Econometrica, Vol. 86, No. 6 (November, 2018), 2049–2082

AGGREGATING ELASTICITIES: INTENSIVE AND EXTENSIVE MARGINS OF WOMEN'S LABOR SUPPLY

ORAZIO ATTANASIO
Department of Economics, UCL, IFS, and NBER

PETER LEVELL
IFS and Department of Economics, UCL

HAMISH LOW
Department of Economics, University of Oxford and IFS

VIRGINIA SÁNCHEZ-MARCOS
Department of Economics, Universidad de Cantabria

We show that there is substantial heterogeneity in women's labor supply elasticities at the micro level and highlight the implications for aggregate behavior. We consider both intertemporal and intratemporal choices, and identify intensive and extensive responses in a consistent life-cycle framework, using US CEX data. Heterogeneity is due to observables, such as age, wealth, hours worked, and the wage level, as well as to unobservable tastes for leisure: the median Marshallian elasticity for hours worked is 0.18, with corresponding Hicksian elasticity of 0.54 and Frisch elasticity of 0.87. At the 90th percentile, these values are 0.79, 1.16, and 1.92. Responses at the extensive margin explain about 54% of the total labor supply response for women under 30, although this declines with age. Aggregate elasticities are higher in recessions, and increase with the length of the recession. The heterogeneity at the micro level means that the aggregate labor supply elasticity is not a *structural* parameter: any aggregate elasticity will depend on the demographic structure of the economy as well as the distribution of wealth and the particular point in the business cycle.

KEYWORDS: Labor supply elasticities, heterogeneity, aggregation, non-separability.

1. INTRODUCTION

THE SIZE OF THE ELASTICITY OF LABOR SUPPLY TO CHANGES IN WAGES has been studied for a long time. Recent debates have focused on the perceived discrepancy between estimates coming from micro studies, which, with a few exceptions, point to relatively small values of such an elasticity, and the assumptions made in macro models, which seem to need relatively large values. Keane and Rogerson (2012) surveyed some of these issues and the papers by Blundell, Bozio, and Laroque (2011), Ljungqvist and Sargent (2011), and Rogerson and Wallenius (2009) contain some alternative views on the debate. To reconcile the micro evidence and the assumptions made in macroeconomics, much attention

Orazio Attanasio: o.attanasio@ucl.ac.uk
Peter Levell: peter_l@ifs.org.uk
Hamish Low: hamish.low@economics.ox.ac.uk
Virginia Sánchez-Marcos: virginia.sanchez@unican.es
We are grateful for a number of useful conversations with Joe Altonji, Richard Blundell, Guy Laroque, Costas Meghir, Richard Rogerson, and Tom Sargent. We received several useful comments from different seminar audiences and during presentation at the NBER Summer Institute and the Society for Economic Dynamics Conference. Attanasio's research was partially funded by an ESRC Professorial Fellowship and by the ESRC Centre for the Microeconomic Analysis of Public Policy at the Institute for Fiscal Studies. Sánchez-Marcos thanks Spanish MCYT for Grant ECO2009-09614.

 https://doi.org/10.3982/ECTA15067

has been given to the distinction between the extensive and intensive margins of labor supply; see, in particular, Chetty, Guren, Manoli, and Weber (2011). Perhaps surprisingly, in this debate, aggregation issues and the pervasive and complex heterogeneity that characterize labor supply behavior have not been given much attention.[1] This paper aims to fill this gap, while making some original methodological contributions and presenting new empirical evidence.

Preferences for consumption and leisure are likely to be affected in fundamental ways by family composition, fertility, and wealth, as well as by unobserved taste 'shocks,' and so heterogeneity in labor supply elasticities in these dimensions is something to be expected. Labor supply elasticities will vary in the cross section and over the business cycle. The key issue, however, is how significant this heterogeneity is and whether it is important at the aggregate level: does it make any sense to talk about *the* elasticity of labor supply as a *structural* parameter? Aggregation issues are likely to be relevant both for the intensive and extensive margin, as we show.

In this paper, we address these issues focusing on women's labor supply. Our approach consists in taking a relatively standard life-cycle model of labor supply to the data. While the essence of the model is relatively simple, we stress two elements that are important for our analysis and that make our contribution novel. First, we consider all the relevant intertemporal and intratemporal margins and choices simultaneously, in particular, consumption and saving, as well as participation and hours of work. This allows for interaction between different decisions. Second, we specify a flexible utility function that allows for substantial heterogeneity, fits the data well, and, at the same time, allows us to make precise quantitative statements. These elements are important because they allow us to address directly the interaction between extensive and intensive margins and to evaluate empirically the importance of aggregation issues and to calculate both micro and macro elasticities.

In evaluating aggregate labor supply elasticities, it is necessary to specify the whole economic environment because, as noted by Chang and Kim (2006), the aggregate response depends on the distribution of reservation wages. On the other hand, an important methodological contribution of our paper is to stress that key components of the model can be estimated using weaker assumptions which closely approximate the overall model structure. We separate our estimation into three steps and specify what assumptions are needed at each step and what variation in the data is used for identification. The first step identifies the within-period preferences over consumption and labor supply at the intensive margin. We use group level variability driven by group or aggregate shocks such as policy reforms, similarly to Blundell, Duncan, and Meghir (1998). These estimates are used to compute within-period Marshallian and Hicksian elasticities, which hold intertemporal allocations constant and are conditional on participation. The second step estimates intertemporal preferences that generate Frisch labor supply elasticities. We estimate these parameters by using the Euler equation for consumption, using synthetic cohorts, similarly to Blundell, Meghir, and Neves (1993) and Attanasio and Weber (1995), and without taking a stance on the determinants of participation and a variety of other issues, such as retirement or the cost of children. Finally, to characterize behavior at the extensive margin, we specify the model fully. In this step, we calibrate key parameters to a number of life-cycle moments, and explicitly aggregate individual behavior, similar

[1] One exception is Keane and Wasi (2016), who showed men's labor supply responses vary substantially with age, education, and the tax structure. Aggregation issues were also discussed in Erosa, Fuster, and Kambourov (2016).

in spirit to Erosa, Fuster, and Kambourov (2016). Labor supply responses to wages in a life-cycle model may change beyond the static response if savings decisions are affected by wages. Our life-cycle elasticities account for these effects and we discuss the circumstances in which static elasticities provide a good approximation to the overall life-cycle response.

We use a flexible specification for utility to allow for observed and unobserved heterogeneity in tastes at both intratemporal and intertemporal margins, and at the same time allowing for possible non-separability of consumption and leisure. Our specification of preferences is much more flexible than generally allowed for in the literature and we show this is important. Classic papers in the micro literature (such as Heckman and Macurdy (1980)) imply a strong relationship between the Frisch intertemporal elasticity and the intratemporal Marginal Rate of Substitution conditions, which, in turn, forces a strict relationship between within-period and intertemporal conditions. Our approach avoids this restriction. In the macro literature, most papers impose additive separability between consumption and leisure, and isoelastic, homothetic preferences that conform to the restrictions for balanced growth, as in Erosa, Fuster, and Kambourov (2016).[2] Here, we show that the isoelastic specification for consumption and hours is strongly rejected by the data. The challenge, therefore, is to work with specifications that allow much more heterogeneity and changes over time.

Estimates of the size of the elasticity of labor supply for women vary considerably. Our estimates, at the median, are not too different from some estimates in the literature. In particular, on the intensive margin, we obtain a median static Marshallian elasticity of 0.18, with the corresponding Hicksian elasticity considerably larger at 0.54, indicating a sizable income effect. For the same median household, the Frisch elasticity for hours is 0.87. At the same time, we document considerable variation in estimated elasticities in the cross section: the Marshallian, for instance, has an inter-decile range of −0.14 to 0.79. As we show, these static Marshallian elasticities are smaller than the responses when we allow savings to adjust.

In comparing our estimates to the literature, we investigated what drives, in our data, differences in results. A key factor is that the size of the estimates depends on the specific estimator and normalization used. When using standard IV or GMM methods, we typically obtain very large estimates when we put wages on the left-hand side of the MRS equation. Instead, we get much smaller estimates when we put consumption or hours worked on the left-hand side. In our baseline estimation, we use methods robust to the normalization, using a method proposed by Fuller (1977), which is a generalization of a LIML approach.

We use the fully specified model to run two experiments: in the first, we evaluate the labor supply response to temporary changes in wages; in the second, we evaluate the response to a change in the entire life-cycle wage profile. The first experiment captures the impact of a temporary tax cut, which has little effect on the marginal utility of wealth, even if the cut is unanticipated. Without an extensive margin, the response would be captured by the Frisch elasticity. Introducing the extensive margin doubles the size of the response, and is particularly important at younger ages when non-participation because of children is prevalent. The second experiment captures the impact of a permanent tax

[2]This assumption is predicated on the perceived need to work with models that match historical trends showing steady secular increases in real wages with little change in aggregate hours. Browning, Hansen, and Heckman (1999) already noted that the fact that the historical trend for aggregate hours is roughly constant hides a large decrease for men and an increase for women.

cut which will change the marginal utility of wealth. The response to the second experiment would be approximated by the static Marshallian elasticity if there was no change in savings behavior. Allowing intertemporal allocations to adjust gives what we call *life-cycle Marshallian and Hicksian elasticities*. These life-cycle elasticities are greater than the static approximations because not all income is spent on non-durable consumption in the period it is earned. However, these life-cycle elasticities are lower than the Frisch responses to temporary changes.

Using the entire model, we can aggregate explicitly individual behavior and study aggregate elasticities that correspond to the concept used in the macro literature. We find an important role for the extensive margin in generating aggregate movements in labor supply. Importantly, in linking the micro and macro analysis of labor supply, we show that what we call the 'aggregate' elasticity changes considerably over the business cycle, and is typically larger in recessions. Moreover, it gets larger in longer recessions. To the best of our knowledge, changes in the elasticity over business cycles have never been discussed.

The closest macroeconomic paper to ours is Erosa, Fuster, and Kambourov (2016), who had similar aims of building aggregate elasticities from men's labor supply behavior over the life-cycle, and of distinguishing the intensive and extensive margins using a fully specified life-cycle model. The focus of our paper is on women's labor supply responses. A second related paper is Guner, Kaygusuz, and Ventura (2012), who modeled heterogeneous married and single households with an extensive margin for women and an intensive margin for both men and women. Their focus was on evaluating different reforms of the U.S. tax system and they abstracted from wage uncertainty. Both papers operate with very specific preference specifications. We discuss the extent to which our results differ from these papers in the conclusions. Among papers using microeconometrics, our paper builds on a long literature starting from MaCurdy (1983) and Altonji (1986), and on Blundell, Meghir, and Neves (1993), who conditioned on the extensive margin, and estimated jointly the within-period decision and the intertemporal decision.

Our exercise is not without important caveats. In much of our analysis, we do not consider the effect of tenure and experience on wages. Such effects can obviously be important, as labor supply choices may change future wages and, therefore, future labor supply behavior, as stressed by Imai and Keane (2004). Keane and Wasi (2016) modeled human capital and found that labor supply elasticities are highly heterogeneous and vary substantially with age, education, and the tax structure. In Appendix F of the Supplemental Material (Attanasio, Levell, Low, and Sánchez-Marcos (2018)), we extend our analysis to introduce returns to experience on the extensive margin. Introducing returns only on the extensive margin means within-period allocations at the intensive margin are unaffected. By contrast, if the return to experience operates on the number of hours (rather than only on participation), we would need to change our analysis substantially.

The rest of the paper is organized as follows. In Section 2, we outline the life-cycle framework. We show how the preference parameters can be mapped into static, intertemporal, and life-cycle elasticities, and discuss the meaning of the different elasticities. In Section 3, we explain the three steps of our empirical strategy to identify the preference parameters and opportunity set. Section 4 describes the data. Section 5 presents the parameter estimates. Section 6 contains the key results of the paper: the implications of our estimates for labor supply elasticities, distinguishing between Marshallian, Hicksian, and Frisch elasticities, and distinguishing static from life-cycle responses. We also report responses on the extensive margin, aggregate responses, and, more generally, the aggregation issues that are central to our paper. Section 7 concludes. The Supplemental Material (Attanasio et al. (2018)) provides supporting evidence.

2. A LIFE-CYCLE MODEL OF WOMEN'S LABOR SUPPLY

To study both the intensive and the extensive margins of women's labor supply, we use a rich model of labor supply and saving choices embedded in a unitary household, life-cycle framework. Both the intensive and extensive margins are meaningful because of fixed costs of going to work related to family composition and because of preference costs specifically related to participation. The intensive choice is over the typical number of hours worked per week, the extensive margin is over whether to work at all in each quarter. Changes at different margins interact and heterogeneity in these responses is important to understand aggregate labor supply responses to changes in wages.

We consider married couples, who maximize the lifetime expected utility of the household, h, and choose consumption and women's labor supply within each period:

$$\max_{c,l} E_t \sum_{j=0}^{T} \beta^j u(c_{h,t+j}, l_{h,t+j}, P_{h,t+j}; z_{h,t+j}, \chi_{h,t+j}, \zeta_{h,t+j}), \tag{1}$$

where c is consumption, l is hours of leisure for women, and P is an indicator of the woman's labor force participation which can affect utility over and above the effect of hours worked. $z_{h,t}$ is a vector of demographic variables (which include education, age, and family composition), $\chi_{h,t}$ and $\zeta_{h,t}$ represent taste shifters. We assume that demographics, $z_{h,t}$, are observable, whereas $\chi_{h,t}$ and $\zeta_{h,t}$ are unobservable to us, but are known to the individual. Leisure for men does not enter the utility function.

The period utility function is given by

$$u(c_{h,t}, l_{h,t}, P_{h,t}) = \frac{M_{h,t}^{1-\gamma}}{1-\gamma} \exp(\varphi P_{h,t} + \pi z_{h,t} + \zeta_{h,t}). \tag{2}$$

The preference aggregator for hours of leisure and consumption, $M_{h,t}$, is

$$M_{h,t}(c_{h,t}, l_{h,t}; z_{h,t}, \chi_{h,t}) = \left(\frac{(c_{h,t}^{1-\phi} - 1)}{1-\phi} + \big(\alpha_{h,t}(z_{h,t}, \chi_{h,t})\big) \frac{(l_{h,t}^{1-\theta} - 1)}{1-\theta} \right). \tag{3}$$

The function $\alpha_{h,t}$ that determines the weight on leisure as a function of demographics is specified as

$$\alpha_{h,t} = \exp(\psi_0 + \psi_z z_{h,t} + \chi_{h,t}). \tag{4}$$

The unknown parameters governing within-period utility over consumption and leisure are ϕ, θ, ψ_0, and ψ_z, with additional parameters governing the full utility specification γ, φ, and π. Our specification allows for non-separability between consumption and leisure at both the intensive and extensive margin. The taste shifter $\chi_{h,t}$ affects within-period utility over consumption and leisure, and the taste shifter $\zeta_{h,t}$ affects intertemporal choices. These are specific to the cohort-education group, known to the individual, and may be correlated. Non-separability between consumption and leisure depends on the value of γ and so cannot be identified from within-period choices alone.

The general specification of utility allows substantial heterogeneity across individuals in intratemporal and intertemporal preferences, across the intensive and extensive margins, and does not impose that the elasticities of intertemporal substitution for leisure and consumption are constant. Heterogeneity arises partly because elasticities will differ by observable characteristics, z, such as education and the presence of children, and partly because elasticities differ at different levels of consumption and hours of work. Our parametric specification gives a log-linear Marginal Rate of Substitution (MRS) and guarantees integrability. Further, our approach is more flexible than alternatives which have less scope for heterogeneity at the intensive margin, and so heterogeneity has to come through the extensive margin and the distribution of reservation wages.

Maximization is subject to the intertemporal budget constraint:

$$A_{h,t+1} = (1+r_{t+1})\big(A_{h,t} + \big(w^f_{h,t}(L-l_{h,t}) - F(a_{h,t})\big)P_{h,t} + y^m_{h,t} - c_{h,t}\big), \tag{5}$$

where $A_{h,t}$ is the beginning of period asset holding, r_t is the risk-free interest rate, F the fixed cost of work, dependent on the age of the youngest child $a_{h,t}$, and L is maximum hours available. Wages for the woman are given by $w^f_{h,t}$, and earnings for the man by $y^m_{h,t}$.

There are no explicit borrowing constraints, but households cannot go bankrupt. Therefore, in each period, households are able to borrow against the minimum income they can guarantee for the rest of their lives. This minimum income is a positive amount because we bound men's income away from zero. Households have no insurance markets to smooth aggregate or idiosyncratic shocks.

We assume that the cost of work has a fixed component and a component that depends on the child care cost needed for the youngest child, whose age is $a_{h,t}$. Denoting with $G(a_{h,t})$ child care services and p their price, we have

$$F(a_{h,t}) = pG(a_{h,t}) + \bar{F}. \tag{6}$$

Women differ in their age at childbirth, but this is assumed to be deterministic and fully anticipated.[3] The fixed cost of work is deterministic and known. The presence of fixed costs and discrete utility costs of participating mean some women decide not to work at all, especially at low levels of productivity. If a woman does not work, she does so by choice, given the offered wage, demographics, taste shifters, and unearned income. By the same token, it is unlikely that if a woman does work, that she will work only very few hours.

Women's wages are given by the following process:

$$\ln w^f_{h,t} = \ln w^f_{h,0} + \ln e^f_{h,t} + v^f_{h,t}, \tag{7}$$

where $e^f_{h,t}$ is the level of human capital at the start of the period. We assume that wage rates do not depend on the number of hours worked in that period, ruling out part-time penalties. This assumption, for women, is consistent with what we observe in our data and with other U.S.-based studies (Hirsch (2005), Aaronson and French (2004)).

In our baseline specification, human capital does not depend on the history of labor supply and is assumed to evolve exogenously according to

$$\ln e^f_t = \iota^f_1 t + \iota^f_2 t^2. \tag{8}$$

Equation (8) implies that decisions on current labor supply do not have a direct effect on continuation values.[4] Therefore, the only linkage across periods is through the decision

[3]In reality, there is, of course, some degree of uncertainty in the realization of households' fertility decisions. We do not consider fertility as a stochastic outcome, as that would increase the numerical complexity of the problem substantively.

[4]In Appendix F of the Supplemental Material, we relax the assumption that there are no returns to experience. We distinguish the cases where returns to experience depend on participation and where returns depend

about total within-period spending. This assumption, combined with the intertemporally additive structure of preferences, implies that standard two-stage budgeting holds so that we can focus on the within-period problem without considering explicitly the intertemporal allocation.

Men always work and their earnings are given by

$$\ln y_{h,t}^{m} = \ln y_{h,0}^{m} + \iota_1^m t + \iota_2^m t^2 + v_{h,t}^{m}. \tag{9}$$

There are initial distributions of wages for women, $w_{h,0}^{f}$, and earnings for men, $y_{h,0}^{m}$. Both women's wages and men's earnings are subject to permanent shocks that are positively correlated, as in MaCurdy (1983) and Abowd and Card (1989):

$$v_{h,t} = v_{h,t-1} + \xi_{h,t}, \tag{10}$$

$$\xi_{h,t} = \left(\xi_{h,t}^{f}, \xi_{h,t}^{m}\right) \sim N\left(\mu_\xi, \sigma_\xi^2\right), \tag{11}$$

$$\mu_\xi = \left(-\frac{\sigma_{\xi f}^2}{2}, -\frac{\sigma_{\xi m}^2}{2}\right) \quad \text{and} \quad \sigma_\xi^2 = \begin{pmatrix} \sigma_{\xi f}^2 & \rho_{\xi f,\xi m} \\ \rho_{\xi f,\xi m} & \sigma_{\xi m}^2 \end{pmatrix}.$$

One period in the model is one quarter. Households choose typical hours of work each week (the *intensive margin*) and this is kept constant across weeks within the quarter, to give within-period hours of work. The *extensive margin* is the decision whether or not to work that quarter. We do not allow individuals to choose how many weeks to work in a quarter.[5] We provide empirical support for this approach in Section 4.2.

Within the dynamic problem just described, households make decisions taking the stochastic processes as given. When considering aggregation, we need to take a stand on the degree of correlations in the shocks different households receive. We assume that households are subject to both idiosyncratic and aggregate shocks, by allowing the shocks that affect individual households at a point in time to be correlated. However, from an individual perspective, households do not distinguish aggregate and idiosyncratic shocks and condition their future expectations only on their own observed wage realizations. Our framework is not a general equilibrium one: we do not construct the equilibrium level of wages (and interest rates). Rather, we study women's aggregate labor supply and its elasticity to wages by simulating a large number of households and aggregating explicitly their behavior.

2.1. *Marginal Rate of Substitution, Marshallian and Hicksian Elasticities*

We use a two-stage budgeting approach and consider the allocation of resources between consumption and hours of leisure within each period. We define within-period resources that are not earned by women as

$$y_{h,t} = \left(A_{h,t} + y_{h,t}^{m} - F(a_{h,t})P_{h,t}\right) - \frac{A_{h,t+1}}{1 + r_{t+1}}. \tag{12}$$

on hours worked. The first two steps of our estimation approach go through in the former case but not in the latter.

[5]This restriction is driven by data limitations. In our data, we observe typical hours per week, whether an individual is working at that point in time, and the number of weeks per year, but we do not observe the number of weeks per quarter that an individual works. We also cannot distinguish the number of days per week, from the number of hours per day, as in Castex and Dechter (2016).

As in Blundell and MaCurdy (1999), $y_{h,t}$ accounts for resources saved into the next period. When taken to the data, this measure of unearned resources implicitly also includes (with a negative sign) durable and other spending not included in consumption c_t, giving the within-period budget constraint:

$$c_{h,t} + w^f_{h,t} l_{h,t} = y_{h,t} + w^f_{h,t} L. \tag{13}$$

For an interior solution with a strictly positive number of hours of work, the first-order condition for within-period optimality implies that the ratio of the marginal utility of leisure to that of consumption, that is, the Marginal Rate of Substitution, equals the after-tax real wage:

$$w^f_{h,t} = \frac{u_{l_{h,t}}}{u_{c_{h,t}}} = \alpha_{h,t} \frac{l^{-\theta}_{h,t}}{c^{-\phi}_{h,t}}. \tag{14}$$

These equations can be used to compute Marshallian and Hicksian labor supply elasticities. The Marshallian and Hicksian elasticities are fundamentally static concepts, as both hold constant the intertemporal allocation of resources.[6] The Marshallian response captures the change in behavior due to a change in the price of leisure and the related change in resources available to spend. This latter income effect arises even if the intertemporal allocation of resources $y_{h,t}$ is held constant, because total resources within the period change with the wage.

In the full dynamic model, when the realized wage is permanently higher than expected, lifetime resources increase, and these extra resources are allocated across periods. The static Marshallian elasticity is a good approximation to the full response if extra resources are spent on non-durable consumption in the period they are earned. To the extent that resources are reallocated, the static Marshallian elasticity only captures part of the labor supply response. If within-period spending is homothetic, and wages have gone up by the same amount in every period, then there may be little change in saving patterns following the wage increase. In this case, the Marshallian elasticity gives a good approximation of the complete life-cycle response. On the other hand, if all extra income from the wage increase is saved to spend in retirement, then there would be no within-period income effect and the response will be closer to a Hicksian compensated response. More generally, how well the static Hicksian and Marshallian elasticities approximate the complete life-cycle responses to compensated and uncompensated wage changes is an open question. In Section 6, we use the full structural model to evaluate how closely the static elasticities approximate the full life-cycle ones.

We differentiate the within-period budget constraint (13) and the MRS equation (14) with respect to wages to get an expression for Marshallian elasticities for hours of work and consumption (omitting subscripts and superscripts for convenience, see Appendix A for details on the derivations):

$$\begin{aligned} \varepsilon^M_h &= \frac{\partial \ln h}{\partial \ln w} = -\left(\frac{\phi w(L-l) - c}{\theta c + \phi w l}\right) \frac{l}{L-l}, \\ \varepsilon^M_c &= \frac{\partial \ln c}{\partial \ln w} = \frac{\theta w(L-l) + wl}{\theta c + \phi w l}. \end{aligned} \tag{15}$$

[6] Blundell and MaCurdy (1999) and Keane (2011) discussed how the static concepts of Marshallian and Hicksian elasticities can be put within the framework of a dynamic life-cycle model through two-stage budgeting, as developed by Gorman (1959) and applied to labor supply by MaCurdy (1981, 1983) and Blundell and Walker (1986).

If preferences were Cobb–Douglas, θ and ϕ would both equal 1; and the Marshallian wage elasticities for consumption and for hours of work would be equal to 1 and 0, respectively, if there were no unearned income or savings. For balanced growth (in women's labor supply), we would require $\phi = 1$. If preferences were a standard CES, $\theta = \phi$. If this value were greater than 1, $\varepsilon_c^M < 1$, and $\varepsilon_h^M < 0$. In Section 6, we show how much heterogeneity is introduced through our more general specification in equations (15) and through allowing for unearned income.

The static Hicksian response nets off the increase in within-period resources due to the wage increase, again holding constant the intertemporal allocation, $y_{h,t}$. We calculate the Hicksian response from the Marshallian elasticities by using the Slustky equation and income elasticities, as would be done in a static labor supply model:

$$\begin{aligned} \varepsilon_h^H &= \left(\varepsilon_l^M - \frac{\partial \ln l}{\partial \ln(c+wl)} \frac{w(L-l)}{(c+wl)}\right)\frac{-l}{L-l} = \frac{-wl^2}{(\theta c + \phi wl)(L-l)}, \\ \varepsilon_c^H &= \varepsilon_c^M + \frac{\partial \ln c}{\partial \ln(c+wl)} \frac{wl}{(c+wl)} = \frac{-c}{\theta c + \phi wl}. \end{aligned} \tag{16}$$

The Marshallian and Hicksian elasticities are the relevant concepts to think about the labor supply responses to permanent changes in wages or taxes. However, as we discuss in Section 6, estimates based on the within-period problem might miss potential intertemporal reallocations that might occur in response to wage changes.

Two additional points are worth noting. First, despite their simplicity, the Marshallian and Hicksian elasticities are nonlinear in c and l: they have the potential of varying greatly across consumers and not aggregating in a straightforward way. Second, for the specification we use, the Marshallian and Hicksian elasticities depend only on ϕ and θ (and on the values of earnings, leisure, and consumption). In particular, they do not depend on intertemporal parameters or on whether the utility function is separable in consumption and leisure, which depends on γ.

2.2. *Frisch Elasticities*

A change in the structure of wages (possibly induced by changes in taxes) may induce a reallocation of resources over time through changes to the time path of hours of work or of the marginal utility of wealth, or both. The Frisch elasticity captures the change over time in hours worked in response to the anticipated evolution of wages, with the marginal utility of wealth unchanged, as the wage change conveys no new information.[7] The Frisch elasticity is therefore the right concept to think about the implications of changes in wages over the business cycle or about temporary changes to taxation.

The expression for the Frisch elasticity for hours of work, derived in Appendix A, is given by[8]

$$\varepsilon_h^F = -\frac{u_c u_{cc}}{u_{cc}u_{ll} - u_{cl}^2}\frac{w}{h}. \tag{17}$$

[7]When wages change stochastically, the response of hours worked is affected by the change in the marginal utility of wealth due to a particular wage realization, whose size depends on how permanent the wage shock is. If the wage shock is temporary, lifetime wealth and the marginal utility of wealth will be approximately unchanged.

[8]Analogous expressions for the consumption Frisch wage elasticities, as well as the interest rate elasticities, can be found in Appendix A.

As is well known, Frisch intertemporal elasticities must be at least as large as Hicks elasticities. Thus, the static elasticities discussed above provide a bound on the intertemporal elasticity, which is particularly useful if data are limited or direct estimation of Frisch elasticities difficult.[9]

In addition to changes in hours, anticipated changes in wages might also change participation. While an elasticity is easily defined when thinking of the intensive margin, the same concept is somewhat vaguer at the extensive margin, especially in the case of the Frisch elasticity, which keeps the marginal utility of wealth constant. We define the extensive-margin Frisch elasticity as the impact of a change in wages on the fraction of individuals that participate, given the distribution of state variables. The extensive margin brings to the forefront aggregation issues: aggregate participation responses to an aggregate shock are bound to depend on the distribution of state variables in the cross section.

3. EMPIRICAL STRATEGY

In this section, we discuss our empirical approach, identification assumptions, and the variability we use in the data. We proceed in three steps, with each successive step identifying a set of structural parameters. In the first step, we consider only the static first-order (MRS) condition that determines within-period optimal allocations, conditional on participation. This first set of parameters can be identified while being agnostic about intertemporal conditions and on life-cycle prospects. In the second step, we identify the parameters that govern the intertemporal allocation of resources using the Euler equation for consumption, making use of additional assumptions. However, we can still identify these parameters without specifying the entire life-cycle environment faced by households. For instance, we can be silent about pension arrangements or the specifics of the wage and earning processes. When estimating the parameters that determine the MRS or those that enter the Euler equation, we use an estimator proposed by Fuller (1977). This choice of estimator turns out to matter for the results we obtain and has advantages over standard methods, as we discuss in Appendix B of the Supplemental Material. Finally, in the third step, we characterize behavior at the extensive margin. This step requires solving the entire model and, therefore, specifying completely the environment in which households operate. We identify the final set of parameters by calibration, matching a set of life-cycle statistics.

3.1. *Intratemporal Margins*

In the first step, we estimate the parameters of the within-period utility function: θ, ϕ, and α. Taking logs of the MRS equation (14), and noticing from equation (4) that $\log\alpha_{h,t}=\psi_0+\psi_z z_{h,t}+\chi_{h,t}$, we obtain

$$\ln w^f_{h,t}=\phi\ln c_{h,t}-\theta\ln l_{h,t}+\psi_z z_{h,t}+\psi_0+\chi_{h,t}, \tag{18}$$

where the vector $z_{h,t}$ includes observable demographic variables.

The econometric estimation of this MRS equation poses two problems. First, the subset of households in which the woman works and the MRS condition holds as an equality is not random. For this selected group, the unobserved heterogeneity term $\chi_{h,t}$ would not

[9]In the context of quasi-linear utility as used by Chetty (2012), the Frisch elasticity equals the Hicks elasticity (and the Marshallian) because there are no wealth effects on hours of work.

average out to zero and might be correlated with the variables that enter equation (18). Second, even in the absence of participation issues, individual wages (and consumption and leisure) are likely to correlate with $\chi_{h,t}$, so that the OLS estimation of equation (18) would result in biased estimates of the structural parameters ϕ and θ. We discuss these two issues in turn.

For participation, we specify a reduced form equation for the extensive margin. Given this participation equation, we use a Heckman-type selection correction approach to estimate the MRS equation (18) only on the households in which the woman works. In particular, we augment the MRS equation with a polynomial in the estimated residuals of the participation equation.[10] Identification requires that some variables that enter the participation equation do not enter the MRS specification: these variables are men's earnings and employment status, and we assume these are uncorrelated with $\chi_{h,t}$.

The fully-specified participation decision depends on a large set of state variables, some of which are not observable. In our 'reduced form,' participation depends only on a subset of these variables. Therefore, our reduced form participation equation is not fully consistent with the complete model we use to characterize participation and, at best, could be considered an approximation of the 'true' participation equation. In Appendix G of the Supplemental Material, we investigate how well this approximation to the full model performs: we estimate MRS parameters using our reduced form empirical strategy on simulated data from the full model. We are able to recover the true parameter estimates and our conclusion is that our reduced form provides an accurate approximation in this context.

The second issue in the estimation of equation (18) is that consumption and hours, as well as our measures of individual wages, obtained dividing earnings by hours, might be correlated with the residual term $\chi_{h,t}$, either because of the possible correlation between tastes for leisure and heterogeneity in productivity or because of measurement error in hours or earnings. To avoid these problems, following the literature on labor supply (such as Blundell, Duncan, and Meghir (1998)), we do not use variation in individual wages to identify the parameters of our equation. Instead, we exploit variation induced by changes in taxation and/or aggregate demand for labor and use changes in cohort-education groups' average wages over time.[11] The Monte Carlo evidence on our MRS estimation in Appendix G shows that both this endogeneity issue and the selection issue have to be taken into account in our context to obtain sensible estimates.

[10]One issue to worry about is the intrinsic nonlinearity of the participation equation. The omission of some state variables could change the properties of the residuals of such a nonlinear equation and, therefore, the shape of the appropriate control function to enter the MRS equation. For this reason, we use a polynomial to model the dependence between the residuals of the participation equation and those of the MRS equation. We assume that $\chi_{h,t} = \beta_0 + \beta_1 e_{h,t} + \beta_2 e_{h,t}^2 + \beta_3 e_{h,t}^3$ and then compute $E[e_{h,t}^s | e_{h,t} > -\Pi Z_{h,t}]$, $s = 1, 2, 3$, where $e_{h,t}$ is the normally distributed residual from the participation equation and $Z_{h,t}$ are the determinants of participation.

[11]Various papers have used variation across education groups; for example MaCurdy (1983) and Ziliak and Kniesner (1999) both used age-education interactions as instruments for wages and hours in their MRS/labor supply conditions. Similarly, Kimmel and Kniesner (1998) used education interacted with a quadratic time trend. One concern with this approach is that individuals with different levels of education might have different preferences for leisure and consumption. Moreover, the composition of education groups has changed substantially over time, particularly for women. In 1980, 19.4% of married women had not attained a high school diploma, and only 18.4% had obtained a college degree in our data. By 2012, these proportions had changed to 9.7% and 36.5%, respectively. These compositional changes may lead to changes in the mix of ability and preferences of workers within each education group over time—making education an invalid instrument.

We use as instruments the interaction of ten-year birth cohort and education dummies with a quintic time trend. Our use of a quintic time trend rather than fully interacted time dummies helps smooth intertemporal movements in wages, consumption, and hours for each of our cohort-education groups.[12]

In our estimating equation, we allow many variables to shift the taste for leisure through an effect on the term $\alpha_{h,t}$ in the CES utility function. The z vector includes: log family size, woman's race, a quartic in woman's age, an indicator for the presence of any child, the number of children aged 0–2, 3–15, and 16–17, the number of individuals in the household 65 or older, region and season dummies, and, most importantly, cohort-education dummies. A corollary of putting variables such as cohort and education dummies in the vector z is that we do not exploit the variation in wages (and leisure and consumption) over these dimensions to identify the structural parameters ϕ and θ. In our estimation, we also control for year dummies, therefore removing year-to-year fluctuations from the variability we use to identify the parameters of interest. The inclusion of year dummies, as in Blundell, Duncan, and Meghir (1998), is needed because aggregate fluctuations change the selection rule year to year in ways that are not fully captured by the selection model we use.[13]

3.2. *Euler Equation Estimation*

The second step of our approach estimates the preference parameters that govern the intertemporal substitutability and non-separability between consumption and leisure, γ, and the non-separability with participation, φ. While in principle we could use either the Euler equation for hours or that for consumption, only one is relevant when coupled with the intratemporal condition (14). If we were to use the Euler equation for labor supply, we would need to consider corner solutions at different points in time (and the dynamic selection problems these involve). Instead, we focus on the Euler equation for consumption, as in Blundell, Meghir, and Neves (1993). In the absence of binding borrowing constraints, the following intertemporal condition holds:

$$E\left[\beta(1+r_{t+1})\frac{u_{c_{h,t+1}}(\cdot)}{u_{c_{h,t}}(\cdot)}\bigg|I_{h,t}\right]=1. \tag{19}$$

The term $I_{h,t}$ denotes the information available to the household at time t.

A natural approach to the estimation of equation (19) is nonlinear GMM. However, as discussed in Attanasio and Low (2004), the small sample properties of nonlinear GMM estimators can be poor in contexts similar to ours. Moreover, given the specification of the utility function and nature of the data, we can only estimate its log-linearized version.

The evolution of the marginal utility of consumption can then be written as

$$\beta(1+r_{t+1})u_{c_{h,t+1}}(\cdot)=u_{c_{h,t}}(\cdot)\epsilon_{h,t+1}, \tag{20}$$

where $\epsilon_{h,t+1}$, whose conditional expectation is 1, is the innovation to the expected discounted marginal utility of consumption. Equation (20) uses the variability in r_t to identify

[12]Using fully interacted cohort-education and year dummies would be equivalent to taking averages within cells defined by year, education, and cohort groups, to use group level rather than individual level variability. Given our sample size, this would result in averages over relatively small cells and, therefore, in very noisy estimates. Using very finely defined and small groups can introduce the very biases grouping is meant to avoid.

[13]We have also run specifications where we do not control for time dummies in the MRS and checked that our results are not affected much by the introduction of the time dummies.

the parameters of $u_c(\cdot)$. Taking the log of equation (20), given utility is given by equation (2):

$$\eta_{h,t+1} = \kappa_{h,t} + \ln \beta + \ln(1 + r_{t+1}) - \phi \Delta \ln c_{h,t+1} - \gamma \Delta \ln(M_{h,t+1}) + \varphi \Delta P_{h,t+1} + \pi \Delta z_{h,t+1}, \quad (21)$$

where $\eta_{h,t+1} \equiv \ln \epsilon_{h,t+1} - E[\ln \epsilon_{h,t+1} | I_{h,t}] + \Delta\zeta_{h,t+1}$ and $\kappa_{h,t} \equiv E[\ln \epsilon_{h,t+1} | I_{h,t}]$.

The identification and estimation of the parameters of this equation depend, obviously, on the nature of the 'residual' term $\eta_{h,t+1}$, which contains expectations errors ($\epsilon_{h,t+1}$), higher order moments, and taste shifters unobservable to the econometrician ($\zeta_{h,t+1}$). Aggregate shocks mean expectation errors may be correlated in the cross section, and average to zero only in the time dimension. Consistency then requires time series variation, as discussed in Attanasio and Low (2004). We construct a long time dimension using a synthetic cohort approach (see Browning, Deaton, and Irish (1985)), defining groups using married couples in ten-year birth cohorts. We assume that the lagged variables used as instruments are uncorrelated with the innovations to the taste shifters $\Delta\zeta_{h,t+1}$. This is trivially true if taste shifters are constant over time or if they are random walks. We maintain one of these two assumptions, a hypothesis that we test in part by checking overidentifying restrictions.

We aggregate equation (21) to be estimated across group g households. For this approach to work, it is necessary that the equation to be estimated is linear in parameters, which would be the case if $M_{h,t}$ were observable. However, $M_{h,t}$ is a nonlinear function of data and unobserved parameters, so that, in principle, it cannot be aggregated within groups to obtain $M_{g,t}$. On the other hand, the parameters that determine $M_{h,t}$ can be consistently estimated using the MRS conditions as discussed in Section 3.1.[14] These estimates can be used to construct consistent estimates of $M_{h,t}$, which can be aggregated across households to give $M_{g,t}$. This gives an equation analogous to equation (21), but where variables are group averages and where all variables on the right-hand side are now observable. We use this procedure to recover the intertemporal preference parameter γ and the participation preference parameter φ. We cannot identify any additional effect of participation that is separable in the utility function. Nor, at this stage, do we know the fixed costs of work and so we cannot identify the extensive margin response to wage changes.

Using group averages on repeated cross sections introduces a number of other econometric problems, linked to the presence of estimation errors in small samples. These issues, as discussed in Deaton (1985), have implications for the choice of instruments and computation of standard errors. Further details of this procedure are discussed in Appendix B of the Supplemental Material.

In principle, the first two steps of our estimation could be followed without making parametric assumptions about the utility function and, instead, estimating leisure and consumption demands directly. However, such an approach would require that the demand functions satisfy integrability conditions. Furthermore, the actual underlying utility function would still need to be recovered to study participation and the extensive margins.

3.3. *Extensive Margins*

The last step of our approach obtains estimates of the remaining model parameters, including the fixed costs of work and child care costs, which drive the extensive margin decision. When considering the extensive margin, it is necessary to solve explicitly the whole

[14] $M_{h,t}$ includes $\chi_{h,t}$ which is unobserved. However, since it is the residual from the MRS equation, it can be included in the calculation of $\alpha_{h,t}$ that is needed to calculate $M_{h,t}$.

dynamic problem. This involves making assumptions on the entire economic environment faced by households over the life-cycle, including both present and future conditions. We solve the model numerically and use the solution to estimate and calibrate the model parameters. To reduce the numerical burden, when simulating the model, we assume a fixed interest rate. As the MRS conditions do not change, this assumption will not change within-period elasticities, but the life-cycle solution of the model and life-cycle elasticities will be affected to the extent that uncertainty about interest rates affects saving. We provide the value functions of the household's problem and details about the numerical solution in Appendix B.

We take as given the estimates of the parameters we obtained from the MRS and the Euler equation, and obtain some parameters from the literature and from direct regressions. We estimate the remaining parameters so that data generated from simulations match key life-cycle aspects of the extensive margin: the participation rate, the participation rate of mothers, and average wage growth of participants (which is endogenous because of selection). Finally, we simulate the model for a large number of individuals to study the properties of individual and 'aggregate' labor supply. We then assess the model's goodness of fit by exploring the life-cycle profiles of several variables as well as participation rates conditioning on individual characteristics and the distribution of hours worked and wages.

4. DATA AND DESCRIPTIVE STATISTICS

We take our data from the Consumer Expenditure Survey (CEX) for the years 1980–2012. In the CEX, households are interviewed up to four times, answering detailed recall questions on expenditures as well as on the demographics, incomes, and labor supply of household members.

We calculate gross hourly wages for individuals using information on the value of each individual's last paycheck, the number of weeks it covered, and the typical number of hours worked per week. Net wages are then calculated by subtracting marginal federal income tax rates generated using the NBER TAXSIM model (Feenberg and Coutts (1993)).[15] We deflate all expenditures, wages, and incomes using the Consumer Price Index. Weekly leisure is calculated by subtracting weekly hours worked from the maximum number an individual has to divide between leisure and labor supply per week (which we set to 100). Participation is defined by employment status at the time of the interview. Consumption covers non-durable goods excluding medical and education spending. We divide quarterly consumption spending by 13 to put it in weekly terms.

Our sample consists of couples with women aged between 25 and 60 and men aged between 25 and 65. We drop those in rural areas; those in the top 1% of the consumption and net wages distribution; those earning less than three-quarters of the national minimum wage in any given year; and those who are employed but who report working less than 5 hours a week. Since labor supply and income questions are (almost always) only asked in the first and last interviews, we drop responses from interviews apart from these two. Our sampling choices leave just under 79,000 households (50,895 where the woman is working). Appendix C of the Supplemental Material presents descriptive statistics on individual characteristics over time.

[15] We are grateful to Lorenz Kueng for making his mapping of the CEX to TAXSIM publically available.

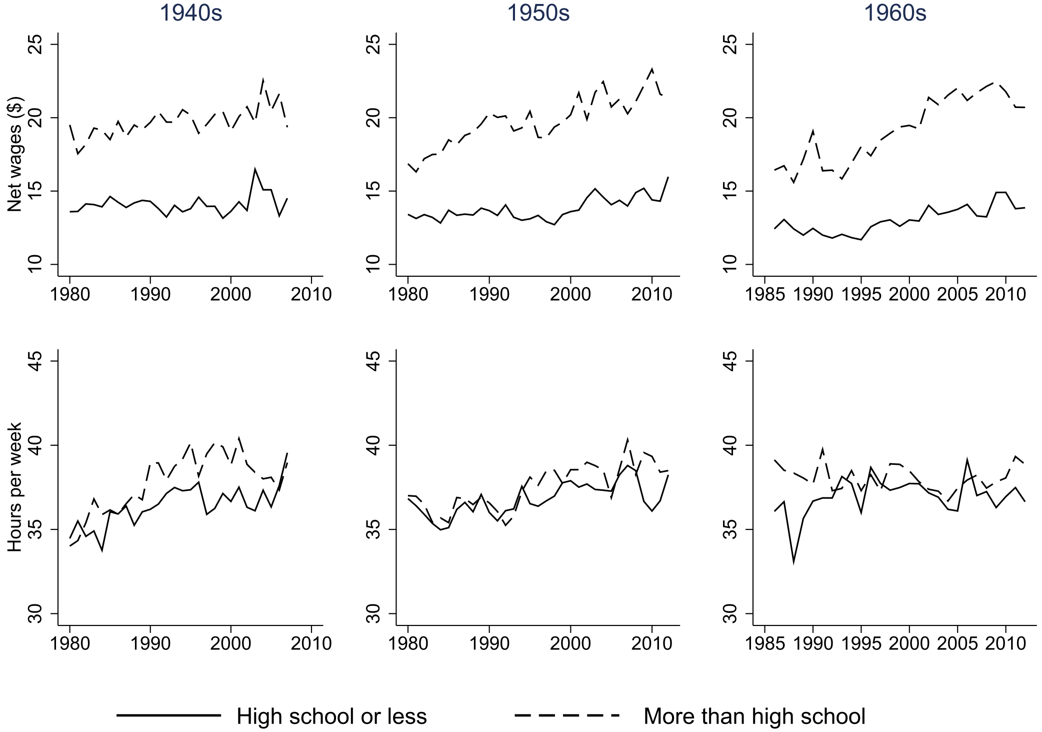


FIGURE 1.—Wages and hours by education group and cohort.

4.1. *Cohort Averages*

We separate households into birth cohorts and examine the evolution of wages and hours by education within each cohort group. In Figure 1, we report patterns for the cohorts born in the 1940s, 1950s, and 1960s, and for women with high school or less and with more than high school.[16]

Within the 1950s cohort, the net wages of those with more than high school education increased from an average of \$16.90 per hour in 1980 to \$21.40 in 2012 (an increase of 27%), while the wages of those with less than high school education only increased by 19% from \$13.40 to \$16.00. Despite this, the bottom row of Figure 1 shows average weekly hours of less educated worked actually increased by more than those from the more educated group (increasing from 36.8 hours per week to 38.2 compared to an increase from 37.4 to 38.5 for those with more than high school education).

4.2. *Individual Variation in Hours and Wages*

In addition to changes in average hours and wages over our sample period, there are two important issues at the individual level: what is the relative importance of the intensive and extensive margins in the raw data and what fraction of individuals are experiencing changes in hours or wages over time.

[16]The advantage of considering the variability over time of a given cohort is that composition is unlikely to change, as it is rare for workers to increase their educational qualifications after age 25.

TABLE I

VARIANCES OF LABOR SUPPLY MEASURES, 2012

	Less Than High School	High School	Some College	Degree or Higher	All
All workers					
Variance (ln hours per week)	0.148	0.117	0.128	0.126	0.126
Variance (ln weeks per year)	0.550	0.271	0.231	0.482	0.367
Covariance (ln hours, ln weeks)	0.031	0.046	0.010	0.028	0.027
Variance (ln annual hours)	0.761	0.479	0.380	0.665	0.546
Working at least 39 weeks (84% of workers)					
Variance (ln hours per week)	0.061	0.040	0.086	0.110	0.086
Variance (ln weeks per year)	0.001	0.003	0.003	0.005	0.004
Covariance (ln hours, ln weeks)	−0.001	0.001	0.002	0.000	0.001
Variance (ln annual hours)	0.059	0.045	0.094	0.115	0.092
Working 52 weeks (69% of workers)					
Variance (ln hours per week)	0.064	0.031	0.068	0.117	0.080

The individual extensive margin decision is whether to incur a fixed cost $F(a_{h,t})$ and participate in the current quarter. We measure this by the stated current employment status. The intensive margin decision is over how many hours to work per week (when working). An additional labor supply response may be through changing weeks worked per quarter. However, we are not able to estimate this margin of adjustment because the CEX asks current workers about the number of weeks they worked over the previous year rather than the previous quarter.

Whether ignoring the margin of the number of weeks worked within a quarter matters, depends on how much of the variance of workers' quarterly hours is driven by differences in weeks worked within a quarter rather than hours per week. Table I decomposes the variance of log annual hours into variation in log annual weeks, variation in log workers' typical weekly hours, and their covariance. The first panel shows this breakdown for the entire sample of workers. The variance in annual weeks worked is around two thirds of the total variance in hours worked. Much of this is likely to be workers not participating for entire quarters: our extensive margin. In the second panel, we restrict the sample to workers who work for more than 39 weeks (and thus could not have been unemployed for a complete quarter). These workers account for 84% of the total, and for them, almost all of the variance in annual hours is a result of differences in hours worked per week, with differences in weeks worked making a negligible contribution. In the third panel, we restrict our sample further to those working exactly 52 weeks per year and notice that even among workers who do not differ in the number of weeks worked, the variance in log hours per week remains substantial (at 0.08).

These results suggest that hours worked per week is the key margin by which workers adjust their quarterly hours. We thus use this measure when estimating our MRS and Euler equations. In Appendix E of the Supplemental Material, we check the robustness of this strategy by showing that our estimates and results are little affected by replacing our current measure with a measure of annual hours worked.

A further question is whether individual workers are able to adjust their weekly hours in response to wage changes, or whether there are market frictions that prevent this.

TABLE II

CHANGES IN WEEKLY HOURS AMONG THE EMPLOYED[a]

Change in Weekly Hours	No Change	1–5 Hrs	6–10 Hrs	11–20 Hrs	>20 Hrs
All workers	53.8%	25.2%	11.9%	6.9%	2.2%
Extent of change in wages:					
<5% wage change	74.9%	17.5%	4.7%	2.3%	0.71%
>5% wage change	47.5%	27.5%	14.0%	8.2%	2.7%

[a] Changes in hours are measured between the second and fifth interviews for individuals who are employed at each interview.

Table II shows the proportion of workers who changed their typical hours from the first to the last CEX interview (a period of nine months). While it is true that most women do not change their hours within this period, a substantial fraction (46%) do. Around a quarter of workers change their weekly hours by 1–5 hours, and 2% change their hours by more than 20 hours.

5. RESULTS: PARAMETER ESTIMATES AND CALIBRATION

In this section, we report estimates of the structural parameters of our model. In Sections 5.1 and 5.2, we report the estimation results obtained using the MRS conditions and the Euler equation. In Section 5.3, we report the calibration of the remaining parameters that govern choices at the extensive margin. In the last subsection, we show how well the complete model fits a number of features of the data that were not used explicitly to obtain the parameter estimates.

5.1. *MRS Estimates*

In Table III, we report the estimates of key parameters for the MRS equation and tests on the quality of our instruments, with results for the participation model reported in Appendix D of the Supplemental Material. We estimate values for θ and ϕ at 1.75 and 0.76, respectively: there is much more curvature in utility on leisure than on consumption. We test the restrictions implied by Cobb–Douglas and standard CES specifications using a wild-cluster residual bootstrap. The Cobb–Douglas specification for preferences, $\phi = \theta = 1$, is rejected at the 5% level (p-value 0.01), while the standard CES specification, $\phi = \theta$, is rejected with a p-value of 0.06.

Table III also reports the coefficients, ψ, on demographic variables in $z_{h,t}$. A larger (positive) value for ψ means, other things equal, a higher marginal utility of leisure and so women will supply fewer hours of work in the market. The positive and significant coefficient on the dummy for having children indicates that the presence of children tends to reduce hours worked, but the effect of children depends on their age. The coefficient on the number of children aged 0–2 is positive and highly significant, on children aged 3–15 the coefficient is positive, but smaller; for older children, the coefficient is negative.

We include three Heckman selection terms corresponding to the first, second, and third moments of the truncated normal distribution (as described in footnote 10). We test the joint significance of these in both our first- and second-stage regressions. These terms are highly significant in each of the first stages, where we are predicting individual consumption, hours, and wages. On the other hand, the selection terms are insignificant in the second stage of the MRS. The Cragg–Donald statistic for weak instruments in our MRS

TABLE III

ESTIMATION OF MRS EQUATION[a]

Parameter	Estimate	(Standard Error)	[95% Confidence Interval]
θ	1.75**	(1.230)	[0.34, 5.12]
ϕ	0.76***	(0.103)	[0.55, 0.95]
Ψ			
ln(*famsize*)	−0.32***	(0.037)	[−0.38, −0.23]
Has kids	0.07***	(0.021)	[0.04, 0.10]
No. of kids 0–2	0.15***	(0.030)	[0.10, 0.22]
No. of kids 3–15	0.06***	(0.017)	[0.04, 0.10]
No. of kids 16–17	−0.02**	(0.011)	[−0.05, 0.00]
Joint tests of selection terms (p-value)			
First stage: ln wage		166.47 (<0.001)	
First stage: ln consumption		311.75 (<0.001)	
First stage: ln leisure		40.83 (<0.001)	
Main equation		0.72 (0.87)	
Cragg-Donald statistic		2.00	
Sargan statistic (p-value)		127.8 (0.66)	

[a] $N = 50{,}895$. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Additional controls: number of elderly (aged over 65) in the household, a quadratic in age, race, region, season, cohort-education interactions, and year dummies. Consumption and leisure are instrumented with the interaction of cohort and education groups and a fifth-order polynomial time trend. Confidence intervals are bootstrapped with 1000 replications allowing for clustering at the individual level.

equation takes a value of 2.00 for 138 instruments, well above the relevant Stock and Yogo (2005) critical level of 1.69, and therefore suggesting that weak instruments are not a problem.[17] The Sargan test does not reject the null of no violation of the over-identifying restrictions.

5.2. *Euler Equation Estimates*

Table IV reports estimates of the Euler equation (21) using group averages. We estimate γ to be 2.07, significantly different from zero at the 10% level, providing evidence that preferences are nonseparable and that consumption and leisure are substitutes ($\gamma = 0$ would imply additively separable preferences over consumption and leisure). Since ϕ, θ, and γ are all positive, the concavity requirements of the utility function are satisfied. The coefficients on the control variables included in the vector z_t are not significant, implying demographics have no role over and above their impact on the relative weight on leisure within-period. The specification in Table IV imposes that φ, the parameter on participation in equation (2), is zero. When we include this term (instrumented with its own lags), the coefficient estimate is negative but not significantly different from zero.

Our instruments are second, third, and fourth lags of $\ln M_{g,t}$ and the lagged real interest rate (defined as the 3-month Treasury Bill rate minus the inflation rate), and we have two endogenous variables $\phi(\Delta \ln c_{g,t} + \ln(1 + r_{t+1}))$ and $\Delta \ln M_{g,t}$. We place the second of these on the left-hand side of the equation. With only one left-hand-side endogenous variable, the Cragg–Donald test for weak instruments is equivalent to a standard F-test of the

[17] The value of 1.69 is given for two endogenous variables and 100 instruments, and given that the critical values for a maximum 5% relative bias for the Fuller estimator are decreasing in the number of instruments, the use of this test statistic is conservative.

TABLE IV

ESTIMATION OF EULER EQUATION[a]

Parameter	Estimate	(Standard Error)	[95% Confidence Interval]
γ	2.07*	(0.656)	[−0.11, 2.60]
$\bar{\kappa} + \ln(\beta)$	0.03	(0.040)	[−0.08, 0.10]
π			
ln(*famsize*)	−0.47	(0.244)	[−0.69, 0.31]
Has kids	0.05	(0.069)	[−0.09, 0.19]
No. of kids aged 0–2	0.22	(0.099)	[−0.05, 0.35]
No. of kids aged 3–15	0.03	(0.038)	[−0.06, 0.09]
No. of kids aged 16–17	0.03	(0.071)	[−0.11, 0.18]
First-stage F-stats (p-values)			
$-\phi(\Delta \ln c_{g,t} + \ln(1 + r_{t+1})$		7.95 (<0.001)	
$\Delta \ln M_{g,t}$		2.08 (0.08)	
Sargan statistic (p-value)		5.70 (0.13)	

[a] $N = 1519$. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Additional controls: season dummies, a quartic in age, the change in the proportion of households in each of four education groups, the change in proportion who are white, and the change in the average number of elderly individuals per household. Instruments are second, third, and fourth lags of $\ln M_{g,t}$, as well as the lagged real interest rate. Confidence intervals are bootstrapped with 1000 replications.

instruments' joint significance in the first-stage regression. The critical values of these F-tests suggest that the instruments are highly correlated with the right-hand side variable (with an F-statistic of 7.95), but less strongly correlated with our choice of left-hand-side variable (with an F-statistic of 2.08). The relevant Stock and Yogo test statistic for having less than a 5% relative bias in our parameter estimates when there are four instruments and one left-hand-side endogenous variable is 7.63. When we carry out a Sargan test for the Euler equation, we fail to reject the null of over-identification (p-value 0.13) as we do for the MRS.

5.3. *Calibration of the Remaining Parameters*

There are three sets of parameters used in the calibration of the full model: those estimated via the MRS conditions and the Euler equation, those coming from external sources, and those that we calibrate using simulations of the full model.

We focus on the cohort of women born in the 1950s, using moments from women age 25–55. We assume that χ and ζ are homogeneous within a cohort. Attanasio, Low, and Sánchez-Marcos (2008) showed that women's labor supply behavior differs substantially across cohorts. The main cause in that paper is differences in costs of child care, but there are also differences in wage processes across cohorts. These differences will lead to different responses across cohorts on the extensive margin and could also lead to differences in the intensive margin because of different levels of consumption and leisure.

Within the 1950s cohort, we assume there are nine different groups of women: one group of women who remain childless for the whole of their lifetime, and eight groups of women who differ by maternity experience. These women exogenously receive two kids but differ in the age at which the first child arrives. To determine when these children are born, we draw on Rendall et al. (2010), who used population and survey data sources to calculate the distribution of maternity age at arrival of the first child for different cohorts of women in various countries.[18] We assume that the second child arrives 2 years after the first.

External Parameters

The complete set of external parameters is reported in Table XIX in Appendix D of the Supplemental Material. We fix the annualized interest rate to equal the average real return on 3-month T-bill at 0.015. The deterministic component of the male earnings process is estimated from the CEX: we take the two parameters of a regression of husband log earnings on age and age squared. The standard deviation of the innovation for husband's earnings, σ_{ξ^m}, is set to be 0.077, consistent with Huggett, Ventura, and Yaron (2011). Further, we estimate an initial standard deviation of husband earnings $\sigma_{\xi_0^m}$ of 0.54. There is limited evidence on the variability of women's wages and/or earnings, and further, since this statistic is highly affected by non-random self-selection into the labor market, we calibrate the parameters that characterize the women's wage process within the model as explained below. Finally, we assume that the correlation coefficient between the two shocks (for husband and wife) ρ is equal to 0.25 as estimated by Hyslop (2001).

As in Attanasio, Low, and Sánchez-Marcos (2008), there are two components to child care costs: the function $G(a_{h,t})$ and the price p. We estimate the function $G(a_{h,t})$ directly from data. For households where the mother is working, we regress total child care expenditure on the age of the youngest child, the age of the oldest child, the number of children, and a dummy equal to 1 if the youngest child is 0. The shape $G(a_{h,t})$ can be derived from the coefficients of this regression function, using the assumption that in our model all women have two children at an interval of two years.[19]

Finally, we assume individuals in this cohort live for 50 years from age 22, with the last 10 in retirement, and that the household receives a pension equal to 70% of the husband's earnings in the final working period.

Calibrated Parameters

There are nine parameters that we calibrate within our decision model: the fixed cost of working, $\bar{F}$; the price of child care, p; the wage gender gap in offered wages, expressed as y_0^f/y_0^m; the standard deviation of the permanent shock to women, σ_{ξ^f}; the standard deviation of the initial wage for women, $\sigma_{\xi_0^f}$; two parameters that determine exogenous wage growth, ι_1^f and ι_2^f; and the base weighting on leisure in the CES utility function, ψ_0, which, together with demographics z and the estimates of ψ_z, determine the total weight on leisure in the utility function. Finally, we calibrate the discount rate β.

The calibration targets are the participation rate of all women, the participation rate of mothers, average hours worked, the observed wage gender gap, the observed variance of wage growth, the observed initial variance in wages, and the observed wage growth at two different stages of the life-cycle. Finally, we target median wealth to median household income ratio as in Low (2005).

[18]Consistent with the distribution for the 1950s cohort, we assume 16% of women are childless, 27% have their first child at the age of 19, 12% at the age of 22, 11% at the age of 24, 5% at the ages of 26, 28, 30, and 32 and, finally, 14% at the age of 34.

[19]Our estimate of $G(a_{h,t})$ combines the cost of the firstborn child along with any subsequent costs associated with additional children who are born later. In this way, any economies of scale in child costs will be captured by $G(a_{h,t})$, but we do not identify separately the marginal cost of extra children.

TABLE V

CALIBRATED PARAMETERS AND TARGETS[a]

Parameters		Value
Constant term weight of leisure	ψ_0	4.20
Child care cost	p	967
Fixed cost of working	$\bar{F}$	468
Offered wage gender gap at age 22	y_0^f/y_0^m	0.72
Standard deviation of permanent shock (Women)	σ_{ξ^f}	0.063
Standard deviation of initial wage (Women)	$\sigma_{\xi_0^f}$	0.50
Exogenous growth in offered wage	ι_1^f	0.052
Exogenous growth in offered wage	ι_2^f	−0.0006
Discount factor (annualized)	β	0.99

Targets	Data	Model
Weekly hours worked	37.2	37.2
Participation rate	0.684	0.679
Participation rate of mothers 0–2	0.538	0.546
Observed wage gender gap	0.720	0.727
Observed variance wage growth (Women)	0.004	0.004
Observed initial variance of wages (Women)	0.14	0.15
Wage growth (if younger than 40)	0.012	0.010
Wage growth (if older than 40)	0.001	0.004
Median wealth to income ratio	1.84	1.80

[a]Statistics for women born in the 1950s and aged 25 to 55. Wage growth is annual.

In Table V, we report the calibrated parameters, and the targeted moments in the data and in the simulations. The monetary fixed cost of working is about 6% of median earnings of women aged 25 to 55. The additional monetary fixed child care cost is up to 13% of median earnings for a child age 0–2. The wage ratio between women and men at age 22 is calibrated to be 0.72. This is needed to match the average observed ratio over the lifetime of 0.72. In addition to the initial wage gender gap, there is a further, exogenous wage gap that opens up through differential wage growth for men and women over the life-cycle. Exogenous wage growth implies that men's wages are on average 77% higher by the age of 45 than at the moment of entering the labor market. By contrast, for women the figure is only 31%. We calibrate the standard deviation of wage innovations for women to be 0.063 and the standard deviation of the initial wages to 0.50.

5.4. *Goodness of Fit*

Our next step is to show whether the model can account for some observed features of women's labor supply behavior that were not explicitly targeted in the calibration. The calibration focused on averages taken over the life-cycle. Our focus here is on life-cycle paths and on the distribution of hours and wages.

Figure 2 shows the life-cycle paths of women's labor supply in the model and in the data, which match well at both the extensive and intensive margins. Table VI reports additional moments on heterogeneity. The model matches the participation of different demographic groups, such as women who have no dependent children, and mothers of

FIGURE 2.—Life-cycle profiles: Baseline model (solid black line) versus data (dashed red line).

children aged 3 to 17. Goldin and Mitchell (2017) showed that, for women born in 1957–1958, the fraction who had worked more than 80% of the years between age 25 and 54 was 0.53, while the fraction who had worked less than 20% was 0.09. In our benchmark economy, the comparable fractions are 0.57 and 0.21. The distribution of observed wages in the model is similar to that in the data, as is the distribution of hours worked, although

TABLE VI

STATISTICS ON HETEREOGENEITY[a]

	Data	Model
Participation rate: mothers with children aged 3–17	0.682	0.688
Participation rate: women without dependent children	0.755	0.692
Average hours worked 10th percentile	20	25
Average hours worked 25th percentile	35	31
Average hours worked 50th percentile	40	38
Average hours worked 75th percentile	40	44
Average hours worked 90th percentile	48	48
Wage 10th percentile	8.16	8.11
Wage 50th percentile	15.05	16.20
Wage 90th percentile	29.23	31.12
Correlation of wages and hours	0.33	0.54

[a]Women without dependent children are women who have never had children and those whose children are over 17.

the fraction of women working an average of 40 hours a week is higher in the data than in the model. Observed wages and the variance of wages are increasing with age in our simulations, consistent with the data. The correlation of wages and hours worked for those employed is 0.33 in the data, compared to 0.54 in the simulations.

Finally, as discussed in Section 3.1 and in Appendix G, using the approximate selection correction to estimate the MRS equation could introduce a bias. To assess the importance of this bias, we take simulated data generated from the complete life-cycle model with taste heterogeneity and estimate the MRS equation using our reduced form procedure which approximates the full model. The estimates of the MRS parameters θ and ϕ used to generate the simulated data are almost identical to those we recover using our reduced form estimation. Given the complexity of the model and of the full-selection process, this is an important validation of the approximation used in the reduced form selection model.

6. LABOR SUPPLY ELASTICITIES

This section provides the key results of the paper. We use the estimates of the model to show implications for various wage elasticities. We start with the static Marshallian and Hicksian elasticities obtained from the MRS parameters. We then move to the Frisch elasticities at the intensive margin using estimates from the Euler equation. Finally, we simulate the full model to obtain elasticities at the extensive margin and the aggregate response of labor supply to changes in wages. When using the full model, first we analyze responses to transitory changes to wages, which do not have wealth effects and so are analogous to the Frisch elasticities; and second, we analyze the effect of shifts in the entire wage profile allowing savings and wealth to change, generating life-cycle Marshallian and Hicksian elasticities.

6.1. *Marshallian and Hicksian Hours Elasticities*

The first two columns in Table VII show how the MRS parameters translate into within-period Marshallian and Hicksian wage elasticities separately for hours of work and for consumption. These elasticities vary according to family characteristics, wages, and the levels of consumption and leisure. We report elasticities at different points of the distribution of Marshallian elasticities to highlight the heterogeneity across individuals.

The median Marshallian hours elasticity is estimated to be 0.18, implying an upward sloping labor supply function. Hicksian elasticities are greater than Marshallian elasticities: for the household with the median Marshallian elasticity, the Hicksian hours elasticity is three times larger at 0.54, indicating large income effects.

The Marshallian and Hicksian elasticities show substantial heterogeneity. The 90–10 range of the Marshallian hours elasticity is 0.93 (from −0.14 to 0.79), while for the Hicksian one it is 0.78 (from 0.38 to 1.16). Differences in hours worked are an important source of variation in both the Hicksian and Marshallian elasticities. Figure 3 plots average elasticities by wages and by hours worked. Those working the fewest hours and those with the lowest wages have the largest proportional response to a wage increase.

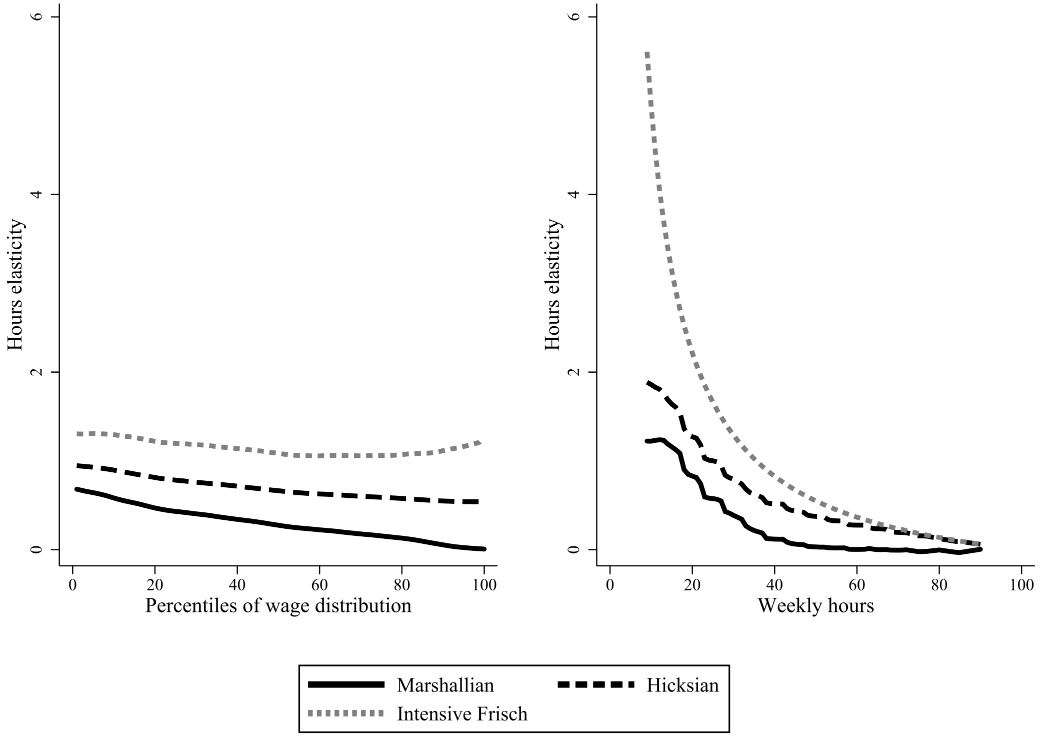
Our median estimates of the Marshallian and Hicksian elasticities are quite small (see Keane (2011) for a survey), and are similar to estimates in the literature obtained using a similar methodology to ours. Blundell, Duncan, and Meghir (1998) estimated values of the Marshallian elasticity ranging from 0.13 to 0.37 and of the Hicksian from 0.14 to 0.44 (depending on the age of the youngest child). The meta-study by Chetty et al. (2011) reports an average Hicksian elasticity (for men and women) of 0.33. Some results

TABLE VII

ELASTICITIES AT PERCENTILES OF MARSHALLIAN DISTRIBUTION[a]

	Wage			Interest rate
	Marshallian	Hicksian	Frisch	Frisch
		Hours worked		*Hours worked*
10th	−0.14 [−0.31,0.00]	0.38 [0.21,0.62]	0.80 [0.25,1.85]	0.78 [0.25,1.61]
25th	0.01 [−0.11,0.13]	0.44 [0.22,0.79]	0.80 [0.24,1.99]	0.76 [0.24,1.68]
50th	0.18 [0.05,0.38]	0.54 [0.24,1.07]	0.87 [0.26,2.29]	0.81 [0.24,1.90]
75th	0.39 [0.16,0.86]	0.69 [0.28,1.49]	1.00 [0.31,2.85]	0.93 [0.31,2.34]
90th	0.79 [0.36,1.65]	1.16 [0.51,2.30]	1.92 [0.57,4.96]	1.82 [0.57,4.07]
		Consumption		*Consumption*
25th	0.82 [[0.68,1.08]	0.43 [0.18,0.87]	0.04 [−0.02,0.50]	−1.17 [−1.83,−0.56]
50th	1.05 [0.94,1.23]	0.52 [0.24,0.98]	0.05 [−0.02,0.57]	−1.19 [−1.84,−0.52]
75th	1.30 [1.14,1.46]	0.61 [0.31,1.06]	0.05 [−0.02,0.63]	−1.20 [−1.84,−0.50]

[a]Elasticities are calculated as averages within five percentage point bands around the 10th, 25th, 50th, 75th, and 90th percentiles of the Marshallian distribution. 95% confidence intervals in square brackets. Confidence intervals are bootstrapped with 1000 replications.



Lines show the distributions of Marshallian, Hicksian and intensive Frisch elasticities smoothed using a local polynomial.

FIGURE 3.—Intensive elasticities. Lines show the distributions of Marshallian, Hicksian, and intensive Frisch elasticities smoothed using a local polynomial.

in the literature, however, report much larger estimates. MaCurdy (1983), for instance, estimated elasticities ranging from 0.74 to 1.43 (for men).

Different studies take different approaches and use different sources of variation to estimate elasticities. We investigated extensively the main reasons for different estimates of labor supply elasticities. Our hypotheses ranged from the type of specification used,[20] to the type of variation in wages that is used to identify the elasticity (i.e., what type of instruments are used), to sample selection rules. To estimate equilibrium conditions such as the MRS equation, researchers often use methods, such as 2SLS and GMM, that are sensitive to the normalization used. Therefore, we also investigated whether the results we obtain depend on which variable is used as a dependent variable. It turns out that the normalization used drives the result in a fundamental fashion, while results are robust to the other hypotheses considered. In particular, we find that IV or GMM estimates obtained using wages as the left-hand-side variable (as in MaCurdy (1983)) result in very large elasticities, while putting hours of leisure on the left-hand side (similarly to Blundell, Duncan, and Meghir (1998), who used hours worked) yields much smaller elasticities. As noted above, we use the Fuller estimator, which is less sensitive to the normalization of the estimating equation than alternative methods. In Appendix D of the Supplemental Material, we report results from GMM estimation with different normalizations.

6.2. *Frisch Hours Elasticity*

We compute Frisch elasticities with respect to wages at the intensive margin using equation (17) and estimates of the Euler equation parameters reported in Section 5.2. We report these elasticities in the third column of Table VII and plot them alongside Hicksian and Marshallian elasticities in Figure 3.

The Frisch elasticity for hours of work is larger than the Hicksian elasticity, as theory would predict. The elasticity also varies in the cross section, rising from 0.8 at the 10th percentile of the Marshallian elasticity to 1.92 at the 90th percentile. The median value is 0.87. It is quite common to find large estimates of the Frisch hours elasticity among married women, and our findings are broadly in line with those of previous studies. Blundell, Pistaferri, and Saporta-Eksten (2016) found a Frisch elasticity for married women of 0.96; Kimmel and Kniesner (1998) estimated a Frisch elasticity of 0.67. Part of the heterogeneity we observe in the Frisch elasticities is due to differences across the life-cycle and in demographics, but, once again, much of it is also due to differences in the level of hours of work. As with the Hicksian and Marshallian elasticities, Figure 3 shows that Frisch hours elasticities are largest for those working the fewest hours.

The elasticity of consumption with respect to anticipated wage changes is small but positive (owing to the fact that consumption and leisure are substitutes). The Frisch elasticity of consumption with respect to the interest rate at the median level of consumption is -1.19.

We compare these results with those obtained when we impose additive separability for preferences over consumption and leisure, as well as when we use a standard CES utility specification in Appendix E of the Supplemental Material. This exercise highlights the importance of adopting a flexible utility specification. A standard CES specification, which is shown to be rejected by the estimation in Section 5.1, leads to similar estimates of Marshallian hours elasticities, but much larger Hicksian and Frisch elasticities. The median Frisch hours elasticity estimated using the more restrictive standard CES specification is

[20]That is, whether one uses consumption to proxy for the marginal utility of wealth or other indicators.

1.33, which is roughly 50% larger than our baseline result. The corollary of this result is that the Frisch elasticity of consumption with respect to the interest rate is much lower: imposing a standard CES forces consumption and leisure to have the same substitution parameters, making consumption less elastic and hours of work more elastic than in our baseline. In addition, the standard CES utility implies much greater non-separability between consumption and leisure: implying a Frisch wage elasticity of consumption of 0.4 compared to 0.05 under our more general utility specification. On the other hand, when we impose additive separability with our general CES specification, the Frisch hours elasticity is very similar to the one we estimate allowing for non-separability.

6.3. *The Extensive Margin, Aggregate Elasticities, and Life-Cycle Responses*

This section discusses labor supply responses at the extensive margin, life-cycle responses, and aggregation issues at different margins and across households. We define the extensive margin elasticity as referring to the change in the percentage of women participating as the wage changes. We calculate how total hours worked by women change as a result of both the extensive and intensive margin responses. This is what we call the 'aggregate response' to a wage change. We also calculate aggregate changes in efficiency units, because women with different levels of productivity may respond differently, as suggested by Figure 3.

We explore responses to two different types of wage changes. First, in Section 6.3.1, we focus on the response to temporary changes in wages, which is relevant for temporary tax changes.[21] We report heterogeneity by age, across the wealth distribution, across demographic groups, and over the business cycle. Then, in Section 6.3.2, we report labor supply responses to changes in the entire life-cycle wage profile, which we call *life-cycle Marshallian* and *life-cycle Hicksian* elasticities. These are interesting for two reasons: first, for thinking about the implications of permanent tax changes or differences in taxes across countries; and, second, for comparing these life-cycle Marshallian and Hicksian elasticities with the static elasticities from the MRS to assess the accuracy of the static approximation.

6.3.1. *Response to Temporary Wage Changes*

Frisch responses are calculated by comparing labor supply at a given age between the baseline economy and a counterfactual economy in which wages are anticipated to be higher at that particular age. The wage difference generates differences in participation rates, differences in hours worked for participants and, therefore, differences in aggregate labor supply. In Table VIII, we report the average response for different age groups. The third column reports the 'extensive response,' calculated as the percentage point change in participation following a one percent increase in the wage. The fourth to sixth columns report different percentiles of the distribution of the intensive margin elasticity at each age, computed by considering only those individuals who participate both in the baseline economy and in the counterfactual economy. Changes in participation also induce changes in the distribution of hours worked that would be reflected in the aggregate response of labor supply. Finally, therefore, the last two columns report the 'aggregate' elasticity: the change in the total number of hours worked and the change in efficiency units of labor, considering both intensive and extensive margins.

[21] We compute responses to both anticipated and unanticipated temporary changes. The results are almost identical because there is very little effect on the marginal utility of wealth, λ, of a temporary change.

TABLE VIII
FRISCH RESPONSES BY AGE[a]

Age Band	Participation Rate (Percent)	Extensive Response (Percent Pt)	Intensive Elasticity			Aggregate Elasticity	
			25th	50th	75th	Hours	Eff Units
25–29	61.61	0.82	0.69	0.85	1.09	1.93	1.44
30–34	70.07	0.63	0.66	0.82	1.11	1.51	1.12
35–39	70.00	0.63	0.64	0.82	1.14	1.49	1.08
40–44	72.05	0.56	0.64	0.85	1.21	1.37	1.01
45–49	69.53	0.59	0.65	0.88	1.25	1.42	1.04
50–55	65.37	0.59	0.67	0.91	1.30	1.49	1.06

[a]The extensive response is the percentage point change in participation in response to a 1% increase in the wage. The aggregate elasticity reports the percentage change in hours corresponding to a percentage change in the wage, accounting for changes at both the extensive and intensive margins.

A first point to notice is the variation in the size of the extensive margin elasticity over the life-cycle. As a consequence, the age composition of the population may have important implications for the aggregate response of labor supply to changes in wages. Early in life, the percentage point response is about 0.82, falling to 0.63 between 30 and 35 and to a minimum of 0.56 for the 40–45 group. The median of the intensive margin elasticity is stable over the life-cycle, at around 0.85;[22] however, the elasticity at the 75th percentile increases substantially with age.

The aggregate elasticity for hours is about 1.45 on average, but again is larger at the start of the life-cycle. The relative importance of the extensive and intensive margins to explaining the macro elasticity varies with age. Before age 30, the intensive margin response contributes approximately 46% of the response in the aggregate. However, by age 50–55, the contribution of the intensive response has increased to 63%. The contribution of the intensive margin is somewhat larger than that of Erosa, Fuster, and Kambourov (2016), who found that the response through the intensive margin contributes about 38% to the aggregate response. This difference is not surprising since the Erosa, Fuster, and Kambourov (2016) calculation is for men, where we see less variability in hours worked, but it highlights the difficulty of aggregating behavior to create a single labor supply elasticity. The aggregate elasticity for efficiency units is smaller than that for hours, but also declines with age.

Household Wealth. In Table IX, we report household responses across the wealth distribution. We calculate the percentiles of household wealth at each age and classify households into quartiles. We find a clear pattern of a decreasing response of the extensive margin with increasing wealth. This is the case at all ages. There is also heterogeneity in the intensive margin elasticity by wealth, with the wealthy being less responsive, but the differences are more moderate than with the extensive margin response. The message from these results is that the distribution of wealth is crucial to understanding the response of aggregate labor supply to changes in wages.

Macroeconomic Conditions. Labor supply responses may change across the business cycle. Differences in the economic environment will lead to differences in the estimated

[22]The comparable value calculated directly from step 2 of the estimation process is 0.86. The similarity of estimates from step 2 and step 3 of the estimation provides further validation of our multi-step approach.

TABLE IX

FRISCH RESPONSES BY HOUSEHOLD WEALTH[a]

Wealth Quartile	Participation Rate (Percent)	Extensive Response (Percent Pt)	Intensive Elasticity (Median)	Aggregate Elasticity
Below $p25$	45.42	1.20	1.20	3.53
$p25$–$p50$	59.25	0.77	1.03	2.07
$p50$–$p75$	76.80	0.39	0.81	1.23
Above $p75$	90.10	0.16	0.66	0.81

[a]The extensive response is the percentage point change in participation in response to a 1% increase in the wage. The aggregate elasticity reports the percentage change in hours corresponding to a percentage change in the wage, accounting for changes at both the extensive and intensive margins.

elasticity for the same underlying preference parameters, as also discussed by Keane and Rogerson (2012). This issue is likely to be relevant particularly for the extensive margin, which is driven by non-convexities in the dynamic problem, such as fixed costs of going to work. If these non-convexities are important, it is likely that a certain sequence of aggregate shocks will tend to bunch (or further disperse) households around the kinks that determine the extensive margin response. As a consequence, different distributions of the state variables will trigger different responses in the aggregate. In particular, whether an economy is in a recession or not may well affect how much individuals are willing to respond to wage growth.

In Table X, we report responses to temporary changes in wages that occur at different points of the business cycle.[23] We report the labor supply response in the first and fourth quarters of the recession. The key finding is that responses are higher in recessions than in the baseline, and further, responses increase with the duration of the recession. From the results in Table IX, the decrease in wealth that households suffer over a recession could be behind the increasing responsiveness of the extensive margin to anticipated changes in wages. Effects may persist beyond the end of the recession, especially if wages or wealth are permanently lower. Both lower wages and lower wealth lead to higher elas-

TABLE X

FRISCH RESPONSES ACROSS THE BUSINESS CYCLE[a]

Business Cycle	Extensive Response (Percent)	Intensive Elasticity (Median)	Aggregate Elasticity	
			Hours	Eff Units
Baseline	0.63	0.86	1.53	1.12
Recession				
First quarter	0.67	0.87	1.61	1.15
Fourth quarter	0.73	0.86	1.71	1.20

[a]The extensive response is the percentage point change in participation in response to a 1% increase in the wage. The aggregate elasticity reports the percentage change in hours corresponding to a percentage change in the wage, accounting for changes at both the extensive and intensive margins.

[23]We define a recession as a situation in which all men and women receive an unexpected negative earnings shock for four consecutive quarters. These wage changes are to the permanent wage and will affect the marginal utility of wealth as well as changing intertemporal incentives. We consider responses to temporary changes in wages at different points in such a recession.

ticities: households who have been hit by recessions earlier in their life are more responsive throughout the remainder of their lives.[24]

Demographics. Finally, we explore the effect of children on the size of the elasticities. Mothers of children aged 0 to 2 are more elastic at the extensive margin (0.82) than mothers of older children (0.68) and childless women show the lowest elasticity (0.57). In contrast, differences in intensive margin elasticities are less pronounced, with mothers of young children being slightly less elastic.

6.3.2. *Life-Cycle Responses to Wage Changes*

In this section, we use our model to compute the response to a change in the entire wage profile, so as to measure the response to a permanent tax change. The life-cycle Marshallian elasticity captures the response of labor supply to changes in wages when savings are allowed to change, that is, when the extra income that arises in period t due to the increased wage does not have to be spent in that period.

The life-cycle Hicksian elasticity arises after netting off the extra lifetime resources from the lifetime budget constraint, in contrast with a static Hicksian response, which would net off the extra resources within period. Life-cycle compensation is calculated as the change in income needed to keep the original bundle of lifetime consumption and hours worked exactly affordable. The change in income from each period that needs to be compensated for is $\Delta w^f_{h,t} * (L - l_{h,t})$. Summing across all periods would give the extra resources from a wage increase that need to be subtracted in a life-cycle context. This compensation can be implemented by imposing either a person-specific lump-sum tax that is equal across periods, or a person-specific lump-sum tax at a given point in time. The choice will matter because uncertainty means the timing of income is important.[25] The alternative to this exact compensation is to do the compensation within a group, or indeed within the whole population as discussed by Keane (2011). This would mean calculating the extra income for all individuals as with the exact calculation, but then redistributing through a common per-period lump-sum payment. This approach does not give exactly the life-cycle Hicksian response because some households will be over-compensated and some under-compensated relative to their individual change in lifetime resources. On the other hand, it may be the right way to calculate the response to a funded tax change. If preferences are quasi-linear, then there are no income effects and so there is no effect on labor supply of any redistribution associated with the lump-sum compensation.

In Table XI, we report the life-cycle Marshallian and Hicksian responses. The first panel shows responses when the Hicksian compensation is common across all individuals. The second panel shows responses when compensation is common within quartiles of the initial wage distribution for women. We compare these life-cycle elasticities with the static elasticities estimated from the MRS. As we argued in Section 2.1 and as emphasized by Meghir and Phillips (2008), life-cycle labor supply responses may be approximated by the static elasticities computed from the MRS.

The median life-cycle Marshallian elasticity for the intensive margin is 0.43, substantially above the 0.18 static Marshallian elasticity. The static elasticities are calculated from the MRS using non-durable consumption, holding constant saving and also, implicitly,

[24]We show this by using our simulations to compare women hit by a recession at age 25 with those not hit by recession. Differences persist throughout their lifetimes. The details of these results are not reported here.

[25]In a model with substantial ex ante and ex post heterogeneity, either form of compensation is computationally costly to calculate.

TABLE XI

LIFE-CYCLE RESPONSES[a]

	Extensive Response (Percent Pt)	Intensive Elasticity			Aggregate Elasticity	
		25th	50th	75th	Hours	Eff Units
			Whole Sample			
Marshallian						
Life-cycle response	0.51	0.28	0.42	0.67	0.91	0.63
Static (MRS)		0.01	0.18	0.39		
Hicksian						
Life-cycle response	0.65	0.42	0.63	0.96	1.26	0.84
Static (MRS)		0.44	0.54	0.69		
			By Quartile of Initial Wage			
Life-cycle Marshallian						
1st quartile	0.62	0.40	0.57	0.80	2.25	1.88
2nd quartile	0.70	0.34	0.48	0.78	1.44	1.21
3rd quartile	0.55	0.32	0.48	0.75	0.97	0.83
4th quartile	0.17	0.22	0.33	0.52	0.46	0.40
Life-cycle Hicksian						
1st quartile	0.66	0.46	0.65	0.87	2.47	2.05
2nd quartile	0.81	0.45	0.62	0.94	1.71	1.43
3rd quartile	0.64	0.48	0.67	0.97	1.25	1.04
4th quartile	0.21	0.41	0.56	0.81	0.71	0.60

[a]The extensive response is the percentage point change in participation in response to a 1% increase in the wage. The baseline participation rate is 67.8%. Within quartiles, the baseline participation rates are 29, 56, 77, and 95%, respectively. The aggregate elasticity reports the percentage change in hours corresponding to a percentage change in the wage, accounting for changes at both the extensive and intensive margins.

durable spending. In a full life-cycle model, however, following a wage increase, savings adjust and individuals reallocate resources across periods. Furthermore, all life-cycle resources are spent, so that we have a broader consumption measure in these calculations. In other words, the extra income from the wage increase is not all spent on nondurables in the period it is earned. Spreading these resources across periods and other goods reduces the amount of extra income and hence the income effect in the period it is earned. This means the life-cycle Marshallian elasticity is more like the static Hicksian elasticity. However, the life-cycle Hicksian elasticity is close to the Hicksian elasticity we estimate with the MRS.

Looking at the responses by quartile in the bottom panel, there is substantial heterogeneity in the size of the life-cycle Marshallian intensive margin response depending on initial conditions, particularly in the extensive margin response. On the other hand, the life-cycle Hicksian elasticity when there is within-quartile compensation does not vary much with the quartile of the initial conditions.[26] The substitution effect is very similar across groups, and it is the income effect which matters more for the heterogeneity in the Marshallian labor supply responses across groups.

[26]We experiment with more finely targeted compensation, in particular making the individual transfer contingent on initial husband earnings and the maternity group, but this does not alter the overall intensive margin response.

6.4. *Elasticities With Returns to Experience*

An important maintained assumption to this point has been the absence of any returns to experience. Imai and Keane (2004) argued that assuming wages are exogenous may introduce a downward bias in estimates of the willingness to substitute intertemporally. Indeed, they presented estimates of such a parameter as high as 3.8 in a model that accounts for returns to labor market experience. We consider as a robustness exercise an alternative framework in which returns to experience accrue to individuals who are participating, but in which returns to experience are not affected by the number of hours worked conditional on participation. Appendix F of the Supplemental Material details the estimation results allowing for returns to experience. Intensive elasticities are similar to our baseline, but the extensive margin response differs: with returns to experience, the current wage is only part of the return to work and so changes to the current wage make little difference to participation. The extensive margin response becomes very small.

7. CONCLUSION

This paper shows that to understand labor supply behavior and to calculate aggregate labor supply elasticities, it is crucial to account for heterogeneity across individuals. To make this point precisely and show its quantitative importance, we estimate a life-cycle model of intratemporal and intertemporal choices over consumption, saving, and work and characterize the response of women's labor supply to different types of wage changes. In estimating such a model, we use a flexible specification of preferences that allows us to test some of the assumptions commonly used in both the macro and labor literature on labor supply.

We find substantial heterogeneity in labor supply responses, and this heterogeneity is prevalent at both the intensive and extensive margins. The median static Marshallian elasticity is 0.18, but has a *90–10* range of −0.14 to 0.79. The corresponding Hicksian elasticity is 0.54, with a *90–10* range of 0.38 to 1.16; and the corresponding Frisch wage elasticity is 0.87, with a *90–10* range of 0.8 to 1.92. The static Marshallian and Hicksian concepts assume there is no intertemporal reallocation of resources in response to a wage change. We use the full life-cycle model to show that these static concepts underestimate the full life-cycle responses, especially for the life-cycle Marshallian response. Finally, over the business cycle we find that the aggregate hours elasticity increases in recessions and more so in longer recessions.

In terms of heterogeneity in the intensive margin responses, the Marshallian, Hicksian, and Frisch elasticities are greatest for those working the least number of hours, those with the lowest wages, and those with the least wealth. For the extensive margin, the response to anticipated wage growth is large for women under 30 and can explain 54% of their labor supply response. This sizable contribution of the extensive margin declines with age. We find some evidence of non-separability between consumption and leisure, but assuming there is separability does not substantially change the distribution of estimates of the Frisch elasticity.

Our preference parameter estimates reject the restrictions required for balanced growth, which are widely used in the macro literature. The curvature on consumption in utility is less than log, and the curvature on hours worked is much greater than the curvature on consumption. This implies individuals are less willing to substitute hours of work over time than they are willing to substitute consumption. Further, the heterogeneity we observe means it is not sensible to talk about a single elasticity to measure how aggregate labor supply responds to wage changes. Instead, we aggregate explicitly from individual behavior to the aggregate in order to understand how economy-wide hours of work change given the demographic and age structure of the economy, the wealth distribution, and the state of the business cycle.

Our results on the importance of the extensive margin in explaining macro elasticities can be compared to others in the literature, especially Erosa, Fuster, and Kambourov (2016) and Guner, Kaygusuz, and Ventura (2012). Our estimates put a greater importance on intensive margin changes in hours worked per week than those papers, but we do find that a substantial fraction of the changes in total hours is due to changes in participation, ranging from 54% to 37%. Erosa, Fuster, and Kambourov (2016) found that the extensive margin is the dominant labor supply response, explaining 62% of the aggregate response. Their model has a similar life-cycle structure to ours, but is focused on men's labor supply, and the conclusion on the importance of the extensive margin is for men, where hours of work are less variable. Guner, Kaygusuz, and Ventura (2012) analyzed the importance of the extensive margin for the aggregate response of labor supply to changes in taxes in a model with heterogeneous married and single households, and with an extensive margin for women as well as an intensive margin for men and women. As with Erosa, Fuster, and Kambourov (2016), they found that the extensive margin for women is a key contributor to the aggregate response to tax reform. The key difference from our framework is their assumption that there is no uncertainty in wages, and this assumption of certainty tends to lead to greater labor supply responses, as shown in Low (2005).

One key point that emerges from our exercise is that aggregate responses of labor supply to changes in wages (both at the intensive and the extensive margin) are not constant: they change with the structure of the population as well as with the state of the economy. This finding is similar to the work of Keane and Rogerson (2012), who argued that there is no contradiction between macro and micro elasticities of labor supply and that they are simply measuring different concepts. Our conclusion is, however, stronger: the macro elasticity is not a structural parameter; it is simply the result of highly nonlinear aggregation which depends on demographic structure as well as the distribution of wealth and the particular point in the business cycle.

REFERENCES

AARONSON, D., AND E. FRENCH (2004): "The Effect of Part-Time Work on Wages: Evidence From the Social Security Rules," *Journal of Labor Economics*, 22, 329–352.[2054]

ABOWD, J. M., AND D. CARD (1989): "On the Covariance Structure of Earnings and Hours Changes," *Econometrica*, 57, 411–445.[2055]

ALTONJI, J. (1986): "Intertemporal Substitution in Labor Supply: Evidence From Micro Data," *Journal of Political Economy*, 94, S176–S215.[2052]

ATTANASIO, O., LEVELL, P., LOW, H., AND SÁNCHEZ-MARCOS, V. (2018): "Supplement to 'Aggregating Elasticities: Intensive and Extensive Margins of Women's Labor Supply'," *Econometrica Supplemental Material*, 86, https://doi.org/10.3982/ECTA15067. [2052]

ATTANASIO, O. P., AND H. LOW (2004): "Estimating Euler Equations," *Review of Economic Dynamics*, 7, 406–435.[2060,2061]

ATTANASIO, O. P., AND G. WEBER (1995): "Is Consumption Growth Consistent With Intertemporal Optimization? Evidence From the Consumer Expenditure Survey," *Journal of Political Economy*, 103, 1121–1157.[2050]

ATTANASIO, O. P., H. LOW, AND V. SÁNCHEZ-MARCOS (2008): "Explaining Changes in Female Labor Supply in a Life-Cycle Model," *American Economic Review*, 98, 1517–1552.[2067,2068]

BLUNDELL, R., AND T. MACURDY (1999): "Labor Supply: A Review of Alternative Approaches," in *Handbook of Labor Economics*, Vol. 3. Elsevier, 1559–1695. [2056]

BLUNDELL, R., AND I. WALKER (1986): "A Life-Cycle Consistent Empirical Model of Family Labour Supply Using Cross-Section Data," *Review of Economic Studies*, 53, 539–558.[2056]

BLUNDELL, R., A. BOZIO, AND G. LAROQUE (2011): "Labor Supply and the Extensive Margin," *American Economic Review*, 101, 482–486.[2049]
BLUNDELL, R., A. DUNCAN, AND C. MEGHIR (1998): "Estimating Labor Supply Responses Using Tax Reforms," *Econometrica*, 66, 827–861.[2050,2059,2060,2071,2073]
BLUNDELL, R., C. MEGHIR, AND P. NEVES (1993): "Labour Supply and Intertemporal Substitution," *Journal of Econometrics*, 59, 137–160.[2050,2052,2060]
BLUNDELL, R., L. PISTAFERRI, AND I. SAPORTA-EKSTEN (2016): "Consumption Inequality and Family Labor Supply," *American Economic Review*, 106, 387–435.[2073]
BROWNING, M., A. DEATON, AND M. IRISH (1985): "A Profitable Approach to Labor Supply and Commodity Demands Over the Life-Cycle," *Econometrica*, 53, 503–544.[2061]
BROWNING, M., L. P. HANSEN, AND J. J. HECKMAN (1999): "Micro Data and General Equilibrium Models," in *Handbook of Macroeconomics*, Vol. 1, ed. by J. B. Taylor and M. Woodford. Elsevier, 543–633, Chapter 8. [2051]
CASTEX, G., AND E. DECHTER (2016): "A Model of Labor Supply, Fixed Costs and Work Schedules," Report, University of New South Wales. [2055]
CHANG, Y., AND S.-B. KIM (2006): "From Individual to Aggregate Labor Supply: A Quantitative Analysis Based on a Heterogeneous Agent Macroeconomy," *International Economic Review*, 47, 1–27.[2050]
CHETTY, R. (2012): "Bounds on Elasticities With Optimization Frictions: A Synthesis of Micro and Macro Evidence on Labor Supply," *Econometrica*, 80, 969–1018.[2058]
CHETTY, R., A. GUREN, D. MANOLI, AND A. WEBER (2011): "Are Micro and Macro Labor Supply Elasticities Consistent? A Review of Evidence on the Intensive and Extensive Margins," *American Economic Review*, 101, 471–475.[2050,2071]
DEATON, A. (1985): "Panel Data From Time Series of Cross-Sections," *Journal of Econometrics*, 30, 189–194.[2061]
EROSA, A., L. FUSTER, AND G. KAMBOUROV (2016): "Towards a Micro-Founded Theory of Aggregate Labor Supply," *Review of Economic Studies*, 83, 1001–1039.[2050-2052,2075,2080]
FEENBERG, D., AND E. COUTTS (1993): "An Introduction to the TAXSIM Model," *Journal of Policy Analysis and Management*, 12, 189–194.[2062]
FULLER, W. A. (1977): "Some Properties of a Modification of the Limited Information Estimator," *Econometrica*, 45, 939–953.[2051,2058]
GOLDIN, C., AND J. MITCHELL (2017): "The New Life Cycle of Women's Employment: Disappearing Humps, Sagging Middles, Expanding Tops," *Journal of Economic Perspectives*, 31, 161–182.[2070]
GORMAN, W. M. (1959): "Separable Utility and Aggregation," *Econometrica*, 27, 469–481.[2056]
GUNER, N., R. KAYGUSUZ, AND G. VENTURA (2012): "Taxation and Household Labour Supply," *Review of Economic Studies*, 79, 1113–1149.[2052,2080]
HECKMAN, J. J., AND T. E. MACURDY (1980): "A Life Cycle Model of Female Labour Supply," *The Review of Economic Studies*, 47, 47–74.[2051]
HIRSCH, B. T. (2005): "Why Do Part-Time Workers Earn Less? The Role of Worker and Job Skills," *ILR Review*, 58, 525–551.[2054]
HUGGETT, M., G. VENTURA, AND A. YARON (2011): "Sources of Lifetime Inequality," *American Economic Review*, 101, 2923–2954.[2068]
HYSLOP, D. R. (2001): "Rising U.S. Earnings Inequality and Family Labor Supply: The Covariance Structure of Intrafamily Earnings," *American Economic Review*, 91, 755–777.[2068]
IMAI, S., AND M. P. KEANE (2004): "Intertemporal Labor Supply and Human Capital Accumulation," *International Economic Review*, 45, 601–641.[2052,2079]
KEANE, M. P. (2011): "Labor Supply and Taxes: A Survey," *Journal of Economic Literature*, 49, 961–1075.[2056, 2071,2077]
KEANE, M. P., AND R. ROGERSON (2012): "Micro and Macro Labor Supply Elasticities: A Reassessment of Conventional Wisdom," *Journal of Economic Literature*, 50, 464–476.[2049,2076,2080]
KEANE, M. P., AND N. WASI (2016): "Labor Supply: The Roles of Human Capital and the Extensive Margin," *The Economic Journal*, 126, 578–617.[2050,2052]
KIMMEL, J., AND T. J. KNIESNER (1998): "New Evidence on Labor Supply: Employment versus Hours Elasticities by Sex and Marital Status," *Journal of Monetary Economics*, 42, 289–301.[2059,2073]
LJUNGQVIST, L., AND T. J. SARGENT (2011): "A Labor Supply Elasticity Accord?" *American Economic Review*, 101, 487–491.[2049]
LOW, H. (2005): "Self-Insurance in a Life-Cycle Model of Labour Supply and Savings," *Review of Economic Dynamics*, 8, 945–975.[2068,2080]
MACURDY, T. (1981): "An Empirical Model of Labour Supply in a Life Cycle Setting," *Journal of Political Economy*, 89, 1059–1085.[2056]

——— (1983): "A Simple Scheme for Estimating an Intertemporal Model of Labor Supply and Consumption in the Presence of Taxes and Uncertainty," *International Economic Review*, 24, 265–289.[2052,2055,2056, 2059,2073]

MEGHIR, C., AND D. PHILLIPS (2008): "Labour Supply and Taxes," IZA WP, 3405. [2077]

RENDALL, M., E. ARACIL, C. BAGAVOS, C. COUET, A. DEROSE, P. DIGIULIO, T. LAPPEGARD, I. ROBERT-BOBÉE, M. RØNSEN, S. SMALLWOOD, AND G. VERROPOULOU (2010): "Increasingly Heterogeneous Ages at First Birth by Education in Southern European and Anglo-American Family-Policy Regimes: A Seven-Country Comparison by Birth Cohort," *Population Studies*, 64, 209–227.[2067]

ROGERSON, R., AND J. WALLENIUS (2009): "Micro and Macro Elasticities in a Life Cycle Model With Taxes," *Journal of Economic Theory*, 144, 2277–2292.[2049]

STOCK, J. H., AND M. YOGO (2005): "Testing for Weak Instruments in Linear IV Regression," in *Identification and Inference for Econometric Models: Essays in Honor of Thomas Rothenberg*, ed. by D. W. K. Andrews and J. H. Stock. Cambridge University Press, 80–108, Chapter 6.[2066]

ZILIAK, J. P., AND T. J. KNIESNER (1999): "Estimating Life-Cycle Labor Supply Tax Effects," *Journal of Political Economy*, 107, 326–359.[2059]

Co-editor Liran Einav handled this manuscript.

Manuscript received 31 January, 2017; final version accepted 15 May, 2018; available online 22 June, 2018.