Aggregating Elasticities: Intensive and Extensive Margins of Women's Labour Supply *

Orazio Attanasio† Peter Levell ‡ Hamish Low § Virginia Sánchez-Marcos¶

August 7, 2018

Abstract

We show that there is substantial heterogeneity in women's labour supply elasticities at the micro level and highlight the implications for aggregate behaviour. 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 labour 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 labour 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.

JEL Codes: J22, D91

Keywords: labour supply elasticities, heterogeneity, aggregation, non-separability

*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.

†UCL, IFS and NBER

‡IFS and UCL

§University of Cambridge and IFS

¶Universidad de Cantabria

1 Introduction

The size of the elasticity of labour 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 (2015) and Keane and Rogerson (2012) survey some of these issues and the papers by Blundell et al. (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 has been given to the distinction between the extensive and intensive margins of labour supply, see, in particular, Chetty et al. (2011). Perhaps surprisingly, in this debate, aggregation issues and the pervasive and complex heterogeneity that characterise labour supply behaviour 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 labour supply elasticities in these dimensions is something to be expected. Labour 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 labour 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 labour supply. Our approach consists in taking a relatively standard life-cycle model of labour supply to the data. Whilst 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 labour 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

[1] One exception is Keane and Wasi (2016) who show men's labour supply responses vary substantially with age, education and the tax structure. Aggregation issues are also discussed in Erosa et al. (2016).

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 labour supply at the intensive margin. We use group level variability driven by group or aggregate shocks such as policy reforms, similar to Blundell et al. (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 labour supply elasticities. We estimate these parameters by using the Euler equation for consumption, using synthetic cohorts, similar to Blundell et al. (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 characterise behaviour 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 behaviour, similar in spirit to Erosa et al. (2016). Labour 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 isoleastic, homothetic preferences that conform to the restrictions for balanced growth, as in Erosa et al. (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 labour 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 sizeable income effect. For the same

[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 et al. (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.

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 normalisation 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 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 labour 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 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 behaviour. 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 behaviour 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 labour supply. Importantly in linking the micro and macro analysis of labour 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 et al. (2016), who have similar aims of building aggregate elasticities from men's labour supply behaviour 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 labour supply responses. A second related paper is Guner et al. (2012), who model heterogeneous married and single households with an extensive margin for women and an intensive margin for both men and women. Their focus is on evaluating different reforms of the US tax system and they abstract 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 et al. (1993), who condition on the extensive margin, and estimate 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 labour supply choices may change future wages and, therefore, future labour supply behaviour, as stressed by Imai and Keane (2004). Keane and Wasi (2016) model human capital and find that labour supply elasticities are highly heterogenous and vary substantially with age, education and the tax structure. In Appendix F, 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 labour 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. An online appendix provides supporting evidence.

2 A life-cycle model of women's labour supply

To study both the intensive and the extensive margins of women's labour supply, we use a rich model of labour 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 work 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 labour supply responses to changes in wages.

We consider married couples, who maximise the lifetime expected utility of the household, h, and choose consumption and women's labour supply within each period.

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

where c is consumption, l is hours of leisure for women, and P is an indicator of the woman's labour 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:

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

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

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

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 both at 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.

Maximisation is subject to the intertemporal budget constraint:

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

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 US-based studies (Hirsch (2005); Aaronson and French (2004)).

In our baseline specification, human capital does not depend on the history of labour 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 labour supply do not have a direct effect on con-

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

tinuation values.[4] Therefore, the only linkage across periods is through the decision 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):

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

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

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

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 realisations. 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 labour supply and its elasticity to wages by simulating a large number of households and aggregating explicitly their behaviour.

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

[4] In Appendix F, 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 on hours worked. The first two steps of our estimation approach go through in 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).

women as:

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

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:

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

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:

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

These equations can be used to compute Marshallian and Hicksian labour 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 behaviour 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 realised 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 labour 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 (see Appendix A for details on the derivations):

[6] Blundell and MaCurdy (1999) and Keane (2011) discuss 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 labour supply by MaCurdy (1981), MaCurdy (1983) and Blundell and Walker (1986).

$$\varepsilon_h^M = \frac{\partial \ln h}{\partial \ln w} = -\left(\frac{\phi w\,(L-l) - c}{\theta c + \phi w l}\right)\frac{l}{L-l} \tag{15}$$

$$\varepsilon_c^M = \frac{\partial \ln c}{\partial \ln w} = \frac{\theta w\,(L-l) + wl}{\theta c + \phi w l}$$

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 labour 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 intertermporal 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 labour supply model:

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

$$\varepsilon_c^H = \varepsilon_c^M + \frac{\partial \ln c}{\partial \ln (c+wl)}\frac{wl}{(c+wl)} = \frac{-c}{\theta c + \phi w l}$$

The Marshallian and Hicksian elasticities are the relevant concepts to think about the labour 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 non-linear 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]

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

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.

[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 realisation, 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.

[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.

Finally, in the third step, we characterise behaviour at the extensive margin. This step requires solving the entire model and, therefore, specifying completely the environment in which households operates. 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:

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

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 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 characterise participation and, at best, could be considered an approximation of the 'true' participation equation. In Appendix G, 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.

[10] One issue to worry about is the intrinsic non-linearity of the participation equation. The omission of some state variables could change the properties of the residuals of such a non-linear 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.

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 labour supply (such as Blundell et al. (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 labour 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.

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 age, an indicator for the presence of any child, the numbers 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 et al. (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]

[11] Various papers have used variation across education groups; for example MaCurdy (1983) and Ziliak and Kniesner (1999) both use age-education interactions as instruments for wages and hours in their MRS/labour supply conditions. Similarly, Kimmel and Kniesner (1998) use 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.

[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.

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 labour 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 et al. (1993). In the absence of binding borrowing constraints, the following intertemporal condition holds:

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

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 non-linear GMM. However, as discussed in Attanasio and Low (2004), the small sample properties of non-linear 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-linearised version.

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

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

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

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

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

The identification and estimation of the parameters of this equation depends, 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 et al. (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 over-identifying 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 non-linear 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.

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 childcare costs, which drive the extensive margin decision. When considering the extensive margin, it is necessary to solve explicitly the whole 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

[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 $M_{h,t}$.

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' labour 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 Consume 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 labour supply of household members.

We calculate gross hourly wages for individuals using information on the value of each individual's last pay cheque, 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 labour 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 labour 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 leaves just under 79,000 households (50,895 where the woman is working). Appendix C presents descriptive statistics on individual characteristics over time.

4.1 Cohort averages

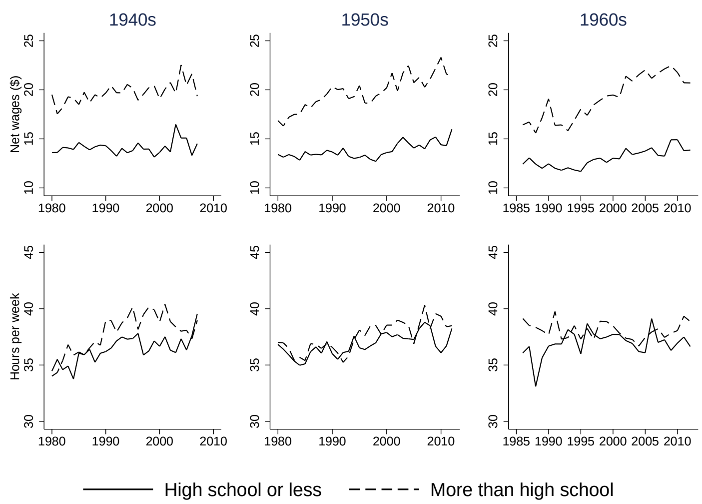
We separate households into birth cohorts and examine the evolution of wages and hours by education within each cohort group. In Figure 1, we report patterns for the cohorts born in the 40's, 50's and

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

60s and for females 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).

Figure 1: Wages and hours by education group and cohort



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.

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 labour 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

[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.

within a quarter rather than hours per week. Table 1 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, 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.

Table 1: Variances of Labour 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

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 2 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.

Table 2: Changes in Weekly Hours among the Employed

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 %

Notes: Changes in hours are measured between the 2nd and 5th interviews for individuals who are employed at each interview.

5 Results: Parameter Estimates and Calibration

In this section, we report estimates of the structural parameters of our model. In subsections 5.1 and 5.2 we report the estimation results obtained using the MRS conditions and the Euler equation. In subsection 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 3, 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. 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 3 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 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 overidentifying restrictions.

Table 3: Estimation of MRS equation

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)	

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.

5.2 Euler equation estimates

Table 4 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 non-separable and that consumption and leisure are substitutes ($\gamma = 0$ would imply additively separable preferences

[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.

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 4 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.

Table 4: Estimation of Euler equation

Parameter	Estimate	(Standard Error)	[95% Confidence Interval]
hline 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)	

N = 1,519. *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.

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 instruments' joint significance in the first stage regression. The critical values of these F-tests suggest that the instruments are highly correlated with the dependent 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 et al. (2008) show that women's labour supply behaviour differs substantially across cohorts. The main cause in that paper is differences in costs of childcare, 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 use 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 19 in Appendix D. We fix the annualized interest rate to equal the average real return on three monthly 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 et al. (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 labour market, we calibrate the parameters that characterise 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 et al. (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 childcare expenditure on the age of the youngest child, the age of the oldest child, the number of children and a dummy equal to one 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

[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.

Table 5: Calibrated Parameters and Targets

Parameters		**Value**
Constant term weight of leisure	ψ_0	4.20
Childcare 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

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

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).

In Table 5, 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 childcare 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 labour 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 labour supply behaviour 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 labour supply in the model and in the data, which match well at both the extensive and intensive margins. Table 7 reports additional moments on

[19] Our estimate of $G(a_{h,t})$ combines the cost of the first born 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.

Figure 2: Life-Cycle Profiles: Baseline Model (solid black line) versus Data (dashed red line)

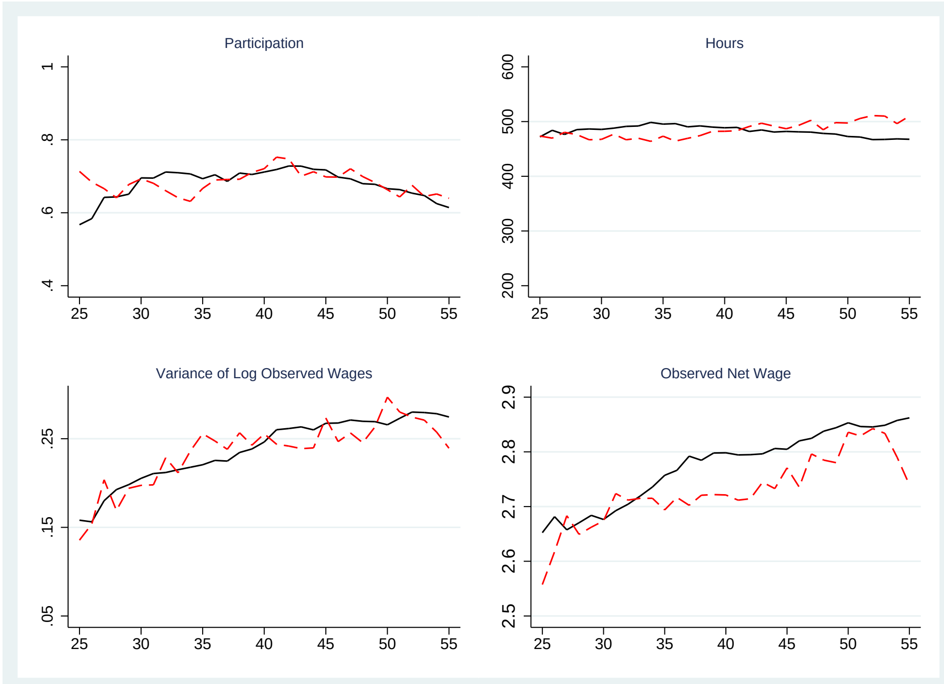


Table 6: Statistics on Hetereogeneity

	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

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

heterogeneity. The model matches the participation of different demographic groups, such as women who have no dependent children, and mothers of children aged 3 to 17. Goldin and Mitchell (2017) show that, for women born in 1957-58, 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 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 Subsection 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 Labour 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 labour supply to changes in wages. When using the full model, first we analyse responses to transitory changes to wages, which do not have wealth effects and so are analogous to the Frisch elasticities; and second, we analyse 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 7 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 labour 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.

Table 7: Elasticities at Percentiles of Marshallian distribution

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

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.

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 et al. (1998) estimate 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 in the literature, however, report much larger estimates. MaCurdy (1983), for instance, estimates 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 labour 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 (that is 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) result in very large elasticities while putting hours of leisure on the left hand side (similar to Blundell et al. (1998), who use hours worked) yields much smaller elasticities. As noted above, we use the Fuller estimator, which is less sensitive to the normalisation of the estimating equation than alternative methods. In Appendix D, we report results from GMM estimation with different normalisations.

6.2 Frisch hours elasticity

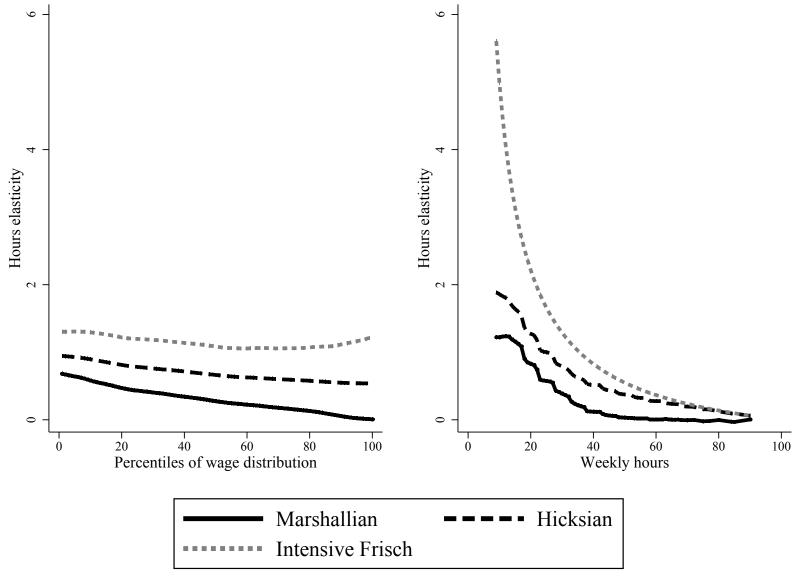
We compute Frisch elasticities with respect to wages at the intensive margin using equation (17) and estimates of the Euler equation parameters reported in section 5.2. We report these elasticities in the third column of Table 7 and plot them alongside Hicksian and Marshallian elasticities in Figure 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 et al. (2016b) find a Frisch elasticity for married women of 0.96; Kimmel and Kniesner (1998) estimate 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

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

Figure 3: Intensive elasticities



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

to the fact 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. 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 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 labour 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 also 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 labour 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 labour 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 labour supply. In Table 8, 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 labour 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 labour, considering both intensive and extensive 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 labour 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

[21] We compute responses to both anticipated and unanticipated the 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 8: Frisch Responses by Age

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

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.

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 Erosa et al. (2016), who find that the response through the intensive margin contributes about 38% to the aggregate response. This difference is not surprising since the Erosa et al. (2016) calculation is for men, where we see less variability in hours worked but it highlights the difficulty of aggregating behaviour to create a single labour supply elasticity. The aggregate elasticity for efficiency units is smaller than that for hours, but also declines with age.

Household Wealth. In Table 9, we report household responses across the wealth distribution. We calculate the percentiles of household's 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 labour supply to changes in wages.

[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 9: Frisch Responses by Household Wealth

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

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.

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

Table 10: Frisch Responses across the Business Cycle

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

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.

In Table 10, we report responses to temporary changes in wages that occur at different points of the business cycle.[23] We report the labour 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 9, 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 elasticities: 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 to measure the response to a permanent tax change. The life-cycle Marshallian elasticity captures the response of labour 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 life-time 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 either by imposing 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), which 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

[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.

[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 detail of these results are not reported here.

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

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 labour supply of any redistribution associated with the lump sum compensation.

In Table 11, 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 emphasized by Meghir and Phillips (2008), life-cycle labour 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, 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 non-durables 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 Marhsallian 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 labour supply responses across groups.

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) argue that assuming wages are exogenous may introduce a downward bias in estimates of the willingness to substitute intertemporally. Indeed, they present estimates of such a parameter as high as 3.8 in a model that accounts for returns to labour market experience. We consider as a robustness exercise an alternative framework in which returns to experience accrue to

[26] We experiment with more finely targetted 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.

Table 11: Life-Cycle Responses

	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

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.

individuals who are participating, but in which returns to experience are not affected by the number of hours worked conditional on participation. Appendix F 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 labour supply behaviour and to calculate aggregate labour 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 characterise the response of women's labour 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 both in the macro and labour literature on labour supply.

We find substantial heterogeneity in labour 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 *90-10* range of 0.38 to 1.16; and the corresponding Frisch wage elasticity is 0.87, with *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 labour 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 labour supply responds to wage changes. Instead, we aggregate explicitly from individual behaviour 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 et al. (2016) and Guner et al. (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 et al. (2016) find that the extensive margin is the dominant labour supply response, explaining 62% of the aggregate response. Their model has a similar life-cycle structure to ours, but is focused on men's labour supply and the conclusion on the importance of the extensive margin is for men where hours of work are less variable. Guner et al. (2012) analyse the importance of the extensive margin for the aggregate response of labour 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 et al. (2016), they find 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 labour supply responses, as shown in Low (2005).

One key point that emerges from our exercise is that aggregate responses of labour supply to changes in wages (both at the intensive and the extensive margin) is not constant: it changes with the structure of the population as well as with the state of the economy. This finding is similar to Keane and Rogerson (2012), who argue that there is no contradiction between macro and micro elasticities of labour 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 non-linear 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.

Abowd, J. M. and D. Card (1989): "On The Covariance Structure of Earnings and Hours Changes," *Econometrica*, 57, 411–445.

Alonso-Borrego, C. and M. Arellano (1999): "Symmetrically Normalized Instrumental-variable Estimation Using Panel Data," *Journal of Business & Economic Statistics*, 17, 36–49.

Altonji, J. (1986): "Intertemporal Substitution in Labor Supply: Evidence from Micro Data," *Journal of Political Economy*, 94, S176–S215.

Attanasio, O. P. (1999): *Handbook of Macroeconomics*, Elsevier, vol. 1, chap. 11, 741–812.

Attanasio, O. P. and H. Low (2004): "Estimating Euler Equations," *Review of Economic Dynamics*, 7, 406–435.

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.

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.

——— (2010): "Consumption and Saving: Models of Intertemporal Allocation and their Implications for Public Policy," *Journal of Economic Literature*, 48, 693–751.

Blundell, R., A. Bozio, and G. Laroque (2011): "Labor Supply and the Extensive Margin," *American Economic Review*, 101, 482–486.

Blundell, R., M. Costa, C. Meghir, and J. Shaw (2016a): "Female Labor Supply, Human Capital and Welfare Reform," *Econometrica*, 84, 17051753.

Blundell, R., A. Duncan, and C. Meghir (1998): "Estimating Labor Supply Responses Using Tax Reforms," *Econometrica*, 66, 827–861.

Blundell, R. and T. MaCurdy (1999): *Handbook of Labor Economics*, Elsevier, vol. 3, chap. Labor Supply: A Review of Alternative Approaches, 1559–1695.

Blundell, R., C. Meghir, and P. Neves (1993): "Labour Supply and Intertemporal Substitution," *Journal of Econometrics*, 59, 137–160.

Blundell, R., L. Pistaferri, and I. Saporta-Eksten (2016b): "Consumption Inequality and Family Labor Supply," *American Economic Review*, 106, 387–435.

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.

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.

Browning, M., L. P. Hansen, and J. J. Heckman (1999): "Micro Data and General Equilibrium Models," in *Handbook of Macroeconomics*, ed. by J. B. Taylor and M. Woodford, Elsevier, vol. 1 of *Handbook of Macroeconomics*, chap. 8, 543–633.

Castex, G. and E. Dechter (2016): "A Model of Labor Supply, Fixed Costs and Work Schedules," *University of New South Wales, mimeo.*

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.

Chetty, R. (2012): "Bounds on Elasticities with Optimization Frictions: A Synthesis of Micro and Macro Evidence on Labor Supply," *Econometrica*, 80, 969–1018.

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.

Davidson, R. and J. G. MacKinnon (2004): *Econometric Theory and Methods*, Oxford University Press.

Deaton, A. (1985): "Panel Data from Time Series of Cross-Sections," *Journal of Econometrics*, 30, 189–194.

Erosa, A., L. Fuster, and G. Kambourov (2016): "Towards a Micro-founded Theory of Aggregate Labor Supply," *Review of Economic Studies*, 83, 1001–1039.

Feenberg, D. and E. Coutts (1993): "An Introduction to the TAXSIM Model," *Journal of Policy Analysis and Management*, 12, 189–194.

Fuller, W. A. (1977): "Some Properties of a Modification of the Limited Information Estimator," *Econometrica*, 45, 939–953.

Goldin, C. and L. F. Katz (2007): "Long-Run Changes in the Wage Structure: Narrowing, Widening, Polarizing," *Brookings Papers on Economic Activity*, 2, 135–167.

Goldin, C. and J. Mitchell (2017): "The New Life Cycle of Womens Employment: Disappearing Humps, Sagging Middles, Expanding Tops," *Journal of Economic Perspectives*, 31, 161182.

Gorman, W. M. (1959): "Separable Utility and Aggregation," *Econometrica*, 27, 469481.

Guner, N., R. Kaygusuz, and G. Ventura (2012): "Taxation and Household Labour Supply," *Review of Economic Studies*, 79, 11131149.

Hahn, J. and J. Hausman (2003): "Weak Instruments: Diagnosis and Cures in Empirical Econometrics," *American Economic Review*, 93, 118–125.

Hahn, J., J. Hausman, and G. Kuersteiner (2004): "Estimation with Weak Instruments: Accuracy of Higher-order Bias and MSE Approximations," *Econometrics Journal*, 7, 272–306.

Hansen, L. P., J. Heaton, and A. Yaron (1996): "Finite-Sample Properties of Some Alternative GMM Estimators," *Journal of Business & Economic Statistics*, 14, 262–280.

Hausman, J., W. K. Newey, T. Woutersen, J. C. Chao, and N. R. Swanson (2012): "Instrumental Variable Estimation with Heteroskedasticity and Many Instruments," *Quantitative Economics*, 3, 211–255.

Hayashi, F. (1987): "Tests for Liquidity Constraints: A Critical Survey," in *Advances in Econometrics, Fifth World Congress*, ed. by T. Bewley, Cambridge University Press, vol. 2, 91–120.

Heckman, J. J. and T. E. Macurdy (1980): "A Life Cycle Model of Female Labour Supply," *The Review of Economic Studies*, 47, 47–74.

Hirsch, B. T. (2005): "Why Do Part-Time Workers Earn Less? The Role of Worker and Job Skills," *ILR Review*, 58, 525–551.

Huggett, M., G. Ventura, and A. Yaron (2011): "Sources of Lifetime Inequality," *American Economic Review*, 101, 2923–2954.

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.

Imai, S. and M. P. Keane (2004): "Intertemporal Labor Supply and Human Capital Accumulation," *International Economic Review*, 45, 601–641.

Iskhakov, F., T. H. Jørgensen, J. Rust, and B. Schjerning (2017): "The Endogenous grid method for discrete-continuous dynamic choice models with (or without) taste shocks," *Quantitative Economics*, 8, 317–365.

Keane, M. P. (2011): "Labor Supply and Taxes: A Survey," *Journal of Economic Literature*, 49, 961–1075.

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.

——— (2015): "Reconciling Micro and Macro Labor Supply Elasticities: A Structural Perspective," *Annual Review of Economics*, 7, 89–117.

Keane, M. P. and N. Wasi (2016): "Labor Supply: the Roles of Human Capital and the Extensive Margin," *The Economic Journal*, 126, 578–617.

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.

Ljungqvist, L. and T. J. Sargent (2011): "A Labor Supply Elasticity Accord?" *American Economic Review*, 101, 487–491.

Low, H. (2005): "Self-insurance in a Life-cycle Model of Labour Supply and Savings," *Review of Economic Dynamics*, 8, 945–975.

MaCurdy, T. (1981): "An Empirical Model of Labour Supply in a Life Cycle Setting," *Journal of Political Economy*, 89, 1059–1085.

——— (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.

Mariano, R. and T. Sawa (1972): "The Exact Finite-Sample Distribution of the Limited-Information Maximum Likelihood Estimator in the Case of Two Included Endogenous Variables," *Journal of the American Statistical Association*, 67, 159–163.

Meghir, C. and D. Phillips (2008): "Labour Supply and Taxes," *IZA WP*, 3405.

Nelson, C. R. and R. Startz (1990): "Some Further Results on the Exact Small Sample Properties of the Instumental Variables Estimator," *Econometrica*, 58, 967–976.

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.

Rogerson, R. and J. Wallenius (2009): "Micro and Macro Elasticities in a Life Cycle Model with Taxes," *Journal of Economic Theory*, 144, 2277–2292.

Sawa, T. (1972): "Finite-Sample Properties of the k-Class Estimators," *Econometrica*, 40, 653–680.

Staiger, D. and J. H. Stock (1997): "IV Regression with Weak Instruments," *Econometrica*, 65, 557–586.

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, chap. 6, 80–108.

Ziliak, J. P. and T. J. Kniesner (1999): "Estimating Life-cycle Labor Supply Tax Effects," *Journal of Political Economy*, 107, 326–359.

Online Appendix A (to section 2): Derivation of Elasticities

Marshallian and Hicksian Elasticities

$$y_t = \left(A_{h,t} + y^m_{h,t} - F\left(a_{h,t}\right) P_{h,t}\right) - \frac{A_{h,t+1}}{1 + r_{t+1}} \tag{22}$$

$$c_t + w_t l_t = y_t + w_t L \tag{23}$$

$$w_{h,t} = \frac{u_{l_{h,t}}}{u_{c_{h,t}}} = \alpha_{h,t} \frac{l^{-\theta}_{h,t}}{c^{-\phi}_{h,t}} \tag{24}$$

Taking the derivative of the budget constraint and the MRS equation and stacking them gives a matrix equation:

$$\begin{bmatrix} 1 & \frac{wl}{c} \\ \phi & -\theta \end{bmatrix} \begin{bmatrix} \frac{\partial \ln c}{\partial \ln w} \\ \frac{\partial \ln l}{\partial \ln w} \end{bmatrix} = \begin{bmatrix} \frac{w(L-l)}{c} \\ 1 \end{bmatrix}$$

This can be inverted to give the Marshallian elasticities in the main text (equation 15):

$$\varepsilon^M_c = \frac{\partial \ln c}{\partial \ln w} = \frac{\theta w\left(L-l\right) + wl}{\theta c + \phi wl} \tag{25}$$

$$\varepsilon^M_l = \frac{\partial \ln l}{\partial \ln w} = \frac{\phi w\left(L-l\right) - c}{\theta c + \phi wl} \tag{26}$$

$$\varepsilon^M_h = \frac{\partial \ln h}{\partial \ln w} = -\left(\frac{\phi w\left(L-l\right) - c}{\theta c + \phi wl}\right) \frac{l}{L-l} \tag{27}$$

To calculate the Hicksian elasticities, we first calculate the income elasticities by differentiating the MRS equation and the budget constraint with respect to income:

$$\varepsilon^y_c = \frac{\partial \ln c}{\partial \ln y} = \frac{\theta y}{\theta c + \phi wl} \tag{28}$$

$$\varepsilon^y_l = \frac{\partial \ln l}{\partial \ln y} = \frac{\phi y}{\theta c + \phi wl} \tag{29}$$

The income elasticity and the marshallian elasticity are then used to calculate the Hicksian elasticity using the Slutsky equation:

$$\begin{aligned} \varepsilon^H_c &= \varepsilon^M_c + \frac{\partial \ln c}{\partial \ln y} \frac{wl}{(c+wl)} \\ \varepsilon^H_l &= \varepsilon^M_l - \frac{\partial \ln l}{\partial \ln y} \frac{w(L-l)}{(c+wl)} \end{aligned}$$

Frisch Elasticities

In this section we provide the formulae for the first and second derivatives that are used to calculate the different elasticities. We define $D = \exp(\pi z + \xi P + \zeta)$ (omitting subscripts for convenience). Then it is easy to show that:

$$u_c(c,l) = DM^{-\gamma}c^{-\phi} \tag{30}$$

$$u_l(c,l) = D\alpha M^{-\gamma}l^{-\theta} \tag{31}$$

$$u_{cl}(c,l) = (-\gamma)DM^{-\gamma-1}\alpha c^{-\phi}l^{-\theta} \tag{32}$$

$$u_{ll}(c,l) = (-\gamma)\frac{u_l(c,l)}{\alpha M}l^{-\theta} - u_l(c,l)\,\theta l^{-1} \tag{33}$$

$$u_{cc}(c,l) = (-\gamma)\frac{u_c(c,l)}{M}c^{-\phi} - u_c(c,l)\,\phi c^{-1} \tag{34}$$

Finally, note that:

$$u_{cl}(c,l) = (-\gamma)u_c(c,l)\,l^{-\theta}\frac{\alpha}{M} = (-\gamma)u_l(c,l)\,c^{-\phi}\frac{1}{M} \tag{35}$$

These expressions can be used to calculate the Frisch elasticities in the paper. The formula for the wage Frisch for intensive margin choices can be derived as follows:

$$\begin{bmatrix} u_{cc} & u_{cl} \\ u_{cl} & u_{ll} \end{bmatrix} \begin{bmatrix} \frac{\partial c}{\partial w} \\ \frac{\partial l}{\partial w} \end{bmatrix} = \begin{bmatrix} 0 \\ u_c \end{bmatrix}$$

$$\begin{bmatrix} \frac{\partial c}{\partial w} \\ \frac{\partial l}{\partial w} \end{bmatrix} = \begin{bmatrix} u_{cc} & u_{cl} \\ u_{cl} & u_{ll} \end{bmatrix}^{-1} \begin{bmatrix} 0 \\ u_c \end{bmatrix}$$

$$\begin{bmatrix} \frac{\partial c}{\partial w} \\ \frac{\partial l}{\partial w} \end{bmatrix} = \frac{1}{u_{cc}u_{ll} - u_{cl}^2} \begin{bmatrix} u_{ll} & -u_{cl} \\ -u_{cl} & u_{cc} \end{bmatrix} \begin{bmatrix} 0 \\ u_c \end{bmatrix}$$

$$\varepsilon_c^F = \frac{w}{c}\frac{\partial c}{\partial w} = -\frac{u_c u_{cl}}{u_{cc}u_{ll} - u_{cl}^2}\frac{w}{c} \tag{36}$$

$$\varepsilon_l^F = \frac{w}{l}\frac{\partial l}{\partial w} = \frac{u_c u_{cc}}{u_{cc}u_{ll} - u_{cl}^2}\frac{w}{l} \tag{37}$$

$$\varepsilon_h^F = \frac{w}{L-l}\frac{\partial(L-l)}{\partial l}\frac{\partial l}{\partial w} = -\frac{u_c u_{cc}}{u_{cc}u_{ll} - u_{cl}^2}\frac{w}{L-l} = -\varepsilon_l\frac{l}{L-l} \tag{38}$$

The formula for the interest-rate Frisch can similarly be derived as follows:

$$\begin{bmatrix} u_{cc} & u_{cl} \\ u_{cl} & u_{ll} \end{bmatrix} \begin{bmatrix} \frac{\partial c}{\partial(1+R_{t+1})} \\ \frac{\partial l}{\partial(1+R_{t+1})} \end{bmatrix} = \begin{bmatrix} u_c \\ u_l \end{bmatrix}$$

$$\begin{bmatrix} \frac{\partial c}{\partial(1+R_{t+1})} \\ \frac{\partial l}{\partial(1+R_{t+1})} \end{bmatrix} = \begin{bmatrix} u_{cc} & u_{cl} \\ u_{cl} & u_{ll} \end{bmatrix}^{-1} \begin{bmatrix} u_c \\ u_l \end{bmatrix}$$

$$\begin{bmatrix} \frac{\partial c}{\partial(1+R_{t+1})} \\ \frac{\partial l}{\partial(1+R_{t+1})} \end{bmatrix} = \frac{1}{u_{cc}u_{ll} - u_{cl}^2} \begin{bmatrix} u_{ll} & -u_{cl} \\ -u_{cl} & u_{cc} \end{bmatrix} \begin{bmatrix} u_c \\ u_l \end{bmatrix}$$

$$\varepsilon_c^{F_R} = \frac{(1+R_{t+1})}{c} \frac{\partial c}{\partial(1+R_{t+1})} = \frac{u_c u_{ll} - u_l u_{cl}}{(1+R_{t+1})(u_{cc}u_{ll} - u_{cl}^2)} \frac{1+R_{t+1}}{c} = \frac{u_c u_{ll} - u_l u_{cl}}{c(u_{cc}u_{ll} - u_{cl}^2)} \tag{39}$$

$$\varepsilon_l^{F_R} = \frac{(1+R_{t+1})}{l} \frac{\partial l}{\partial(1+R_{t+1})} = \frac{u_l u_{cc} - u_c u_{cl}}{(1+R_{t+1})(u_{cc}u_{ll} - u_{cl}^2)} \frac{1+R_{t+1}}{c} = \frac{u_l u_{cc} - u_c u_{cl}}{c(u_{cc}u_{ll} - u_{cl}^2)} \tag{40}$$

$$\varepsilon_h^{F_R} = \frac{(1+R_{t+1})}{L-l} \frac{\partial(L-l)}{\partial l} \frac{\partial l}{\partial(1+R_{t+1})} = -\frac{u_l u_{cc} - u_c u_{cl}}{(1+R_{t+1})(u_{cc}u_{ll} - u_{cl}^2)} \frac{1+R_{t+1}}{L-l} = -\varepsilon_l^{F_R} \frac{l}{L-l} \tag{41}$$

Online Appendix B (to section 3): Estimation Strategy and Solution Method

Fuller's estimator

When estimating the parameters that determine the MRS or those that enter the Euler equation, we use first order conditions to derive restrictions on the data to identify structural parameters. Although these sets of conditions are different, as one set is static in nature and one set is dynamic, they are of a similar nature, in that they can be reduced to an expression of the type

$$E\left[h(X;\theta)\mathcal{Z}\right] = 0 \tag{42}$$

where $h(\cdot)$ is a function of data X and parameters, θ, and is linear in the vector of parameters. The vector $\mathcal{Z}$ contains observable variables that will be assumed to be orthogonal to h. The nature of the instruments that deliver identification depends on the nature of the residual h and, as we discuss below, is different when we estimate the MRS conditions or the Euler equations. However, in both cases, we exploit a condition such as (42).

In equation (42), one needs to normalise one of the parameters to 1. In the context of the MRS equation (18), for example, we set the coefficient on $\ln w_{h,t}$ to 1, but we could have set the coefficient on $\ln l_{h,t}$, or that on $\ln c_{h,t}$ to be 1. A well-known issue with many estimators in this class is that in small samples they are not necessarily robust to the normalisation used. A number of alternative estimators that avoid this issue are available, ranging from LIML-type estimators, to the estimator discussed in Alonso-Borrego and Arellano (1999), to the iterated GMM proposed by Hansen et al. (1996). We use the estimator proposed by Fuller (1977) to estimate both our MRS and Euler equations. This estimator is a modified version of LIML with an adjustment that is designed to ensure that it has finite moments. Roughly speaking, it can be thought of as a compromise between LIML and 2SLS (being closer to LIML when the sample size is large relative to the number of instruments). While this estimator is not completely normalisation free, it is much less sensitive to the choice of normalisation than estimators such as 2SLS and GMM.

An additional advantage of the Fuller estimator is that it is known to have better bias properties than estimators such as 2SLS, when instruments are relatively weak. In section 3, we test the strength of our instruments comparing the values of the Cragg-Donald test statistic to the relevant entries of the table supplied in Stock and Yogo (2005).[27] For the Fuller estimator that we employ, these critical values are typically lower than those for 2SLS, and, unlike 2SLS, they are decreasing in the number of instruments used.

[27] The Cragg-Donald statistic is usually used to provide a test of underidentification. Stock and Yogo (2005) propose using it as a test of instrument relevance as well.

Euler Equation Estimation with Repeated Cross-Sections

We estimate the intertemporal parameters using the Euler equation. We need a long time series because, even under rational expectations, expectations errors do not necessarily average out to zero (or are uncorrelated with available information) in the cross section, but only in the time series: expectation errors may be correlated with available information in the cross section in the presence of aggregate shocks. See the discussion in Hayashi (1987), Attanasio (1999), or Attanasio and Weber (2010). We also need to 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 can in part test by checking over-identifying restrictions.

We estimate equation (21) using the Consumer Expenditure Survey (CEX). Although the CEX covers many years, each household is only observed for a few quarters and so we use a synthetic cohort approach (see Browning et al. (1985)): we aggregate equation (21) over groups with constant membership and follow the average behaviour of the variables of interest (or their non-linear transformation) for such groups. A time series of quarterly cross sections can be used to construct consistent estimates of these aggregates and, in this fashion, use a long time period to estimate the parameters of the Euler equation and test its validity.

We define groups using married couples in ten year birth-cohorts. The assumption of constant membership of these groups might be questioned at the beginning and at the end of the life-cycle for a variety of reasons, including differential rates of family formation, differential mortality and so on. To avoid these and other issues, we limit our sample to households whose husband is aged between 25 and 65 and where wives are aged between 25 and 60.[28]

Having identified groups, we aggregate equation (21) to be estimated across group g households. For this approach to work, however, 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 non-linear 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.[29] These estimates can be used to construct consistent estimates of $M_{h,t}$, which can be aggregated across households to give $M_{g,t}$.

We can obtain consistent estimates of the grouped variables from the time series of cross sections,

[28] If credit constraints are binding, the Euler equation will not be holding as an equality. The youngest consumers are excluded because they are more likely to be affected by this issue. For older consumers, in addition to changes in labour force participation and family composition, health status also changes in complex ways that may be difficult to capture with the taste shifters that we have been considering.

[29] $M_{h,t}$ 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 $M_{h,t}$.

giving the group average log-linear Euler equation:

$$\widetilde{\eta}_{g,t+1} = \overline{\kappa} + \ln\beta + \ln(1+r_{t+1}) - \phi\Delta\overline{\ln c_{g,t+1}} + -\gamma\Delta\ln(\overline{\widehat{M}_{g,t+1}}) + \varphi\Delta\overline{P_{g,t+1}} + \pi\Delta\overline{z_{g,t+1}} \tag{43}$$

The residual term $\widetilde{\eta}_{g,t+1}$ now includes, in addition to the average of the expectation errors and of the changes in taste shifters, several other terms: (i) a linear combination of the difference between the population and sample averages at time t and $t+1$ for all the relevant variables (induced by the fact that we are considering sample means rather than population means for group g); (ii) the difference between the (consistently) estimated $M_{g,t}$ and its actual value (induced by estimation error in the parameters of the MRS); (iii) the difference between the innovation over time to the average value of $\kappa_{g,t}$, which we have denoted with the constant $\overline{\kappa}$.

All the variables on the right hand side of equation (43) are observable. We can therefore use this equation to estimate the parameters of interest. However, the instruments need to be uncorrelated with $\widetilde{\eta}_{g,t+1}$.[30] The covariance structure of the $\widetilde{\eta}_{g,t+1}$ is quite complex: the contemporaneous covariance of $\widetilde{\eta}_{g_i,t+1}$ and $\widetilde{\eta}_{g_j,t+1}$ is not, in general, zero, as aggregate shocks have effects that correlate across different groups. When computing the variance-covariance matrix of the estimates, this structure should be taken into account. Whilst it is in principle possible, given our assumptions, to estimate the variance-covariance matrix of $\widetilde{\eta}_{g,t+1}$ from estimated parameters, in practice it turns out to be cumbersome, as there is no guarantee that, in small samples, these estimates are positive-definite. Given these difficulties, we follow a different and, as far as we know, novel approach, based on bootstrapping our sample, with a structure consistent with the basic assumptions of our model. We describe the bootstrapping procedure in detail in the next subsection.

Bootstrap Procedure

We bootstrap standard errors and confidence intervals for both our MRS and Euler equations.

The two step Heckman-selection procedure for estimating the MRS coefficients can be bootstrapped in the standard way. Bootstrapping results for our Euler equation requires a slightly more complicated procedure however. This is because we aggregate our data into cohort groups and then implement an IV procedure. Taking Z_t as a vector of exogenous variables, and X_t and Y_t as endogenous variables (with Y_t as our dependent variable) we can reformulate our approach as estimating the equations

$$\begin{aligned} X_t &= \Pi Z_t + v_t \\ Y_t &= X_t\beta + u_t \end{aligned}$$

[30] As noted by Deaton (1985) and discussed extensively in the context of the CEX by Attanasio and Weber (1995), the use of sample rather than population averages for all the 'group' variables induces an MA(1) in the residuals, because of the sampling variation in the rotating panel structure. We need to assume that the instruments are not correlated with the (average) estimation error of the $M_{h,t}$ or with the innovations to the higher moments of the expectation errors $(\kappa_{g,t} - \overline{\kappa})$. This last assumption is discussed in Attanasio and Low (2004).

where v_t is a vector of errors in our first stage. These can be thought of as economic shocks which may have a complicated structure. For instance they may be correlated across time for a given cohort, or may have an aggregate component which is correlated across cohorts for a given time period. Errors may also be correlated across the equations for different exogenous variables Z_t. We will wish to preserve these correlations when we implement our bootstrap procedure. In order to do this, we attempt to construct the variance-covariance matrix of the residuals v. Rather than filling in all possible cross-correlations in this matrix, we calculate the following moments for each cohort c, and equation i

$$\begin{aligned}
&var(v^{i,c}) \\
&cov(v_t^{i,c}, v_{t-1}^{i,c}) \\
&cov(v_t^{i,c}, v_t^{j,c}) \\
&cov(v_t^{i,c}, v_t^{i,k})
\end{aligned}$$

Setting all other correlations to zero. Thus we impose for instance that there is zero correlation between $v_t^{i,c}$ and $v_{t-1}^{i,k}$. Unfortunately, there is no guarantee that this matrix will be positive definite. In our procedure we therefore apply weights to the non-zero elements of our 'off-diagonal' matrices - which give the covariances across different cohorts for the same equation - and to our 1st autocovariances for residuals for the same cohort and same equation. The weights we apply to these are the maximum that ensure the resulting matrix is positive definite: in our case they are both set at 0.23.

Once we have this matrix we can Cholesky decompose it to obtain a vector of orthogonalised residuals

$$\Omega = vv' = \epsilon CC'\epsilon'$$

We then draw from the orthogonalised residuals, premultiply them by C and then add them to ΠZ_t to reconstruct the endogenous variables (including Y). We then reestimate our second stage equation to obtain a new set of estimates for β.

The values of Z_t in our case will depend on the results we obtain from our MRS equation, so in each iteration of our bootstrap we resample with replacement from from our disaggregated data, re-run the MRS equation, reaggregate to obtain the cohort averages which make up Z_t and then make a draw from our residuals.

Solution Method

Households have a finite horizon and so the model is solved numerically by backward recursion from the terminal period. At each age we solve the value function and optimal policy rule, given the current state variables and the solution to the value function in the next period. This approach is standard. The complication in our model arises from the combination of a discrete choice (to participate or not) and a continuous choice (over saving). This combination means that the value function will not necessarily be concave. We briefly describe in this appendix how we deal with this potential non-concavity. An alternative would be to follow the method in Iskhakov et al. (2017).

In addition to age, there are four state variables in this problem: the asset stock, the permanent component of earnings of the husband, $v^m_{h,t}$, the permanent component of wife's wage, $v^f_{h,t}$,and the experience level of the wife. We discretise both earning and wage variables and the experience level, leaving the asset stock as the only continuous state variable. Since both permanent components of earnings are non-stationary, we are able to approximate this by a stationary, discrete process only because of the finite horizon of the process. We select the nodes to match the paths of the mean shock and the unconditional variance over the life-cycle. In particular, the unconditional variance of the permanent component must increase linearly with age, with the slope given by the conditional variance of the permanent shock. Our estimates of the wage variance are for annual shocks, but the model period is one quarter. We reconcile this difference by imposing that each quarter an individual receives a productivity shock with probability 0.25, and this implies that productivity shocks occur on average once a year.

Value functions are increasing in assets A_t but they are not necessarily concave, even if we condition on labour market status in t. The non-concavity arises because of changes in labour market status in future periods: the slope of the value function is given by the marginal utility of consumption, but this is not monotonic in the asset stock because consumption can decline as assets increase and expected labour market status in future periods changes. By contrast, in Danforth (1979) employment is an absorbing state and so the conditional value function will be concave. Under certainty, the number of kinks in the conditional value function is given by the number of periods of life remaining. If there is enough uncertainty, then changes in work status in the future will be smoothed out leaving the expected value function concave: whether or not an individual will work in $t+1$ at a given A_t depends on the realisation of shocks in $t+1$. Using uncertainty to avoid non-concavities is analogous to the use of lotteries elsewhere in the literature.

The choice of participation status in t is determined by the maximum of the conditional value functions in t. In our solution, we impose and check restrictions on this participation choice. In particular, we use the restriction that the participation decision switches only once as assets increase, conditional on permanent earnings and experience. When this restriction holds, it allows us to interpolate behaviour across the asset grid without losing our ability to determine participation status. We therefore define a reservation asset stock to separate the value function and the choice of consumption made when participating from the value function and choice of consumption made when not participating. There are some regions of the state space where individuals are numerically indifferent between working and not working. Since we solve the model by value function iteration, it does not matter which conditional value function we use in these regions.

In solving the maximisation problem at a given point in the state space, we use a simple golden search method. Note that in addition to the optimal total expenditure, the optimal amount of leisure is computed in each period by solving the MRS condition. We solve the model and do the calibration assuming this process is appropriate and assuming there is a unique reservation asset stock for each point in the state space, and then check ex-post.

There are no non-concavities due to borrowing constraints in our model because the only borrowing constraint is generated by the no-bankruptcy condition which is in effect enforced by having infinite marginal utility of consumption at zero consumption.

Finally, we include here the value functions of the household problem. In each period, if the woman chooses to participate, the value function is given by

$$\begin{gathered} V_{h,t}^{1}\left(A_{h,t},\upsilon_{h,t}\right)= \\ \max_{c_{h,t},l_{h,t}}\left\{u\left(c_{h,t},l_{h,t},P_{h,t}=1\right)+\ \ \beta E_{t}\left[\max\left\{\begin{array}{c} V_{h,t+1}^{0}\left(A_{h,t+1},\upsilon_{h,t+1}\right) \\ V_{h,t+1}^{1}\left(A_{h,t+1},\upsilon_{h,t+1}\right)\end{array}\right\}\right]\right\} \end{gathered} \tag{44}$$

Note that the state variable $\upsilon_{h,t}$ is a vector containing the woman and the man's productivity type. If she chooses not to participate, the value function is given by,

$$\begin{gathered} V_{h,t}^{0}\left(A_{h,t},\upsilon_{h,t}\right)= \\ \max_{c_{h,t}}\left\{u\left(c_{h,t},P_{h,t}=0\right)+\ \ \beta E_{t}\left[\max\left\{\begin{array}{c} V_{h,t+1}^{0}\left(A_{h,t+1},\upsilon_{h,t+1}\right) \\ V_{h,t+1}^{1}\left(A_{h,t+1},\upsilon_{h,t+1}\right)\end{array}\right\}\right]\right\} \end{gathered} \tag{45}$$

The decision of whether or not to participate in period t is determined by comparing $V_{h,t}^{0}\left(A_{h,t},\upsilon_{h,t}\right)$ and $V_{h,t}^{1}\left(A_{h,t},\upsilon_{h,t}\right)$. The participation choice, the hours choice and the consumption choice in t determines the endogenous state variable (assets) at the start of the next period.

Online Appendix C (to section 4): Data Sources and Descriptive Statistics

As discussed in the paper, most of the data are from the CEX. One important exception are the data on the real interest rate. We define this variable as the 3 month T-Bill rate (on a quarterly basis) minus the rate of growth in the CPI. The source for the T-Bill rate is from the St Louis Fed (https://fred.stlouisfed.org/series/TB3MS).

In for Table 12 presents descriptive statistics at the individual level using data from three particular years (1980, 1995 and 2012). Married women have seen large changes in their wages, hours and patterns of employment over our sample period. Employment rates increased from 60% in 1980 to 69.8% in 1995 before falling back to 61.9% in 2012.

Table 12: Descriptive statistics for married women, 1980, 1995 and 2012

		1980	1995	2012
Demographics	No. of children	1.25	1.15	1.17
Education	% Less than high school	19.4	12.3	9.7
	% High school	44.1	36.8	25.3
	% Some college	18.1	25.3	28.5
	% Degree or higher	18.4	25.5	36.5
Hours (workers)	All	35.2	37.5	38.4
	Less than high school	34.9	37.4	34.2
	High school	35.2	36.2	38.6
	Some college	35.0	36.7	37.1
	Degree or higher	35.5	39.7	39.5
Hourly net wages ($ 2016)	All	15.58	16.63	18.95
	Less than high school	12.16	11.23	11.33
	High school	14.22	13.41	14.62
	Some college	16.62	16.41	17.28
	Degree or higher	19.30	22.26	23.20
% Employed	All workers	60.0	69.8	61.9
	% Workers part-time	28.4	23.7	20.6
Sample sizes	All	2,199	2,064	2,026
	Workers	1,318	1,441	1,254

Part-time is defined as working less than 35 hours per week.

Table 12 also shows wage levels over the three years. Average real wages increased over this period, though with marked differences across different education groups. The wages of those with less than high school education actually fell slightly from \$12.16 in 1980 to \$11.33 in 2012. By contrast, married women with a college degree or higher saw a 20% increase in their wages between 1980 and 2012 (from \$19.30 to \$23.20). This increase in the education premium has been attributed to skill-biased technological change which outstripped the supply of educated workers (Goldin and Katz, 2007).

Changes in hours worked across education groups appear to mirror these patterns. While all education groups worked very similar hours in 1980, by 2012 those with a college degree were working on average five hours more per week than those with less than high school education, although the fraction with a college degree has markedly increased over the period.

Online Appendix D (to section 5)

Alternative methods of estimating the MRS

In this appendix we discuss results from alternative MRS specifications. For comparison with later results, we present a more complete set of parameter estimates from our baseline MRS specification in Table 14. First, we present results for the selection probit we run prior to estimating our MRS equation. Husband's earnings are strongly negatively correlated with participation.

Table 13: Selection Probit Results

Log earnings of husband	-0.164***	(0.007)
Husband employed	-1.929***	(0.064)
No. of Elderly HH members	0.023	(0.026)
Log family size	-0.110***	(0.022)
Wife: White	-0.015	(0.014)
Age	-0.056	(0.042)
Age^2	0.001	(0.001)
Age^3 /1000	0.003	(0.018)
Age^4 /10000	-0.003*	(0.001)
Has kids	-0.034	(0.018)
No. of kids aged 0-2	-0.515***	(0.014)
No. of kids aged 3-15	-0.167***	(0.008)
No. of kids aged 16-17	0.071***	(0.017)
North East	-0.004	(0.015)
Mid-West	0.119***	(0.014)
South	0.035**	(0.013)

N= 78,674. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$ Standard errors in parentheses. Additional controls for season and year dummies and cohort-education interactions.

Estimation method and normalisation

We start by considering the issue of how the MRS is normalised. Recall that our MRS relationship is

$$\ln w_{h,t} = \psi_0 + \psi z_{h,t} - \theta \ln l_{h,t} + \phi \ln c_{h,t} + \upsilon_{h,t} \tag{46}$$

As Keane (2011) notes, this is not a labour supply equation but an equilibrium condition in which wages, leisure and consumption are all endogenous. All three variables are potentially correlated with the error term $\upsilon_{h,t}$ and so there is no natural choice of the dependent variable.

Despite this, we find that, when conventional methods are used, results can be highly sensitive to whether wages, leisure or consumption are placed on the left hand side of the MRS equation. Table 15 shows results from estimating ϕ and θ using GMM under the three different possible normalisations.

Table 14: Baseline MRS estimates

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			
Age	0.05**	(0.02)	[0.01,0.09]
Age^2	-0.0005	(0.0007)	[-0.002,0.001]
Age^3/1000	-0.01	(0.01)	[-0.03,0.01]
Age^4/10000	0.002**	0.0007	[0.0002,0.003]
North East	0.01	(0.03)	[-0.02,0.08]
Mid West	-0.05**	(0.01)	[-0.07,-0.02]
South	-0.11***	(0.02)	[-0.18,-0.09]
White	-0.04	(0.03)	[-0.09,0.04]
No. elderly HH members	0.02	(0.02)	[-0.02,0.05]
ln($famsize$)	-0.32***	(0.037)	[-0.38,-0.23]
Has kids	0.07***	(0.021)	[0.04, 0.12]
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]
Constant (Ψ_0)	4.70	(4.94)	[-1.19,18.52]
Heckman selection terms			
e_1	0.07	(0.167)	[-0.18, 0.48]
e_2	0.05	(0.172)	[-0.21, 0.51]
e_3	0.01	(0.052)	[-0.08, 0.13]

N = 50,895. *p<0.10, ** p<0.05, *** p<0.01. Additional controls for season and year dummies and cohort-education interactions. Confidence intervals are bootstrapped with 1000 replications allowing for clustering at the individual level.

We include results form our baseline specification in the first column. The implied parameter estimates and elasticities vary a great deal across these different approaches. When wages are selected as the left-hand side variable, elasticites are relatively large. When leisure is the dependent variable, they are much smaller. Very similar considerations apply to the estimation of our Euler equation.

Differences of this kind can emerge in IV estimation in 2SLS and GMM estimation when the instruments chosen are relatively weak. Indeed, Hahn and Hausman (2003) propose using the differences in parameters implied by 2SLS estimates run under different normalisations as a test of instruments' strength.

Various papers have discussed possible remedies for cases when strong instruments are not available (Hahn and Hausman, 2003; Hausman et al., 2012). One possible solution is the use estimators such as Limited Information Maximum Likelihood (LIML) rather than 2SLS, which is known to have poor

Table 15: MRS Estimates using GMM

	Fuller	GMM		
Dependent variable:	Wages	Wages	Leisure	Consumption
Parameters				
θ	1.75** [0.34,5.12]	0.46* [−0.04,0.61]	13.8 [−120.13,186.11]	0.13 [−0.54,0.58]
ϕ	0.76*** [0.55,0.95]	0.61*** [0.48,0.66]	0.17 [−3.44,2.78]	1.38*** [1.24,1.74]
Wage elasticities at median				
Marshallian	0.18 [0.05,0.38]	0.55 [0.50,1.16]	0.09 [0.00,0.12]	−0.17 [−0.41,−0.08]
Hicksian	0.54 [0.27,1.29]	1.19 [1.10,2.25]	0.11 [−0.01,0.14]	0.77 [0.59,1.10]

N = 50,895. *p<0.10, ** p<0.05, *** p<0.01. Controls as in Table 3. Elasticities are calculated as averages within a 5 percent band of the 50th percentile of the Marshallian distribution. 95% confidence intervals in square brackets. Confidence intervals are bootstrapped with 1000 replications.

bias properties in such circumstances (Staiger and Stock, 1997; Nelson and Startz, 1990). Using the notation from Davidson and MacKinnon (2004), for the case where

$$\begin{aligned} y &= Z\beta_1 + Y\beta_2 + u = X\beta + u \\ Y &= \Pi W + v \end{aligned}$$

where Z is a matrix of exogenous variables, Y a matrix of endogenous variables, and $W = [Z, W_1]$ (with W_1 being a matrix of instruments). Matrices X and W are $n \times k$ and $n \times l$ respectively (with $l \geq k$). In general, so-called k-class estimators such as OLS, 2SLS, and LIML can be written in the form

$$\hat{\beta}^{\text{LIML}} = (X'(I - kM_W)X)^{-1}X'(I - kM_W)y \tag{47}$$

where $M_W = I - W(W'W)W'$. In the case of OLS $k = 0$, and in the case of 2SLS $k = 1$. In the case of LIML we use

$$k = k_{\text{LIML}} = \frac{(y - Y\beta_2)'M_Z(y - Y\beta 2)}{(y - Y\beta_2)'M_W(y - Y\beta 2)} \tag{48}$$

While LIML is often found to have better bias properties than 2SLS, it has long been recognised that conventional normalisations of LIML do not have finite moments (Mariano and Sawa, 1972; Sawa, 1972), and simulation exercises have shown that this can add considerable volatility to empirical estimates (Hahn et al., 2004). As a result Hahn et al. (2004) recommend the use of either jack-knifed

2SLS or the modification of LIML proposed by Fuller (1977). For this latter estimator, we replace k in equation (47) with

$$k_{\text{Fuller}} = k_{\text{LIML}} - \frac{\lambda}{(n-k)} \tag{49}$$

where λ here is a parameter chosen by the researcher, to obtain a value for $\hat{\beta}^{\text{Fuller}}$. We choose a value of one for this as suggested by Davidson and MacKinnon (2004) as it yields estimates that are approximately unbiased. The resulting estimator is guaranteed to have bounded moments in finite samples Fuller (1977). Since the adjustment to LIML is smaller when $(n-k)$ is large, the Fuller estimator will be closer to LIML when sample sizes are large relative to the number of instruments. In our case, the Fuller estimator can be thought of as a compromise between 2SLS and LIML, as it adjusts the value of k we use downwards slightly towards one.

As well as its superior bias properties, the Fuller estimator has the advantage that is much less sensitive than GMM or 2SLS to the choice of the dependent variable, as Table 16 shows. Both the elasticity and parameter estimates obtained using alternative normalisations of the Fuller estimator are very similar to our baseline results.

Table 16: MRS Estimates with Different Dependent Variables

	Dependent variable		
	Wages	Leisure	Consumption
Parameters			
θ	1.75** [0.34,5.12]	1.84* [−0.43,5.38]	1.75* [−0.00,4.60]
ϕ	0.76*** [0.55,0.95]	0.76*** [0.53,0.95]	0.77*** [0.58,0.95]
Wage elasticities at median			
Marshallian	0.18 [0.05,0.38]	0.17 [0.06,0.37]	0.18 [0.07,0.42]
Hicksian	0.54 [0.24,1.07]	0.53 [0.23,0.95]	0.54 [0.27,1.29]

N = 50,895. *p<0.10, ** p<0.05, *** p<0.01. Controls as in Table 3. Elasticities are calculated as averages within a 5 percent band of the 50th percentile of the Marshallian distribution. 95% confidence intervals in square brackets. Confidence intervals are bootstrapped with 1000 replications.

Alternative instruments

In Table 17 we show results using alternative choices of instruments. We show results using GMM (with wages as the dependent variable) and the Fuller estimator described above, in both cases using a *full* set of cohort-education-year interactions as used in Blundell et al. (1998). This approach is similar to the approach we adopt for our main results but interacts cohort-eduction dummies full set of year effects rather than a polynomial in time trends. The estimates we obtain from fully adopting the Blundell et al. (1998) approach are very similar to our main results, though somewhat less precise.

The sensitivity of our results to the choice of instruments is on the whole quite small when we compare it to the differences that can arise from the choice of estimation method. Just as we find for our main set of results, the hours elasticities estimated using the GMM estimator with wages as the dependent variable are substantially larger than those using the Fuller estimator when using the alternative instrument set.

Table 17: MRS Estimates using Alternative Instruments

	Fuller	GMM
Parameters		
θ	1.93 [−9.09,11.58]	0.08 [−0.20,0.19]
ϕ	0.76*** [0.42,1.03]	0.52*** [0.41,0.52]
Wage elasticities at median		
Marshallian	0.17 [−0.84,1.13]	1.08 [1.04,2.29]
Hicksian	0.51 [−1.91,2.65]	1.97 [1.82,3.85]

N = 50,895. *p<0.10, ** p<0.05, *** p<0.01. We use a full set of cohort-education-year dummies as instruments, following Blundell, Duncan and Meghir (1998). Controls as in Table 3. Elasticities are calculated as averages within a 5 percent band of the 50th percentile of the Marshallian distribution. 95% confidence intervals in square brackets. Confidence intervals are bootstrapped with 1000 replications.

Alternative samples

Table 18 shows how our MRS results are affected by alternative sample selection choices. Column (1) presents results when we exclude those individuals who report working exactly 40 hours a week. The justification of this experiment is that these individuals may be affected by some kind of friction that does not allow them to adjust their hours worked as desired. Such frictions would mean that the MRS condition that we exploit to recover ϕ and θ need not hold. Excluding these observations, we obtain greater estimates of our Marshallian and Hicksian hours elasticities (at 0.45 and 0.72 respectively). These values are however somewhat imprecisely estimated and the confidence bands that surround them include our baseline estimates.

In Column (2) we show results when we exclude individuals working less than 20 hours per week (with an appropriate adjustment to our selection correction). We consider results from this specification because there may be certain frictions that prevent individuals working fewer hours than this, which would again lead to potential violations of the MRS condition. Excluding these observations delivers somewhat lower elasticity estimates, but again the estimates are imprecise.

Table 18: MRS Estimates using alternative samples/hours measures

	Exc. 40 hours (1)	Exc. <20 hours (2)	Born 1925-1965 (3)	Ann. hours (4)
Parameters				
θ	1.52 [−3.16,5.69]	2.81 [−2.48,9.81]	2.08*** [0.68,4.66]	2.30* [−1.30,7.00]
ϕ	0.42* [−0.05,0.92]	0.76*** [0.33,1.03]	0.56*** [0.42,0.82]	0.78*** [0.53,1.01]
Wage elasticities at median				
Marshallian	0.45 [−0.45,1.41]	0.13 [0.02,0.30]	0.27 [0.09,0.62]	0.13 [−0.09,0.41]
Hicksian	0.72 [−1.29,2.57]	0.39 [−0.14,1.43]	0.53 [0.27,1.09]	0.42 [−0.25,1.24]
N	26,060	47,743	39,057	50,895

*p<0.10, ** p<0.05, *** p<0.01. Specification (1) excludes individuals who work exactly 40 hours. Specification (2) excludes those working less than 20 hours (part-time workers). Specification (3) only includes individuals from cohorts with the most similar labour supply choices over the life-cycle. Elasticities are calculated as averages within a 5 percent band of the 50th percentile of the Marshallian distribution. 95% confidence intervals in square brackets. Confidence intervals are bootstrapped with 1000 replications.

Finally, in Column (3) we consider only those ten-year birth cohorts with the most similar labour-supply behaviour over the life-cycle. In particular we exclude those born before 1925 as they tend to work fewer hours at older ages than other cohorts, and those born after 1975, as less-educated individuals born after this date tend to have lower employment rates than other earlier cohorts at the same ages. Using this sample, we obtain a Marshallian elasticity of 0.27 and a Hicksian 0.53. While the Marshallian elasticity estimated from this sample is slightly higher than our baseline estimates, the Hicksian elasticity is essentially unchanged.

Alternative definitions of hours

In Column (4) of Table 18 we consider how elasticity estimates are affected when we use an alternative measure of hours of leisure. The measure we use here is

$$\text{leisure} = \frac{5200 - \text{hours per week} \times \text{weeks worked per year}}{52} \tag{50}$$

This measure accounts for the observed variation in weeks worked per year in addition to variation in hours worked per week across workers.

The elasticities resulting from this exercise are in general lower but on the whole similar to than those in our baseline specification, with a Marshallian elasticity of 0.13 and a Hicksian elasticity of 0.42. The value of θ is larger than in our main results (at 2.30), and much less precisely estimated. The value of ϕ is essentially unchanged.

External Parameters for the Calibration

Table 19 reports the complete set of estimated and external parameters used in the calibration. The first panel reports the estimated parameters from Tables 3 and 4 above. The second panel reports parameters which come from external sources.

Table 19: External Parameters

Estimated Parameters (from first-order conditions)		
Curvature on leisure	θ	1.75
Curvature on consumption	ϕ	0.76
Curvature on utility	γ	2.07
Exogenous Parameters		
Interest Rate (annual)	r	0.015
Regression Log Wage on Age and Age^2 (Men)	ι_1^m, ι_2^m	0.0684, -0.00065
Husband and Wife Wage Correlation	ρ	0.25
Standard Deviation of Permanent Shock (Men)	σ_{ξ^m}	0.077
Standard Deviation of Initial Wage (Men)	$\sigma_{\xi_0^m}$	0.54
Length of Life (in years)	T	50
Length of Working Life (in years)	R	40

Online Appendix E (to section 6): Results for CES and additive separability

In this Appendix we discuss results for alternative specifications of our utility function. In particular we consider results from a standard CES utility function (where we impose that θ=ϕ), and one where we impose additive separability between consumption and leisure (i.e $\gamma = 0$).

Table 20 presents parameter estimates when we impose the restrictions implied by CES utility. Under this functional form for utility, we get a slightly larger value of ϕ and a much lower value of θ than we obtain from our preference specification (at 0.83 compared to 0.76 and 1.75 that we obtain for ϕ and θ respectively in Table 3). We also obtain a slightly larger value of γ however (at 3.04 compared to 2.07 for our less restrictive utility function).

Table 20: Parameter values

	CES
ϕ	0.83 [0.66,0.97]
θ	0.83 [0.66,0.97]
γ	3.04 [0.64,4.27]

Taken together, the CES parameter estimates imply that utility is less concave in leisure, and hence that labour supply elasticities are greater. We show the elasticities implied by these estimates in Table 21. While Marshallian hours elasticities for the CES specification are only greater at the upper end of the distribution, the estimated Hicksian and Frisch hour elasticities are roughly 50% larger. The CES estimates also imply a more substantial degree of non-separability between consumption and leisure. The Frisch elasticity of consumption with respect to predictable wage increases has a median of around 0.4 compared to 0.05 from our main estimates. This reflects both a greater sensitivity of the marginal utility of consumption to changes in leisure and the fact that leisure responses to given wage changes will in general be greater under these preferences. Finally we note that, the interest rate Frisch elasticity at the median is much lower than in our baseline specification.

Table 21 also shows Frisch elasticities for our preference specification in the case where we impose additive separability for preferences over consumption and leisure (that is we impose that $\gamma = 0$). This necessarily sets the Frisch consumption responses to wage changes to zero. It turns out that Frisch hours elasticities are very similar to those estimated when we allow for non-separability in our main specification. This reflects the fact that when, as we find, the parameters θ and ϕ are small and α large, then the numerator and denominator in formulae for Frisch elasticities given in equations (36) and (37) will be dominated by the term M_t. Consequently, the impact of small changes in γ will be limited.

When additive separability is imposed, the Frisch elasticity is identical - a direct result of setting $u_{cl} = 0$ in expressions (39) and (40). The estimated Frisch elasticity of consumption with respect to the interest rate (now simply given by $-1/\phi$) also falls relative to our baseline results, from a median value of -1.19 in our baseline results to -1.31.

Table 21: Elasticities at Percentiles of Marshallian distribution: CES

	$\gamma = 0$		CES			
	Wage	Interest rate	Wage			Interest rate
	Frisch	Frisch	Marshallian	Hicksian	Frisch	Frisch
	Hours worked		*Hours worked*			
10th	0.84 [0.23,3.17]	0.84 [0.23,3.17]	-0.24 [-0.30,-0.11]	0.48 [0.41,0.60]	1.08 [0.97,1.47]	0.83 [0.63,1.32]
25th	0.83 [0.23,3.15]	0.83 [0.23,3.15]	-0.04 [-0.13,0.12]	0.60 [0.51,0.76]	1.16 [1.06,1.56]	0.85 [0.65,1.35]
50th	0.90 [0.25,3.40]	0.90 [0.25,3.40]	0.21 [0.10,0.42]	0.77 [0.67,0.98]	1.33 [1.23,1.75]	0.93 [0.71,1.47]
75th	1.04 [0.29,3.93]	1.04 [0.29,3.93]	0.54 [0.39,0.82]	1.04 [0.89,1.32]	1.66 [1.53,2.17]	1.13 [0.86,1.78]
90th	1.98 [0.55,7.50]	1.98 [0.55,7.50]	1.11 [0.88,1.55]	1.62 [1.39,2.06]	2.71 [2.51,3.57]	1.89 [1.44,2.99]
	Consumption		*Consumption*			
25th	0.00 [$-$,$-$]	-1.31 [-1.81,-1.05]	0.91 [0.82,1.08]	0.63 [0.54,0.80]	0.32 [0.12,0.53]	-0.59 [-0.93,-0.45]
50th	0.00 [$-$,$-$]	-1.31 [-1.81,-1.05]	1.07 [0.97,1.27]	0.72 [0.62,0.91]	0.37 [0.15,0.62]	-0.58 -0.91,-0.44]
75th	0.00 [$-$,$-$]	-1.31 [-1.81,-1.05]	1.23 [1.12,1.44]	0.80 [0.68,1.01]	0.42 [0.17,0.69s]	-0.57 [-0.90,-0.43]

Elasticities are calculated as averages within 5 percent bands of the 10th, 25th, 50th and 75th and 90th percentiles of the Marshallian distribution. 95% confidence intervals in square brackets. Confidence intervals are bootstrapped with 1000 replications.

Online Appendix F: Returns to Experience

We recalibrate parameter values: the fixed cost of working, $\bar{F}$, child care price, p, the offered wage gender gap and ψ_0. In addition to these parameters, we also need to calibrate the parameter that characterises human capital accumulation function and its depreciation rate.[31] As in the baseline, we identify these parameters by targeting participation rate of women, the participation rate of mothers, the average hours worked, the observed wage gender gap, the observed wage growth at early ages, and the observed depreciation of wages during non-participation (we take this figure from Attanasio et al. (2008)). We report the calibrated parameters in Table 22 and compare them to the baseline. In the context of returns to experience, where there is a strong incentive to work to reap future returns, a much larger childcare cost is required in order to reduce participation and match participation statistics.

Analogously to Figure 2, Figure 4 shows life-cycle profiles in the simulations and in the data; and Table 23 reports additional statistics on the distribution of hours and of wages. There are some differences between the model with returns to experience and the baseline. First, there is a decline in the participation profiles at ages beyond 35. These patterns are not observed either in the data or in the baseline model. Second, very few women change their participation decisions. For example, the fraction of women who worked in all previous periods at the age of 52 is 57%, which compares to 40% in the economy without returns to experience. Third, the childcare cost that is needed here to keep women out of the labour market during childbearing is substantially higher because of the incentive to accumulate labour market experience. In particular the monetary fix childcare cost is up to 76% of median earnings of a women aged 25 to 55.

Response to Temporary Wage Changes

In Table 24, we report the labour supply responses in the economy with returns to experience. The key finding is that, in contrast to the economy without returns to experience, the extensive margin response is close to zero and, as a result, the aggregate elasticity is about half of the one in the baseline economy (reproduced in the final column). In the return to experience economy, there is a strong incentive to participate to obtain the return to experience. The larger childcare cost of participating that is estimated in this economy alongside the strong incentive to participate implies that changes in the current wage makes little difference to the incentive to participate. As expected, the size of the intensive margin response is similar to the one in the economy without returns to experience. Our results here are in line with Imai and Keane (2004) who argue that the response of labour supply to transitory changes in wages may be mitigated when there are returns to experience. Our results show

[31] Note this is only one parameter in contrast to the two parameters ι_1^f and ι_2^f for the exogenous wage growth that were used in the baseline economy.

Table 22: Baseline economy: Calibrated Parameters and Targets

Parameter Name		Values	
		Ret to Exp	Baseline
Constant term weight of leisure	ψ_0	4.13	4.20
Childcare Cost	p	5820	967
Fixed Cost of Working	$\bar{F}$	315	468
Offered Wage Gender Gap at age 22	y_0^f/y_0^m	0.78	0.74
Standard Deviation of Permanent Shock (Women)	σ_{ξ^f}	0.063	0.063
Standard Deviation of Initial Wage (Women)	$\sigma_{\xi^f,0}$	0.50	0.50
Exogenous growth in offered wage	ι_1^f	-	0.052
Exogenous growth in offered wage	ι_2^f	-	-0.0006
Women's Human Capital Tech	ν	0.003	-
Discount Factor (annualized)	β	0.99	0.99
Depreciation rate	δ	0.017	-

Targets	Data	Ret to Exp	Baseline
Weekly hours worked	37.2	37.5	37.2
Participation Rate	0.684	0.690	0.679
Participation Rate of Mothers	0.538	0.544	0.546
Observed Wage Gender Gap	0.720	0.716	0.727
Observed Variance Wage Growth (Women)	0.004	0.005	0.004
Observed Initial Variance of Wages (Women)	0.14	0.15	0.15
Wage Growth (if younger than 40)	0.012	0.013	0.010
Wage Growth (if older than 40)	0.001	0.013	0.004
Median wealth to income ratio	1.84	1.82	1.80
Observed Depreciation Rate	-0.050	-0.040	0.02

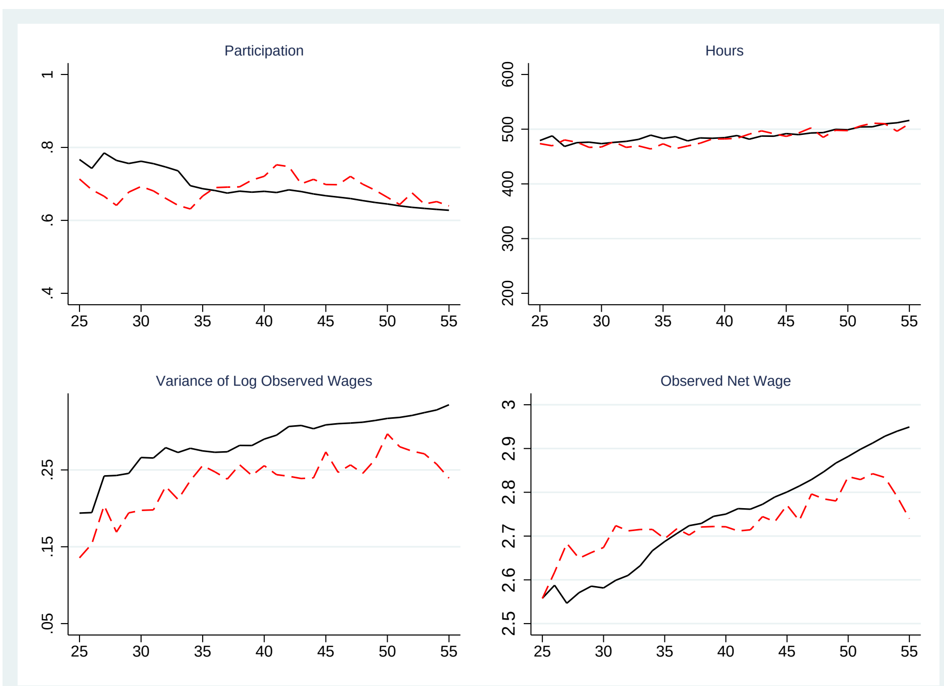
Statistics for women born in the 1950s and aged 25 to 55. Wage growth and depreciation rate are annual.

Table 23: Returns to Experience: Statistics on Hetereogeneity

	Data	Model
Participation Rate Mothers with Children Aged 3-17	0.682	0.672
Participation Rate Childless Women	0.755	0.724
Average Hours Worked 10th Percentile	20	21
Average Hours Worked 25th Percentile	35	31
Average Hours Worked 50th Percentile	40	40
Average Hours Worked 75th Percentile	40	46
Average Hours Worked 90th Percentile	48	50
Wage 10th Percentile	8.16	7.43
Wage 50th Percentile	15.05	15.58
Wage 90th Percentile	29.23	31.71

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

Figure 4: Life-Cycle Profiles: Baseline Model (solid black line) versus Data (dashed red line)



that the response of the participation margin to a transitory anticipated change in the wage (for given preferences on the intensive margin) may be very different depending on the assumption that is made about the nature of the wage growth over the life-cycle (exogenous or endogenous). The extra wage that is provided by an anticipated increase in the wage in a particular period is a small fraction of the total return to participate in that period (in particular at early ages) and then it has a small impact on the participation decision.

Table 24: Returns to experience: Frisch Changes

	Extensive Response	Intensive Elasticity 25th	50th	75th	Agg Hours Elasticity	Baseline
25-29	0.02	0.65	0.81	1.15	0.91	1.85
30-34	0.04	0.63	0.79	1.17	0.91	1.48
35-39	0.03	0.63	0.78	1.17	0.90	1.45
40-44	0.03	0.61	0.79	1.19	0.89	1.35
45-49	0.04	0.60	0.77	1.19	0.88	1.39
50-55	0.07	0.58	0.75	1.09	0.86	1.45

The extensive response is the percentage point change in participation in response to a 1% increase in the wage. The aggregate hours 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.

It may well be that the small response of the extensive margin labour supply that we find is related to the simple model of return to experience we have considered. Whether returns to experience operate in a more subtle manner through intensive margins and the number of hours is a question we leave for future research. If that is the case, we would need to change substantially the estimation methods we used in the core of the paper.

One possibility, of course, is that returns to tenure are important for some occupations and/or skill levels and not for others. In such a case, it would be necessary to introduce an additional dimension of heterogeneity that would make the aggregation issues we have repeatedly stressed even more salient.[32]

Life-Cycle Responses to Changes in Wage Profiles

Finally, in Table 25 we report the extensive, intensive margin and the macro responses to an increase in the entire wage profile of 10% for both husband and wife. In this case the response both at

[32] Alternatively it could be that returns to experience depend on hours worked. Blundell et al. (2016a) show that these returns are close to zero for part-time work.

the extensive and the intensive margin is very similar in the economy with and without returns to experience.

Table 25: Labour supply changes, Marshallian

	Extensive Response	Intensive Elasticity 25th	50th	75th	Agg Hours Elasticity
Ret to experience	0.53	0.25	0.40	0.77	0.99
Baseline	0.51	0.28	0.42	0.67	0.91

The extensive response is the percentage point change in participation in response to a 1% increase in the wage. The aggregate hours 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.

Online Appendix G (to section 3): Selection Correction

In this Appendix we consider an extension of the full model that we calibrate in section 5.3. We allow for taste shocks χ and ζ in the utility function. We solve and simulate this economy and explore the ability of our empirical strategy in section 3 to recover the preference parameter values that are assumed in the simulations ($\phi = 0.76$ and $\theta = 1.75$). We discretise both χ and ζ.

In the first column of Table 26 we report the OLS estimates of the MRS equation using simulated data. We estimate $\phi = 0.52$ and $\theta = 1.97$. These are clearly biased with respect to the assumed parameter values. As discussed in section 3, there are two reasons for the bias. First, regressors are endogenous since consumption and leisure are correlated with the error term in the MRS equation, and, second, there is non-random self-selection of women into the labour market. In order to address the first issue we solve and simulate the economy for several cohorts of women that differ in the average wage they face. We use the variation in wages across cohorts to instrument consumption and leisure in the MRS equation and we provide the estimates in the second column of Table 26. Estimated parameter values are still biased with respect to the assumed parameter values. In order to address the consequences of non-random selection we estimate a probit of the participation decision in which we include as exclusion restrictions a cohort dummy and the log of husband earnings (the same as what we used in the data). We then include the selection correction terms as regressors of the MRS equation and report the results in the third column of the Table. In this case estimated parameter values are very close to the assumed parameter values.

Table 26: MRS estimates with simulated data

	OLS	IV	IV+Selection Correction
ϕ	0.52***	0.61***	0.76***
	(0.000130)	(0.00229)	(0.0666)
θ	1.97***	2.20***	1.71*
	(0.000561)	(0.00528)	(0.902)

Standard errors in parentheses.$*p < 0.10$, $**p < 0.05$,$***p < 0.01$