{\gamma - 1}$$

$$\text{(112)} \quad h_{it} > 0 \quad \textit{iff}$$

$$\frac{\varepsilon_{2it} - \varepsilon_{1it}}{\gamma - 1} < -f_i - X_{it}\frac{\theta}{\gamma - 1} + Z_{it}\frac{\phi}{\gamma - 1} - \frac{\ln[\rho(1+r)]}{\gamma - 1}t + \ln\ H_{\max},$$

where

$$\text{(113)} \quad f_i \equiv \frac{1}{\gamma - 1}\{\ln \lambda_{i0} + \eta_{2i} - \eta_{1i}\}.$$

Here f_i is an individual specific fixed effect which subsumes the marginal utility of wealth term λ_{i0} as well as the individual effects in tastes for work and productivity.

Under the assumptions of the model (i.e., perfect foresight, no borrowing constraints), the fixed effect f_i is time invariant. It captures everything relevant from periods outside of time t for the woman's labor supply decision at time t. For example, in this model it is not necessary to control for current or potential future earnings of a married woman's husband *explicitly* because it is captured by λ_{i0}. Consider a married woman whose husband has a high income level. But at time t he becomes unemployed. This event will have no affect on λ_{i0} because by assumption it was anticipated and should have already been built in. The same argument applies to indicators for unemployment or hours of work.[125]

While these assumptions seem extreme if taken literally, it is not clear a priori they are necessarily a bad *approximation* to reality, or that they provide a worse approximation than a static model in which women make decisions based only on the current income of the husband. For instance, consider a woman whose husband is in a high wage occupation that is also cyclically volatile. Is it plausible she would substantially revise her perceived lifetime wealth every time his earnings drop in a recession, and alter her labor supply plans as well?

To estimate the model, Heckman and MaCurdy (1980, 1982) assume the stochastic terms ε_{1it} and ε_{2it} are jointly normal and serially uncorrelated, and set $H_{\max} = 8760$. They then estimate the hours and participation equations (111)–(112) jointly with the wage

[125] Indeed, in principle in this model it is not even necessary to control explicitly for whether a woman is married, as the woman's marriage history is also built into λ_{i0}. For instance, a single woman is assumed to anticipate the earnings potential of any husband she will eventually marry. Marriage can only enter the model because it shifts tastes for work, not because it alters perceived lifetime wealth.

equation (110b) by maximum likelihood. The data consist of 30 to 65 year-old married white women from the 1968–75 waves of the PSID. 672 women meet the selection criteria, but to estimate the fixed effects f_i only women who work at least once can be used, leaving 452.[126]

The variables in the wage equation (X_{it}) are "potential" experience (age-education-6) and its square, and the local unemployment rate. Time invariant covariates (like education) cannot be included, as the wage equation contains a fixed effect.

The variables included as taste shifters (Z_{it}) are number of children, children less than 6, the wife's age,[127] and an indicator for if the husband is retired or disabled. Motivated by Mincer (1962), Heckman and MaCurdy also include a measure of "other" family income (i.e., income of the husband and other family members), and the number of hours the husband is unemployed. As noted earlier, transitory changes in the husband's income or employment should *not* affect the woman's labor supply decisions under the assumptions of the model. Thus, if these variables show up as taste shifters, it may indicate misspecification, perhaps due to violation of the perfect foresight or no borrowing constraint assumptions.[128] Or, as noted earlier, it may simply be that leisure of the husband and wife are nonseparable.

Heckman and MaCurdy (1982) estimate $\gamma = -1.44$, which implies a Frisch elasticity of leisure of $1/(\gamma - 1) = -0.41$. Converting to a Frisch labor supply elasticity, and noting that mean hours worked in the sample is about 1,300, we have

$$\begin{aligned} \frac{\partial \ln h_{it}}{\partial \ln w_{it}} &= \frac{\partial \ln h_{it}}{\partial \ln L_{it}} \frac{\partial \ln L_{it}}{\partial \ln w_{it}} \\ &= \frac{L_{it}}{H_{\max} - L_{it}} \frac{1}{1-\gamma} \approx \frac{L_{it}}{h_{it}}(0.41) \\ &= \frac{7460}{1300}(0.41) = 2.35. \end{aligned}$$

This is quite a large value compared to most of the estimates we saw for men.

The other results are mostly standard. Tastes for home time are increasing in number of children, especially children less than 6. Husband unemployment hours are marginally significant and negative. This may suggest the presence of borrowing constraints or failure of the perfect foresight assumption, or it may simply imply that husband time at home increases the wife's tastes for work. The coefficient on other income is quantitatively large, but only significant at the 20 percent level. Heckman and MaCurdy (1982) interpret these results as "less favorable toward the permanent income hypothesis" than those in their 1980 paper.[129]

[126] If a woman never works, we can see from equation (112) that the likelihood of her history is maximized by sending f_i to $-\infty$. Heckman and MaCurdy (1980) report results with and without adjusting the likelihood function to account for this sample section criterion, but find it makes little difference.

[127] Age may capture the time variable in (111)–(112) so that its coefficient is interpretable as an estimate of $\ln[\rho(1 + r)]/(\gamma - 1)$. But it may also affect tastes for work directly.

[128] This is analogous to the literature on testing for borrowing constraints by including current income in consumption Euler equations. Of course, significance of current income in the consumption Euler equation does not necessarily imply the existence of borrowing constraints. It may also arise if leisure and consumption are not separable, due to the fact that income is obviously highly correlated with leisure.

[129] As in MaCurdy (1981), Heckman and MaCurdy (1980) conduct a second stage where they regress the fixed effects on various determinants of lifetime wealth. Using (110b) and (113), we can obtain ($\ln \lambda_{i0} - \eta_{1i}$). That is, the marginal utility of wealth minus the fixed effect in tastes for leisure. Heckman and MaCurdy (1980) find this composite is reduced by wife's education. We would expect education to increase lifetime wealth, thus reducing λ_{i0}, both by increasing own and potential husband's earnings. But the effect of education on tastes for leisure (η_{1i}) is an empirical question. The result implies either that education increases taste for leisure, or, if it reduces it, that this effect is outweighed by the income effect.

7.1.1 *Accounting for Fixed Costs of Work—Cogan (1981), Kimmel and Kniesner (1998)*

The Heckman and MaCurdy (1980, 1982) papers, as well as earlier work in a static framework by Heckman (1974), have been criticized because, while allowing for a participation decision, they did not accommodate fixed costs of work. Within a static model, Cogan (1981) showed that ignoring fixed costs can lead to severe bias in estimates of labor supply functions. To see the problem, consider the simple quasi-linear utility function

$$(114)\quad U = C + \beta\frac{(\overline{H}-h)^{1+\gamma}}{1+\gamma} = (wh+N-F) + \beta\frac{(\overline{H}-h)^{1+\gamma}}{1+\gamma},$$

where N represents nonlabor income and F represents fixed costs of working (e.g., child care costs). The equation for optimal hours *conditional* on working is simply

$$(115)\qquad h^* = \overline{H} - \left(\frac{w}{\beta}\right)^{\frac{1}{\gamma}}.$$

In the absence of fixed costs, the reservation wage would be obtained simply as

$$(116)\quad h^* > 0 \;\Rightarrow\; \overline{H} - \left(\frac{w}{\beta}\right)^{\frac{1}{\gamma}} > 0 \;\Rightarrow\; w > \beta\overline{H}^{\gamma}$$

However, as Cogan (1981) points out, it is inappropriate to use marginal conditions to derive the participation decision rule in the presence of fixed costs. Instead, we must compare the utilities conditional on working and not working:

$$(117)\quad U(h^*) = w\left[\overline{H} - \left(\frac{w}{\beta}\right)^{\frac{1}{\gamma}}\right] + N - F + \frac{\beta}{1+\gamma}\left[\left(\frac{w}{\beta}\right)^{\frac{1}{\gamma}}\right]^{1+\gamma}$$

$$U(0) = N + \frac{\beta}{1+\gamma}\left[\overline{H}\right]^{1+\gamma}.$$

Now the decision rule for working is $U(h^*) > U(0)$, which can be expressed as

$$(118)\quad h^* = \left[\overline{H} - \left(\frac{w}{\beta}\right)^{\frac{1}{\gamma}}\right] > \frac{F}{w} + \frac{1}{w}\frac{\beta}{1+\gamma} \times \left\{\overline{H}^{1+\gamma} - \left[\left(\frac{w}{\beta}\right)^{\frac{1}{\gamma}}\right]^{1+\gamma}\right\} \equiv h_R > 0.$$

It is instructive to compare (116), which simply says the person works if desired hours are positive ($\overline{H} - (w/\beta)^{1/\gamma} > 0$), with (118), which says a person only works if optimal hours cross a positive threshold value h_R, which Cogan (1981) refers to as reservation hours. Inspection of the right-hand side of the inequality in (118) gives a good intuition for what the threshold entails: optimal hours conditional on working must be high enough to cover fixed costs, plus an additional term which equals the monetized value of the lost utility from leisure.

Thus, with fixed costs, the labor supply function is discontinuous, jumping from zero to h_R when the reservation wage is reached. The specifications in Heckman (1974) and Heckman and MaCurdy (1980, 1982) are not consistent with such behavior. Another key point is that both costs of working (F) and tastes for work (β) enter the participation equation, while only β enters the labor supply equation. Hence, it is possible that a variable like young children may affect fixed costs of work but not tastes for work. Then, it would affect participation decisions but not labor supply conditional on participation.

To estimate labor supply behavior given fixed costs, Cogan (1981) jointly estimates a labor supply function as in (115), a reservation hours function as in (118), and an offer wage function. In contrast, Heckman's (1974) approach is to jointly estimate a labor supply function (115), a participation equation based on marginal conditions as in (116), and an offer wage function.

Cogan (1981) compares both approaches using data on married women aged 30 to 34 from the 1967 National Longitudinal Survey of Mature Women. In the sample, 898 wives worked and 939 did not. The labor supply and reservation hours functions both include the wife's education and age, number of young children, and husband's earnings. Cogan estimates that fixed costs are substantial (about 28 percent of average annual earnings), and that a young child raises fixed costs by about a third. He finds that ignoring fixed costs leads to severe overestimates of labor supply elasticities (conditional on work). Cogan's labor supply function implies a Marshallian elasticity of 0.89 at the mean of the data, compared to 2.45 when using the Heckman (1974) approach. The Hicks elasticities are 0.93 versus 2.64.

Cogan also shows, however, that elasticities can be misleading in this context. A 10 percent increase in the offer wage to the average nonworking woman would *not* induce her to enter the labor market. But a 15 percent increase would induce her to jump to over 1,327 hours. Then, an additional 15 percent wage increase would "only" induce a further increase of 180 hours, or 13.6 percent. [Note: this is still a large increase, consistent with a Marshallian elasticity of $13.6/15 = 0.90$].

An important aspect of Cogan (1981) is that he pays close attention to how the model fits the distribution of hours. This is unusual in the static literature, as the focus tends to be on estimating elasticities rather than simulating behavior.[130] Cogan finds the model without fixed costs cannot explain how few people work at very low hours levels. Indeed, it has to predict that many women do work at low hours levels in order to also predict the large fraction of women who do not work at all. As Cogan describes, this leads to a flattening of the labor supply curve, which exaggerates wage elasticities (see his figure 2). The model with fixed costs provides a much better fit to the data and does not have this problem.

Kimmel and Kniesner (1998) extend the Heckman and MaCurdy (1982) analysis to include fixed costs. That is, they estimate a labor supply equation analogous to (111) jointly with a participation decision rule and an offer wage function. We can write the system as

$$(119)\quad \ln h_{it} = f_{hi} + e_F \ln w_{it} + \alpha_h Z_{it} + \varepsilon_{hit}$$

$$(120)\quad P(h_{it} > 0) = F(f_{pi} + \beta \ln w_{it} + \alpha_p Z_{it}).$$

Here (119) is a Frisch labor supply function. The fixed effect f_{hi} captures the marginal utility of wealth, along with heterogeneity in tastes for work. Equation (120) gives the probability of participation and F is a cumulative distribution function (which Kimmel and Kniesner 1998 assume to be normal,

130 The only exceptions I have come across are van Soest, Woittiez, and Kapteyn (1990) and Keane and Moffitt (1998). Both papers note that it is rare to observe people working very low levels of hours (the former paper looking at men, the latter looking at single mothers). Van Soest, Woittiez, and Kapteyn (1990) capture this by building in a job offer distribution where few jobs with low levels of hours are available. Keane and Moffitt (1998) build in actual measures of fixed costs of working (e.g., estimates of child care costs).

giving a probit). The fixed effect f_{pi} captures not just the marginal utility of wealth and tastes for work, but *also* individual heterogeneity in fixed costs of work.

Following Cogan (1981), the existence of fixed costs breaks the tight link between the parameters in the participation and labor supply equations that we saw in (111)–(112). Thus, there is no necessary relationship between the parameters e_F and α_h in (119) and β and α_p in (120). In this framework, e_F is the conventional Frisch elasticity of labor supply conditional on employment. But we now introduce a Frisch participation elasticity given by

$$(121) \quad e_P = \frac{\partial \ln P(h_{it} > 0)}{\partial \ln w_{it}} = \beta \frac{F'(\cdot)}{F(\cdot)}.$$

Kimmel and Kniesner (1998) estimate this model using data on 2,428 women from the Survey of Income Program Participation (SIPP), 68 percent of them married. Triannual interview information was collected in May 1983 to April 1986, giving nine periods of data. The variables included in Z_{it} are marital status, children, education, and a quadratic in time. The model is estimated in two stages. In stage one, wages are predicted for workers and nonworkers using Heckman's (1976) two-step procedure. The use of predicted wages serves three purposes: (i) to deal with measurement error, (ii) to fill in missing wages, and (iii) to deal with possible endogeneity of wages (which would arise, e.g., if women with high unobserved tastes for work tend to have high wages). Variables that appear in the wage equation but not in Z_{it} are race and a quadratic in age (potential experience). In stage two, they estimate (119)–(120).

The estimates imply a Frisch elasticity of 0.66 for employed women, and a Frisch participation elasticity of 2.39. Average hours for the entire population are given by $\overline{h} = P\overline{h}_e$ where $\overline{h}_e$ is average hours of the employed and P is the percentage employed. Thus we have

$$\frac{\partial \ln \overline{h}}{\partial \ln w} = \frac{\partial \ln P}{\partial \ln w} + \frac{\partial \ln \overline{h}_e}{\partial \ln w} = 0.66 + 2.39 = 3.05.$$

Kimmel and Kniesner (1998) also obtain results for men, and find $e_F = 0.39$ and $e_P = 0.86$ so that $e_F + e_P = 1.25$. Thus, the results suggest that: (i) the participation elasticity is much larger than the hours elasticity for both women and men, and (ii) the overall elasticity is quite a bit larger for women than men (although the 1.25 value for men is still larger than most results in table 6). These results provide some justification for models of female labor supply that focus primarily on the participation decision (see below).

7.1.2 *Accounting for Human Capital—Altug and Miller (1998)*

Altug and Miller (1998) occupies a position in the female labor supply literature analogous to the paper by Shaw (1989) in the male literature. That is, they extend the life-cycle model of Heckman and MaCurdy (1980, 1982) to include human capital accumulation (i.e., learning-by-doing). But they also incorporate fixed costs of work, state dependence in tastes for leisure, and aggregate shocks. Thus, they combine ideas from Heckman and MaCurdy (1980, 1982), Shaw (1989), Cogan (1981), and Altug and Miller (1990).

As in Shaw (1989), the first step in Altug and Miller (1998) is to estimate how wages depend on work experience. They specify a wage function of the form

$$(122) \quad \tilde{w}_{it} = \omega_t \nu_i \, \gamma(Z_{it}) \exp(\varepsilon_{it})$$

$$\Rightarrow \quad \ln \tilde{w}_{it} = \ln \omega_t + \ln \nu_i + \ln \gamma(Z_{it}) + \varepsilon_{it}.$$

Here Z_{it} is a vector containing work experience and other characteristics of person i at time t. $\gamma(Z_{it})$ is a function mapping Z_{it} into skill. ν_i is the time-invariant skill endowment of person i. ω_t is a skill rental price (determined in equilibrium). A key assumption is that ε_{it} is purely measurement error. If (122) is expressed as a log wage equation, the $\ln \nu_i$ are individual fixed effects while the $\ln \omega_t$ are time dummies. Given that ε_{it} is measurement error, no selection bias problem arises if we estimate (122) by OLS, provided we include fixed effects.[131]

Altug and Miller (1998) estimate (122) using PSID data from 1967 to 1985. They require that women be in a PSID household for at least 6 consecutive years and that they be employed for at least two years (so that the fixed effects $\ln \nu_i$ can be estimated). This gives a sample of 2,169 women. The data from 1967 to 1974 is used to form indicators of lagged participation and lagged hours, while 1975 to 1985 is used for estimation.

The estimates imply that labor market experience, particularly recent experience, has a large effect on current wages. For instance, a person who worked the average level of hours for the past four years would have current offer wages about 25 percent higher than someone who had not worked. Interestingly, the lagged participation coefficients are negative while lagged hours coefficients are positive. The implication is that low levels of hours do not increase human capital: one has to work about 500 to 1,000 hours to avoid depreciation of skill.

The time dummies from (122) are estimates of the rental price of skill. This falls in the recession years of 1975 and 1980–82, while rising in 1977, 1983, and 1985. Thus, the wage is pro-cyclical. Average wages among all women in the sample are slightly more pro-cyclical than the estimated rental rates. This suggests a compositional effect whereby people with high $\ln \nu_i$ tend to enter during booms. This is consistent with the mild pro-cyclical bias in aggregate wage measures for males found by Keane, Moffitt, and Runkle (1988).

Altug and Miller (1998) assume a current period utility function of the form

$$(123)\quad U_{it} = \alpha_{it}\,\eta^{-1}\,C_{it}^{\eta} + d_{it}\{U_0(X_{it}) + U_1(Z_{it}, h_{it}) + \varepsilon_{1it}\} + (1 - d_{it})\varepsilon_{0it}.$$

The first term is a CRRA in consumption. d_{ir} is an indicator for positive hours, $U_0(\cdot)$ is the fixed cost of work and $U_1(\cdot)$ is the disutility of labor. X_{it} is a vector of demographics that shift tastes for leisure and fixed costs of work. Z_{it} includes X_{it} along with lagged labor supply decisions that are allowed to shift tastes for leisure. ε_{1it} and ε_{0it} are stochastic shocks to tastes for the work and nonwork options, respectively. These may be interpreted as shocks to the fixed cost of work and the value of home time. Additive separability and the distributional assumptions on ε_{1it} and ε_{0it} play a key role in the estimation procedure, as we'll see below.

As in Altug and Miller (1990), the authors assume that there is no idiosyncratic risk. So, using (69), we obtain the following expression for the marginal utility of consumption

$$(124)\qquad \alpha_{it}\,C_{it}^{\eta-1} = \lambda_{it} = \eta_i\,\lambda_t$$

$$\Rightarrow \quad \ln C_{it} = \frac{1}{\eta - 1}\{\ln \eta_i + \ln \lambda_t - \ln \alpha_{it}\}.$$

Again, the λ_t are aggregate shocks and the η_i capture a person's (time-invariant) position

131 Altug and Miller (1998) actually estimate the wage equation in first differences, and use GMM to gain efficiency by accounting for serial correlation in the errors. They note Thomas A. Mroz (1987) found selection corrections have little impact on estimates of fixed effects wage equations for women. Similarly, Keane (1993a) found selection corrections have little impact on estimates of fixed effects occupational wage equations for males.

in the wealth distribution.[132] To obtain an estimable equation let $\ln \alpha_{it} = X_{it}\beta + \varepsilon_{cit}$ where X_{it} and ε_{cit} are observed and unobserved (exogenous) taste shifters for consumption. Then (124) is estimated by fixed effects. Altug and Miller include household size, age, children and region in X_{it} and the aggregate shocks are estimated as time dummies. The equation is estimated on the same PSID sample as above (recall that the PSID contains only food consumption). As we would expect, the estimated values of λ_t are high in the recession years of 1975 and 1980–82.

In the final step, Altug and Miller (1998) estimate the first order condition for hours jointly with a participation condition that allows for fixed costs of work. As in Shaw (1989), the FOC for hours is complex because the marginal utility of leisure is not simply equated to the current wage times the marginal utility of consumption. An additional term arises because working today increases future wages and alters future disutilities from work. I'll refer to this term as the "expected future return to experience." The situation here is more complex than in Shaw (1989) because of nonparticipation—i.e., work today may increase probabilities of future participation (an effect not present in Shaw (1989), where men work with probability one). Altug and Miller handle this problem using the Hotz and Miller (1993) algorithm.

First, given estimates of (122) and (124), we can back out estimates of the individual effects ν_i and η_i. Second, use nonparametric regression to estimate participation probabilities conditional on the state variables ν_i and η_i, work history, and demographics (age, education, marital status, race, children, age, and region). Denote these estimates $p_{1it}(S_{it})$ where S_{it} is the vector of state variables.[133] Third, assume the ε_{1it} and ε_{0it} in (123) are iid extreme value, and that they are the only source of randomness in current period payoffs from working versus not working. Given this, the differences in expected values of working versus not working are simply $V_{1it}(S_{it}) - V_{0it}(S_{it}) = \ln[p_{1it}(S_{it})/(1 - p_{1it}(S_{it}))]$, so the value functions at any state can be backed out from the conditional choice probabilities calculated in step 2 . This allows one to express the "expected future return to experience" as a simple function of the conditional participation probabilities (and their derivatives with respect to h_{it}).[134]

It is important to see what is ruled out here. There can be (i) no stochastic variation in the marginal utility of leisure and (ii) no individual level productivity shocks affecting wages. Such additional sources of randomness would preclude obtaining simple expressions for the expected future return to experience. And consumption and leisure must be separable in utility, so the stochastic term in tastes for consumption does not influence labor supply decisions. Thus, the extreme value error and additive separability assumptions are crucial.

In this final estimation step, the parameters to be estimated describe the fixed costs of work $U_0(X_{it})$ and the disutility of labor $U_1(Z_{it}, h_{it})$. Unfortunately, the results are problematic. The estimated $U_1(Z_{it}, h_{it})$ is convex in hours, so the estimated first order condition implies no interior solution.

132 Alternatively, in a social planner's problem, η_i is the inverse of the social planner's weight on person i.

133 It is important *not* to include the aggregate prices λ_t and ω_t in these regressions. Agents are assumed not to know future realizations of these prices, and so cannot condition on them when forming expected future payoffs.

134 Specifically, the "return to experience" term can be written as a function of differences in the expected values of working versus not working in future states, $V_{1it}(S_{it}) - V_{0it}(S_{it})$, as well as probabilities of working in future states, $p_{1it}(S_{it})$, and their derivatives with respect to current hours. See Altug and Miller (1998) equations 6.8 and 6.9, which give the final simple expressions for the labor supply and participation equations. [And notice how the idea here is similar to that in equation (89).]

And the fixed costs are very imprecisely estimated. These results may stem in part from the restrictiveness of the assumption of no stochastic variation in tastes for work.

7.2 *The "Life-Cycle Consistent" Approach—Blundell, Duncan, and Meghir (1998)*

So far I have discussed approaches that involve estimating the MRS condition for optimal hours. I now turn to the "life-cycle consistent" approach, where one estimates labor supply equations that condition on the full income allocated to a period (MaCurdy 1983 method #2). Recall that Blundell and Walker (1986) estimated a life-cycle consistent model of labor supply of married couples. They use data on couples where both the husband and wife work, and estimation is done jointly with a probit for whether the wife works (to control for selection into the sample). In section 6.2.1, I discussed their results for men, and here I turn to their results for women. In sharp contrast to Heckman and MaCurdy (1982) and Kimmel and Kniesner (1998), they obtain an (average) Frisch elasticity for women of only 0.033. The Hicks elasticity is 0.009. Based on the figures in their paper, I calculate an income effect of –0.206 (at the mean of the data) and a Marshallian elasticity of –0.197. (Limitations of this paper, especially treating consumption as exogenous, were discussed earlier).

More recently, Blundell, Duncan, and Meghir (1998) applied the life-cycle consistent approach to married women in the FES from 1978–92. U.K. tax rates fell substantially over the period, and the basic idea of the paper is to exploit this variation to help identify labor supply elasticities. As the authors describe, the decline in rates caused different cohorts to face different tax rate paths. Relative wages for different education groups also changed markedly.

The basic idea of the paper is as follows: Imagine grouping the data by cohort and education—i.e., for each cohort/education level, construct group means of hours and wages in each year. Then subtract group and time means from these quantities. The key assumption in Blundell, Duncan, and Meghir (1998) is that any residual variation in wages (after taking out group and time means) is exogenous. Their leading example of what might cause such residual variation in wages for a group is tax changes that affect it differentially from other groups. (Another example is exogenous technical change that affects groups differently.) Their key assumption rules out labor *supply* shifts *within* any of the groups over time (e.g., tastes for leisure can vary by cohort or education, but not within an education/cohort group over time).

They also assume that taking out time means purges both hours and wages of all groups from the influence of aggregate shocks. This seems like a strong assumption, as time affects (like the business cycle) may well affect different education/skill groups differently. In this regard, see the earlier discussion of Angrist (1991) and Altug and Miller (1990).

The simplest way to think about using the grouped data is to regress the group mean of hours on the group mean of wages, after purging these means of group and time effects. An equivalent approach is to use the individual data and proceed in two steps. First, regress after-tax wages on time/group interaction dummies, and obtain the residuals. Second, regress hours on the after-tax wage, time and group dummies, and the wage residual. Note: we want the wage coefficient to be identified by wage variation by group over time. The wage residual captures other sources of variation, as the first stage controls for time/group interactions.[135]

[135] An alternative computational approach to taking out group and time means is to regress the group mean of hours on the group mean of wages and a complete set of time and group dummies. Then the wage effect is identified purely from the wage variation not explained by aggregate time or group effects. The advantage of the more involved two-step procedure is that the coefficient on the residual provides a test of exogeneity of wages.

The authors also try to deal with the compositional effects of changes in participation rates on the mean of the error term in the labor supply equation (e.g., a higher wage may induce women with higher tastes for leisure to enter the market). So they include an inverse Mills ratio term that is a function of the group/time participation rate, $M(P_{gt})$. The labor supply equation that Blundell, Duncan, and Meghir (1998) actually estimate has the form:

$$(125)\ h_{it} = \beta \ln w_{it}(1-\tau_{it}) + \gamma[C_{it} - w_{it}(1-\tau_{it})h_{it}] + X_{it}\phi + d_g + d_t + \delta_w R_{wit} + \delta_c R_{cit} + M(P_{gt}) + e_{it}.$$

The second term is virtual nonlabor income allocated to period t (see discussion of MaCurdy 1983 method #2). X_{it} is a vector of demographics (i.e., dummies for children in various age ranges). d_g and d_t are the group and time dummies. R_{wit} and R_{cit} are residuals from first stage regressions of wages and virtual income on group/time interactions. And $M(P_{gt})$ is a Mills ratio to control for participation rates. Estimation of (125) is by OLS. But, with the inclusion of R_{wit} and R_{cit}, the procedure is equivalent to IV, with the group/time interactions as overidentifying instruments.[136] The identifying assumption is that the main equation (125) does not contain group/time interactions (e.g., no group specific trends in tastes for work).

[136] As I noted earlier, Blundell and Walker (1986) treated virtual income as exogenous. I questioned this on the grounds that, in the first stage of the two stage budgeting process, we might expect households to allocate more virtual income to periods when tastes for work are low. I should note, however, that Blundell, Duncan, and Meghir (1998) find that R_{cit} is insignificant in (125), suggesting that endogeneity of virtual income may not be a problem.

To implement this procedure Blundell, Duncan, and Meghir (1998) group the FES into two education groups (legal minimum versus additional education) and four cohorts (people born in 1930–39, 1940–49, 1950–59 and 1960–69), giving eight groups. They screen the data to include only 20 to 50 year old women with employed husbands. This gives 24,626 women of whom 16,781 work. Only workers are used to estimate (125) while the full sample is used to form the $M(P_{gt})$. One detail is that 2,970 of these women are within a few hours of a kink point in the tax schedule. Blundell, Duncan, and Meghir drop these women from the data and construct additional Mills ratio terms to deal with the selection bias this creates. In the first stage, they find the group/time interactions are highly significant in the wage and virtual income equations.

The estimates of (125) imply an uncompensated wage elasticity at the mean of the data of 0.17 and a compensated elasticity of 0.20. In a sensitivity test, the Blundell, Duncan, and Meghir also report results where, in the first stage, the overidentifying instruments are five parameters that describe the tax rules interacted with group dummies. This reduces the number of instruments relative to the case where the group dummies were fully interacted with time dummies. It also means that only variation in wages and virtual income specifically induced by tax changes is used to identify the labor supply elasticities. The estimates give an uncompensated elasticity of 0.18 and an essentially zero income effect. Thus, results are little affected.

7.3 *"Approximate Reduced Form" Approach (Fertility)—Moffitt (1984)*

The paper by Moffitt (1984) departs from those reviewed so far in three key ways. First, it focuses only on the discrete participation decision, ignoring choice of hours. Second, it treats work and fertility choices as being made jointly, rather than treating fertility as exogenous. Third, wages are endogenous in that they depend on work experience. 1984 technology would have made fully structural estimation of such a complex model infeasible. Instead, Moffitt estimates an "approximate reduced form" of the structure outlined above. Thus, his work forms a link to the "full solution" approaches that I discuss in the next section.

At the beginning of marriage, a couple plans the future path of the wife's labor supply and fertility. The husband's income stream, along with other sources of nonlabor income, is taken as exogenous (perfect foresight). The woman's labor supply and fertility plans depend on this exogenous nonlabor income stream (Y), along with her "permanent wage" or "skill endowment," denoted w^*_{1i}, which she also knows with certainty. Let $\ln w^*_{1i} = Z_i\,\eta + \mu_{wi}$ where Z_i is a vector of observed determinants of initial skill (e.g., education, race, parent education) for woman i, and μ_{wi} is the unobserved part of the skill endowment. The structural model generates the following approximate decision rules for fertility (B_{it}) and work (S_{it})

$$\begin{aligned} B^*_{it} &= a_0 + a_1 f(t) + a_2 \ln w^*_{1i} \\ &\quad + a_3 Y_i + a_4 X_i + a_5 B_{i,t-1} \\ &\quad + \mu_{Bi} + u_{it} \\ S^*_{it} &= b_0 + b_1 f(t) + b_2 \ln w^*_{1i} \\ &\quad + b_3 Y_i + b_4 X_i + b_5 B_{i,t-1} \\ &\quad + \mu_{Si} + v_{it}. \end{aligned}$$

The woman chooses to have a child if the latent variable $B^*_{it} > 0$, and to work if $S^*_{it} > 0$. The parameters a_2 and b_2 determine how a woman's skill endowment affects her probabilities of having children and working, respectively. Similarly, a_3 and b_3 determine the influence of the present value of exogenous nonlabor income. Variables in X_i affect tastes for fertility/work (i.e., education, race, birth cohort). a_1 and b_1 capture effects of marriage duration. Parameters a_0 and b_0 will capture fixed costs of work/fertility. Lagged births ($B_{i,t-1}$) affect current tastes and/or fixed costs (but lagged work does not). Finally, μ_{Bi} and μ_{Si} are permanent unobserved heterogeneity in tastes/fixed costs of fertility/work, while u_{it} and v_{it} are transitory shocks.

A key point is that the skill endowment $\ln w^*_{1i}$ is not observed. Thus, Moffitt (1984) infers it from observed wages, using the wage function

$$\begin{aligned} \ln w_{it} &= \ln w^*_{1i} + \gamma \sum_{\tau=1}^{t-1} S_\tau - \delta(t-1) + \varepsilon_{it} \\ \Rightarrow \ln w_{it} &= Z_i\,\eta + \gamma \sum_{\tau=1}^{t-1} S_\tau - \delta(t-1) \\ &\qquad + \mu_{wi} + \varepsilon_{it}. \end{aligned}$$

Here γ captures the effect of work experience and δ captures skill depreciation. The term ε_{it} is a stochastic shock to wages. Substituting the wage equation into the B^*_{it} and S^*_{it} equations, we obtain the reduced form

$$\begin{aligned} B^*_{it} &= a_0 + a_1 f(t) + a_2(Z_i\eta) \\ &\quad + a_3 Y_i + a_4 X_i + a_5 B_{i,t-1} \\ &\quad + (a_2 \mu_{wi} + \mu_{Bi} + u_{it}) \\ S^*_{it} &= b_0 + b_1 f(t) + b_2(Z_i\eta) \\ &\quad + b_3 Y_i + b_4 X_i + b_5 B_{i,t-1} \\ &\quad + (b_2 \mu_{wi} + \mu_{Si} + v_{it}). \end{aligned}$$

Two key points are essential to note here. First, the transitory wage error ε_{it} is *not* included in these reduced form decision rules (Furthermore, u_{it} and v_{it} are assumed uncorrelated with ε_{it}). This can be rationalized in two ways: either (i) ε_{it} represents only measurement error, or (ii) work decisions are made before transitory wage draws are revealed. Either way, this means the results will not be informative about effects of transitory wage changes on labor supply.

Second, identification of permanent wage effects on labor supply/fertility (b_2 and a_2) requires Z_i to contain at least one variable that does not affect tastes (X_i). Playing this role are parent education and year-of-marriage, the latter to capture productivity growth over time.[137]

Moffitt (1984) estimates the wage equation jointly with the reduced form decision rules for fertility and work via maximum likelihood (assuming errors are normal). Note that the unobserved skill endowments μ_{wi}, which are treated as random effects, enter all three equations. Thus, joint estimation corrects for (i) selection bias and (ii) endogeneity of work experience in the wage equation, both of which arise because those with higher μ_{wi} are more likely to work (accumulating more work experience) and to have observed wages.[138]

Moffitt estimates the model using married women from the NLS Young Women sample who were 14–24 in 1968. The sample covers the years 1968–75. Simulations of his estimated model imply that the long run (uncompensated) elasticity of life-cycle labor supply with respect to a permanent wage increase is 1.25. The elasticity of fertility is –0.25. It is important to note that the labor supply elasticity reported here differs from those reported earlier: it is a "long run" response that accounts for how the wage change alters fertility. This key point will come into play in many of the models discussed in the next section.

7.4 *Female Labor Supply—Full Solution Structural Methods*

7.4.1 *Participation and Human Capital—Eckstein and Wolpin (1989)*

The first paper to adopt a full solution approach to modeling female labor supply was Eckstein and Wolpin (1989). Indeed, it is the first paper to model labor supply of any group using a discrete choice dynamic programming (DCDP) approach (provided we maintain a distinction between labor supply models and job search models such as Wolpin 1987). The paper looks at work decisions by married women in the NLS Mature Women's cohort.

The main focus of the paper is on how the decision to work today affects wages and tastes for work in the future. Thus, it considers three of the four issues that I stated at the outset were central to the female labor supply literature: (i) participation decisions and how they are influenced by fixed costs, (ii) human capital, and (iii) state dependence in tastes for work.[139] In order to make estimation feasible (particularly given 1989 technology), Eckstein and Wolpin (1989) make some key simplifying assumptions. First, they ignore saving and assume a static budget constraint. Second, they ignore the choice of hours of work and treat labor supply as a discrete work/no-work decision.

This set of decisions is notable, as it illustrates well the different paths the male and female life-cycle labor supply literatures have

[137] Note that cohort enters X_i while year-of-marriage enters Z_i. This restriction is debatable—it seems at least as natural to use birth cohort to capture productivity change as year-of-marriage. Also, the two variables are presumably highly correlated. Thus, I expect it is the parent education variables that primarily identify a_2 and b_2.

[138] This approach is only valid if the transitory wage errors do not influence labor supply decisions. Otherwise it would be necessary to also accommodate correlation between ε_{it} and u_{it} and v_{it}. The only two correlations allowed are between μ_{Bi} and μ_{Si} and between u_{it} and v_{it} —the errors in the fertility and work equations.

[139] The fourth issue, which is not yet addressed here, is the attempt to treat marriage and fertility as endogenous.

taken. The literature on males has emphasized decisions about hours and savings, which Eckstein and Wolpin (1989) ignore. But work on males has usually ignored participation, human capital, and state dependence, which Eckstein and Wolpin stress. This is not a value judgment on either literature, but simply an observation about what aspects of behavior researchers have found it most essential to model in each case. The emphasis on participation, human capital, and state dependence explains why the female literature came to use DCDP methods several years earlier than the male literature, as these features are difficult to handle using Euler equation methods.

A third key simplifying assumption is that Eckstein and Wolpin do not model marriage or fertility. To avoid having to model fertility decisions, the paper looks only at women who were at least 39 years old in 1967 (and so for the most part past child bearing age). The number of children affects fixed costs of work, but it is treated as predetermined. And marriage is taken as exogenously given. Including marriage and fertility as additional choice variables would not have been feasible given 1989 technology. But, as we'll see, incorporating them as choice variables has been the main thrust of the subsequent literature.

Eckstein and Wolpin (1989) assume a utility function of the form:

$$(126)\quad U_t = C_t + \alpha_1 p_t + \alpha_2 C_t p_t + \alpha_3 X_t p_t + \alpha_4 N_t p_t + \alpha_5 S_t p_t.$$

Here p_t is an indicator of labor force participation, X_t is work experience (a sum of lagged p_t), N_t is a vector of children in different age ranges, and S_t is schooling. The budget constraint is

$$(127)\quad C_t = w_t p_t + Y_t^H - cN_t - bp_t,$$

where w_t now stands for the potential earnings of the wife and Y_t^H is the annual income of the husband (assumed exogenous).[140] That utility is linear in consumption has some important consequences. First, substitution of (127) into (126) makes clear that we cannot separately identify fixed costs of work b and monetary costs of children c from the disutility of work α_1 or effects of children on the disutility of work α_4. So Eckstein and Wolpin normalize $b = c = 0$.

The second implication is that the model will not exhibit income effects on labor supply unless consumption and participation interact in (126). For instance, if $\alpha_2 = 0$ then husband's income will have no impact on the wife's labor supply. But a clear pattern in the data is that women with higher income husbands work less (see, e.g., Mincer 1962). For the model to capture this, it must have $\alpha_2 < 0$. This implies consumption and leisure are compliments in utility. This illustrates a limitation of the model, as a negative income effect and consumption/leisure complimentarity are conceptually distinct phenomena.

Eckstein and Wolpin (1989) assume a Mincer-type log earnings function (linear in schooling, quadratic in work experience), with both a stochastic productivity shock and measurement error. These are the only stochastic terms in the model, as there are no shocks to tastes for work. This simplifies the solution to the dynamic programming problem, which takes the form of a set of reservation wages (which are a deterministic function of age, experience and other state variables).[141] The

140 Annual earnings if the woman works are assumed to equal 2,000 times the hourly wage rate, regardless of how many hours the woman actually works. This is necessitated by the 1/0 nature of the work decision.

141 Eckstein and Wolpin (1989) also assume that husband earnings is a deterministic function of husband age, a fixed effect, and a schooling/age interaction. If there were taste shocks or shocks to husband earnings they would have to be integrated out in solving the DP problem.

decision rule for participation is to work if the offer wage exceeds the reservation wage. Measurement error accounts for cases where women are observed to make decisions that violate this condition.

Eckstein and Wolpin estimate the model by maximum likelihood using data on 318 white married women from the NLS Mature Women cohort. The NLS interviewed them only eleven times in sixteen years from 1967–82, so in some cases it was difficult to construct complete work histories. To be in the sample a woman had to have at least four consecutive valid years of data on employment, and have a spouse present in every interview. The data set contained 3,020 total observations, 53 percent for working years. The discount factor is fixed at 0.952.

One interesting aspect of the estimates is they show substantial selection bias in OLS wage equation estimates. The OLS schooling coefficient is 0.08 while the model estimate is 0.05. The experience profile is initially less steep but also less strongly concave than implied by OLS. The estimates imply that 85 percent of observed wage variation is measurement error.[142]

As expected, Eckstein and Wolpin find that children (especially young children) have a negative effect on tastes for work. State dependence is imprecisely estimated, but it implies experience reduces tastes for work. Schooling reduces tastes for work as well. However, both effects are heavily outweighed by positive effects of experience and schooling on wages.

Eckstein and Wolpin also find $\alpha_2 < 0$, so husband income reduces the wife's work. Consider a woman at age 39 with fifteen years of work experience, twelve years of schooling, no children and a husband with \$10,000 in annual earnings (which is close to the mean in the data). The baseline prediction of the model is that she will work 5.9 years out of the 21 years through age 59, or 28 percent of the time. If his earnings increase 50 percent, the model predicts her participation rate will drop to 14 percent, a 50 percent decrease. So the elasticity of the participation rate with respect to nonlabor income is roughly –1.0. Converting this to an income effect, and noting that the mean wage in the data is \$2.27 per hour and that work is assumed to be 2,000 hours per year, we obtain $ie = (wh/I)e_I = [(2.27)(2000)/10{,}000](-1) = -0.45$.

Unfortunately, Eckstein and Wolpin do not simulate how an exogenous change in the wage rate (due to a shift in the rental rate or tax rate) affects labor supply. However, as schooling is exogenous, and the effect of school on tastes for work is very small, we can approximate this using the schooling coefficient. Consider the same representative woman described above, and assume her education level is increased from twelve to sixteen. An extra four years of schooling raises the wage rate roughly 22 percent at the mean of the data. The Eckstein–Wolpin model predicts this will increase her participation rate from age 39 to 59 by 108 percent. Thus, the (uncompensated) elasticity of the participation rate with respect to the wage is roughly 5.0.

Finally, Eckstein and Wolpin (1989) report a detailed description of how their model fits labor force participation rates, conditional on twenty-eight experience/age cells (see their table 5). In general the model provides a very good fit to the data. As I noted earlier, there are very few papers in the static literature, or the literature on dynamic models

142 Note that the measurement error in wages cannot be estimated using wage data alone in the absence of multiple measures. But joint estimation of a wage equation and a labor supply model does allow measurement error to be estimated, as true wage variation affects behavior while measurement error does not. Of course, any estimate of the extent of measurement error so obtained will be contingent on the behavioral model.

based on first order conditions, that examine model fit.[143] In contrast, as we will see below, since Eckstein and Wolpin (1989), careful examination of model fit has become standard practice in the DCDP literature. This situation has presumably arisen because the focus of the former literatures is estimation of parameters or elasticities, while the focus of the DCDP literature is on model simulations under baseline versus policy change scenarios. And it is only natural to compare baseline simulations to the actual data. But clearly it should be standard practice to assess model fit in all econometric models (including static models, reduced form models, etc.).

7.4.2 *Extensions to Make Marriage and Fertility Endogenous*

The next paper in the DCDP literature on female labor supply did not appear until Wilbert van der Klaauw (1996). He extends Eckstein and Wolpin (1989) to make marriage a choice. Thus women have up to four options in each period—the cross product of work and marriage choices. Also, van der Klaauw models decisions starting from when a woman left school, which may be as young as 14. So obviously he cannot treat fertility as given. Thus, he models arrival of children as a stochastic process, where arrival probabilities depend on state variables (marital status, age, race, and education). This is common practice in DCDP modeling—i.e., to take variables one believes are endogenous, but which one does not wish to model explicitly as a choice (either for computational reasons or because they are not the focus of the analysis), and treat them as generated by a stochastic process that depends on state variables.[144]

The model is in many other ways similar to Eckstein and Wolpin (1989). There is a static budget constraint, with utility linear in consumption. Consumption is again interacted with participation, to enable the model to explain why women with high income husbands work less. Specifically, utility conditional on participation and marriage (p_t, m_t) is given by

$$(128) \quad U_{pm,t} = a_{1t}\, m_t + (a_{2t} + a_{3t}\, m_t) p_t + (\beta_1 + \beta_2\, p_t + \beta_3\, m_t) C_{pm,t} + \varepsilon_{pm,t}.$$

Note that tastes for marriage a_{1t} are allowed to depend on demographics, children, and lagged marriage. Marriage (m_t) is also interacted with consumption, so it can shift the marginal utility of consumption. The effects of demographics, children, and lagged participation on tastes for work are captured by letting a_{2t} and a_{3t} depend on these variables.

Recall that in Eckstein and Wolpin (1989) a woman got utility from total household consumption. Here, she consumes her own income plus a fraction of the husband's income (depending on her work status), so she gets utility from private consumption. A single woman has a probability of receiving a marriage offer each year. A potential husband is characterized only by his mean wage, which depends on the *woman's* characteristics (reflecting marriage market equilibrium), a transitory wage draw, and transitory shocks to utility of the married states (i.e., $\varepsilon_{pm,t}$ for $m = 1$). The latter capture any nonpecuniary aspects of the marriage offer.

[143] As noted earlier, the only exceptions I have found in the static literature are Cogan (1981), van Soest, Woittiez, and Kapteyn (1990) and Keane and Moffitt (1998). In the literature on dynamic models based on first order conditions, it is not possible to examine model fit, as one cannot simulate data from the model.

[144] This procedure does have limitations. For instance, in this model, an exceptionally good wage draw for the woman (or an exceptionally bad wage draw for her husband) could not induce a woman to delay childbearing.

It is worth noting that this is only a search model of marriage in a trivial sense. There is no match-specific component to the marriage; a husband does not come with a permanent component to his earnings level, which could make him a "good draw" given a woman's demographics. Nor is there any permanent component to the utility level he provides. Thus, a woman has no reason to decline a marriage offer in the hope of a better one. Her only reasons for delay are (i) the mean of the husband income distribution is increasing with a woman's age, and (ii) transitory aspects of offers vary over time. This setup substantially reduces the computational burden of estimation, as there is no "husband type" variable that must be included in the state space. But at the same time it renders the model rather uninformative for assessing the effect of permanent differences in husband income on the wife's labor supply, as all permanent differences are a deterministic function of the wife's own characteristics.

The woman's own wage offer function includes standard covariates like education, a quadratic in experience, race, age, and region. It also includes a lagged participation indicator, which lets more recent work experience be more important (see Altug and Miller 1998). An unusual aspect of the specification, however, is that it is specified in levels, with an additive error. This is also true of the husband's wage function. Given this setup, when these functions are substituted into the budget constraint to get the choice $(p_t,\ m_t)$-specific consumption level, $C_{pm,t}$, and this in turn is substituted into the utility function (128), each of the four alternatives has an additive error that consists of the $\varepsilon_{pm,t}$ plus a function of the female and male wage equations errors. I'll denote these four composite errors as $e_{pm,t}$ for $p = 0, 1$, $m = 0, 1$.

The key computationally aspect of van der Klaauw (1996), which enabled him to handle making marriage a choice, is that he assumes the four additive choice-specific errors are iid extreme value. This lets him adopt Rust's (1987) closed form solution method, which makes solving the DP problem and forming choice probabilities much faster. The extreme value assumption is hard to evaluate. There is much evidence suggesting wage errors are approximately log normal. But we have little to go on when choosing a distribution for taste shocks, let alone sums of taste and wage shocks. What does appear very strong is the within period *iid* assumption: The $e_{pm,t}$ for $p = 0, 1$ and $m = 1$ contain common taste for marriage and husband income draws, and the $e_{pm,t}$ for $m = 0, 1$ and $p = 1$ contain common shocks to the wife's wage and tastes for work. So we would expect the four errors to be correlated.[145]

The model is estimated on PSID data from 1968 to 1985. The sample includes 548 females aged 12 to 19 in 1968, so complete work and marital histories can be constructed. They are 29 to 36 by the end of the sample period. The terminal age is set at 45 to reduce computational burden. As the discount factor is set at 0.85, this may be innocuous.[146] It is assumed that $p_t = 1$ if the woman worked at least 775 hours in a year, but, as in Eckstein and Wolpin (1989), the work choice is assumed to entail 2,000 hours regardless of actual hours. This approximation is necessitated by the discrete nature of the work decision.

The model is estimated in stages. In the first stage, the "reduced form" model (with the wage equations substituted into (128)) is estimated via Rust's (1987) method. In the second stage, the wage equations are estimated using employment and marriage decision rules from the reduced form model to implement a selection correction. In the

145 An idea that might prove useful here is the generalized extreme value distribution (see Peter Arcidiacono 2005).

146 E.g., if a person leaves school at age 22, then the number of periods is $45 - 22 = 23$ and $(0.85)^{23} = 0.024$.

third stage, a minimum distance estimator (see Gary Chamberlain 1984) is used to uncover the structural parameters. All reported model simulations are for the "reduced form" model.

The utility function estimates imply that children reduce the utility from participation while lagged work increases it. The wage equation estimates are a bit difficult to compare to prior literature as they are in levels. For instance, they imply that a year of schooling raises a woman's earnings by \$1,379 per year. As mean earnings in the data are \$13,698 per year, this is roughly 10 percent at the mean of the data. A year of schooling also raises potential husband's earnings by \$1,266 per year (versus a mean of \$19,800) or 6.4 percent. This suggests that an important part of the return to schooling for women comes through the marriage market.[147]

Van der Klaauw (1996) presents considerable evidence on model fit. It provides a good fit to the proportion of women who are working/married conditional on years since leaving school, to marriage rates by age, and to hazard functions for marriage and divorce. It also provides a good fit to the proportion of women making each of the four marital status/work choices conditional on work experience and age.

Van der Klaauw uses the model to simulate the impact of exogenous increases in annual offer wages and husband offer wages. A \$1,000 wage increase (7.3 percent on average) leads to a 2.5 year (26 percent) increase in work experience by age 35, implying an uncompensated labor supply elasticity of roughly 3.6. It is notable, however, that this is not comparable to a conventional Marshallian elasticity that holds all else fixed. In particular, the wage increase causes a 1 year increase in average years to first marriage, and a 1.3 year decrease in total years of marriage. The fall in marriage is part of what induces the increase in labor supply.[148]

There is another large time gap until the next significant paper in the DCDP literature on female labor supply, which is Marco Francesconi (2002). In contrast to van der Klaauw (1996), he extends Eckstein and Wolpin (1989) to make fertility a choice. He also allows for both full-time and part-time work. Thus, women have six choices in each annual period (after age 40 only the three work options are available). Francesconi allows for separate full and part-time wage functions (so the model can explain lower wages for part-time work), and he lets full and part-time experience have separate effects on offer wages. Thus, the model has three endogenous state variables: number of children, and part-time and full-time experience.

Marriage is taken to be exogenous and the model begins when a woman first gets married. The terminal period is age 65. Women are assumed to make decisions based on *expected* husband income. As in Eckstein and Wolpin (1989), women get utility from total household consumption, net of fixed costs of work and costs of children. There is again a static budget constraint, with utility linear in consumption. Utility conditional on

147 The estimates imply that a married woman who works receives 34 percent of husband income. Unfortunately, the share if she does not work is not identified. As we see from (128), if a married woman does not work her utility from consumption is $(\beta_1 + \beta_3)$ times her share of husband income. Only this product is identified in the model.

148 Van der Klaauw (1996) simulates that a \$1,000 (5 percent) increase in husband offer wages reduces mean duration to first marriage by 1 year, increases average years of marriage (by age 35) by 2.3 years, and reduces average years of work by 2.6 years (27 percent). These are very large income effects, but they are not comparable to standard income effect measures, as they refer to changes in husband offer wages, not actual husband wages. Also, it is not clear how much credence to give to these figures: as noted earlier, all permanent differences in husband income in the model are generated by differences in the wife's own characteristics

the part-time and full-time work and fertility choices (p_t, f_t, n_t) is given by

$$(129)\quad U_{pfn,t} = C_{pfn,t} + a_{1t}\, p_t + a_{2t} f_t + (a_3 + \varepsilon_t^n)(n_t + N_{t-1}) + a_4(n_t + N_{t-1})^2 + \{\beta_1\, p_t\, C_{pfn,t} + \beta_2 f_t\, C_{pfn,t} + \beta_3\, n_t\, C_{pfn,t}\} + \{\beta_4\, p_t\, n_t + \beta_5 f_t\, n_t\}.$$

Tastes for part and full-time work, a_{1t} and a_{2t}, are a function of children, work experience and schooling. Tastes for children $(N_t = N_{t-1} + n_t)$ depend on the stochastic term ε_t^n. In the first term in curly brackets, consumption is interacted with all the choice variables (p_t, f_t, n_t). This allows husband income to affect work and fertility decisions. In the second term in curly brackets, work and fertility decisions are interacted. This enables the model to capture the fact that newborn children greatly reduce workforce participation.

The stochastic terms in the model are in the full and part-time wage equations and in tastes for children. There are no shocks to tastes for work. Thus, as in Eckstein and Wolpin (1989), he assumes wages are measured with error to account for observations where women work at wages below the reservation wage. Given that the model contains six choices and three stochastic terms, the evaluation of the Emax function integrals is difficult. Francesconi uses a simulation method similar to that of Keane and Wolpin (1994) to approximate them (see section 6.3.2.1).[149] The three dimensional choice probability integrals are also simulated.

Francesconi (2002) also follows van der Klaauw (1996) by assuming mean income of the husband is purely a function of a woman's characteristics (i.e., age at marriage, education, age). This reduces the size of the state space, as no husband specific characteristic (e.g., a husband skill endowment) need be included in the state vector. But, as a result, effects of husband income on the wife's behavior can only be identified to the extent we invoke some exclusion restrictions, such that certain characteristics of the wife affect only the husband wage and not the wage or tastes of the wife. In fact, the husband wage function includes the wife's age, age at marriage and education/age of marriage interactions, and all of these variables are excluded from the wife's wage function and from her taste parameters.

Finally, Francesconi (2002) extends earlier DCDP models of female labor supply by following the procedure in Keane and Wolpin (1997) to allow for unobserved heterogeneity. Specifically, he allows for three discrete types of women in terms of their skill endowments (the intercepts in the offer wage functions) and in tastes for children (a_3 and a_4).

The model is estimated on 765 white women from the NLS Young Women Survey who were interviewed sixteen times (in twenty-four years) from 1968–91. To be included in the sample, a woman must be at least 19 and be married to the same spouse for the whole sample period.[150] Part-time is

[149] However, unlike Imai and Keane (2004), the state space here is small enough that Francesconi can simulate the Emax function at every state point (there is no need to interpolate between points). The state space is small because Francesconi assumes only the number of children, not their ages, enters the state. If children of different ages had different effects on labor supply, the state space would grow astronomically. Francesconi can capture that newborns have a larger effect on labor supply than older children, because newborns are a current choice variable. Thus, they do not enter the state, as they are no longer newborns in the next period. But allowing, e.g., children aged 1–5 to have a different effect than children aged 6–17 would lead to a great increase in complexity.

[150] This is a subsample of a group of 1,783 women who were married at least once during the period.

defined as 500 to 1,500 hours and full-time is defined as 1,500+ hours. The discount factor is fixed at 0.952. Unlike the multistep procedure in van der Klaauw (1996), decision rules and wage functions are estimated jointly, and wage errors are log normal.

The wage function estimates imply a year of schooling raises full-time offer wages by 8.4 percent and part-time offer wages by 7.6 percent (intermediate between the Eckstein–Wolpin and van der Klaauw results). Full-time experience has a large positive effect on full-time offer wages, while part-time experience has a much smaller effect. Experience effects on part-time wages are generally small. Measurement error accounts for 63 percent of the variance of observed wages. At the mean of the data, an extra year of school raises husband wages by 11 percent, consistent with the finding of van der Klaauw (1996) that much of the return to schooling for women comes through the marriage market. The interaction terms between consumption, work and fertility (β_1, β_2, β_3) are all negative. This generates negative income effects on labor supply and fertility. Also, the high-skill type has relatively low tastes for children.

Francesconi (2002) shows the model provides a good fit to all six annual choice options up to twenty-four years after marriage (which corresponds to age 47 on average). This is where the observed data ends. He also fits a static model (i.e., discount factor set to 0) and finds that it too provides a good fit to the in-sample data. But the models differ dramatically in their out-of-sample predictions. The static model predicts women's labor supply will increase sharply after about age 47 and into the 60s. The DCDP model implies work will stay flat and then drop slowly in the 60s. The latter prediction is much closer to what we observe in the CPS.[151] The problem with the static model is that it explains low participation rates as resulting from the presence of children, so when children leave participation rises sharply. The dynamic model is able to counteract this effect with a declining return to human capital investment as one approaches the terminal period (see the earlier discussion of Imai and Keane 2004).

Finally, Francesconi conducts several simulations of how permanent wage changes affect labor supply. For example, consider an average woman with two years of full-time work experience at the time of marriage. In the baseline simulation of the model, she works for 6.8 out of the 11 years from age 30 to 40. Now consider an increase in the log wage function intercept (i.e., in the rental price of skill) that increases offer wages by roughly 10.5 percent. This increases full-time work by roughly 60 percent, implying an elasticity of labor supply with respect to rental price of skill of roughly 5.6. Note however that this is somewhat of an exaggeration, as some of the increase in full-time must come from reduced part-time work. Unfortunately, Francesconi does not report the drop in part-time work that accompanies this experiment.

7.4.3 *Human Capital, Marriage, Fertility, and Welfare—Keane and Wolpin (2010)*

The next two papers I discuss are Keane and Wolpin (2007, 2010). These papers use approximate solution methods developed in Keane and Wolpin (1994), and new estimation methods developed in Keane and Wolpin (2001), to estimate a model of female life-cycle behavior that is considerably richer than previous models in the literature. Both marriage and fertility are treated as choices, and both full and part-time work options are available. Schooling is also a choice. An important feature of the data not accommodated in prior dynamic models is that many

[151] Neither model captures the sharp decline in participation in the 60s due to retirement. But to be fair, neither model incorporates any features designed to explain retirement behavior (such as pensions or Social Security).

single women with children participate in welfare programs. Thus, welfare participation (when eligible) is also a choice.

In the model, women begin making decisions at age 14, and the terminal period is age 65. The fertile period is assumed to last until age 45. During that time, women have up to thirty-six choice options in each period.[152] Afterwards they have up to eighteen options. The decision period is assumed to be six months up until age 45. This is a compromise between the length of a school semester and the child gestation period. After age 45, the decision period is one year (as the fraction of women who either attend school or have children after 45 is negligible).

Given that behavior of children as young as 14 is being modeled, it is essential to consider the role of parental coresidence and parental income support. But, as this is not a focal point of the model, the authors do not treat living with parents as a choice. Instead, consistent with practices we have discussed earlier, both the probability of coresidence and parental transfers are treated as stochastic processes that depend on a person's state variables.

One fundamental difference from van der Klaauw (1996) and Francesconi (2002) is that marriage is a true search process. Each period a woman may receive a marriage offer that consists of: (1) the mean wage of the husband, and (2) a taste for marriage draw (which captures nonpecuniary aspects of the match). The mean wage of the potential husband is drawn from a *distribution* that depends on a woman's characteristics, such as her schooling and skill level. The distribution may produce either a good or a bad draw, and the husband specific mean wage remains fixed for the duration of the marriage if an offer is accepted. In this setup, a woman has an incentive to reject marriage offers while waiting for a husband with a high mean wage. And a husband fixed effect becomes part of the state space.

Another fundamental difference from prior work is nonstationarity. That is, welfare rules change over time and differ by state. So each cohort of women (defined by the calendar time period when they reach age 14) in each state faces a different sequence of welfare rules. This creates serious problems: First, each cohort of women in each state faces a different dynamic optimization problem (raising computational burden). Second, one must make an assumption about how women forecast future rules. Third, the rules are complex, making it hard to characterize them. (Similar problems would arise in modeling progressive taxation).

Keane and Wolpin (2007, 2010) deal with these problems as follows: First, they develop a simple five parameter function that characterizes the welfare benefit rules in each State and each year quite accurately. Second, they assume women use a state-specific vector autoregression in these five parameters to predict future rules. Third, they only use data from five large states, so as to reduce the number of DP problems that must be solved in estimation. This enables them to use other states for out-of-sample validation.

Keane and Wolpin assume a woman receives disutility from a variable that measures "nonleisure" time (h_t). This is a sum of work hours, a fixed time cost of work, time spent in school, time required to collect welfare, and time needed to care for children.[153] The authors estimate weights on activities other than work to allow other time uses to entail more/less disutility than market work time. A woman's consumption is a share of

[152] The choice set differs across women for several reasons. For instance, only unmarried women with children under 18 can participate in welfare, and working while on welfare is not an option if the offer wage rate is high enough that income would exceed the eligibility level. Also, girls under 16 cannot choose marriage.

[153] Childcare time is, in turn, is a weighted sum of time required to care for children in different age ranges.

total (net) household income. Utility is quadratic in h_t and linear in consumption. As in papers discussed earlier, consumption is interacted with h_t. The estimated coefficient is negative, so consumption and leisure are compliments. This induces negative income effects on labor supply and fertility.[154]

Women also receive utility/disutility from children, pregnancy, marriage, school, and welfare participation. Utility from pregnancy is a polynomial in age. As expected, it becomes a large negative as women approach 45, consistent with greater risks of pregnancy at older ages. The disutility from welfare enables the model to explain nonparticipation by eligible women (see Moffitt 1983). The utility function coefficient on each of the five choice variables (hours, pregnancy, marriage, school, and welfare) consists of a constant plus a stochastic taste shock. This enables the model to generate a nonzero probability for any observed choice.

The model allows for unobserved heterogeneity in the form of six types of women who have different vectors of constants on the five choice variables (different tastes), and different intercepts in the wage functions (different skills). The model includes observed heterogeneity as well: the skill/taste parameters are allowed to differ across states and ethnic groups (blacks, whites, Hispanics). Finally, the utility function includes interactions of full and part-time work, school, marriage, and welfare participation with lagged values of these variables, to capture state dependence in tastes for these choice options.[155]

The model is estimated on data from the NLSY79, which contains women aged 14 to 21 in 1979. The data span the years 1979–91, so the maximum age is 33. The states used in estimation are California, Michigan, New York, North Carolina, and Ohio. To be included in the sample, a woman must reside in the same state for the whole period, which screens out 30 percent. This leaves data on roughly 2,800 women.[156] The annual discount factor is fixed at 0.93.

Estimates of the log wage function imply that (at the mean of the data) an extra year of school raises wages by 9.1 percent. And 84 percent of the variance of wages is due to measurement error (the true log wage standard deviation is 0.17). The experience coefficients imply that the first year of full-time work raises wages by 2.6 percent, and the experience profile peaks at 36 years. Lagged full-time work raises the current offer wage by 7 percent; lagged part-time raises it by 3 percent. Ceteris paribus, part-time wages are 10 percent lower than full-time (Both this and a fixed cost of work were needed to explain the relatively low prevalence of part-time). Blacks and Hispanics have lower offer wages than whites, by 13 percent and 6 percent, respectively).

In the husband offer wage function, the coefficient on the woman's skill endowment (i.e., the intercept in her wage function) is 1.95, implying strong assortative mating. An extra year of education for the woman raises the husband offer wage by 3 percent, and women receive 55 percent of household income. So, as in van der Klaauw (1996) and

154 Keane and Wolpin (2007, 2010) add additional interactions to let marriage and children shift the degree of complimentarity between consumption and leisure. This would have been irrelevant in the papers discussed previously, as they do not model labor supply, marriage, and fertility jointly. The estimates imply marriage and children both reduce the degree of complimentarity between consumption and leisure, but do not eliminate it.

155 The utility function includes some additional terms to capture detailed features of the data. Full and part-time are interacted with school to capture that people who work while in school tend to work part-time. They are also interacted with high school, as part-time work is far more prevalent in high school. Pregnancy is interacted with school to capture that women rarely go to school when pregnant. Tastes for school, marriage, and pregnancy are allowed to shift at key ages (16, 18, 21). And there is a linear time trend (across cohorts) in tastes for marriage.

156 Keane and Wolpin (2002) provide a more detailed description of these data.

Francesconi (2002), much of the return to schooling comes through the marriage market. However, black and Hispanic women have much lower husband offer wages than whites (by 30 percent and 14 percent, respectively).

Estimates of the utility function parameters are interesting in that they show no significant differences in tastes for leisure, school, or welfare participation between black, white, and Hispanic women. But black and Hispanic women do get more utility from children. And black (Hispanic) women get less (more) utility from marriage. There is clear evidence of state dependence in tastes for school, part-time work, and full-time work.

Keane and Wolpin (2007) provide a good deal of evidence on the fit of the model, and assess how well it predicts behavior in the holdout state of Texas. The model performs rather well in these tests, including providing better predictions than some alternative reduced-form models. Keane and Wolpin (2010) use the model for a variety of policy experiments. These focus on (i) factors accounting for differences among blacks, whites, and Hispanics in choice behavior, (ii) effects of changing welfare rules, and (iii) effects changing offer wages. Here I focus on the labor supply simulations. For instance, for women in the 22.5 to 25.5 age range, a 5 percent increase in the skill rental price causes average weekly hours to increase by 14 percent, from 25.8 hours to 29.4 hours. This implies a labor supply elasticity of roughly 2.8.

Recall that the model has six types of women, which we can rank order by skill level from type 1s (highest skill endowment) to type 6s (lowest). Type 6s account for the majority of welfare participants. Keane and Wolpin (2010) report experiments where they increase the offer wage by 5 percent for each type separately. The wage elasticities are inversely proportional to skill level, ranging from only 0.6 for type 1s to 9.2 for type 6s. Thus, the overall elasticity of 2.8 is deceptive with regard to behavior of various subsets of the population.

For type 6 women, the 5 percent wage increase has a dramatic impact on many aspects of behavior. For instance, for whites of type 6, the percent working at ages 22 to 29.5 increases from 34 percent to 50 percent (a 47 percent increase). But also notable is that mean schooling increases from 11.5 to 12 years, the high school drop out rate drops from 42 percent to 24 percent, welfare participation drops from 25 percent to 20 percent, and incidence of out-of-wedlock teenage births drops from 3.4 percent to 2.8 percent. All these behavioral changes (more education, fewer teenage pregnancies, less welfare participation) contribute to the increase in labor supply. In contrast, type 1s already complete a high level of schooling, rarely have children at young ages, do not participate in welfare, and participate in the labor market at a high rate. Thus, there are fewer ways that a wage increase can affect them. In summary, these results indicate that wage elasticities of labor supply for low skilled women are much greater than for high skilled women.

Finally, I emphasize that a weakness of all the empirical work discussed in sections 7.1 and 7.3–7.4 is that it does not explicitly account for taxes in the estimation. A partial exception is Keane and Wolpin (2010), who account for taxes/transfers at low income levels.

7.5 *Assessing the Impact of Tax Reforms—Eissa (1995, 1996a)*

In two influential papers, Eissa (1995, 1996a) assessed the impact of the Economic Recovery Tax Act of 1981 (ERTA81) and Tax Reform Act of 1986 (TRA86) on labor supply of married women. In contrast to the work I have described previously, she does not seek to estimate structural labor supply models, but instead relies on a "difference in differences" (DD) approach. The idea of DD is to compare the behavior of a "treatment" group that was substantially affected by tax changes with that of a "control" group that was little affected.

Eissa assumes that married women condition their labor supply on husband's income, and nonlabor income of the household. These determine the marginal tax rate on her first dollar of earnings. As both ERTA81 and TRA86 flattened the tax structure, the reduction in marginal tax rates was greater for women with high income husbands. For instance, to evaluate ERTA81, Eissa (1996a) uses women whose husbands earned at least \$50,000 as the treatment group, and those whose husbands earned 30k–50k as the controls. Between 1980 and 1984 (as ERTA was phased in), the former group experienced an increase of 12.3 percent in the after-tax share $(1 - \tau)$, while the latter group experienced only a 5.4 percent increase.[157]

In the baseline, 1980, the participation rates of the two groups were 41.9 percent and 56.3 percent. By 1984, they had grown to 49.9 percent and 61.8 percent, respectively. Thus, the participation rate of the women with high-income husbands grew 19 percent, while for those with lower-income husbands it grew 9.7 percent. The idea of DD is to view the change for the control group (9.7 percent) as capturing common time factors affecting both groups. Then, we can factor these out, and calculate the elasticity of participation with respect to the tax change as $e_T = [19.0 - 9.7]/[12.3 - 5.4] = 1.35$.

Two things are worth noting here. First, I use the notation e_T to denote that this is an elasticity with respect to this particular tax change. It does not correspond to the Marshall or Hicks elasticity per se. Indeed, Blundell and MaCurdy (1999, p. 1613) note the DD approach uncovers a weighted average of the Marshall and Hicks, reducing to the Marshall if the tax has no differential income effects. In the present context, much of the income effect of the tax cut is likely to operate through the husband's after-tax income (e.g., if this effect were greater for women with higher income husbands, we would understate the Marshallian elasticity).

Second, the noncomparability of the treatment and control groups, made clear by their different baseline participation rates (41.9 percent versus 56.3 percent) raises two issues: The more fundamental is, if the groups differ to begin with, can we safely assume they are affected in the same way by policy changes? The more prosaic is, with different baseline rates, there is no unique way to calculate the DD estimator—e.g., we could adopt the view that the control group captures the percentage *point* (not percent) change that occurs due to common factors. As participation grew by 8.0 percentage points for the treatments and 5.5 points for the controls, we obtain $8.0 - 5.5 = 2.5$ percentage points as the tax effect. This implies participation growth of $0.025/0.419 = 6.0$ percent, so we have $e_T = 6.0/[12.3 - 5.4] = 0.87$. This is the participation elasticity that Eissa (1996a) actually reports, but I see no reason to choose this over the 1.35 figure. Accounting for the higher wage growth of the more educated women (2.2 percent greater) reduces these two figures to $e_T = 6.0/9.3 = 0.65$ and $e_T = 9.3/9.3 = 1.0$, respectively.

Eissa also calculates an hours elasticity of about 0.60. Baseline hours (conditional on working) are fairly similar for two groups, so this result is not so sensitive to how the DD estimator is constructed. The implied *total* annual hours elasticity is about 1.25 to 1.60.

Eissa (1995) uses a similar procedure to analyze the response to TRA86. I focus on her results using women with husbands at

157 Eissa (1996a) uses the March CPS for all her calculations. A limitation of these data is that tax deductions are not recorded and must be imputed. I discussed this problem with measuring taxes in section 4.

the 99th percentile of the income distribution as the treatment group and husbands at the 90th percentile as the controls. The reform increased the after-tax wage by 29.1 percent for the treatment group and 12.3 percent for the control group, implying a 16.8 percent relative wage increase for the treatments.[158]

In the baseline, 1983–85, the participation rates were 46.4 percent and 61.1 percent for the two groups. By 1989–91 they had risen to 55.4 percent and 65.6 percent, respectively. Thus, the participation rate of the women with high-income husbands grew 19.5 percent, while that for women with lower-income husbands grew only 6.5 percent. The DD estimate of the elasticity of participation with respect to the tax change is thus $e_T = [19.0 - 6.5]/16.8 = 0.74$. Or, if we take the alternative approach of looking at percentage points, we get $e_T = \{[9.0 - 4.5]/0.464\}/16.8 = 0.58$. For the elasticity of *total* annual hours, we get $e_T = [34.5 - 14.7]/16.8 = 1.18$ using percent changes, while using level changes we get $e_T = [\{206 - 129\}/596]/16.8 = 0.77$.

Focusing on total annual hours, it is interesting that the range of DD estimates for ERTA81, which is 1.25 to 1.60, is higher than that for TRA86, which is 0.77 to 1.18. Eissa (1995) notes that TRA86 was designed to be roughly revenue and distributionally neutral. As high-end rates were cut, a variety of exemptions benefiting high-income people were also eliminated. Thus, she argues the reform was designed to generate a compensated substitution effect. In contrast, the high-end rate cuts of ERTA81 were uncompensated. So it is surprising that TRA86 generated a smaller response. Despite this puzzle, Eissa's results seem to strongly suggest that the response of married women to taxes is quite elastic.[159]

7.6 *Summary of the Female Labor Supply Literature*

A problem arises in summarizing labor supply elasticity estimates for women because the nature of what is estimated differs greatly across studies. I discussed several studies that estimate what might be called "short run" elasticities, holding marriage, fertility, and work experience fixed. But Moffitt (1984) and the DCDP models estimate "long run" elasticities that allow, depending on the study, experience, fertility, marriage, and/or education to adjust to wage changes. And Eissa (1995, 1996a) estimates responses to particular tax law changes. Nevertheless, table 7 attempts to summarize these varied elasticity estimates.

A reasonable assessment of table 7 is that labor supply elasticity estimates for women are generally quite large. DCDP models give uniformly large "long run" elasticities ranging from 2.8 to 5.6. The life-cycle models of Heckman and MaCurdy (1982) and Kimmel and Kniesner (1998) give large Frisch elasticities (2.35 to 3.05). The Marshallian elasticity of 0.89 obtained by Cogan (1981) in a static model is also quite large.[160] Thus, nine of the eleven listed studies obtain large female labor supply elasticities (of various types). Only Blundell and Walker (1986) and Blundell, Duncan, and Meghir (1998) find generally small elasticities. This may

[158] This is actually the increase in the after tax rate $(1 - \tau)$. Eissa (1995) presents evidence that wages did not grow significantly for the 99th percentile group relative to the 95th percentile group.

[159] Notably, Eissa did not try to use a DD approach to analyze effects of these tax reforms on men. As she notes, it is far more difficult to find a plausible control group for men. Eissa (1996b) does compare behavior of high versus low education men (the former having higher wages and therefore being more affected), but she does not claim to find the results very convincing (given the obvious noncomparability of these groups).

[160] Note this is an elasticity for hours conditional on working. It is unfortunate that Cogan does not report a participation elasticity. Given his estimates, this would presumably have been much larger.

TABLE 7
Summary of Elasticity Estimates for Women

Authors of study	Year	Marshall	Hicks	Frisch	Uncompensated (dynamic)	Tax response
Static, life-cycle and life-cycle consistent models						
Cogan	1981	0.89[a]				
Heckman-MaCurdy	1982			2.35		
Blundell-Walker	1986	−0.20	0.01	0.03		
Blundell-Duncan-Meghir	1998	0.17	0.20			
Kimmel-Kniesner	1998			3.05[b]		
Moffitt	1984				1.25	
Dynamic structural models						
Eckstein-Wolpin	1989				5.0	
Van der Klauuw	1996				3.6	
Francesconi	2002				5.6	
Keane-Wolpin	2010				2.8	
Difference-in-difference methods						
Eissa	1995, 1996a					0.77–1.60[b]

Notes:
[a] = Elasticity conditional on positive work hours.
[b] = Sum of elasticities on extensive and intensive margins.

be because these studies consider the labor supply response of working women to wage changes, while the other nine studies incorporate participation choices.

In summary, we see the female labor supply literature has emphasized participation, human capital, fertility, and marriage. Papers that have attempted to model fertility and/or marriage as choices have ignored savings choices so as to achieve computational tractability. There is as of yet no model of female life-cycle behavior that includes savings along with human capital, fertility, and marriage. This is an important avenue for future research, but a difficult one, as it is not clear how to sensibly model savings outside of a household context.

My survey of female labor supply has been narrower than that for men, as I share with Mincer (1962) the view that it is difficult to think sensibly about labor supply of women without adopting a life-cycle perspective. Also, in the interest of space I have largely ignored the work on effects of welfare programs on labor supply of single mothers (see, e.g., Moffitt 1983, 1992; Keane and Moffitt 1998), *except* where welfare is integrated into a life-cycle framework (Keane and Wolpin 2010). Notably, single mothers may be the group for whom static models are most useful: they are likely liquidity constrained and have small returns to work experience. Hence, the literature on single mothers has emphasized methodological issues different from those discussed here, and to do it justice would require another major section (see Moffitt 1992 and Blundell and MaCurdy 1999 for surveys).

8. *Conclusion and Suggestions for Future Work*

My review suggests that labor supply of men may be more elastic than conventional wisdom suggests. When I simply average the Hicks elasticity across twenty-two well-known studies of males, I obtain 0.31. Several studies have shown that such a value is sufficient to induce substantial efficiency losses from progressive income taxation. Furthermore, if one weighs studies by features I argued are desirable, such as (i) use of direct rather than ratio wage measures, or (ii) accounting for human capital, one gets a larger value of the Hicks elasticity.

For women, most studies find very large labor supply elasticities. This is especially true of papers that calculate "long run" elasticities—meaning some combination of fertility, marriage, work experience, and education are allowed to respond to wage changes, rather than being held fixed. Across five such studies that I examine, long-run elasticities average 3.6.

In this survey, I sought to avoid bias toward any one methodological perspective. I discussed static and "life-cycle consistent" labor supply models, reduced form and structural dynamic models, and natural experiment methods. But one bias is notable: Out of hundreds of papers on labor supply, I chose to review ones that (i) pushed the methodological frontier in some way, (ii) are very well-known/influential, or (iii) both. Of course this is a small subset of the whole universe of papers on labor supply. An interesting (but difficult) question is whether the distribution of results in the universe of papers differs significantly from that in the subset of relatively prominent papers selected here.

Econometricians face many difficult problems in estimating labor supply elasticities (see section 4). Clearly, no paper deals with *all* of them. In my view, the most obvious gap in the literature is a dynamic model that includes the participation margin, human capital and progressive taxation (PH&P) simultaneously. Recent work claims that ignoring one or more of these features caused prior work to understate labor supply elasticities (see Rogerson and Wallenius 2009, Imai and Keane 2004, Aaronson and French 2009). It would be of great interest to incorporate all three features in one model, as ignoring any one may exaggerate the importance of the others.[161]

Another limitation of existing work is that few papers treat labor supply of couples as a joint decision (Blundell and Walker 1986 is an exception). Consideration of family labor supply has important implications for optimal tax calculations (see Patricia Apps and Ray Rees 2009). But this is a difficult area as results depend on one's model of family decision making.

Another gap is that few papers model dynamics and also consider the tax treatment of asset income (Ziliak and Kniesner 1999, 2005 are exceptions). Even among those that do, poor quality of data on nonlabor income is a serious concern (see Blomquist 1996).

Finally, all the papers I discussed ignore equilibrium effects of taxes on wages. There are a few equilibrium models in the literature (see, Heckman, Lance Lochner, and Christopher Taber 1998, Donghoon Lee 2005, Lee and Wolpin 2006, Keane and John E. Roemer 2009) but none account for PH&P simultaneously. These papers typically adopt simplistic models of either the

[161] Incorporating taxes in dynamic models is difficult, as the econometrician must model how agents forecast future taxes. Keane and Wolpin (2007, 2010) developed methods that can handle this problem, but they only applied them to forecasting welfare rules.

demand or supply side (or both) to make estimation feasible. But faster computers will make equilibrium modeling easier.

References

Aaronson, Daniel, and Eric French. 2009. "The Effects of Progressive Taxation on Labor Supply When Hours and Wages Are Jointly Determined." *Journal of Human Resources*, 44(2): 386–408.

Aguirregabiria, Victor, and Pedro Mira. 2010. "Dynamic Discrete Choice Structural Models: A Survey." *Journal of Econometrics*, 156(1): 38–67.

Altonji, Joseph G. 1986. "Intertemporal Substitution in Labor Supply: Evidence from Micro Data." *Journal of Political Economy*, 94(3): S176–215.

Altug, Sumru, and Robert A. Miller. 1990. "Household Choices in Equilibrium." *Econometrica*, 58(3): 543–70.

Altug, Sumru, and Robert A. Miller. 1998. "The Effect of Work Experience on Female Wages and Labour Supply." *Review of Economic Studies*, 65(1): 45–85.

Angrist, Joshua D. 1991. "Grouped-Data Estimation and Testing in Simple Labor-Supply Models." *Journal of Econometrics*, 47(2–3): 243–66.

Apps, Patricia, and Ray Rees. 2009. *Public Economics and the Household*. Cambridge and New York: Cambridge University Press.

Arcidiacono, Peter. 2005. "Affirmative Action in Higher Education: How Do Admission and Financial Aid Rules Affect Future Earnings?" *Econometrica*, 73(5): 1477–1524.

Ashenfelter, Orley, and James J. Heckman. 1973. "Estimating Labor Supply Functions." In *Income Maintenance and Labor Supply*, ed. G. Cain and H. Watts, 265–78. Chicago: Markham.

Atkinson, Anthony B., Thomas Piketty, and Emmanuel Saez. 2011. "Top Incomes in the Long Run of History." *Journal of Economic Literature*, 49(1): 3–71.

Bajari, Patrick, and Ali Hortacsu. 2005. "Are Structural Estimates of Auction Models Reasonable? Evidence from Experimental Data." *Journal of Political Economy*, 113(4): 703–41.

Ben-Porath, Yoram. 1967. "The Production of Human Capital and the Life Cycle of Earnings." *Journal of Political Economy*, 75(4 Part 1): 352–65.

Bernal, Raquel, and Michael P. Keane. 2011. "Child Care Choices and Children's Cognitive Achievement: The Case of Single Mothers." *Journal of Labor Economics*, 29(3): 459–512.

Blomquist, Soren. 1983. "The Effect of Income Taxation on the Labor Supply of Married Men in Sweden." *Journal of Public Economics*, 22(2): 169–97.

Blomquist, Soren. 1985. "Labour Supply in a Two-Period Model: The Effect of a Nonlinear Progressive Income Tax." *Review of Economic Studies*, 52(3): 515–24.

Blomquist, Soren. 1996. "Estimation Methods for Male Labor Supply Functions: How to Take Account of Nonlinear Taxes." *Journal of Econometrics*, 70(2): 383–405.

Blomquist, Soren, Matias Eklöf, and Whitney Newey. 2001. "Tax Reform Evaluation Using Non-parametric Methods: Sweden 1980–1991." *Journal of Public Economics*, 79(3): 543–68.

Blomquist, Soren, and U. Hansson-Brusewitz. 1990. "The Effect of Taxes on Male and Female Labor Supply in Sweden." *Journal of Human Resources*, 25(3): 317–57.

Blundell, Richard, Alan Duncan, and Costas Meghir. 1998. "Estimating Labor Supply Responses Using Tax Reforms." *Econometrica*, 66(4): 827–61.

Blundell, Richard, and Thomas MaCurdy. 1999. "Labor Supply: A Review of Alternative Approaches." In *Handbook of Labor Economics, Volume 3A*, ed. Orley Ashenfelter and David Card, 1559–1695. Amsterdam; New York and Oxford: Elsevier Science, North-Holland.

Blundell, Richard, Luigi Pistaferri, and Ian Preston. 2002. "Partial Insurance, Information and Consumption Dynamics." Institute for Fiscal Studies Working Paper 02/16.

Blundell, Richard, and Ian Walker. 1986. "A Life-Cycle Consistent Empirical Model of Family Labour Supply Using Cross-Section Data." *Review of Economic Studies*, 53(4): 539–58.

Boskin, Michael J. 1973. "The Economics of Labor Supply Functions." In *Income Maintenance and Labor Supply*, ed. G. Cain and H. Watts, 163–81. Chicago: Markham.

Bover, Olympia. 1989. "Estimating Intertemporal Labour Supply Elasticities Using Structural Models." *Economic Journal*, 99(398): 1026–39.

Brewer, Michael, Emmanuel Saez, and Andrew Shepard. 2010. "Means-Testing and Tax Rates on Earnings." In *Dimensions of Tax Design: The Mirrlees Review*, ed. Stuart Adam et al., 90–173. Oxford and New York: Oxford University Press.

Browning, Martin, Angus Deaton, and Margaret Irish. 1985. "A Profitable Approach to Labor Supply and Commodity Demands over the Life-Cycle." *Econometrica*, 53(3): 503–43.

Burtless, Gary, and Jerry A. Hausman. 1978. "The Effect of Taxation on Labor Supply: Evaluating the Gary Negative Income Tax Experiments." *Journal of Political Economy*, 86(6): 1103–30.

Chamberlain, Gary. 1984. "Panel Data." In *Handbook of Econometrics, Volume 2*, ed. Zvi Griliches and Michael D. Intriligator, 1247–1318. Amsterdam: Elsevier.

Chetty, Raj, John N. Friedman, Tore Olsen, and Luigi Pistaferri. 2009. "Adjustment Costs, Firm Responses, and Labor Supply Elasticities: Evidence from Danish Tax Records." National Bureau of Economic Research Working Paper 15617.

Cogan, John F. 1981. "Fixed Costs and Labor Supply." *Econometrica*, 49(4): 945–63.

Colombino, Ugo, and Daniela Del Boca. 1990. "The Effect of Taxes on Labor Supply in Italy." *Journal of Human Resources*, 25(3): 390–414.

Deaton, Angus. 1982. "Model Selection Procedures, or Does the Consumption Function Exist?" In *Evaluating the Reliability of Macroeconomic Models*, ed. Gregory C. Chow and Paoli Corsi, 43–69. New York: Wiley.

Diamond, Peter A. 1998. "Optimal Income Taxation: An Example with a U-Shaped Pattern of Optimal Marginal Tax Rates." *American Economic Review*, 88(1): 83–95.

Eckstein, Zvi, and Kenneth I. Wolpin. 1989. "Dynamic Labour Force Participation of Married Women and Endogenous Work Experience." *Review of Economic Studies*, 56(3): 375–90.

Eissa, Nada. 1995. "Taxation and Labor Supply of Married Women: The Tax Reform Act of 1986 as a Natural Experiment." National Bureau of Economic Research Working Paper 5023.

Eissa, Nada. 1996a. "Labor Supply and the Economic Recovery Tax Act of 1981." In *Empirical Foundations of Household Taxation*, ed. Martin Feldstein and James M. Poterba, 5–32. Chicago and London: University of Chicago Press.

Eissa, Nada. 1996b. "Tax Reforms and Labor Supply." In *Tax Policy and the Economy, Volume 10*, ed. James M. Poterba, 119–51. Cambridge and London: MIT Press.

Eklöf, Matias, and Hans Sacklén. 2000. "The Hausman–MaCurdy Controversy: Why Do the Results Differ across Studies? Comment." *Journal of Human Resources*, 35(1): 204–20.

Erdem, Tulin, Michael P. Keane, T. Sabri Oncu, and Judi Strebel. 2005. "Learning About Computers: An Analysis of Information Search and Technology Choice." *Quantitative Marketing and Economics*, 3(3): 207–46.

Fang, Hanming, and Michael P. Keane. 2004. "Assessing the Impact of Welfare Reform on Single Mothers." *Brookings Papers on Economic Activity*, 1: 1–95.

Francesconi, Marco. 2002. "A Joint Dynamic Model of Fertility and Work of Married Women." *Journal of Labor Economics*, 20(2): 336–80.

French, Eric. 2005. "The Effects of Health, Wealth, and Wages on Labour Supply and Retirement Behaviour." *Review of Economic Studies*, 72(2): 395–427.

Geweke, John, and Michael P. Keane. 2000. "An Empirical Analysis of Earnings Dynamics among Men in the PSID: 1968–1989." *Journal of Econometrics*, 96(2): 293–356.

Geweke, John, and Michael P. Keane. 2001. "Computationally Intensive Methods for Integration in Econometrics." In *Handbook of Econometrics, Volume 5*, ed. James J. Heckman and Edward Leamer, 3463–3568. Amsterdam; London and New York: Elsevier Science, North-Holland,.

Goeree, Jacob K., Charles A. Holt, and Thomas R. Palfrey. 2003. "Risk Averse Behavior in Generalized Matching Pennies Games." *Games and Economic Behavior*, 45(1): 97–113.

Haley, William J. 1973. "Human Capital: The Choice between Investment and Income." *American Economic Review*, 63(5): 929–44.

Hall, Robert E. 1973. "Wages, Income and Hours of Work in the U.S. Labor Force." In *Income Maintenance and Labor Supply*, ed. G. Cain and H. Watts, 102–62. Chicago: Markham.

Hausman, Jerry A. 1980. "The Effect of Wages, Taxes, and Fixed Costs on Women's Labor Force Participation." *Journal of Public Economics*, 14(2): 161–94.

Hausman, Jerry A. 1981. "Labor Supply." In *How Taxes Affect Economic Behavior*, ed. Henry J. Aaron and Joseph A. Pechman, 27–71. Washington, D.C.: Brookings Institution Press.

Hausman, Jerry A. 1985a. "The Econometrics of Nonlinear Budget Sets." *Econometrica*, 53(6): 1255–82.

Hausman, Jerry A. 1985b. "Taxes and Labor Supply." In *Handbook of Public Economics, Volume 1*, ed. Alan J. Auerbach and Martin Feldstein, 213–63. New York and Oxford: North-Holland.

Heckman, James J. 1974. "Shadow Prices, Market Wages, and Labor Supply." *Econometrica*, 42(4): 679–94.

Heckman, James J. 1976. "Estimates of a Human Capital Production Function Embedded in a Life-Cycle Model of Labor Supply." In *Household Production and Consumption*, ed. Nestor E. Terleckyj, 225–64. Cambridge, Mass.: National Bureau of Economic Research.

Heckman, James J., Lance Lochner, and Christopher Taber. 1998. "Explaining Rising Wage Inequality: Explorations with a Dynamic General Equilibrium Model of Labor Earnings with Heterogeneous Agents." *Review of Economic Dynamics*, 1(1): 1–58.

Heckman, James J., and Thomas MaCurdy. 1980. "A Life Cycle Model of Female Labour Supply." *Review of Economic Studies*, 47(1): 47–74.

Heckman, James J., and Thomas MaCurdy. 1982. "Corrigendum on a Life Cycle Model of Female Labour Supply." *Review of Economic Studies*, 49(4): 659–60.

Heckman, James J., and Thomas MaCurdy. 1986. "Labor Econometrics." In *Handbook of Econometrics, Volume 3*, ed. Zvi Griliches and Michael D. Intriligator, 1917–77. Amsterdam; Oxford and Tokyo: North-Holland.

Heckman, James J., and Robert J. Willis. 1977. "A Beta-logistic Model for the Analysis of Sequential Labor Force Participation by Married Women." *Journal of Political Economy*, 85(1): 27–58.

Hotz, V. Joseph, and Robert A. Miller. 1993. "Conditional Choice Probabilities and the Estimation of Dynamic Models." *Review of Economic Studies*, 60(3): 497–529.

Imai, Susumu, and Michael P. Keane. 2004. "Intertemporal Labor Supply and Human Capital Accumulation." *International Economic Review*, 45(2): 601–41.

Keane, Michael P. 1993a. "Individual Heterogeneity and Interindustry Wage Differentials." *Journal of Human Resources*, 28(1): 134–61.

Keane, Michael P. 1993b. "Nominal-Contracting Theories of Unemployment: Evidence from Panel Data." *American Economic Review*, 83(4): 932–52.

Keane, Michael P. 1995. "A New Idea for Welfare Reform." *Federal Reserve Bank of Minneapolis*

Quarterly Review, 19(2): 2–28.

Keane, Michael P. 2009. "Simulated Maximum Likelihood Estimation Based on First-Order Conditions." *International Economic Review*, 50(2): 627–75.

Keane, Michael P. 2010. "Structural vs. Atheoretic Approaches to Econometrics." *Journal of Econometrics*, 156(1): 3–20.

Keane, Michael P. 2011. "Income Taxation in a Life-Cycle Model with Human Capital." CEPAR Working Paper 2011–17.

Keane, Michael P., and Robert Moffitt. 1998. "A Structural Model of Multiple Welfare Program Participation and Labor Supply." *International Economic Review*, 39(3): 553–89.

Keane, Michael P., Robert Moffitt, and David E. Runkle. 1988. "Real Wages over the Business Cycle: Estimating the Impact of Heterogeneity with Micro Data." *Journal of Political Economy*, 96(6): 1232–66.

Keane, Michael P., and John E. Roemer. 2009. "Assessing Policies to Equalize Opportunity Using an Equilibrium Model of Educational and Occupational Choices." *Journal of Public Economics*, 93(7–8): 879–98.

Keane, Michael P., and David E. Runkle. 1992. "On the Estimation of Panel-Data Models with Serial Correlation When Instruments Are Not Strictly Exogenous." *Journal of Business and Economic Statistics*, 10(1): 1–9.

Keane, Michael P., and Kenneth I. Wolpin. 1994. "The Solution and Estimation of Discrete Choice Dynamic Programming Models by Simulation and Interpolation: Monte Carlo Evidence." *Review of Economics and Statistics*, 76(4): 648-72.

Keane, Michael P., and Kenneth I. Wolpin. 1997. "The Career Decisions of Young Men." *Journal of Political Economy*, 105(3): 473–522.

Keane, Michael P., and Kenneth I. Wolpin. 2001. "The Effect of Parental Transfers and Borrowing Constraints on Educational Attainment." *International Economic Review*, 42(4): 1051–1103.

Keane, Michael P., and Kenneth I. Wolpin. 2002. "Estimating Welfare Effects Consistent with Forward-Looking Behavior, Part II: Empirical Results." *Journal of Human Resources*, 37(3): 600–622.

Keane, Michael P., and Kenneth I. Wolpin. 2007. "Exploring the Usefulness of a Nonrandom Holdout Sample for Model Validation: Welfare Effects on Female Behavior." *International Economic Review*, 48(4): 1351–78.

Keane, Michael P., and Kenneth I. Wolpin. 2010. "The Role of Labor and Marriage Markets, Preference Heterogeneity, and the Welfare System in the Life Cycle Decisions of Black, Hispanic, and White Women." *International Economic Review*, 51(3): 851–92.

Killingsworth, Mark R. 1983. *Labor Supply*. Cambridge and New York: Cambridge University Press.

Killingsworth, Mark R., and James J. Heckman. 1986. "Female Labor Supply: A Survey." In *Handbook of Labor Economics, Volume 1*, ed. Orley Ashenfelter and Richard Layard, 103–201. Amsterdam; Oxford and Tokyo: North-Holland.

Kimmel, Jean, and Thomas J. Kniesner. 1998. "New Evidence on Labor Supply: Employment versus Hours Elasticities by Sex and Marital Status." *Journal of Monetary Economics*, 42(2): 289–301.

Kosters, Marvin. 1969. "Effects of an Income Tax on Labor Supply." In *The Taxation of Income from Capital*, ed. A. Harberger and M. Baily, 304–24. Washington, D.C.: Brookings Institution Press.

Lee, Donghoon. 2005. "An Estimable Dynamic General Equilibrium Model of Work, Schooling, and Occupational Choice." *International Economic Review*, 46(1): 1–34.

Lee, Donghoon, and Kenneth I. Wolpin. 2006. "Intersectoral Labor Mobility and the Growth of the Service Sector." *Econometrica*, 74(1): 1–46.

MaCurdy, Thomas. 1981. "An Empirical Model of Labor Supply in a Life-Cycle Setting." *Journal of Political Economy*, 89(6): 1059–85.

MaCurdy, Thomas. 1983. "A Simple Scheme for Estimating an Intertemporal Model of Labor Supply and Consumption in the Presence of Taxes and Uncertainty." *International Economic Review*, 24(2): 265–89.

MaCurdy, Thomas. 1992. "Work Disincentive Effects of Taxes: A Reexamination of Some Evidence." *American Economic Review*, 82(2): 243–49.

MaCurdy, Thomas, David Green, and Harry J. Paarsch. 1990. "Assessing Empirical Approaches for Analyzing Taxes and Labor Supply." *Journal of Human Resources*, 25(3): 415–90.

McCallum, Bennett T. 1976. "Rational Expectations and the Estimation of Econometric Models: An Alternative Procedure." *International Economic Review*, 17(2): 484–90.

Meghir, Costas, and David Phillips. 2010. "Labor Supply and Taxes." In *Dimensions of Tax Design: The Mirrlees Review*, ed. Stuart Adam et al., 202–74. Oxford and New York: Oxford University Press.

Mincer, Jacob. 1962. "Labor Force Participation of Married Women: A Study of Labor Supply." In *Aspects of Labor Economics*, 63–106. Princeton: Princeton University Press.

Mirrlees, James A. 1971. "An Exploration in the Theory of Optimum Income Taxation." *Review of Economic Studies*, 38(114): 175–208.

Moffitt, Robert. 1983. "An Economic Model of Welfare Stigma." *American Economic Review*, 73(5): 1023–35.

Moffitt, Robert. 1984. "Profiles of Fertility, Labour Supply and Wages of Married Women: A Complete Life-Cycle Model." *Review of Economic Studies*, 51(2): 263–78.

Moffitt, Robert. 1992. "Incentive Effects of the U.S. Welfare System: A Review." *Journal of Economic Literature*, 30(1): 1–61.

Mroz, Thomas A. 1987. "The Sensitivity of an Empirical Model of Married Women's Hours of Work to Economic and Statistical Assumptions." *Econometrica*, 55(4): 765–99.

Pencavel, John H. 1986. "Labor Supply of Men: A

Survey." In *Handbook of Labor Economics, Volume 1*, ed. Orley Ashenfelter and Richard Layard, 3–102. Amsterdam; Oxford and Tokyo: North-Holland.

Pistaferri, Luigi. 2003. "Anticipated and Unanticipated Wage Changes, Wage Risk, and Intertemporal Labor Supply." *Journal of Labor Economics*, 21(3): 729–54.

Rogerson, Richard, and Johanna Wallenius. 2009. "Micro and Macro Elasticities in a Life Cycle Model with Taxes." *Journal of Economic Theory*, 144(6): 2277–92.

Rust, John. 1987. "Optimal Replacement of GMC Bus Engines: An Empirical Model of Harold Zurcher." *Econometrica*, 55(5): 999–1033.

Saez, Emmanuel. 2001. "Using Elasticities to Derive Optimal Income Tax Rates." *Review of Economic Studies*, 68(1): 205–29.

Saez, Emmanuel. 2010. "Do Taxpayers Bunch at Kink Points?" *American Economic Journal: Economic Policy*, 2(3): 180–212.

Saez, Emmanuel, Joel Slemrod, and Seth H. Giertz. 2009. "The Elasticity of Taxable Income with Respect to Marginal Tax Rates: A Critical Review." National Bureau of Economic Research Working Paper 15012.

Sargent, Thomas J. 1978. "Estimation of Dynamic Labor Demand Schedules under Rational Expectations." *Journal of Political Economy*, 86(6): 1009–44.

Shaw, Kathryn L. 1989. "Life-Cycle Labor Supply with Human Capital Accumulation." *International Economic Review*, 30(2): 431–56.

Skinner, Jonathan. 1987. "A Superior Measure of Consumption From the Panel Study of Income Dynamics." *Economics Letters*, 23(2): 213–16.

Triest, Robert K. 1990. "The Effect of Income Taxation on Labor Supply in the United States." *Journal of Human Resources*, 25(3): 491–516.

van der Klaauw, Wilbert. 1996. "Female Labour Supply and Marital Status Decisions: A Life-Cycle Model." *Review of Economic Studies*, 63(2): 199–235.

van Soest, Arthur, Isolde Woittiez, and Arie Kapteyn. 1990. "Labor Supply, Income Taxes, and Hours Restrictions in the Netherlands." *Journal of Human Resources*, 25(3): 517–58.

Wales, T. J., and A. D. Woodland. 1979. "Labour Supply and Progressive Taxes." *Review of Economic Studies*, 46(1): 83–95.

Weiss, Yoram, and Reuben Gronau. 1981. "Expected Interruptions in Labour Force Participation and Sex-Related Differences in Earnings Growth." *Review of Economic Studies*, 48(4): 607–19.

Wolpin, Kenneth I. 1987. "Estimating a Structural Search Model: The Transition from School to Work." *Econometrica*, 55(4): 801–17.

Ziliak, James P., and Thomas J. Kniesner. 1999. "Estimating Life Cycle Labor Supply Tax Effects." *Journal of Political Economy*, 107(2): 326–59.

Ziliak, James P., and Thomas J. Kniesner. 2005. "The Effect of Income Taxation on Consumption and Labor Supply." *Journal of Labor Economics*, 23(4): 769–96.

This article has been cited by:

1. Michael Keane, , Richard Rogerson. 2012. Micro and Macro Labor Supply Elasticities: A Reassessment of Conventional Wisdom. *Journal of Economic Literature* **50**:2, 464-476. [Abstract] [View PDF article] [PDF with links]