Labor Supply: Are the Income and Substitution Effects Both Large or Both Small?[1]

Miles S. Kimball and Matthew D. Shapiro
University of Michigan and NBER

revised May 16, 2003

[1]This research was funded by National Institute on Aging grant P01-AG10179. We are grateful to F. Thomas Juster for help in designing the survey instrument, to Robert Barsky for collaboration in the early stage of this project, and to John Bound and Gary Solon for important suggestions. Ana Quinones and Brahima Coulibaly provided outstanding research assistance.

Abstract

Labor Supply: Are the Income and Substitution Effects Both Large or Both Small?

Labor supply is unresponsive to permanent changes in the wage rate. Hence, income and substitution effects cancel, but are they both close to zero or are they both large? The paper develops a theory of labor supply that imposes the restriction that income and substitution effects cancel. The theory takes into account optimization over time, fixed costs of going to work, and interaction of labor supply decisions within the household. The paper then applies this theory to experimental survey evidence on the response of labor supply to a large wealth shock. The evidence implies that the constant marginal utility of wealth (Frisch) elasticity of labor supply is about one.

Miles S. Kimball
Department of Economics
and Survey Research Center
University of Michigan
Ann Arbor MI 48109-1220
and NBER
tel. 734 764-2375
mkimball@umich.edu

Matthew D. Shapiro
Department of Economics
and Survey Research Center
University of Michigan
Ann Arbor MI 48109-1220
and NBER
tel. 734 764-5419
shapiro@umich.edu

1 Introduction

One of the best-documented regularities in economics is that—when they affect all members of a household proportionately—large, permanent differences in the real wage induce at most modest differences in the quantity of labor supplied by a household. This is true across households, across countries, and across time. The standard explanation is that the substitution and income effects of a permanently higher real wage are of approximately the same size; that is, the motivation to give up leisure to take advantage of a higher real wage is roughly cancelled out by the extra freedom to pursue leisure afforded by the higher income that the higher real wage provides. This explanation has broad support among economists because it has the merit of accounting for a wide range of data with one restriction on the utility function.

Among those economists who agree with the view that the income and substitution effects of a permanent increase in the real wage are approximately equal, there is much less agreement about whether the income and substitution effects are both large or both small. The size of the substitution effect is closely related to the elasticity of labor supply with respect to fluctuations in the real wage that are too short-lived to have substantial income effects. The size of the substitution effect is also a key factor in the magnitude of distortions induced by labor-income taxation and by other government policies that affect the margin between consumption and leisure or consumption and work. Hence, having a good estimate of the elasticity of labor supply has very broad and significant implications for understanding economic fluctuations and for assessing the effects of changes in public policy.

Drawing with a broad brush, one can paint the picture that macroeconomists, trying to explain substantial cyclical movements in labor hours in the face of modest cyclical movements in the real wage, see evidence that the substitution effect is large. Labor economists, looking at regressions of labor hours on fluctuations in the real wage or regressions of labor hours on the variation in the real wage over the life cycle, see evidence that the substitution effect is small. Direct evidence on the size of the substitution effect is muddied by several difficulties with the evidence.

- It is hard to find temporary, exogenous movements in the real wage that could identify movements in labor supply;
- Fluctuations in the shadow wage within a long-term relationship between firms and workers can look quite different from fluctuations in the observed wage in standard

data series; and

- o Many workers may face constraints on their labor hours imposed by their employers, so that they are not able to respond freely to variations in the real wage.

These issues affect the evidence contemplated both in the macroeconomics and labor economics literatures.

In this paper we propose an alternative to directly inferring the substitution effect from the relationship between wages and labor supply. The equality of income and substitution effects implies that one can infer the size of the substitution effect from the size of the income effect. Thus, we estimate the income effect and use that estimate, together with restrictions from a theory of labor supply, to infer the substitution effect.

We estimate the size of the income effect using a module designed by us in the Health and Retirement Study (HRS) which asks respondents to imagine what they would do if they won a sweepstakes that would pay them an amount equal to last year's family income every year as long as they live. We analyze this data using a structural model of household labor supply that imposes the restriction that income and substitution effects cancel. The model is based on the dynamic optimization problem of the household. It has several other features needed to capture important features of behavior. First, it allows for nonseparabilities between consumption and labor. These can account for a drop in observed consumption at retirement because working increases the marginal utility of consumption. Second, it integrates the decisions of married partners about consumption and labor supply. Finally, to match the observed fact that few people work less than 20 hours per week, our structural model allows for fixed utility costs of going to work. This final feature is very important for the analysis of the labor supply response to the sweepstakes because many households report that they would quit work entirely rather than smoothly reducing hours.

Section 2 of the paper presents this theory. Section 3 discusses how to implement it to infer elasticities of labor supply from the response to wealth shocks. Section 4 applies this framework to inferring labor supply elasticities from the survey responses to winning the sweepstakes. Section 5 shows how to relate these estimates to alternative measures of the labor supply elasticity. Section 6 discusses estimates of labor supply elasticities and other relevant parameters from the literature and discusses the pros and cons of the survey and econometric approaches. Section 7 discusses the implications of our estimates.

2 Theoretical Framework for Labor Supply

This section develops a theory of labor supply in a life-cycle, permanent income setting. It has the following main features that distinguish it from other theoretical frameworks.

1. The long-run elasticity of labor supply is zero.

2. Utility is nonseparable in consumption and labor.

3. Working incurs a fixed as well as a variable cost in utility terms.

4. For married couples, labor supply decisions are integrated.

A zero long-run elasticity of labor supply is a more precise statement of what we referred to in the introduction informally as "income and substitution effects canceling." Nonseparability between consumption and labor and nonconcavity of the utility function due to fixed costs of going to work require rederiving basic labor supply relationships in that more general context. To begin with, we need to derive the functional form implications of cancellation between income and substitution effects in this more general context. Then we need to derive the structural model of labor supply when consumption and labor are not additively separable and there are fixed costs of work.

These features of the theory complement each other in significant ways. Basu and Kimball (2002) give an extended argument that allowing for nonseparability between consumption and labor is especially important in a context where one wishes to maintain cancellation between income and substitution effects and gives references to a number of papers that have found direct evidence for nonseparability between consumption and labor. Moreover, the nonseparability has an important interaction with the nonconvexity in labor supply. The nonconvexity causes workers to move abruptly from work to non-work. The nonseparability between consumption and labor allows for jumps in consumption across these labor market transitions even when the household is maximizing utility intertemporally, anticipates the labor-market transition, and faces no borrowing or lending constraints.

2.1 Long-run labor supply elasticity close to zero

2.1.1 Motivation

What is the evidence that convinces a wide range of economists that the long-run labor supply elasticity is approximately zero?

First, looking across the cross-section of countries, there is remarkably little variation in hours of work across a very broad range of income levels.[1] Except for very poor countries, hours of work cluster around forty hours per week. Though there is a slight negative relationship between the level of per capita income and hours per week, that relationship is very weak: a tripling of per capita income perhaps yields a reduction of an hour or two in average weekly hours.[2]

Second, there is very little time trend in the hours worked within a country. Pencavel quotes Klein and Kosobud (1961) that the constancy of family labor supply since World War II is one of the "great ratios of economics."[3] There is only a modest decline in hours compared to the very dramatic increases in income and consumption.

Third, the cross-sectional evidence on labor supply suggests that it is very hard to explain differences in amount of work across individuals by differences in their wage rates. That is, low income individuals tend to work about the same number of hours as high income individuals. Even when this extensive literature finds cross-sectional evidence of a non-zero long-run labor supply elasticity, the point estimate is typically very small. This evidence is discussed further in Section 6.2.

2.1.2 General implications for functional form

It is useful to embody the evidence for approximate cancellation of income and substitution effects or an approximately zero long-run labor supply elasticity in a restriction on utility that, for simplicity, imposes a long-run labor supply elasticity that is exactly zero. If households spent exactly their income in each period, what we would mean by equality of income and substitution effects would be that multiplying both the real wage and the amount of non-labor income—and thereby consumption—by the same positive constant would leave the marginal rate of substitution between consumption and labor consistent with that higher real wage. To allow for a nontrivial consumption-saving decision, and possible non-concavities in the utility function, it is convenient to define a zero long-run elasticity of labor supply directly in terms of tradeoffs between consumption and labor.

[1] See Abel and Bernanke (2001, p. 84), who take this as one of the stylized facts of economic growth.

[2] Of course, there is heterogeneity in hours. Hours of work in Europe are lower than in the United States. These differences are not, however, a function of per capita income, and instead likely arise from differences in tax systems or differences in taste. Note that taxes and transfers can easily change the after-tax wage to consumption ratio and so generate differences in labor hours even if the long-run labor supply elasticity is zero.

[3] See Pencavel, 1986, p. 10. There are trends in labor supply within family, i.e., that women are working more and that men are working less. Per capita income has increased by a factor of over three while average weekly hours have fallen from 40.3 to 34.2 hours per week.

Labor Supply Function: We derive a labor supply function $\mathbf{N^s(W/C)}$ that depends only on the *ratio* of the real wage to consumption: W/C, or in the case of a couple, the ratios of both partners' real wages to household consumption, W_1/C and W_2/C. Many of the differences in weekly hours obvious to casual empiricism can indeed be associated with an obvious difference in the ratio W/C. For example, the long hours of doctors no doubt owe something in many cases to high levels of debt coming out of medical training that depress consumption C in relation to the real wage W. The high hours young lawyers in big law firms are willing to put up with may owe something to law-school debts but probably even more to the very high implicit wage that arises from the effect of additional effort on the probability of making partner in the law firm. The average male retirement age has been declining over the last few decades. This is what one would expect if increased female earnings have added more to household consumption, thereby causing male W/C to fall.

In terms of utility theory, let felicity (the instantaneous utility function) be given by $U(C, \mathcal{N})$, where C is total consumption expenditure by the household and $\mathcal{N}$ is a vector of the labor hours of various members of the household. To impose a zero long-run labor supply elasticity, formally, we assume preferences scale symmetric in consumption.

Preferences Scale Symmetric in Consumption: $U(C, \mathcal{N})$ exhibits scale symmetry in consumption if

$$U(C', \mathcal{N}') = U(C, \mathcal{N})$$

implies that for any positive α,

$$U(\alpha C', \mathcal{N}') = U(\alpha C, \mathcal{N}).$$

In other words, the utility function allows the difference between the labor vector $\mathcal{N}$ and the new labor vector $\mathcal{N}'$ to be compensated by a given proportional increase in consumption $\frac{C'}{C} = \frac{\alpha C'}{\alpha C}$, regardless of the initial level of consumption. Thus, the marginal rate of substitution between consumption and labor is proportional to the original level of consumption, where we can look at the marginal rate of substitution between points far apart as well as points that are only an infinitesimal distance apart.

Besides scale symmetry in consumption, in order to establish a convenient representation of the utility function, we need the completeness condition that for any $\mathcal{N}$ and $\mathcal{N}'$ in a feasible set Γ which includes the zero vector, there is a proportional change in consumption that can compensate for that difference in the labor vector:

$$\forall \mathcal{N}, \mathcal{N}' \in \Gamma, \ \exists C' \ \text{ s.t.}$$

$$U(C', \mathcal{N}') = U(1, \mathcal{N}).$$

Proposition 1: *If $U(C, \mathcal{N})$ is scale symmetric in consumption and satisfies the completeness condition above, then*

$$U(C, \mathcal{N}) = \Omega(Ce^{-f(\mathcal{N})}) = \Omega(e^{\ln(C) - f(\mathcal{N})})$$

for some pair of real-valued functions Ω and $f(\mathcal{N})$.

Proof: For each $\mathcal{N}$, let $f(\mathcal{N})$ be the solution to

$$U(e^{-f(\mathcal{N})}, 0) = U(1, \mathcal{N}).$$

That is, relative to not working, the labor vector $\mathcal{N}$ has the same effect on utility as multiplying consumption by e^{-f}. Such a solution exists by the completeness condition. Then by scale symmetry in consumption,

$$U(C, \mathcal{N}) = U(Ce^{-f(\mathcal{N})}, 0) = \Omega(Ce^{-f(\mathcal{N})})$$

where Ω is defined by

$$\Omega(x) = U(x, 0).$$

Remarks: In practice, it is convenient to model utility as homogeneous in consumption. This corresponds to

$$\Omega(x) = -\frac{(1-\alpha)}{\alpha} x^{-\alpha/(1-\alpha)},$$

where α indexes the degree of complementarity between consumption and labor. (We will discuss this nonseparability more below.) The overall utility function becomes

$$U(C, \mathcal{N}) = -\frac{(1-\alpha)}{\alpha} C^{-\alpha/(1-\alpha)} e^{[\alpha/(1-\alpha)]f(\mathcal{N})}$$

when $\alpha \neq 0$ and $\Omega(x) = \ln(x)$ or

$$U(C, \mathcal{N}) = \ln(C) - f(\mathcal{N})$$

when $\alpha = 0$. This form of the utility function—multiplicatively separable between consumption and the labor vector when $\alpha \neq 0$ and additively separable between consumption and the labor vector when $\alpha = 0$—is called the *King-Plosser-Rebelo* (1988) form of the utility function. It arose originally in the context of finding a utility function consistent with a steady growth path for an economy. The additional features here are (1) having a labor vector for a household instead of a single summary measure of labor for a representative consumer and (2) allowing for the possibility of nonconcavity of the utility function because of fixed utility costs of working. The theory above about how to impose a zero long-run labor supply elasticity is general enough to handle both of those features.

2.2 Structural model of labor supply

Consider a household maximizing

$$E_0 \int_0^\infty e^{-\rho t} U(C, \mathcal{N}, \nu) dt,$$

where C is total consumption expenditure by the household, $\mathcal{N}$ is a vector of the labor of the various potential workers in the household and ν is an exogenous vector of demographic variables which evolves stochastically over time.[4] In particular, the death of a household member or the departure of a household member at maturity generates a change in ν. The household can freely borrow and lend at the real interest rate r and faces fair annuity and life insurance markets which allow the household to smooth its marginal utility of consumption $U_C(C, \mathcal{N}, \nu) = \lambda$ over time and over all stochastic changes in ν. If, in addition, the real interest rate r is equal to the utility discount rate ρ, optimization requires the household to keep the marginal utility of consumption λ constant over time as well as over stochastic changes in ν. Note that λ is not only the marginal utility of consumption, it is also the costate variable giving the marginal value of real wealth in the household's dynamic control problem.

[4]Our model of the family is unitary. Blundell, et al. (2001) consider a non-unitary model that, like our model, also allows for discrete choice of whether or not to work. They find some support for their non-unitary model, but also cannot reject the unitary model.

We abstract from all other kinds of uncertainty besides uncertainty in the household demographic variables in ν—except for the very unlikely (and totally uninsured) hypothetical events we asked respondents to contemplate in our survey questions. These hypothetical events, in particular winning a sweepstakes, *would* change the value of λ. We will consistently put a hat ($\hat{\ }$) over variables to indicate the values they would take *after* one of the events described in our survey questions. Values without a hat are the original values a household would choose in the absence of such an event.

In our choice of functional form, the most important consideration is to build in the observed zero long-run labor supply elasticity. In line with the results of the previous section, extended to allow for dependence of felicity on demographics, this leads to

$$U(C, \mathcal{N}, \nu) = -\frac{(1-\alpha)}{\alpha} C^{-\alpha/(1-\alpha)} [\psi(\nu) + \alpha g(\mathcal{N})]^{1/(1-\alpha)} \quad , \tag{1}$$

where $\psi(\nu) = e^{\alpha f(0,\nu)}$ and $g(N) = \alpha^{-1}[e^{\alpha f(N,\nu)} - e^{\alpha f(0,\nu)}]$. Obtaining the limiting utility function as $\alpha \to 0$ requires adding the exogenous quantity $\frac{\psi(\nu)}{\alpha} + \psi(\nu) \ln(\psi(\nu))$ to the utility function in (1) and a messy application of L'Hopital's rule. This yields the intuitive result

$$U(C, \mathcal{N}) = \psi(\nu) \ln(C) - g(\mathcal{N}). \tag{2}$$

As discussed above, α measures the degree of complementarity between consumption and labor—*or equivalently, the degree of substitutability between consumption and leisure.* $\alpha = 0$ corresponds to additive separability, while $\alpha = 1$ corresponds to perfect substitutability between consumption and leisure. $\alpha \in (0, 1)$ corresponds to consumption being a partial substitute for leisure.[5]

Given $\alpha \in [0, 1)$, disliking additional labor is equivalent to $g(\mathcal{N})$ being an increasing function of $\mathcal{N}$. We assume that the household acts as a unit, maximizing the joint utility of its members. For any given total amount of household expenditure, the utility from consumption depends on how many people that expenditure is spread over. The function $\psi(\nu)$ can be thought of as a household equivalency scale. It determines the level of consumption needs when no one in the household is working. The flexibility of the function ψ allows one to deal with whatever degree of returns to scale there are in household consumption.

Aside from the assumptions embodied in the form above—a zero long-run elasticity of

[5]Basu and Kimball (2002) argue that the degree of complementarity between consumption and labor theoretically should be linked to the elasticity of intertemporal substitution for consumption, but because the real interest rate is assumed constant here, intertemporal substitution in consumption plays no real role in this analysis. Therefore, we focus on the degree of complementarity between consumption and labor that does matter for our analysis and calibrate α accordingly.

labor supply and constancy of the elasticity of intertemporal substitution—when there is more than one potential worker in the family, we consider the convenient benchmark of a $g(\mathcal{N})$ function that is additively separable in the labor of the various potential workers in the family:

$$g(\mathcal{N}) = \sum_i g_i(N_i).$$

As an obvious normalization,

$$g_i(0) = 0.$$

The best way to explain our solution method is to think first of what one would do if the utility function were additively separable between consumption and labor, i.e., $\alpha = 0$. The current-value Hamiltonian for the logarithmic case (2) would be

$$H = \psi(\nu)\ln(C) - g(\mathcal{N}) + \lambda[\mathcal{W} \cdot \mathcal{N} - C + rA + \Pi],$$

where $\mathcal{W}$ is a vector of wage rates for various household members, A is the value of financial assets earning real interest rate r, and Π is the non-labor, non-interest income of the household. Because of the additive separability between consumption and labor, to study labor supply in this logarithmic case, one could focus on the optimization subproblem

$$\max_{\mathcal{N}} \quad \lambda\mathcal{W} \cdot \mathcal{N} - g(\mathcal{N}),$$

or equivalently, for each potential worker i, after dividing through by λ one could focus on the optimization subproblem

$$\max_{N_i} \quad W_i N_i - \lambda^{-1} g_i(N_i) \tag{3}$$

for each N_i.

We find that when $\alpha \neq 0$, so that consumption and labor are not additively separable, by maximizing first over consumption conditional on labor quantities, we obtain an optimization subproblem for labor supply that is an exact counterpart of the optimization subproblem (3), but with $\lambda^{\alpha-1}$ in place of λ^{-1}.

When $\alpha \neq 0$, the household maximizes the current value Hamiltonian

$$H = \left(\frac{1-\alpha}{\alpha}\right) C^{-\alpha/(1-\alpha)}[\psi(\nu) + \alpha g(\mathcal{N})]^{1/(1-\alpha)} + \lambda[\mathcal{W} \cdot \mathcal{N} - C + rA + \Pi].$$

As emphasized in the introduction, a key element of the analysis is that $g_i(N_i)$ will typically embody a fixed utility cost of working, such as the utility cost of commuting time. Therefore, $[\psi(\nu) + g(\mathcal{N})]^{1/(1-\alpha)}$ will not usually be concave in $\mathcal{N}$. Because of the non-concavity in $\mathcal{N}$, the solution method that works best is to solve for the optimal value of consumption expenditure C and substitute this value back into the Hamiltonian before tackling the more difficult task of maximization over $\mathcal{N}$. We rely on Pontryagin's Maximum Principle that—even when non-concave—the Hamiltonian must be maximized over all of the control variables as a necessary condition for maximization.

The first-order condition for optimal consumption is

$$\frac{\partial H}{\partial C} = C^{-1/(1-\alpha)}[\psi(\nu) + \alpha g(\mathcal{N})]^{1/(1-\alpha)} = \lambda.$$

Solving for consumption,

$$C = \lambda^{\alpha-1}[\psi(\nu) + \alpha g(\mathcal{N})] = \lambda^{\alpha-1}[\psi(\nu) + \alpha \sum_i g_i(N_i)]. \tag{4}$$

Define *baseline consumption* B by

$$B = \lambda^{\alpha-1}\psi(\nu) \tag{5}$$

and *job-induced consumption* J_i for each potential worker i in the household by[6]

$$J_i = \alpha\lambda^{\alpha-1}g_i(N_i). \tag{6}$$

Then total consumption equals baseline consumption plus the job-induced consumption of each worker in the household:

$$C = B + \sum_i J_i. \tag{7}$$

[6]We will consistently maintain the assumption that $1 > \alpha \geq 0$. However, if α *were* negative—something that is a *theoretical* possibility—job-induced consumption would be negative. This highlights the fact that J_i is "job-induced consumption" in a very broad sense. J_i includes every way in which working makes one choose to consume more or less.

Because it represents every interaction between work and consumption, *job-induced consumption* must be construed quite broadly. It includes both (1) *work-related consumption* (such as childcare, transportation to and from work, the extra expense of food at work, and the extra expense of clothes suitable for work), and (2) *extra time-saving consumption* (such as easy-to-prepare foods at home, house-cleaning and house-repair services, and household conveniences). Baseline consumption and the job-induced consumption for each worker play a key role in the analysis at a later stage.

Substituting the underlying expression for optimized consumption into the Hamiltonian, the Hamiltonian maximized over C—which we can call $\bar{H}$—is

$$\begin{aligned}
\bar{H} &= -\frac{(1-\alpha)}{\alpha} C^{-\alpha/(1-\alpha)} [\psi(\nu) + \alpha g(\mathcal{N})]^{1/(1-\alpha)} - \lambda C + \lambda[\mathcal{W} \cdot \mathcal{N} + rA + \Pi] \\
&= \frac{\lambda^{\alpha}[\psi(\nu) + \alpha g(\mathcal{N})]}{-\alpha/(1-\alpha)} - \lambda^{\alpha}[\psi(\nu) + \alpha g(\mathcal{N})] + \lambda[\mathcal{W} \cdot \mathcal{N} + rA + \Pi] \\
&= \lambda \left\{ rA + \Pi - \frac{\lambda^{\alpha-1}\psi(\nu)}{\alpha} + \mathcal{W} \cdot \mathcal{N} - \lambda^{\alpha-1} g(\mathcal{N}) \right\}
\end{aligned}$$

By Pontryagin's maximum principle—which operates with exactness in continuous time—the household maximizes H even though H is non-concave. Maximizing H over both C and $\mathcal{N}$ is the same as maximizing $\bar{H}$ over $\mathcal{N}$. So the vector of labor hours $\mathcal{N}$ solves

$$\max_{\mathcal{N}} \quad \lambda \left\{ rA + \Pi - \frac{\lambda^{\alpha-1}\psi(\nu)}{\alpha} + \sum_i \left[W_i N_i - \lambda^{\alpha-1} g_i(N_i) \right] \right\}. \tag{8}$$

Maximizing the expression in (8) requires solving the optimization subproblems

$$\max_{N_i} \quad W_i N_i - \lambda^{\alpha-1} g_i(N_i) \tag{9}$$

for each N_i. Thus, it is as if there were a *money-metric disutility* of work $\lambda^{\alpha-1} g_i(N_i)$ for each potential worker i. Note that job-induced consumption J_i is always equal to α times the money-metric disutility of work $\lambda^{\alpha-1} g(N_i)$.

The structure of (9), the optimization subproblem for N_i, is the same regardless of the shape of $g_i(N_i)$. But to get any further in the analysis we need more structure on $g_i(N_i)$. While leaving the functional form $v_i(\cdot)$ below for the variable disutility of labor supply general as long as possible, we will now commit to (1) a fixed utility cost of working as the source of nonconcavity in the utility function, and (2) a multiplicative *work aversion* parameter as one of the dimensions in which individuals differ from one another. Let

$$g_i(N_i) = \begin{cases} 0 & \text{if } N_i = 0 \\ M_i[F_i + v_i(N_i)] & \text{if } N_i > 0 \end{cases}$$

or, defining

$$\chi(N_i) = \begin{cases} 0 & \text{if } N_i = 0 \\ 1 & \text{if } N_i > 0 \end{cases},$$

$$g_i(N_i) = \chi(N_i) M_i [F_i + v_i(N_i)]. \tag{10}$$

Here M_i is the *work aversion* parameter, F_i is a positive number that models the fixed utility cost of going to work and v_i is a function satisfying $v_i(0) = 0$, $v_i'(N) > 0$, $v_i''(N) > 0$ and $v_i'(168) = \infty$. (N_i is measured in weekly hours; 24 hours a day, seven days a week is 168 hours a week.)

Substituting in for $g_i(N_i)$, the maximization subproblem (9) becomes

$$\max_{N_i} \quad W_i N_i - \chi(N_i) \lambda^{\alpha-1} M_i [F_i + v_i(N_i)]. \tag{11}$$

If $N_i > 0$, the first order necessary condition for optimal N_i in (11) is

$$W_i = \lambda^{\alpha-1} M_i v_i'(N_i). \tag{12}$$

Call the solution to this first order condition N_i^*. Then

$$N_i^* = v_i'^{-1}\left(\frac{\lambda^{1-\alpha} W_i}{M_i}\right). \tag{13}$$

Because this is a necessary condition for an optimum when $N_i > 0$, the optimal N_i must be either N_i^* or 0. To determine whether the optimal N_i is N_i^* or 0, we need to compare the value of the criterion function

$$W_i N_i - \chi(N_i) \lambda^{\alpha-1} M_i [F_i + v_i(N_i)] \tag{14}$$

at $N_i = 0$ with its value at $N_i = N_i^*$ to see which one is greater.

When $N_i = 0$, the criterion function (14) is equal to zero. When $N_i = N_i^*$, the criterion function (14) is equal to

$$\lambda^{\alpha-1}M_i\left\{\frac{\lambda^{1-\alpha}W_iN_i^*}{M_i}-F_i-v_i(N_i^*)\right\}=\lambda^{\alpha-1}M_i\left\{N_i^*v_i'(N^*)-v_i(N_i^*)-F_i\right\}. \tag{15}$$

The right-hand side of (15) is greater than or equal to zero if and only if

$$N_i^*v_i'(N_i^*)-v_i(N_i^*)\geq F_i.$$

Define the cutoff value for labor, $N_i^{\#}$, by

$$N_i^{\#}v_i'(N_i^{\#})-v_i(N_i^{\#})=F_i. \tag{16}$$

Since $\frac{d}{dN_i}[N_iv_i'(N_i)-v_i(N_i)]=N_iv_i''(N_i)>0$ whenever $N_i>0$, $N_iv_i'(N_i)-v_i(N_i)$ is an increasing function of N. This, plus the assumption that $v_i'(168)=\infty$, guarantees that there is a unique solution to (16) between 0 and 168. In terms of the cutoff value $N_i^{\#}$ that solves equation (16), the rule for optimal labor supply can be written as

$$N_i=\begin{cases}N_i^* & \text{if } N_i^*>N_i^{\#}\\ 0 & \text{if } N_i^*<N_i^{\#}\\ \text{either } N_i^* \text{ or } 0 & \text{if } N_i^*=N_i^{\#}\end{cases}, \tag{17}$$

where N_i^* is given by

$$N_i^*=v_i'^{-1}\left(\frac{\lambda^{1-\alpha}W_i}{M_i}\right). \tag{18}$$

In practice, we use empirical evidence to determine the cutoff value $N_i^{\#}$ directly rather than to determine F_i. From this point of view, one can see equation (16) the other way around as a mapping from the cutoff value $N_i^{\#}$ to the fixed cost F_i. In terms of $N_i^{\#}$, the function $g_i(N_i)$ is given by

$$g_i(N_i)=\chi(N_i)M_i[v_i(N_i)-v_i(N_i^{\#})+N_i^{\#}v_i'(N_i^{\#})]. \tag{19}$$

In our empirical implementation in Section 4, we use $N_i^{\#}\equiv N^{\#}$ of 19 hours per week for all individuals.

We now have all the elements of the theory to infer elasticities from wealth shocks. In the next section, we use this theory to provide estimates of the Frisch labor supply elasticity.

3 Implementing the Theory

We use the theory developed in Section 2 to show how to develop a parametric estimate of the labor supply elasticity based on the survey responses.

The Frisch or λ-constant labor supply elasticity η_i^λ is given by differentiating (18) with respect to the wage, that is,

$$\eta_i^\lambda \equiv \frac{\partial N_i^*/\partial W_i}{N_i^*/W_i}. \tag{20}$$

At this point it is useful to specialize the analysis by choice of functional form for the function v. In particular, suppose that it has the constant elasticity form

$$v_i(N) = \frac{1}{1+\frac{1}{\eta_i}} N^{1+\frac{1}{\eta_i}}, \tag{21}$$

so

$$\eta_i^\lambda = \eta_i.$$

The constant elasticity of labor supply is a convenient formulation.[7] We have explored other functional forms, but since they do not alter the message of the paper we do not discuss them here.

The conceptual experiment posed by our survey is to consider how labor supply and consumption respond to a shock to wealth. Our theory provides a framework for studying these responses.

As noted above, let variables without hats denote their values before the wealth shock and variables with hats denote their values after the wealth shock. That is, C, J_i N_i, and λ are the values of total consumption, job induced consumption, labor, and marginal utility of wealth prior to the wealth shock and $\hat{C}$, $\hat{J}_i$, $\hat{N}_i$, and $\hat{\lambda}$ are their values after the wealth shock. Conditional on not quitting, the change in labor supply after the wealth shock is

$$\frac{\hat{N}_i}{N_i} = \left(\frac{\hat{\lambda}^{1-\alpha}}{\lambda^{1-\alpha}}\right)^{\eta_i}. \tag{22}$$

The wage W_i and the work aversion parameter M_i cancel out because they are unchanged by the wealth shock. Equation (22) shows the relationship between the change in labor and

[7] It is trivial and affects nothing to modify this form of $v_i(N)$ near $N = 168$ to satisfy the technical condition that $v_i'(168) = \infty$. In practice, no one has hours anywhere near 168 and the work aversion parameter gets calibrated to keep labor hours in the observed range.

the marginal utility of wealth. The more elastic is labor supply, the less marginal utility needs to move down (and so the less consumption needs to move up) to generate a given decline in labor.

Our theoretical framework allows us to relate marginal utility to observables. Substituting the expression (5) for baseline consumption B into (22) yields

$$\frac{\hat{N}_i}{N_i} = \left(\frac{\hat{B}}{B}\right)^{-\eta_i}. \tag{23}$$

Given data on labor, consumption, and job-induced consumption before and after the wealth shock, one can infer a value of η_i from the above equation. In the case of quits,

$$\frac{N_i^{\#}}{N_i} = \left(\frac{\hat{B}}{B}\right)^{-\eta_i^{\#}} \tag{24}$$

provides a lower-bound on the elasticity where $\hat{B}$ is baseline consumption evaluated at zero hours.

Hence, given observations on the change in labor and the change in baseline consumption from a wealth shock, one can infer the elasticity of labor supply consistent with these choices by inverting (23) as follows

$$\eta_i = -\frac{\ln(\hat{N}_i/N_i)}{\ln(\hat{B}/B)} \tag{25}$$

and for those who quit,

$$\eta_i^{\#} = -\frac{\ln(N_i^{\#}/N_i)}{\ln(\hat{B}/B)}. \tag{26}$$

In the additively separable case, baseline consumption equals actual consumption; but a key theme of the analysis in this paper is to take into account non-separabilities. Recall that baseline consumption is defined as

$$B = C - \sum_i J_i \tag{27}$$

where C is total consumption and J_i is job-induced consumption for each worker in the household.

To infer the labor supply elasticity from the joint movements in consumption and labor given above, we must calculate job-induced consumption. Recall that job-induced consumption is given, for each worker J_i in the household, by

$$J_i = \alpha\lambda^{\alpha-1} g_i(N_i). \tag{6}$$

Substituting in the expression for $g_i(N_i)$ from equation (19) and noting that $\chi(N_i) = 1$ when the individual is working yields

$$J_i = \alpha\lambda^{\alpha-1} M_i[v_i(N_i) - v_i(N_i^{\#}) + N_i^{\#} v_i'(N_i^{\#})].$$

Now substitute $\lambda^{\alpha-1}M_i = W_i/v'(N_i)$ from the first-order condition for labor (12). This yields

$$J_i = \begin{cases} 0 & \text{if } N_i = 0 \\ \frac{\alpha W_i}{v'(N_i)}[v_i(N_i) - v_i(N_i^{\#}) + N_i^{\#} v_i'(N_i^{\#})] & \text{if } N_i \geq N_i^{\#}. \end{cases} \tag{28}$$

The identical expression gives $\hat{J}_i$ as a function of $\hat{N}_i$, that is,

$$\hat{J}_i = \begin{cases} 0 & \text{if } \hat{N}_i = 0 \\ \frac{\alpha W_i}{v'(\hat{N}_i)}[v_i(\hat{N}_i) - v_i(N_i^{\#}) + N_i^{\#} v_i'(N_i^{\#})] & \text{if } N_i \geq N_i^{\#}. \end{cases} \tag{29}$$

Evaluated at the constant-elasticity functional form (21), the nonzero expression for J_i is

$$J_i = \frac{\alpha W_i N_i}{1 + \frac{1}{\eta_i}} \left[1 + \frac{1}{\eta_i} \left(\frac{N_i^{\#}}{N_i} \right)^{1+\frac{1}{\eta_i}} \right] \tag{30}$$

and similarly for $\hat{J}_i$. Consider the case when $N_i = N^{\#}$, that is, the individual just overcomes the fixed cost of going to work. In that case, job-induced consumption is just $\alpha W_i N_i$. As labor increases above the cutoff, the ratio of job-induced consumption to total labor income falls—the more so the lower is η_i. Note also that the limit of J_i is zero as η_i goes to zero.

Our data consist of consumption and labor supply, before and after the wealth shock, that is, $C, N_i, \hat{C}$, and $\hat{N}_i$. Given these data, equation (20), equation (27), and equations (28) and (29) subject to the functional form (21) for $v_i(N_i)$ yield a value for the labor supply elasticity η_i. Because of the dependence of job-induced consumption on the labor supply elasticity, we must solve for η_i numerically.

Labor supply is fairly straightforward to measure. The core HRS asks about labor hours N_i. Our module on the sweepstakes asks about labor hours $\hat{N}_i$ after the wealth shock. Measuring consumption is much less straightforward. In the HRS, consumption is not measured directly apart from some experimental measures. For both the baseline consumption C and the post-sweepstakes consumption $\hat{C}$ we use the budget constraint and assume unconstrained intertemporal utility maximization (baseline-consumption smoothing) to infer consumption. In a static, one-period problem, this amounts to saying that however many hours labor is reduced after the sweepstakes, consumption falls by the wage times the reduction in hours. The intertemporal version of this makes equal reductions in baseline consumption each year until the death of one spouse (with everything scaled down after the household drops to one member) and appropriate reductions in job-induced consumption each year until retirement. Appendix A details this computation.[8]

Note that the calculation of the elasticity depends on the *change* in consumption. Our procedure might get the level wrong, but it should be much more accurate for the change.

4 Survey Measures of the Labor Supply Response to a Wealth Shock

In this section we discuss a survey question that we have designed and implemented to gauge the response of labor supply to a wealth shock. The survey was designed to yield estimates of the labor supply parameter η that is central to the theoretical discussion in the previous sections of this paper. The survey was implemented as an experimental module on the Health and Retirement Study. In this section, we discuss the design of the survey and how the survey responses can be mapped into the theory that we outline. In the next section, we discuss the results of the survey and estimate the value of the labor supply parameter based on them.

4.1 Survey Instrument

The module begins with questions about whether or not the respondent and his or her spouse or partner work for pay. These responses are used to put the respondent into the

[8] In the survey, we attempted to elicit other uses for the wealth windfall other than increasing consumption or decreasing work, e.g., giving away to charity. This question was not successful and we decided to infer consumption as described in the text. Taking into account other uses of the windfall would increase the implied labor elasticity, i.e., consumption would rise less as a result of the sweepstakes, so any change in labor would be a response to a smaller increase in the W/C ratio.

appropriate branch of the module.[9] The sweepstakes is described to all respondents as follows:

> Suppose you won a sweepstakes that will pay you [and your (husband/wife/partner)] an amount equal to your current family income every year for as long as you [or your (husband/wife/partner)] live. We'd like to know what effect the sweepstakes money would have on your life.

Working respondents are then asked about their labor supply responses to winning the sweepstakes. First, the survey asks whether the workers would quit work entirely after winning the sweepstakes. Second, for those who would not quit, the survey asks whether they would reduce hours and if so, by how much. For single workers and for married workers whose spouse does not work, the questions are as follows:

> "Would you quit work entirely?"
>
> If the answer is "no" to quitting, the survey asks "Would you work fewer hours?"
>
> If the answer is "yes" to fewer hours, the survey asks "How many fewer hours?"

Married respondents who do not work, but whose spouse does work, are asked for a proxy response for their spouse. When both members of the couple work, the respondent is asked the questions for both the respondent and the spouse.

The survey also includes follow up questions about whether the respondents who do not quit would look for a more pleasant job at reduced pay and about how they would use the extra income. These questions do not play a role in our main analysis. As discussed in footnote 8, taking into account alternative uses of the windfall from the sweepstakes would lead to higher estimates of the labor supply elasticity.

4.2 Survey Implementation

The sweepstakes module is an experimental module on the Health and Retirement Study (HRS). The HRS is a panel dataset of individuals 50 years and older and their spouses.[10] Our module gives us the data on the labor supply response of individuals to winning the sweepstakes. The main HRS data provides us with extensive information on income,

[9] The questionnaire for the module is available online at the HRS WWW site http://hrsonline.isr.umich.edu.

[10] The HRS is a nationally-representative sample except that blacks and residents of Florida are oversampled. See the HRS WWW site for detailed information about the HRS.

current hours of work, and so on, that allow us to make the computations outlined in Section 3. Our module was on HRS 1994 (Wave 2) and HRS 1996 (Wave 3). The baseline interview of the HRS is conducted in person. Subsequent waves, including those containing the module, are conducted by telephone. Experimental modules are asked of a randomly selected subset of HRS households. They are placed at the end of the survey so that their experimental nature does not affect responses to the main HRS questionnaire.

The HRS data has both an individual and household structure. We make use of both the individual and household structure in our analysis. As discussed above, the sweepstakes question is asked of only one respondent in the household. The respondent is asked to give a proxy response for his or her spouse, if applicable.

Across the two waves, the module was asked of 2,660 respondents. Of these, 2,069 had spouses and 531 were single. Hence, there were 4,669 potential individuals in our data set. Of these 1789 were not working according to their response in the module. We lose 1,115 observations owing to missing data on the main HRS (e.g., age, hours, wages) and 377 observations owing to missing or invalid responses to the module. This leaves us with 1,388 observations for the analysis.

4.3 Labor Supply Response to Winning the Sweepstakes

4.3.1 Overall results

The responses to the sweepstakes question imply very high labor supply elasticities. Table 1 shows the basic results by marital status. Each entry in the table is the fraction of individuals in each type of family (expressed as a percent) who would make the labor supply choices indicated in the rows. The bottom row of the table gives the number of observations.

- Somewhat over half the individuals report that they would quit work after winning the sweepstakes that pays their current family income for life.
- A substantial minority—about one-fifth of individuals—would not change their hours of work at all.
- The remaining individuals would reduce their hours of work.

Regardless of the parametric model, the high fraction of those who quit will imply a very high elasticity of labor supply in aggregate. The high fraction of those who quit also highlights the importance of taking into account nonconvexities in labor supply.

4.3.2 Flexibility of hours

Not quitting work could be consistent with unconstrained maximization as long as workers are far from the quit/not-quit margin. It could also be affected by whether a respondent can flexibly adjust hours at a job since many employers mandate the workweek as one of the features of the job. The HRS asks whether hours of work are flexible and if not, whether the worker works more or less than desired. The majority of workers in the HRS report that their hours are not flexible. This inflexibility could affect the response to the sweepstakes question. In particular, workers in inflexible jobs might be compelled to quit in response to a wealth shock even though they would prefer to cut hours were their job flexible.

Table 2 shows the labor supply responses to winning the sweepstakes as a function of flexibility of hours. Of the 1,388 respondents, 29.8 percent (414 respondents) report having hours that are flexible down, 64.3 percent (892 respondents) report hours that are inflexible downward, and 5.9 (82 respondents) do not give an answer.[11] There is only a small excess propensity to quit work after winning the sweepstakes for those workers with inflexible hours. Moreover, the fraction of inflexible workers who say they would reduce their hours is about the same as for those with flexible workweeks. Hence, inflexibility should not have a major impact on our results. As we discuss in section 6.3, this ability of respondents to give answers that break free of the actual constraints on behavior they face in their current situation is an important potential benefit of the experimental survey approach. Apparently, our respondents do embrace the counterfactual of hours flexibility in answering our question about the sweepstakes, as we hoped they would.

The right side of Table 2 cuts the data in terms of differences between desired and actual hours. For those workers who say their hours are inflexible, the HRS asks them whether they would like to work more or less, and by how much. The responses to the sweepstakes question are given according to whether desired hours are greater, equal, or less than actual hours. The "equal" category includes those who are flexible, and those who are inflexible, but still work their desired schedule. There is no discernible difference in the rate of quitting among those whose actual hours equal desired hours and those whose actual hours exceed desired hours, i.e., those who have a repressed desire to reduce hours. This latter group is somewhat more likely to reduce hours rather than leave them unchanged. The smaller group that has repressed desire to increase hours has more noticeably different

[11] A majority also report not having upward flexibility. This margin is not relevant for the sweepstakes response.

behavior, i.e., is less likely to quit and is more likely to have no change in hours. The message of this tabulation is that the constraints on changing hours are at least as likely to push down the estimate of the elasticity as to increase it. Though binding constraints on reducing hours are more common than constraints on increasing them, constraints on increasing hours have a bigger effect on the answer to the sweepstakes question.

4.3.3 Family structure, sex, and age

Marital status has a substantial effect on our estimate of the labor supply elasticity. As shown in Table 1, single individuals are less likely to quit and more likely to leave hours unchanged or reduce hours than married workers. Among married workers, single earners are less likely to quit or reduce hours than dual earners. The structural estimates of the labor supply elasticities presented in the next section account for features of family structure that could account for differences in behavior across household type. Moreover, differences in circumstances, e.g., age and income, could explain the differences in behavior by marital status. We will also investigate these factors in what follows.

The last column of Table 1 shows the individual responses for the dual earners. Table 3 shows the joint distribution of labor supply responses for dual-earner households. Both members of the couple quitting is by far the most frequent response. Having one member quit and the other not change is the next most common response. The wife is more likely to be the worker who quits. When both members of a couple have interior reductions in hours, the magnitude of the reductions is positively correlated. These results are consistent with a model that features fixed costs of working, though coordination between spouses also appears to be a factor.

Table 4 examines the response to the sweepstakes by sex and household type. For all types of households, females are more likely to quit than males, with proportionate reductions in no change and reducing hours. The differences are much more pronounced, however, for the married couples.[12] This finding suggests that family labor supply decisions rather than sex *per se* accounts for most differences in labor supply decisions between men and women.

For the sample overall, Table 5 shows there is surprisingly little heterogeneity by age in quitting work as a response to winning the sweepstakes. Older workers are less likely to cut hours conditional on not quitting than younger workers. The increase in the "no change"

[12]Note that single-earner females are quite similar to dual-earner females in their behavior. In this age cohort, many single-earner females will have a recently-retired husband, so their situation is similar to dual-earner females whose husbands also quit in response to the sweepstakes.

response with age likely indicates endogenous selection. Those who are still working in their sixties are more likely to like their jobs for their own sakes.

For the responses by age in particular, but also by other covariates, there are important interactions with family status. After we discuss the estimation of the behavioral parameters by individual, we will be able to analyze them allowing for such interactions. This issue is taken up in the next section.

4.4 Estimates of Labor Supply Elasticities from Responses to Wealth Shocks

In this section we discuss estimating labor supply elasticities. While we apply them to the evidence from our survey about the labor supply response to winning the sweepstakes, these techniques could be applied to studying other wealth shocks.

In section 3 we described how to produce an individual-specific estimate of the elasticity of labor supply. In this section, we present estimates based on the method described there using the HRS data and the responses to the sweepstakes questions. In addition to these data, we also need to provide a value of the nonseparability parameter α. We have no way of identifying this parameter with our data, so we set it to a plausible value and explore alternatives. We present estimates with α equal to 0.3 as a baseline. In section 6.1 we discuss the empirical literature relevant for calibrating this parameter. We also consider cases with α equal to 0.5 to illustrate the effect of a high degree of nonseparability and of 0.1 to illustrate what happens when the utility function is close to separable.

Table 6 presents the ratio of job-induced consumption to total consumption before and after winning the sweepstakes—that is, J/C and $\hat{J}/\hat{C}$—for the three values of α and for various responses to the sweepstakes question. Consumption C is a household-level variable. J and $\hat{J}$ are the sum of the two individual values of job-induced consumption for dual-earner households. The estimates of the ratio before and after winning the sweepstakes both depend on the individuals' response to winning the sweepstakes. Those who quit or reduce hours substantially have relatively high job-induced consumption before winning the sweepstakes. Given the proportionality between job-induced consumption and the money-metric disutility of work, these are precisely those who would be forecast to reduce work more or to quit work after a positive wealth shock because going to their jobs is relative costly. In contrast, those who would not adjust hours at all have zero job-induced consumption because they have zero elasticity of labor supply.

For those who quit, the job-induced consumption in the last row of Table 6 is a lower bound. We address the problem of estimating population averages in presence of censoring later in this section.

The second column of Table 6 shows the ratio after winning the sweepstakes, that is, after adjusting hours and consumption according to the response of the individual to the sweepstakes question. Overall, job-induced consumption falls dramatically as a fraction of total consumption. Individuals mostly work less. Job-induced consumption is zero for those who quit because they are not working at all. For those who continue to work after the lottery, total consumption increases dramatically while job-induced consumption remains at zero (for those who do not adjust hours) or falls (for those who reduce hours).

The pairs of columns with alternative values of α show the effect of varying the degree of non-separability. For $\alpha = 0.5$, a higher degree of non-separability, job-induced consumption is higher. For $\alpha = 0.1$, a lower degree of non-separability, job-induced consumption is lower. For a zero value of α, it would be zero.

Table 7 presents the estimates of the utility-function parameters η for the three values of α. We calculate these parameters for each individual in our sample based on their labor-supply response to the sweepstakes questions and what it implies about the change in their total and job-induced consumption. See equation (25). The parameter η is zero for those who do not change hours at all. For those who quit, shown in the last row, we only know the lower bound $\eta^{\#}$ given in equation (26). For those who reduce hours, we have a non-zero point estimate of η.

For α equal to 0.3, the median value of the elasticity η for those who reduce hours is 0.59. Among those who reduce hours, there is substantial variation in η depending on how much they reduce hours in response to the proportionate wealth shock.

Again, for those who quit, we only have a lower bound on η. For α equal to 0.3, the median value of the lower bound is is 0.88.

The estimate of the elasticity depends on the nonseparability of consumption of labor parameterized by α. The higher the degree of non-separability, the lower is the value of the labor supply elasticity η implied by a given reduction in labor supply in response to the wealth shock. The higher the nonseparability, the higher the job-induced consumption for the level of labor supply before the wealth shock. A given reduction in labor in response to the wealth shock corresponds to a lower elasticity the higher is job-induced consumption.

The value of α has noticeable effects on the magnitudes of the estimates of the elasticity, but not so great as to make the key findings of the paper depend critically on the value of α as it varies over a range of plausible values. For α equal to 0.5, the median lower bound is 0.80; the median value of the bound for the elasticity is 0.97 for α equal to 0.1.

Individuals have high levels of labor supply elasticity as evidenced by the high fraction who quit and the substantial reductions in hours among those who do not. Quitting represents very elastic behavior. Note that this conclusion does not depend on knowing the magnitude of the fixed cost of going to work. It depends only on the empirically validated claim that if hours fall below some critical value (19 in our calculations) that few individuals work. The quitters, because they are so common and because they have very high elasticities, will tend to dominate the aggregate calculation. The workers who use the intensive margin are also informative. Most make substantial adjustments, which leads to high estimates of their labor supply elasticity.

A further lesson of this research is that there is substantial heterogeneity in the labor supply elasticity. Differences in preferences will generate different responses to shocks and policy interventions. We emphasize this point elsewhere for risk tolerance and intertemporal substitution. See Barsky, et al. (1997).[13] Notwithstanding the substantial heterogeneity in labor supply responses, it is useful to provide a summary, average measure of the elasticity of labor of supply. For the majority of individuals who quit work after winning the sweepstakes, we have only a lower bound on their elasticity, so simple sample statistics do not convey the central tendency of the results. To provide this central tendency, we present estimates using the standard truncated normal regression model, that is, we assume that

$$\eta_i = X_i' \beta + \epsilon_i \tag{31}$$

describes the individual Frisch elasticity, where the i subscript denotes the individual. The covariates X_i include a constant and observables that might account for heterogeneity. The ϵ_i is a mean zero disturbance that accounts for unobserved heterogeneity.[14] We observe the individual elasticity if the individual does not quit, and the lower bound $\eta_i^{\#}$, otherwise. Under the assumption that ϵ_i is distributed normally, this is simply a Tobit model with known, variable truncation points.

Table 8 presents the estimates of this censored regression model for the individual elasticities for the case where α is equal to 0.3. The first column of Table 8 reports the

[13] Note that the heterogeneity documented here is for the elasticity of labor supply. There is also substantial heterogeneity in the *level* of labor supply: individuals who appear quite similar based on observables work different amounts. Our model takes into account this heterogeneity through the work aversion parameter M_i. These fixed effects, however, drop out of the calculation of the elasticity. See equation (22).

[14] Those who do not change hours at all have an exactly zero elasticity. Hence, a logarithmic model is inappropriate.

specification with a constant only. The results are not surprising in light of Table 7. The mean elasticity is 1.004 with a very small standard error. This point estimate is higher than the median lower bound for the quitters. Given that quitters make up much of the sample, their behavior dominates the results. The standard error of ϵ, denoted σ, is 0.765. This sizeable estimate reflects the substantial heterogeneity in the individual elasticities discussed above. This heterogeneity is observed for the non-quitters and censored for the quitters.[15]

Table 8, Column (2), adds in controls for household type. Households with single individuals are the excluded category. The married households, particularly those with two earners, have higher labor supply elasticities than single workers. The dual-earners have elasticities 0.379 greater than singles on average and 0.239 (0.379 - 0.140) greater than married single earners. Recall that our analytic framework accounts for the most basic features of family structure. The family budget constraint integrates the joint labor market decisions of members of couples and these decisions account for fixed costs of working. Specifically, it takes into account the differential hours and wages of dual-earners and that having one spouse quit is not as drastic as having both quit. Taking these factors into account, though important, does not fully account for the effects of family structure. A richer parameterization could account for differences in behavior by family structure. For example, the leisure of members of couples might have nonseparabilities that cause them to quit simultaneously. For single-earner couples, these non-separabilities could well be weaker than for dual-earner couples given that *ex ante* only one member of the couple is working. Similarly, single individuals might socialize more through their jobs, and thus have a fixed benefit of working that partially offsets the fixed cost of working. Our model permits one to speculate about these factors (and in future work estimate them) because it controls for how income affects family labor supply decisions.

By being imbedded in a life-cycle setting, the model accounts for the differential impact of winning the sweepstakes based on age. Age *per se* should not effect the labor supply elasticity. The results show that including controls for age do not explain the dispersion in the labor supply elasticity. Table 8, Column (3), includes controls for roughly equally-sized age groups (with the youngest group omitted). Age is not economically or statistically significant. Hence, at least over the relatively narrow age range of the HRS, our model can account for behavior without having the elasticity be age-dependent.

The last column of Table 8 includes controls for income quintile (lowest quintile omit-

[15] We have done the same calculations for the high and low job-induced consumption cases. For $\alpha = 0.5$, the mean labor supply elasticity is 0.908; for $\alpha = 0.1$, it is 1.125.

ted). Though the estimated parameters bounce around somewhat, there is no systematic pattern of labor supply elasticity as a function of income. Thus, there is no clear relationship between income and labor supply elasticity. Again, it is a victory for the theory that income does not directly explain the estimate of a preference parameter.

Table 9 considers how other variables covary with labor supply. Column (1) adds sex and race. Females have somewhat more elastic labor supply than do men, but the difference is quite small compared with the finding of the literature (discussed below) that the labor supply elasticity is much larger for women than for men. The model accounts for factors other than preferences that can make women more likely to quit work. These include working few hours and therefore being both closer to the quit margin and having relatively high job-induced consumption. Moreover, to the extent that wives earn less than husbands, the model accounts for the differential impact of wives quitting on household consumption. Hence, it can account for substantial differences in behavior between men and women without attributing this behavior to a substantial difference in a utility-function parameter. In contrast to sex, race has essentially no explanatory power for labor supply elasticity. As with income, there is no good reason to expect race to directly affect preferences.

Recall that respondents were asked to respond to the sweepstakes question for their working spouses. It is possible that the uses of a proxy response could systematically affect the results. Table 9, Column (2) adds a dummy variable for whether the individual had such a proxy response. The coefficient of the dummy is small and statistically insignificant, which indicates that the use of the proxy responses do not systematically bias the findings.

As a final attempt to find correlates that predict the estimated elasticities, Table 9, Column (3) includes dummies for educational attainment (with the "some high school" category excluded). Those with post-graduate education have significantly lower labor supply elasticities than those who did not attend college.[16] The finding that individuals in the highest education class have substantially lower labor supply elasticities is intriguing. It could be related to the choices that induced that level of education in the first place. Clearly, unobserved factors, e.g. ability, affect educational choices, so these factors may well be the correlates of labor supply elasticities. In any case, it is a salient and striking finding that those who chose to get more education have relatively low labor supply elasticities.

[16]Recall that heterogeneity in the *level* of labor supply is differenced out in our estimates.

5 Alternative Measures of Labor Supply Elasticity: Theory and Evidence

Given our assumption of a zero long-run elasticity of labor supply, there is a direct relationship between the Frisch labor supply elasticity and other familiar labor supply elasticities. In this section we discuss how to translate our estimates of the Frisch labor supply elasticity into estimates of the marginal expenditure share of leisure ℓ, the consumption-constant labor supply elasticity η^C, and the utility-constant labor supply elasticity η^U. We also show how to derive implications for the uncompensated labor supply elasticity η^X for one member of a dual-earner household (and for a single or single-earner when consumption is not equal to labor income initially). In the constant elasticity functional form we use in this paper, $\eta^\lambda = \eta$, but all local relationships in this section hold more generally. Appendix B gives the derivation of the results that are stated in this section.

5.1 Review of Elasticity Concepts

Let us briefly review the economic significance of these labor supply elasticity concepts. In a frictionless world, the *marginal expenditure share of leisure* ℓ is the fraction of an extra dollar a household would spend on reducing labor hours to get more leisure. Given normality of leisure so that it is positive, ℓ is the *absolute value of the local marginal propensity to earn.*

The *utility-constant* labor supply elasticity η^U is the theoretical substitution effect, that is, the percentage increase in labor hours when the household moves along an indifference curve to a point with a 1 percent higher slope, representing a 1 percent higher real wage. When income and substitution effects cancel in the simple case where consumption equals labor income, the utility-constant labor supply elasticity equals the absolute value of the local marginal propensity to earn, $\eta^U = \ell$.

We noted above that when labor supply is scale-symmetric in consumption, labor supply can be represented as a function $N^s(W/C)$. The *consumption-constant* labor supply elasticity η^C is the elasticity of this function with respect to the ratio W/C. The consumption-constant labor supply elasticity equals the Frisch elasticity in the additively separable (α equal zero) case. The consumption-constant labor supply elasticity is especially useful for comparative steady-state analysis of permanent changes in marginal tax rates.

Finally, in a frictionless world the *marginal-utility-of-wealth-constant* labor supply elasticity or *Frisch* labor supply elasticity η^λ represents the sensitivity of labor supply to a temporary increase in the real wage. That is, if the real wage increases temporarily by 1 percent, labor supply would increase for the same period of time by η^λ percent. In the frictionless world, the Frisch elasticity governs intertemporal substitution in labor supply and is tightly linked to the effect of the real interest rate on labor supply. Macroeconomic models often are calibrated in terms of the Frisch elasticity. It is the elasticity concept relevant for understanding the response of labor to transitory, cyclical changes in the real wage.

The raw data from our survey is the change in labor supply as a result of a specified wealth shock. The *raw marginal propensity to earn*, MPE, is the fraction labor income falls relative to the income flow from the sweepstakes, i.e., the wage times the hours reduction divided by the annual income. For those who do not adjust hours it is zero. For those who quit, it equals one in absolute value for individuals whose labor income is the sole source of household income. For others, its absolute value is between zero and one. The MPE is typically used as a nonparametric measure of wealth effects. [See Imbens, Rubin, and Sacerdote (2001), for example.]

Table 10 shows for different values of the raw MPE how these different labor supply elasticity concepts relate to each other numerically in the simple case where a single-earner household consumes its labor income each period. The uncompensated elasticity is zero by construction in this specification. (This simplification allows us to calculate the elasticities analytically. The income equals consumption restriction, which would hold for liquidity-constrained households, makes the model a static one.) The relevant formulas are discussed below. What leaps out of the table is that while η^U (the "substitution effect" in its pure form) is numerically equal to the marginal expenditure share of leisure ℓ (equal to the absolute value of the local marginal propensity to earn) when income and substitution effects cancel, the numerical values of the consumption-constant elasticity η^C and Frisch labor supply elasticity η^λ can be substantially larger and—at the upper end—quite sensitive to the size of the income effect as measured by ℓ.

The top panel of Table 10 gives the estimates for the separable case where $\alpha = 0.0$. The bottom panel gives them for the non-separable case with $\alpha = 0.3$ that we focus on in Section 4. The comparison of these tables highlights how non-separability affects the key parameters. There is little effect on the local MPE. There are however, substantial effects on the elasticities. For a given MPE, increasing the non-separability parameter α increases the consumption-constant elasticity and decreases the Frisch elasticity. (See the discussion of Table 11 for further discussion of the effect of α on the estimates.)

5.2 Formulas and Estimates

In this section, we present the formulas for the local marginal propensity to earn, consumption-constant labor supply elasticity and utility-constant labor supply elasticity. See Appendix B for the derivation of these formulas. Because of the importance of dual-earner families empirically, it is important to have formulas that are valid for dual-earner households as well as single and single-earner households.

We then use these formulas to construct individual-specific measures based on our survey. Table 11 presents averages of these measures based on the same statistical model used in Tables 8 and 9. Before turning to the structural estimates of these parameters, we consider the raw data. The MPE is -0.373 on average. It is not very different across household types. The table shows individuals MPEs for members of dual-earner households. The dual-earner MPE at the household level is the sum of the two individual MPEs. The household MPEs are much larger for the dual earners than the single or single-earners. The female dual-earner MPE is relatively low, since females contribute a relatively low fraction of family income in these data. Hence, even in studying the raw data, the lessons of the model concerning joint labor supply and the importance of relative wages within the household becomes apparent.

The Local Marginal Propensity to Earn: In a dual-earner household, the absolute value of the local marginal propensity to earn for individual i is given by

$$\ell_i = \eta_i^\lambda \left(\frac{W_i N_i}{C + (1-\alpha) \sum_j \eta_j^\lambda W_j N_j} \right).$$

In a single or single-earner household, this reduces to

$$\ell = \eta^\lambda \left(\frac{WN}{C + (1-\alpha)\eta^\lambda WN} \right).$$

The local marginal propensity to earn is greater in absolute value than the raw MPE. The parameters are evaluated at the initial level of hours. Convexity of the Engel curve mandates that the local value will be greater than the raw MPE, which averages over a large change in labor input.[17] Our results imply that, absent frictions, individuals would

[17] Even aside from our particular functional form, we consider the argument for a convex interior-solution portion of the Engel curve strong (with N on the horizontal axis and C on the vertical axis). Imagine gradually increasing the amount of non-labor income. As labor hours fall, it seems plausible that—*up to the point where fixed costs lead a worker to quit*—inducing each hour of reduced labor will require at least as great a percentage increase in consumption as the previously reduced hour. This makes the interior-

spend on average almost 60 percent of a wealth shock in reduced labor input.

The Utility-Constant Labor Supply Elasticity: In a dual-earner household the utility-constant labor supply elasticity is given by

$$\eta_i^U = \eta_i^\lambda \left(\frac{C}{C + (1-\alpha) \sum_j \eta_j^\lambda W_j N_j} \right).$$

In a single or single-earner household, this reduces to

$$\eta^U = \eta^\lambda \left(\frac{C}{C + (1-\alpha)\eta^\lambda W N} \right).$$

Note that when consumption is equal to labor income, η^U is equal to the local marginal propensity to earn in the single and single-earner case, but not in the dual-earner case. However, when η_1^U and η_2^U are *averaged* with labor income weights, the value is equal to the *sum* of the local marginal propensities to earn times the consumption to total labor income ratio:

$$\frac{\sum_j \eta_j^U W_j N_j}{\sum_j W_j N_j} = \frac{C}{\sum_j W_j N_j} [\sum_j \ell_j].$$

This means that in dual-earner households, it is the *sum* of the marginal propensities to earn one should look at to guess the average sizes of the other labor supply elasticities associated with the substitution effect.

The mean utility-constant labor supply elasticity is 0.793. As with the Frisch elasticity analyzed in greater detail in Tables 8 and 9, the estimated elasticity varies with household type and to a much lesser extent by sex. Except in the single-earner, consumption-equals-labor-income case, η^U will not equal exactly ℓ. For the single households, they are very close. In the dual-earner households, ℓ is substantially smaller than η^U because the money for reducing labor hours must be divided between two spouses, reducing the size of the individual ℓ.

The Consumption-Constant Labor Supply Elasticity: In a dual-earner household, the consumption constant labor supply elasticity, holding the wage ratio between members of the household constant, is

solution portion of the Engel curve at least as convex as the Engel curve for a constant-semielasticity form of v. Of course, the fixed costs effectively make the Engel curve concave around $N^{\#}$, but we can still get a relatively high lower bound on the elasticity when people quit.

$$\eta_i^C = \eta_i^\lambda \left(\frac{C}{C - \alpha \sum_j \eta_j^\lambda W_j N_j} \right).$$

In a single or single-earner household, this reduces to

$$\eta^C = \eta^\lambda \left(\frac{C}{C - \alpha \eta^\lambda W N} \right).$$

In the separable case with $\alpha = 0$, the consumption-constant and Frisch elasticity are the same. The consumption-constant elasticity will be greater than the Frisch elasticity for positive values of α.

For the value α of 0.3 in Table 11, the consumption-constant elasticity exceeds the Frisch elasticity by about fifty percent. What accounts for the relationships among these elasticities? The estimates of ℓ—and the closely linked η^U and η^C—come from the slope of the Engel curve at the initial value of N. Given η^C, the formula relating η^λ to η^C algebraically pushes η^λ down for a given value of η^C, since consumption must increase to keep λ fixed as the real wage increases and this increased consumption reduces the rise in labor hours from the higher wage.

Why is ℓ—and therefore η^U and η^C—higher when α is higher? The discrete change in C and N is treated as data here. It is the curvature of the Engel curve that determines the relationship between the global MPE and the local MPE. A convex Engel curve means that the slope at the initial value of N will be flatter (yielding a higher local MPE) than the slope of the secant on the Engel curve that determines the global MPE. A higher value of α makes the Engel curve more convex because as non-labor income (from the sweepstakes winnings) increases, reductions in N hold down the increases in C substantially at first, but as the marginal disutility of working falls, the effect of hours reductions on job-induced consumption becomes much less important. Therefore, consumption rises more rapidly with increases in non-labor income at low levels of N where variations in job-induced consumption lose most of their force.

5.3 Uncompensated Labor Supply Elasticity

The uncompensated labor supply elasticity η_{ii}^X for household member i is defined as the effect of an increase in the real wage W_i when all of the additional labor income is devoted to additional consumption and the other household member's real wage is unchanged. The uncompensated elasticity is often what ends up being estimated in econometric studies because of the difficulty of holding utility, consumption, or marginal utility constant in an empirical specification.

The definition of η^U and η^C above involve varying both wages proportionately.[18] But in the case of the uncompensated elasticity η^{X}, it is interesting to look at what happens when only one wage changes holding the other wage constant, since this is what is typically estimated econometrically. In the dual-earner household, let us call the individual we are focusing on individual 1. The uncompensated labor supply elasticity for individual 1, varying only W_1 is

$$\eta^{\mathrm{X}}_{11} = \eta^{\lambda}_1 \left(\frac{C - W_1 N_1 + (1-\alpha)\eta^{\lambda}_2 W_2 N_2}{C + (1-\alpha)\sum_j \eta^{\lambda}_j W_j N_j} \right)$$

In a single or single-earner household, this reduces to

$$\eta^{\mathrm{X}} = \eta^{\lambda} \left(\frac{C - WN}{C + (1-\alpha)\eta^{\lambda} WN} \right) = \ell \left(\frac{C - WN}{WN} \right).$$

Even for single and single-earner households, devoting the extra labor income to consumption is not quite the same as increasing consumption by the same *ratio* as labor income or by the same ratio as the wage itself, so scale symmetry in consumption does not lead to a precisely zero uncompensated labor supply elasticity unless consumption is equal to labor income initially. For a dual-earner household, the fact that the other spouse's wage is held constant makes it possible for the uncompensated labor supply elasticity to be far from zero even when scale symmetry in consumption holds.

Using these formulas, Table 11 shows the average values of the uncompensated labor supply elasticities for males and females in each family structure (allowing for the censoring when an individual quits). The average value is 0.327. While not zero, this value is substantially smaller than the various compensated measures. Hence, our quite substantial compensated elasticities are consistent with the much lower uncompensated elasticities that researchers have come to expect to find in data.

For singles, whether male or female, our theory implies an uncompensated elasticity close to zero. For single earners, the higher average level of non-labor income makes the men's elasticity modestly positive and women's somewhat higher. The uncompensated

[18] Because of our separability assumption, this is not an issue for η^{λ}_i; we do not need to specify the behavior of the other wage to think about η^{λ}_i.

elasticity is largest for dual earners, especially women.

These average values for the uncompensated labor supply elasticity are well in line with the standard empirical finding that women have positive elasticities . For men, the standard empirical estimates of the uncompensated labor supply elasticity are slightly negative, where these numbers are modestly positive.[19]

6 Experimental Survey versus Econometric Evidence

This section discusses how our approach compares to the literature. The first subsection concerns estimates of how much consumption drops at retirement, which bear directly on the parameter α. The second subsection discusses econometric evidence concerning labor supply elasticities. The final subsection compares the survey and econometric approaches to quantifying preference parameters.

6.1 Quantifying α

There are various ways other research could identify α. As we note above, it is related to the amount that observed consumption C falls when a worker quits working *holding lifetime resources constant*. There is a variety of evidence in the literature that shows that consumption does fall after retirement.

Hamermesh (1984), using the Retirement History Study, shows that consumption early in retirement is about 15 percent higher than can be sustained by remaining lifetime resources, so that consumption must fall later in retirement (and might have changed further at the point of retirement.)

Banks, Blundell, and Tanner (1998) present estimates based on a cohort-level analysis of British consumption expenditure data. They find a 35.2 percent drop in consumption

[19]This may point to some degree of departure from strict scale symmetry in consumption. Fortunately, our main conclusions are robust to small departures from exact scale symmetry in consumption. It would be difficult to fully analyze the effects of large wealth shocks (like the sweepstakes we asked people to contemplate) without the structure we imposed with the help of the assumption of exact scale symmetry in consumption. But we show in a Technical Appendix (Kimball and Shapiro, 2003) that *given particular values of the local marginal propensities to earn*, as one could theoretically get from looking at the effects of small wealth shocks, the inference from those local marginal propensities to earn to the size of η^{λ}, η^{C} and η^{U} is not very sensitive to small departures from exact scale symmetry in consumption. Because we are analyzing the effects of wealth shocks, we believe that the estimated local marginal propensities to earn are likely to be quite robust to modest departures from exact scale symmetry in consumption. Therefore, we find it reassuring to know that the mapping from the local marginal propensities to earn to the other elasticities is not very sensitive to modest departures from exact scale symmetry in consumption.

unconditional on income and other covariates except unemployment. See Banks, Blundell, and Tanner (1998, Table 1, Column 2). This unconditional drop is probably their estimate most analogous to our parameter α. Their figure 3 also shows substantial cumulative drops in consumption growth around retirement, which lead to a permanently lower post-retirement level of consumption.

Bernheim, Skinner, and Weinberg (2001), using the Panel Study of Income Dynamics, find reductions of about 30 percent in consumption in the two years following retirement. The reductions are sharpest for low wealth households. For high wealth and income households, there are no such reductions. (See their Figure 4.) The interaction of wealth with the change in consumption growth supports their interpretation of the finding as representing the effects of inadequate wealth accumulation.

Hurd and Rohwedder (2003) also find a drop in consumption of about 15 percent, but used a novel component of the Health and Retirement study to present evidence that it is anticipated. This anticipated drop is consistent with what our non-separable formulation predicts. Hurd and Rohwedder have a similar interpretation of their findings, i.e., that job-induced consumption declines after retirement. Ameriks, Caplin, and Leahy (2002) have similar findings from a survey of TIAA-CREF participants.

Hence, the value of α of 0.3 that we highlight in the presentation of the results is within the range of estimates of how much consumption falls at retirement, though it is on the high side of the range.[20] On the other hand, if the drop in consumption at retirement does owe to a lack of foresight, as suggested by some authors, then the evidence on changes of consumption at retirement may overstate the value of α, making the calculations with $\alpha = 0.1$ more relevant.

6.2 Evidence on Labor Supply

6.2.1 Labor supply elasticity

Though at broad brush, the econometric evidence on labor supply elasticities from the cross-section and the life-cycle support the notion that elasticities are small, the extensive literature is subject to varying interpretations. This section has a brief survey of empirical evidence on labor supply.

Overall, there is very little evidence that male labor supply responds at all to wages (see Pencavel, 1986, Card, 1994) in the cross-section. If anything, the uncompensated labor

[20]Basu and Kimball (2002), using aggregate time series to estimate the nonseparability of consumption and labor, find a somewhat larger value of α.

supply elasticity η^X is negative; that is, the income effect exceeds the substitution effect. The simple average of estimates reported by Pencavel is -0.12 (p. 69). Card (1994, p. 57) suggests that the finding of weakly higher hours for more highly educated workers calls into question the conventional wisdom of a negative relationship between wages and hours. Yet, in any case, whether the estimates show a positive or negative relationship of hours and wages, the estimated effects are very small [see Technical Appendix (Kimball and Shapiro, 2003)].

Estimates for women tend to be positive, often substantially so. (See Heckman and Killingsworth, 1986). Evidence surveyed by Blundell and MaCurdy (1999, Table 1) based on models of non-linear budget sets show a similar pattern of zero uncompensated elasticities for men and uncompensated elasticities close to one for women. Hence, apart from married men, where we find a higher elasticity than is typically found in the literature, our estimates of the uncompensated elasticity reported in Table 11 are broadly consistent with what is found in the literature. Moreover, as our theory and findings make clear, marital status is more important than sex in understanding labor supply elasticity. We have not been able to detect a consensus view from the labor economics literature on the role of marital status in the empirical estimates of labor supply elasticity.

The intertemporal approach to estimating labor supply elasticities [MaCurdy (1981), Altonji (1986), Browning, Deaton, and Irish (1985), Ziliak and Kniesner (1999)] generates estimates of the Frisch labor supply elasticity.[21] The estimates range for a low of around zero in Browning, Irish, and Deaton (1985) to positive but small in MaCurdy (1981). Lee (2001) shows, however, that MaCurdy's estimates are biased down owing to the weak instrument problem. Lee's point estimates using LIML and a longer sample period to overcome the econometric problems with 2SLS are centered around one-half. In work that accounts for the participation margin, Blundell, Meghir, and Neves (1993) find Frisch elasticities of about 0.5 to 1 depending on number of children for married women taking into account the participation margin. Hence, our estimate of one is high relative to most estimates, but given biases in the data that attenuate the coefficients of wages (see next section), our estimates are not wholly out of line with what this evidence indicates.[22]

Mulligan (1998) finds larger intertemporal elasticities than the earlier work. He does so

[21] These studies—though they assume separability between labor and consumption—are akin to ours in producing an estimated parameter that corresponds to a structural utility function parameter. Altonji in particular takes the cancellation of income and substitution effects, i.e., wages and consumption having equal and opposite coefficients in his specification, as the benchmark for judging the estimates.

[22] Moreover, both Altonji (1986) and Card (1994) note the poor quality or absence of consumption data necessary to produce estimates in this framework.

by including older workers, who have declining wages and declining hours, in the sample. Other authors typically exclude these workers.[23] Mulligan's use of this variation in the data is controversial. See the discussions by Hall (1998) and Pischke (1998). This controversy highlights the difficulties with finding suitable variation in the data to identify structural labor-supply parameters.

Recent work in labor economics has looked for circumstances where workers face temporary wage fluctuations and where they have flexibility in responding to them. Oettinger (1999) finds that baseball park vendors respond quite elastically to changes in their effective wages arising from level of attendance at games. Farber (2003) finds that taxicab drivers' labor supply responds strongly to high-frequency variation in their implicit wage.[24] Fehr and Götte (2002) also find elastic behavior in a field experiment varying the wages of bicycle messengers. Hence, this recent work based on experimental or quasi-experimental variation in wages finds a higher elasticity than is typically found in econometric studies.

6.2.2 Home production

One strand of the home production literature, initiated by Benhabib, Rogerson and Wright (1991) among others and surveyed in Greenwood, Rogerson and Wright (1995), has argued that the substitution effect is larger than the income effect. This literature accounts for the facts discussed in Section 2 by productivity in home production being equal to productivity in market production—an assumption that has the same effect as assuming preference shocks that exactly cancel the effects permanent changes in the real wage would otherwise have on labor hours. In order to explain all three facts just mentioned, productivity in home production must have an almost perfect correlation with market productivity internationally, over time and cross-sectionally. We consider this unlikely. But the much weaker assumption that productivity in home production has a significant positive correlation with market productivity over the life cycle [see Rupert, Rogerson, and Wright (2000)] can significantly raise one's estimate of the Frisch elasticity of labor supply. In general, as Greenwood, Rogerson and Wright (1995) point out, models of home production are equivalent to models with the usual reduced-form utility functions, but with what look like preference shocks to the reduced-form utility function interpretable as changes

[23] Mulligan (p. 91) explains the low comovement of hours and wages early in the lifecycle, when wages are rising sharply but hours are relatively flat, as the result of on-the-job training.

[24] Camerer, Babcock, Loewenstein, and Thaler (1997), in an early contribution to this emerging literature, suggest that taxi drivers target daily income rather than responding to implicit wages in deciding how much to work. Farber raises questions about their statistical methodology that call this specific result into question.

to factors like the household production technology. Unless there are strong restrictions on the sources and structure of these shifters of the reduced-form utility function, all the arguments we give about the shape of the reduced-form utility function go through even after considering household production.

If instead, the substitution effect is larger than the income effect as one strand of the home production literature argues, the Frisch elasticity implied by our data would be larger than what is given in our tables.

6.2.3 Marginal propensity to earn

Evidence on the MPE from the literature is also disparate. Pencavel cites estimates ranging from zero to -0.7, though he discounts the larger ones. Ashenfelter and Heckman (1973) report a median estimate of -0.27, not too far from what we find. These estimates are typically based on the role of non-labor income in labor supply, which is typically hard to measure and relatively small. Evidence from tax changes, especially the British studies discussed by Pencavel (1986) and Blundell and MaCurdy (1999), show larger MPEs.

In constrast, negative income tax experiments typically find small MPEs. Given that income and substitution effects go in the same direction in these experiments, this evidence points away from elastic behavior.

There are a few studies that look at how wealth windfalls affect labor supply. Joulfaian and Wilhelm (1994) study the role of inheritances on consumption and labor supply. They find that inheritances have only a small effect on consumption and an even smaller effect on labor supply. There are reasons to believe, however, that these estimates are attenuated relative to true utility function parameters. As they point out, but cannot fully correct for, inheritances may be anticipated. Moreover, they are typically very small, so small as to not overcome the frictions affecting behavior discussed in the next section.

In an important study that closely parallels our design, Imbens, Rubin, and Sacerdote (2001) survey state lottery winners about their labor supply. This study has the important features that we advocate, i.e., an experimental design with large shocks (for at least some of the respondents). They also take into account the fundamental nonlinearity owing to quits by including a quadratic term.[25] Their headline results seem to suggest relatively

[25]It has some limitations owing to the nature of the data, e.g., there is a small sample of winners with large winnings, the authors made a conscious decision not to survey the winners about their spouses' behavior (fn. 6), and those who buy lottery tickets are unlikely to be a random cross section of the population. Yet, as with our study, the large, exogenous signal in their data should help compensate for other limitations.

low MPEs, but these include values of zero for the non-workers. Additionally, unlike our structural estimates that take into account the differences in the cost of quitting as a function of time to retirement, it is hard to compare their results across ages. The raw MPE in the Imbens, Rubin, and Sacerdote results for the 55 to 65 age group, which substantially overlaps the HRS sample, is -0.291 ($= -0.124 - 0.167$, p. 791), which is only a little smaller than our comparable figure of -0.373. They find smaller MPEs for younger workers. Hence, the findings from the hypothetical sweepstakes are quite similar to those of the actually lottery winners. They also find little difference between men and women. As the discussion in the previous section shows, these ranges of estimates imply high structural elasticities of labor supply.

6.3 Econometric versus Experimental Survey Methodology

This paper offers a different type of evidence based on response to survey questions that pose thought experiments. Though this kind of survey evidence has its own set of difficulties, it overcomes some substantial problems with the econometric approach. Econometric identification has a number of rigorous requirements that are very difficult to meet in practice.

First, there must be sufficient signal in the data to identify the parameter of interest. We argue in our introduction that the signal for identifying the long-run elasticity of labor supply is very strong. In contrast, the available signal for identifying the income and substitution effects separately is typically quite weak. Pencavel (1986) and Card (1994) both emphasize how frictions make it very difficult for actual behavior to reveal preference parameters. Blundell and MaCurdy (1999, p. 75) stresses that even in social experiments, it may be difficult for individual behavior to overcome frictions. The experiments last only for several years. Moreover, they are not sufficiently large to affect the economic and social context in which the experimental subjects operate. The problem of low signal also affects studies of income effects. Many shocks to non-labor income or wealth used to identify wealth effects are simply too small to overcome the frictions that impinge on labor supply decisions.

Second, measurement error and other data problems are a source of attenuation biases and other biases in econometric analyses. In the analysis of labor supply, the difficulty in measuring the wage heads the list of measurement difficulties. Difficulties arise as pure measurement issues, for example, when individuals make errors in reporting their wages in surveys or when wage *rates* must be inferred from periodic flows of income divided into periodic hours. Even when a wage rate is accurately measured, the spot wage paid might not be allocative for current labor supply. Workers in all but the shortest-term employment relationships are likely to vary effort in line with the needs of the employer in return for a compensation package that delivers appropriate wages over a fairly long horizon. This averaging out of effort and compensation cannot keep wages from adjusting to changes in effort in the long run, but can certainly operate over a horizon of a year or two. Finally, identification of labor supply relies on being able to decompose wage changes into temporary and permanent shifts.

Measurement issues with non-human wealth are perhaps even more severe than those for wages, especially in the datasets used to study labor supply. Most workers have little non-labor income, so this variable has little signal. Changes in measured non-labor income tend to be subject to substantial noise. Moreover, whatever changes there are in this variable might not correspond to wealth shocks, e.g., they might be anticipated, so they would not result in a contemporaneous change in labor supply even were there no frictions. Hence, either cross-sectionally or in the time series, there is little signal in non-labor income to explain variation in labor supply.

Third, econometric identification requires exogenous variation in the data. Labor supply choices, however, are typically made jointly with other economic decisions—particularly the level of consumption. Moreover, the unobserved factors jointly affecting labor supply, wages, and non-human wealth can easily confound results. Much of the labor literature seeks exogenous instruments or natural experiments to overcome these problems. In addition, panel data techniques are frequently used to control for time-invariant confounding factors. Both instrumental variables techniques and panel data techniques make significant demands for data—such as availability of instruments and a sufficiently long panel—and on the econometric model—such as the exogeneity of instruments and the constancy of unobserved effects. The recent work that uses field experiments or quasi-experimental variation in wages does suggest that elastic behavior is obscured in econometric studies using more conventional data.

Our survey approach is *designed* to overcome these difficulties with the econometric approach. It parallels the experimental approach in building in features to the empirical design that are meant to get directly at the object of analysis—in this case labor supply elasticities.

- The survey approach confronts the respondent with an exogenous shock. Winning the hypothetical sweepstakes is entirely exogenous. All respondents in the sample receive the same treatment.

- o In our design, the shock confronting the respondent is large. Specifically, it is meant to overcome the inertia in behavior in the face of shocks caused by frictions in markets and in individual behavior. While an actual experiment could achieve this same outcome in principle, delivering wealth shocks equal to lifetime income would not be practical.

- o The survey approach has much more modest data requirements than the econometric approach. In particular, the survey approach does not rely on having tightly accurate measures of wages and of non-labor income.

- o The survey approach yields a "within" estimator. It differences out unobserved factors as does the panel approach. The survey approach, however, has the advantage that it does the differencing *at a point in time* instead of across time as in the panel approach. Hence, unlike the panel approach, there is no need to assume that unobserved effects are time invariant in order to control for them.

Hence, the survey approach provides a valuable complement to the econometric approach. As noted above, the econometric approach has difficulties inherent to analyzing data on actual realizations of economic data. For some purposes, it may be desirable to have the labor supply estimates reflect those limitations, for example, by reflecting frictions that confound estimates of preference parameters. For example, consider a one-month reduction in the payroll tax rate. Even if the labor supply elasticity is high, most economists would expect a payroll tax holiday to have a negligible effect on labor supply owing to the difficulties of changing work schedules for a month or two. For some purposes, e.g., a reduced-form analysis of similar shocks to policy, this low effective elasticity might be the correct answer: it describes what might be the actual response to a particular type of policy given the frictions faced by firms and workers. But for purposes of identifying a preference parameter that controls the labor supply response in general, this estimate would be very misleading.[26] Pencavel is quite pessimistic about the ability of standard approaches to yield

[26]Even when frictions are important, the underlying preference parameter can be central to understanding what is going on. For example, it is likely that many of movements in hours that occur over the business cycle are initiated by employers rather than by workers. Because of the potential for institutional coordination, frictions may be much smaller for employer-initiated hours changes than for worker-initiated hours changes. If frictions for employer-initiated hours changes are indeed small in a business cycle context, the underlying labor-supply elasticity might be central to understanding the cyclical hours decisions of a firm that has a long-term relationship with its workers and does not want to alienate them by excessive variation in hours.

estimates of preference parameters because of the complexity and frictions of behavior in actual markets:

> I am not suggesting that the preferences of workers have nothing to do with their market work decisions, only that what I call below the canonical model may not be the most useful characterization of the way in which preferences and opportunity come together to determine outcomes in the labor market. (Pencavel, 1986, p. 6.)

Our survey approach is meant to overcome the canonical difficulties faced by the empirical labor literature in estimating preference parameters *per se*.

The survey approach is not, of course, without its own set of difficulties. Economists typically complain that responses to hypothetical questions are unreliable. While we admit that hypothetical responses are subject to errors, these errors must be balanced against the advantages of the approach. Apart from errors introduced by respondents, our approach has all the advantages of an exogenous experimental treatment, which is totally uncorrelated with any and all unobserved factors that one might otherwise worry about. While the experimental survey approach has potential problems, it overcomes problems that pervade virtually all of the existing empirical literature. Hence, it is necessary to balance the potential problems with hypothetical questions against the benefits of the experimental design. Moreover, there is a declining marginal product argument to considering evidence from experimental surveys. The existing empirical literature, though vast, has not firmly established estimates of the key preference parameters we seek to estimate by our approach. New techniques are needed to overcome the well-known limitations of existing empirical approaches and the data that are available to implement them. The experimental survey approach has this potential to provide distinctive evidence about preference parameters.[27]

[27] There are other potential difficulties with surveys, emphasized more by researchers in cognitive psychology and in survey methodology, than in economics. These include biases arising from the framing and ordering of survey questions and from the mode of the survey. Our survey has the advantage of being implemented as part of a well-established and professionally conducted survey of a representative sample of the HRS's target population. Our survey instrument was subject to editing by staff of the Survey Research Center (SRC) that is expert in the technical and cognitive aspects of surveys, and to pretesting using standard SRC procedures. While these procedures do not eliminate problems of framing and mode effects, our survey did have the advantage of being implemented by survey professionals as part of an ongoing study. Further work along these lines is needed to quantify mode and framing effects.

7 Implications and Conclusions

This paper uses a theoretical framework for relating income and substitution effects in labor supply. The framework accounts for optimization over time, for the joint behavior of members of couples, and for fixed costs of working. The motivation for linking income and substitution effects is that it is very difficult to find data that reliably identify substitution effects from observed variation in wages. Our strategy is to use reliably estimated income effects from an experimental setting with large shocks to infer substitution effects.

Our estimate of the constant marginal utility of wealth (Frisch) labor supply elasticity is about one. This estimate is high relative to conventional estimates from the labor supply literature. The estimate is not as high as used in some of the macroeconomics literature. Prescott (1986) assumes a Frisch elasticity of at least 2 in an additive-separable specification. Our estimate goes a long way toward filling the gap between what the microeconomic evidence appears to say and what a dynamic general equilibrium model needs in order to deliver the observed degree of labor fluctuations. Although there has been substantial work in dynamic general equilibrium modeling since the early real business cycle models to account for fluctuations without extreme assumptions about labor supply elasticity, the typically calibrated elasticities are similar to those in Prescott (1986).

High estimates of labor supply elasticities also have implications for public finance. With parameters as large as we find, the long-run effects of taxation on labor supply are likely to yield very substantial distortions and dead-weight burdens.

Finally, this work carries a methodological lesson. It is very difficult to infer preference parameters from data generated by actual economies. The main variation in the data is not exogenous. Even when instruments are available, the signal in the data may be small relative to the rigidities and frictions determining behavior. The experimental survey approach holds the promise of designing studies to infer key preference parameters. That said, we should acknowledge the converse: rigidities and frictions may mean that preference parameters do not translate into behavior in a simple way. Elasticity estimates such as ours do not necessarily predict significant responses to small, temporary changes in wages precisely because of the frictions and rigidities that make econometric inference difficult.

Appendices

A Imputing the Change in Baseline Consumption from the Life-Cycle Model

Consumption before winning the sweepstakes is imputed from income and an age-appropriate saving rate. Baseline consumption before winning the sweepstakes is then imputed from $B = C - \sum_i J_i$. Baseline consumption after winning the sweepstakes is imputed by adding the change to baseline consumption implied the life-cycle model:

$$\Delta B = \hat{B} - B.$$

This appendix lays out how we impute ΔB from our data. We will use the fact that equations that hold before winning the sweepstakes hold after winning the sweepstakes once hats have been added to the appropriate variables.

If we use B_t to designate baseline consumption t years in the future and $B_0 = B$ to designate baseline consumption now, equation (5) implies that

$$B_0 = \lambda_0^{\alpha-1}\psi(\nu_0)$$

and

$$B_t = \lambda_t^{\alpha-1}\psi(\nu_t).$$

As mentioned above, $r = \rho$ and the fair annuity and life insurance markets imply that $\lambda_t = \lambda_0$. Therefore,

$$B_t = \frac{B_0\psi(\nu_t)}{\psi(\nu_0)}. \tag{32}$$

The fair annuity and life insurance markets allow one to focus on expected present values. But it is important not to double-count. Therefore we define Υ as non-labor, non-interest income. The lifetime budget constraint looks like

$$
\begin{aligned}
A_0 \quad &+ \quad E_0 \int_0^\infty e^{-rt} \Upsilon(\nu_t) dt + \sum_i \int_0^{R_i - a_i} e^{-rt} \frac{p_i(a_i + t)}{p_i(a_i)} W_{i,t} N_{i,t} dt \\
&= \quad E_0 \int_0^\infty e^{-rt} C_t dt + Q \\
&= \quad E_0 \int_0^\infty e^{-rt} [B_t + \sum_i J_{i,t}] dt + Q \\
&= \quad \frac{B_0}{\psi(\nu_0)} E_0 \int_0^\infty e^{-rt} \psi(\nu_t) dt + \sum_i \int_0^{R_i - a_i} e^{-rt} \frac{p_i(a_i + t)}{p_i(a_i)} J_{i,t} dt + Q
\end{aligned}
$$

where A_0 is the current net worth of the household, a_i is the current age of worker i, R_i is the retirement age for worker i (for now assumed exogenous), $p_i(a_i + t)/p_i(a_i)$ is the probability of living to age $a_i + t$ conditional on having lived to age a_i, $W_{i,t}N_{i,t}$ is labor income for worker i conditional on living to time t, and Q is the expected present value of bequests and other gifts from the household.[28]

Now let us take the "first-difference" of the two extreme ends of equation (33): the after-sweepstakes values of each term minus the original values of each term. The only change to the sum of initial net worth and the expected present value of exogenous non-interest, non-labor income is the expected present value of the sweepstakes winnings. The text of the survey questions states that the sweepstakes pays an amount equal to last year's total family income as long as you [or your partner] live. Given the low rate of inflation during the relevant sample period, we assume that the respondents interpret this to mean that the sweepstakes pays the same *real* amount every year. Thus, denoting the expected present value of the sweepstakes winning by $\mathcal{L}$ and last year's total family income by Y,

$$
\mathcal{L} = Y \int_0^\infty e^{-rt} \frac{p_1(a_1 + t)}{p_1(a_1)} dt \tag{33}
$$

for a single respondent and

$$
\mathcal{L} = Y \int_0^\infty e^{-rt} \left[\frac{p_1(a_1 + t)}{p_1(a_1)} + \frac{p_2(a_2 + t)}{p_2(a_2)} - \frac{p_1(a_1 + t)}{p_1(a_1)} \frac{p_2(a_2 + t)}{p_2(a_2)} \right] dt \tag{34}
$$

for a couple, using the approximation of independence in mortality.

Substituting in $\mathcal{L}$ for the change in $A_0 + E_0 \int_0^\infty \Upsilon(\nu_t) dt$

[28]Both here and in the corresponding expressions below, when a household member works is paid for less than 52 weeks per year, both the labor income $W_{i,t}N_{i,t}$ and the job-induced consumption $J_{i,t}$ need to be multiplied by the actual number of weeks worked per year divided by 52.

$$\begin{aligned}\mathcal{L} &+ \sum_i \int_0^{R_i-a_i} e^{-rt}\frac{p_i(a_i+t)}{p_i(a_i)} W_{i,t}[\hat{N}_{i,t}-N_{i,t}]dt \\ &= \frac{(\hat{B}_0-B_0)}{\psi(\nu_0)} E_0 \int_0^{\infty} e^{-rt}\psi(\nu_t)dt + \sum_i \int_0^{R_i-a_i} e^{-rt}\frac{p_i(a_i+t)}{p_i(a_i)}[\hat{J}_{i,t}-J_{i,t}]dt + [\hat{Q}-Q].\end{aligned}$$

Solving for $\Delta B = \hat{B}_0 - B_0$ using the notation Δ more generally for changes from the original situation to the after-sweepstakes situation,

$$\begin{aligned}\Delta B &= \frac{\psi(\nu_0)}{E_0 \int_0^{\infty} e^{-rt}\psi(\nu_t)dt} \cdot \left\{ \mathcal{L} + \sum_i \int_0^{R_i-a_i} e^{-rt}\frac{p_i(a_i+t)}{p_i(a_i)} W_{i,t}[\Delta N_{i,t}]dt \right. \\ &- \left. \sum_i \int_0^{R_i-a_i} e^{-rt}\frac{p_i(a_i+t)}{p_i(a_i)} \Delta J_{i,t}dt - \Delta Q \right\}.\end{aligned}$$

The income effect of the sweepstakes is the impetus for any reduction in labor supply. Thus, for a given change in labor hours reported by the respondents, the larger is ΔB, the smaller the estimate of the labor supply elasticity. This is important to keep in mind as we make the necessary assumptions to operationalize equation (35). For instance, we begin by assuming that $\Delta Q = 0$—no change in bequests. This biases the estimate of the labor supply elasticity downwards as compared to the likely increase in bequests as a result of the winning the sweepstakes. The assumption of constant real sweepstakes payments also biases the labor supply elasticity downward if the respondent's interpretation of the survey question is constant nominal payments. The other rough and ready assumptions we make in order to operationalize (35) are $\Delta R_i = 0$ and the simplifying pair of assumptions

$$W_{i,t}\Delta N_{i,t} = W_{i,0}\Delta N_{i,0} = W_i[\hat{N}_i - N_i]$$

and

$$\Delta J_{i,t} = \Delta J_{i,0} = \hat{J}_i - J_i$$

for the years leading up to retirement. If the household would gradually adjust labor hours downward as retirement neared, the constant elasticity of labor supply assumption implies that ΔN and the closely linked but smaller value of ΔJ would gradually get smaller instead of staying the same size—up to the point where the fixed costs led to retirement.

In itself, this tends to bias the labor supply elasticity estimate upwards. But the earlier retirement due to wining the sweepstakes would result in increased values of ΔN, coupled with smaller increases in ΔJ, and so tend to bias the labor supply elasticity estimate downwards. Modeling both of these factors precisely is beyond the scope of this paper, because it requires calibration of the evolution of the aversion to work parameter M with age (the subject of related work), but we believe that the bias from ignoring earlier retirement is larger than the bias from ignoring the smaller absolute reductions in hours as initial hours fall towards retirement. Thus, we defend the simpler calculations we make as reasonable conservative benchmarks, even though they are not precisely accurate.

The only other assumption necessary is a functional form for $\psi(\nu_t)$, the household equivalence scale at time t. Our households have only one or two adults.[29] For single-person households, we normalize $\psi(\nu_t) = 1$. For dual-person households, we set $\psi(\nu_t) = 2^{0.7}$ based on the evidence on scale-economies in household consumption reviewed by Citro and Michael (1995, p. 176).

B Relationships Among Local Labor Supply Elasticities

This appendix gives a very brief background for the equations for various elasticities in terms of the Frisch labor supply elasticity in the context of the functional form used in this paper.[30] We will derive these equations for the dual-earner case. The single earner case is easy to obtain by setting one of the wages to zero.

One key equation for determining elasticities is (13):

$$N_i^* = v'^{-1}\left(\frac{\lambda^{1-\alpha}W_i}{M_i}\right).$$

Inverted, (13) implies

$$v'(N_i) = \frac{\lambda^{\alpha-1}W_i}{M_i}. \tag{35}$$

Totally log-differentiated, (13) implies

[29] We do not make adjustments for children, which are infrequent in the HRS sample.

[30] See the Technical Appendix (Kimball and Shapiro, 2003), for a demonstration of how these formulas hold in much greater generality.

$$d\ln(N_i) = \eta_i^\lambda[d\ln W_i + (1-\alpha)d\ln\lambda] \tag{36}$$

where $d\ln M_i = 0$ and

$$\eta_i^\lambda = \frac{v'(N_i)}{N_i v''(N_i)}.$$

Note that setting $d\ln\lambda = 0$ yields

$$d\ln(N_i) = \eta_i^\lambda d\ln W_i.$$

Also, combining equations (4), (10) and (16),

$$C = \lambda^{\alpha-1}\{\psi(\nu) + \alpha\sum_j[v(N_j) - v(N^\#) + N^\# v'(N^\#)]\}. \tag{37}$$

Totally differentiating (37) and using (35) yields

$$d\ln C = -(1-\alpha)d\ln\lambda + \alpha\lambda^{\alpha-1}\sum_j \frac{M_j v'(N_j)}{C}dN_j = -(1-\alpha)d\ln\lambda + \alpha\sum_j \frac{W_i N_i}{C}d\ln N_i. \tag{38}$$

Treating W_1, W_2 and the initial level of C, N_1 and N_2 as data, and α, η_1^λ and η_2^λ as known parameters, the objective is to calculate the size of $d\ln N_i$ in response to particular changes in $d\ln W_1$, $d\ln W_2$ and $d\ln\lambda$. Equations (36) and (38) are used in calculating every elasticity below.

To find ℓ_i, set $d\ln W_1 = d\ln W_2 = 0$ in (36) and calculate

$$\ell_i = \frac{-W_i dN_i}{dC - \sum_j W_j dN_j} = \eta_i^\lambda\left(\frac{W_i N_i}{C + (1-\alpha)\sum_j \eta_j^\lambda W_j N_j}\right). \tag{39}$$

In finding both η^U and η^C, we set $d\ln W_1 = d\ln W_2 = d\ln W$. This equal proportional change in both wages gives an elasticity concept appropriate for thinking about the macroeconomic labor supply elasticity. The Technical Appendix (Kimball and Shapiro, 2003) discusses other elasticity concepts. For η_i^U, the additional equation is

$$dC = W_1 dN_1 + W_2 dN_2. \tag{40}$$

In words, (40) says that the household moves along an indifference surface. For η_i^C, the additional equation is

$$dC = 0.$$

In both cases, straightforward but tedious algebra yields the expressions for η_i^U and η_i^C given above in the main text.

Finally, to find η_{11}^{X}, set $d \ln W_2 = 0$ and use the additional equation

$$dC = d(W_1 N_1 + W_2 N_2) = N_1 dW_1 + W_1 dN_1 + W_2 dN_2.$$

Again, straightforward but tedious algebra yields the expression given in the text for η_{11}^X.

References

Abel, Andrew B. and Ben S. Bernanke. 2001. *Macroeconomics*, 4th ed. Addison-Wesley.

Altonji, Joseph G. 1986. "Intertemporal Substitution in Labor Supply: Evidence from Micro Data." *Journal of Political Economy* 94, Issue 3, Part 2: S176-S215. June.

Ameriks, John, Andrew Caplin, and John Leahy. 2002. "Retirement Consumption: Insights from a Survey." NBER Working Paper No. 8735.

Ashenfelter, Orley and James Heckman. 1973. "Estimating Labor-Supply Functions," In *Income Maintenance and Labor Supply*. Glen Cain and Harold Watts, eds. Chicago: Markham.

Banks, James, Richard Blundell and Sarah Tanner. 1998. "Is There a Retirement-Savings Puzzle?" *American Economic Review* 88, Issue 4: 769-788. September.

Barsky, Robert B., F. Thomas Juster, Miles S. Kimball and Matthew D. Shapiro. 1997. "Preference Parameters and Behavioral Heterogeneity: An Experimental Approach in the Health and Retirement Study." *Quarterly Journal of Economics* 112(2): 537-579.

Basu, Susanto and Miles Kimball. 2002. "Long-Run Labor Supply Elasticity and the Elasticity of Intertemporal Substitution for Consumption." University of Michigan working paper.

Benhabib, Jess, Richard Rogerson, and Randall Wright. 1991. "Homework in Macroeconomics: Household Production and Aggregate Fluctuations." *Journal of Political Economy* 99: 1166-1187, December.

Bernheim, B. Douglas, Jonathan Skinner, and Steven Weinberg. 2001. "What Accounts for the Variation in Retirement Wealth Among U.S. Households?" *American Economic Review* 91(4): 832-857.

Blundell, Richard, Pierre-Andre Chiappori, Thierry Magnac and Costas Meghir. 2001. "Collective Labor Supply: Heterogeneity and Nonparticipation." Institute for Fiscal Studies, London Working Paper 01/19.

Blundell, Richard and Thomas MaCurdy. 1999. "Labor Supply: A Review of Alternative Approaches." In Orley Ashenfelter and David Card, eds. *Handbook of Labor Economics* Vol. 3A. Amsterdam: North-Holland.

Blundell, Richard, Costas Meghir and Pedro Neves. 1993. "Labour Supply and Intertemporal Substitution." *Journal of Econometrics* 59: 137-160.

Browning, Martin, Angus Deaton and Margaret Irish. 1985. "A Profitable Approach to Labor Supply and Commodity Demands over the Life Cycle". *Econometrica*, 59: 503-44.

Camerer, Colin, Linda Babcock, George Loewenstein, and Richard Thaler. 1997. "Labor Supply Of New York City Cabdrivers: One Day At A Time." *Quarterly Journal of Economics* 112(2): 407-441.

Card, David. 1994. "Intertemporal Labour Supply: An Assessment." In *Advances in Econometrics*: Sixth World Congress, ed. C.A. Sims.

Citro, Constance F. and Robert T. Michael, eds. 1995. *Measuring Poverty: A New Approach*. National Academies of Science: Washington, DC.

Greenwood, Jeremy, Richard Rogerson and Randall Wright. 1995. "Household Production in Real Business Cycle Theory." In Thomas F. Cooley, ed., *Frontiers in Business Cycle Research*. Princeton, NJ: Princeton University Press.

Farber, Henry S. 2003. "Is Tomorrow Another Day? The Labor Supply of New York City Cab Drivers." Industrial Relations Section, Princeton University, Working Paper No. 473.

Fehr, Ernst and Lorenz Götte. 2002. "Do Workers Work More if Wages are High? Evidence from a Randomized Field Experiment." Institute for Empirical Research in Economic, University of Zurich, Working Paper No. 125.

Hall, Robert E. 1998. "Discussion." In Ben S. Bernanke and Julio Rotemberg, eds. *NBER Macroeconomics Annual 1998*. Cambridge: MIT Press.

Hamermesh, Daniel S. 1984. "Consumption During Retirement: The Missing Link in the Life Cycle." *Review of Economics and Statistics* 66(1): 1-7

Heckman, James and Mark R. Killingsworth. 1986. "Female Labor Supply: A Survey" In Orley C. Ashenfelter, and Richard Layard, eds., *Handbook of Labor Economics*, Vol. 1. Amsterdam: North Holland.

Hurd, Michael and Susan Rohwedder. 2003. "The Retirement-Consumption Puzzle: Anticipated and Actual Declines in Spending at Retirement." NBER Working Paper No. 9586, Cambridge.

Imbens, Guido, Donald B. Rubin, and Bruce I. Sacerdote. 2001. "Estimating the Effect of Unearned Income on Labor Earnings, Savings, and Consumption: Evidence from a Survey of Lottery Players." *American Economics Review* 91(4): 778-794.

Joulfaian, David and Mark O. Wilhelm. 1994. "Inheritance and Labor Supply." *Journal of Human Resources* 29(4): 1205-1234.

Kimball, Miles S. and Matthew D. Shapiro. 2003. "Technical Appendix: General Relationships Among Local Labor Supply Elasticities: " Unpublished, University of Michigan. Available at http://www.econ.lsa.umich.edu/~shapiro/.

King, Robert G., Charles I. Plosser, and Sergio R. Rebelo. 1988. "Production, Growth and Business Cycles: I. The Basic Neoclassical Model." *Journal of Monetary Economics* 21: 195-232.

Klein, L.R. and R. F. Kosobud. 1961. "Some Econometrics of Growth: Great Ratios of Economics." *Quarterly Journal of Economics* 75(2): 173-98.

Lee, Chul-In. 2001. "Finite Sample Bias in IV Estimation Of Intertemporal Labor Supply Models: Is the Intertemporal Substitution Elasticity Really Small?" *Review of Economics and Statistics* 83(4): 638–646.

MaCurdy, Thomas E. 1981. "An Empirical Model of Labor Supply in a Life-Cycle Setting." *Journal of Political Economy* 89: 1059-1085.

Mulligan, Casey B. 1998. "Substitution over Time: Another Look at Life-Cycle Labor Supply." In Ben S. Bernanke and Julio Rotemberg, eds. *NBER Macroeconomics Annual 1998*. Cambridge: MIT Press.

Oettinger, Gerald S. 1999. "An Empirical Analysis of the Daily Labor Supply of Stadium Vendors." *Journal of Political Economy* 107(2): 360-392.

Pencavel, John. 1986. "Labor Supply of Men: A Survey." In Orley C. Ashenfelter, and Richard Layard, eds., *Handbook of Labor Economics*, Vol 1. Amsterdam: North Holland.

Pischke, Jörn-Steffen. 1998. "Discussion." In Ben S. Bernanke and Julio Rotemberg, eds. *NBER Macroeconomics Annual 1998*. Cambridge: MIT Press.

Prescott, Edward C. 1986. "Theory Ahead of Business Cycle Measurement." *Federal Reserve Bank of Minneapolis Quarterly Review* 10, Fall.

Rupert, Peter, Richard Rogerson, and Randall Wright. 2000. "Homework in Labor Economics: Household Production and Intertemporal Substitution." *Journal of Monetary Economics* 46(3): 557-579.

Ziliak, James P. and Thomas J. Kniesner. 1999. "Estimating Life Cycle Labor Supply Tax Effects." *Journal of Political Economy* 107(2): 326-359

Table 1
Labor Supply Responses to Winning the Sweepstakes
(Percent of Responses)

Change in labor	Total	Single	Single-earner, married	Dual-earner, married
No change	21.3	31.4	27.4	14.5
Reduce hours	22.5	32.4	18.8	22.0
By less than 10 percent	0.4	1.0	0.4	0.1
10-25 percent	5.3	7.2	4.6	5.2
26-49 percent	9.3	11.1	8.5	9.3
50 percent	6.1	11.1	3.9	5.9
more than 50 percent	1.4	1.9	1.3	1.4
Quit	56.3	36.2	53.8	63.5
Number	1388	207	457	724

Note: Percent of responses in column. Totals may not add to 100.0 due to rounding. "Reduce hours" are all respondents who reduce hours. Next rows show those reducing hours by particular percentages.

Table 2
Labor Supply Responses to Winning the Sweepstakes, by Flexibility in Labor Hours
(Percent of Responses)

Change in labor	Flexible	Inflexible	Missing	Desired versus actual hours		
				Actual exceeds desired	Actual equals desired	Desired exceeds actual
No change	23.2	19.7	28.1	16.5	22.1	27.5
Reduce hours	21.5	21.7	35.4	26.4	20.9	24.2
By less than 10 percent	0.5	0.3	0.0	1.2	0.0	0.8
10-25 percent	6.0	4.6	9.8	4.5	5.6	5.8
26-49 percent	8.9	9.0	14.6	13.5	7.7	10.0
50 percent	4.8	6.5	7.3	4.5	6.4	7.5
more than 50 percent	1.2	1.4	3.7	2.7	1.2	0.0
Quit	55.3	58.5	36.6	57.1	57.0	48.3
Number	414	892	82	333	935	120

Note: Percent of responses in column. Totals may not add to 100.0 due to rounding. "Reduce hours" are all respondents who reduce hours. Next rows show those reducing hours by particular percentages.

Table 3
Labor Supply Responses to Winning the Sweepstakes, Dual Earners
(Frequency of Responses)

	Wife:							
Change in labor	No change	By less than 10 percent	By 10-25 percent	By 26-49 percent	By 50 percent	By more than 50 percent	Quit	Total
Husband:								
No change	34	0	2	8	10	0	64	118
By less than 10 percent	0	0	0	0	0	0	2	2
10-25 percent	2	0	14	8	2	0	24	50
26-49 percent	8	0	0	20	12	2	36	78
50 percent	2	0	0	12	12	2	18	46
More than 50 percent	2	0	0	0	0	2	8	12
Quit	44	0	10	8	4	2	350	418
Total	92	0	26	56	40	8	502	724

Table 4
Labor Supply Responses to Winning the Sweepstakes, by Sex
(Percent of Responses)

	Total		Single		Single-earner, married		Dual-earner, married	
Change in labor	male	female	male	female	male	female	male	female
No change	22.7	19.8	34.8	30.4	29.4	24.5	16.3	12.7
Reduce hours	25.3	19.8	32.6	32.3	23.0	12.8	26.0	18.0
By less than 10 percent	0.1	0.6	0.0	1.2	0.0	1.1	0.3	0.0
10-25 percent	6.6	4.1	8.7	6.8	5.9	2.7	6.9	3.6
26-49 percent	10.9	7.7	15.2	9.9	10.4	5.9	10.8	7.7
50 percent	5.6	6.5	6.5	12.4	4.5	3.2	6.4	5.5
more than 50 percent	1.9	1.0	2.2	1.9	2.2	0.0	1.7	1.1
Quit	52.0	60.3	32.6	37.3	47.6	62.8	57.7	69.3
Number	677	711	46	161	269	188	362	362

Note: Percent of responses in column. Totals may not add to 100.0 due to rounding. "Reduce hours" are all respondents who reduce hours. Next rows show those reducing hours by particular percentages.

Table 5
Labor Supply Responses to Winning the Sweepstakes, by Age
(Percent of Responses)

	Age (years)				
Change in labor	28-53	54-56	57-58	59-61	62-71
No change	17.2	18.7	21.3	24.8	25.7
Reduce hours	25.3	27.2	22.1	19.0	16.5
By less than 10 percent	0.3	0.9	0.4	0.0	0.0
10-25 percent	6.4	4.8	6.8	4.6	3.9
26-49 percent	9.8	12.7	8.0	7.8	6.8
50 percent	6.4	7.3	6.0	5.2	4.9
more than 50 percent	2.4	1.5	0.8	1.3	1.0
Quit	57.4	54.1	56.6	56.2	57.8
Number	296	331	249	306	206

Note: Percent of responses in column. Totals may not add to 100.0 due to rounding. "Reduce hours" are all respondents who reduce hours. Next rows show those reducing hours by particular percentages.

Table 6
Ratio of Job-Induced Consumption to Total Consumption
Before (J/C) and After ($\hat{J}/\hat{C}$) Winning Sweepstakes
(Median)

	α = 0.3		α = 0.5		α = 0.1	
Change in labor	Before	After	Before	After	Before	After
No change	0.0	0.0	0.0	0.0	0.0	0.0
Reduce hours	0.119	0.041	0.190	0.067	0.042	0.014
By less than 10 percent	0.017	0.007	0.029	0.011	0.006	0.002
10-25 percent	0.075	0.027	0.120	0.043	0.026	0.009
26-49 percent	0.126	0.042	0.201	0.068	0.044	0.015
50 percent	0.146	0.052	0.233	0.086	0.051	0.017
more than 50 percent	0.189	0.057	0.299	0.095	0.066	0.019
Quit*	0.161	0.0	0.255	0.0	0.057	0.0

*For quitters, the *J/C* ratios are lower bounds.
Note: Median individual estimate. α = nonseparability parameter. “Reduce hours” are all respondents who reduce hours. Next rows show those reducing hours by particular percentages.

Table 7
Individual-Specific Elasticity of Labor Supply η
(Median)

	η			N
Change in labor	α = 0.3	α = 0.5	α = 0.1	
No change	0.0	0.0	0.0	295
Reduce hours	0.59	0.56	0.64	312
By less than 10 percent	0.09	0.09	0.09	5
10-25 percent	0.28	0.27	0.29	74
26-49 percent	0.58	0.54	0.62	129
50 percent	0.81	0.75	0.88	84
more than 50 percent	1.00	0.93	1.08	20
Quit*	0.88	0.80	0.97	781

*For quitters, the elasticities are the lower bounds, $\eta^{\#}$.

Note: Median individual estimate. η = Frisch labor supply elasticity.
α = nonseparability parameter. "Reduce hours" are all respondents who reduce hours. Next rows show those reducing hours by particular percentages.

Table 8
Average Labor Supply Elasticity: Censored Regression Estimates

	(1)	(2)	(3)	(4)
Constant	1.004	0.763	0.761	0.767
	(0.034)**	(0.059)**	(0.078)**	(0.064)**
Dual-earner, married		0.379	0.384	0.335
		(0.067)**	(0.069)**	(0.078)**
Single-earner, married		0.140	0.142	0.111
		(0.069)*	(0.070)*	(0.073)
Age 54 to 56 years			-0.006	
			(0.066)	
Age 57 to 58 years			-0.017	
			(0.077)	
Age 59 to 61 years			0.030	
			(0.076)	
Age 62 years and older			-0.022	
			(0.081)	
Second income quintile				-0.094
				(0.076)
Third income quintile				0.196
				(0.087)*
Fourth income quintile				0.079
				(0.086)
Fifth income quintile				-0.031
				(0.086)
σ	0.765	0.752	0.752	0.747
	(0.020)**	(0.020)**	(0.020)**	(0.020)**
Observations	1388	1388	1388	1388

Note: Censored normal regressions. The parameter σ is the standard error of the disturbance in the censored regression model. Dependent variable is the Frisch labor supply elasticity η calculated at α equal to 0.3. The parameter σ is the standard error of the disturbance in the censored regression model. Censoring is at individual-specific cutoffs $\eta^{\#}$ for those who quit after winning the sweepstakes. Excluded category for household type is single. Excluded category for age is 53 years and younger. Excluded category for income quintile is the first quintile.
*Significant at the 5 percent level.
**Significant at the 1 percent level.

Table 9
Average Labor Supply Elasticity: Censored Regression Estimates

	(1)	(2)	(3)
Constant	0.655	0.653	0.721
	(0.069)**	(0.069)**	(0.086)**
Dual-earner, married	0.425	0.399	0.412
	(0.069)**	(0.073)**	(0.069)**
Single-earner, married	0.196	0.167	0.190
	(0.072)**	(0.076)*	(0.071)**
Female	0.121	0.123	0.117
	(0.047)*	(0.047)**	(0.047)*
Black	0.062	0.063	0.037
	(0.079)	(0.079)	(0.078)
Proxy response by spouse		0.055	
		(0.049)	
High school diploma			0.015
			(0.070)
Some college			-0.098
			(0.082)
College degree			-0.102
			(0.092)
More than college degree			-0.226
			(0.087)**
σ	0.752	0.753	0.746
	(0.020)**	(0.020)**	(0.020)**
Observations	1388	1388	1386

Note: Censored normal regressions. Dependent variable is the Frisch labor supply elasticity η calculated at α equal to 0.3. The parameter σ is the standard error of the disturbance in the censored regression model. Censoring is at individual-specific cutoffs $\eta^{\#}$ for those who quit after winning the sweepstakes. Excluded category for household type is single. Excluded category for education is less than high school graduate.
*Significant at the 5 percent level.
**Significant at the 1 percent level.

Table 10
Marginal Propensity to Earn and Alternative Labor Supply Elasticities: Calibrated Values in a Static Model

Raw Marginal Propensity to Earn (|MPE|)	Marginal Expenditure Share of Leisure (ℓ)	Consumption -constant Elasticity (η^C)	Frisch Elasticity (η^λ)
	α=0.0		
0.00	0.00	0.00	0.00
0.05	0.07	0.08	0.08
0.10	0.14	0.16	0.16
0.15	0.21	0.26	0.26
0.20	0.28	0.38	0.38
0.25	0.34	0.51	0.51
0.30	0.40	0.67	0.67
0.35	0.46	0.86	0.86
0.40	0.52	1.09	1.09
0.45	0.58	1.36	1.36
0.50	0.63	1.71	1.71
0.55	0.68	2.15	2.15
0.60	0.73	2.72	2.72
0.65	0.78	3.50	3.50
0.70	0.82	4.59	4.59
	α=0.3		
0.00	0.00	0.00	0.00
0.05	0.07	0.08	0.08
0.10	0.14	0.17	0.16
0.15	0.21	0.27	0.25
0.20	0.28	0.39	0.35
0.25	0.35	0.54	0.47
0.30	0.42	0.72	0.59
0.35	0.49	0.94	0.74
0.40	0.55	1.23	0.90
0.45	0.62	1.61	1.08
0.50	0.68	2.13	1.30
0.55	0.74	2.90	1.55
0.60	0.81	4.15	1.85
0.65	0.87	6.51	2.20
0.70	0.93	12.67	2.64

Notes. See text for description of the static model. The utility-constant elasticity, η^U, equals the marginal expenditure share of leisure (ℓ). The uncompensated elasticity, η^X, is zero by construction.

Table 11
Marginal Propensity to Earn and Alternate Labor Supply Elasticities:
Average Censored Regression Estimates

	Raw Marginal Propensity to Earn (|MPE|)	Marginal Expenditure Share of Leisure (ℓ)	Utility-constant Elasticity (η^U)	Consumption-constant Elasticity (η^C)	Frisch Elasticity (η^λ)	Uncompensated Elasticity (η^X)
All	0.373	0.581	0.793	1.499	1.004	0.327
Single	0.358	0.588	0.561	1.267	0.743	-0.008
Single Earner	0.405	0.617	0.742	1.265	0.912	0.169
Dual Earner	0.338	0.519	0.908	1.689	1.156	0.555
Single, Male	0.340	0.564	0.545	1.360	0.731	-0.004
Single, Female	0.363	0.594	0.566	1.228	0.745	-0.009
Single Earner, Male	0.407	0.605	0.667	1.199	0.837	0.089
Single Earner, Female	0.393	0.638	0.881	1.374	1.052	0.309
Dual Earner, Male	0.366	0.562	0.856	1.656	1.106	0.412
Dual Earner, Female	0.285	0.438	0.969	1.705	1.210	0.706

Note: Estimates based on complete model described in main text. Estimates of mean parameters using censored normal estimates (see text). Elasticities are calculated at α equal to 0.3.