The Evolution of the Wage Elasticity of Labor Supply over Time

Todd Elder, Steven J. Haider, and Cody Orr
Department of Economics
Michigan State University

September 2020

The wage elasticity of labor supply is arguably one of the most fundamental parameters in economics. Despite the central role of this parameter, little is known about how it has changed over time. We examine the evolution of the labor supply elasticity using cross-sectional methods. We find robust evidence that the labor supply elasticities for married and single men and women have increased modestly over the last two decades. For women, this finding is a substantial departure from earlier evidence. We also contribute to the literature on the robustness of discrete-choice labor supply models, focusing on assumptions that could affect our findings. Our results suggest that the estimated trends are remarkably similar across a variety of specifications.

1. Introduction

The wage elasticity of labor supply is arguably one of the most fundamental parameters in economics. This parameter captures how a percent change in the net wage, perhaps induced by a change in taxes, retirement credits, or productivity, affects the percent change in labor supply, measured by the supply of workers or work hours. With such a parameter in hand, and especially with knowledge of how it varies with observable characteristics, one can readily forecast how labor supply will respond to policy changes that affect the wage. Moreover, the magnitude of the wage elasticity has implications for numerous questions, such as optimal taxation, the causes of business cycles, and the role of human capital formation, among others.

Despite the central role of this parameter, there exists limited evidence about how the wage elasticity of labor supply has changed over time. Several papers have documented changes in women's elasticities, documenting a strong decline for both married (Heim, 2007; Blau and Kahn, 2007; Kumar and Liang, 2016) and single (Bishop et al., 2009) women. A more recent meta-analysis similarly documented declining elasticities for women in both Europe and the US (Bargain and Peichl, 2016), but the latest estimates used in the meta-analysis are from 2005. Moreover, Bargain and Peichl (2006) conjecture that changes in how researchers have estimated labor elasticities account for much of the intertemporal variation in the estimates. There is little formal analysis of the trend in the labor supply elasticities for men.

In this paper, we set out to directly examine how the wage elasticity of labor supply varies over time. To do so, we rely on standard models of labor supply. We first provide a theoretical discussion about why we might expect the elasticities to evolve. We then provide empirical evidence using state-of-the-art cross-sectional methods based on structural discrete choice methods. These methods, which are routinely used to examine wage elasticities, allow us to isolate the wage variation we wish to exploit. At the same time, we critically examine the importance of some of the assumptions used in applying these methods. In doing so, we contribute to the methodological literature that assesses the sensitivity of these models to different modeling assumptions (see Loffler, Peichl, and Stiegloch, 2018, for an excellent recent example).

We have two primary findings. First, we find robust evidence that the labor supply elasticities for married and single men and women have increased modestly over the last two decades. For women, this finding is a substantial departure from earlier evidence. Second, we contribute to the literature on the robustness of discrete choice empirical labor supply models, focusing on the role of key statistical and economic assumptions. Our results suggest that while the levels of the estimated elasticities vary across studies, the trends in those estimates are remarkably insensitive to the imposition of key assumptions.

2. Background and Literature Review

Numerous integrative surveys have summarized the vast literature on labor supply elasticities; see Blundell and MaCurdy (1999) and Keane (2011) for just two examples. In this section, we briefly lay out the basic issues related to our question, heavily relying on the exposition in Keane (2011).

2.1. The Key Elasticities

In a static model, two key wage elasticities arise. The first is the Marshallian (or uncompensated) elasticity, e_M, which measures how hours worked h change with a change in wage w:

(1) $e_M = \frac{w}{h}\frac{\partial h}{\partial w}.$

The second is the Hicksian (or compensated) elasticity, e_H, which measures how hours change with a change in wage, holding utility u constant:

(2) $e_H = \frac{w}{h}\frac{\partial h}{\partial w}\Big|_u.$

Letting e_I denote the income elasticity of labor supply and s denote the share of income accounted for by labor income, these elasticities are related to each other based on the Slutsky equation:

(3) $e_M = e_H + \frac{s}{1-s} e_I.$

Under the assumption that leisure is a normal good, the income elasticity is less than zero, which implies that the Hicksian elasticity is greater than the Marshallian elasticity, i.e., $e_H > e_M$.

In a dynamic model, these two elasticities measure the effect of a permanent (or parametric) change in wages – the Marshallian elasticity when the permanent wage change is uncompensated, and the Hicksian elasticity when the permanent wage change is compensated. In addition, a third elasticity arises. The Frisch (or intertemporal) elasticity measures how labor supply responds to known wage changes between two periods. We denote this elasticity as e_F. Based on our standard assumptions of diminishing marginal utility, we further know that the Frisch elasticity is greater than the other two (i.e., $e_F > e_H > e_M$).

Which elasticity is relevant depends on the policy being considered. For example, the effect of the imposition of a tax to fund a public good might be best approximated by using the Marshallian elasticity, whereas the imposition of a tax to fund a universal income transfer might be best approximated by the Hicksian elasticity. In contrast, the evaluation of a change in tax rate with age, such as that with the Social Security retirement earnings test, might be best approximated with the Frisch elasticity.

Regardless of which elasticity is relevant, maintained assumptions and empirical specifications play a potentially central role in the resulting estimates. Key issues include the simultaneity of wages and hours, the treatment of taxes, measurement error in wages and nonlabor income, the treatment of missing wages for nonworkers, the possibility of non-separabilities between consumption and leisure and over time, the choice of control set and what is considered to be exogenous, and sources of dynamics such as human capital accumulation. Blundell and MaCurdy (1999) and Keane (2011) include detailed discussions of these and other econometric issues. The literature is large and complex, and given this complexity, empirical studies often must make compromises at various points in their analysis.

2.2. Why Might Elasticities Vary Over Time?

The intent of studying how labor supply elasticities have changed over time is to capture how labor supply itself changes. Within the canonical model where jobs are only indexed by wage and workers can freely choose their hours, two explanations arise. One candidate is that secular changes in wages or non-labor income will affect the Marshallian elasticity through the income effect. Another potential candidate within the standard model, but one that is arguably unsatisfying, is that the tastes for work have changed. We consider such an explanation to be a last resort.

Numerous shortcomings of the canonical model have long been recognized. Two shortcomings that are often directly addressed in papers that use discrete-choice models are the possibility that there are fixed costs to working and that the availability of certain hours-wage bundles may be restricted. Changes in the fixed costs to working or in the restrictions of hours-wage bundles could change the measured elasticities, both of which we will consider directly.

Another shortcoming of the canonical model that has long been recognized, but one that is less frequently addressed in the empirical literature, is that non-wage job attributes influence labor supply decisions. Atrostic (1982) presents estimates of labor supply elasticities that account for such non-wage attributes and concludes that labor supply estimates tend to be larger and of opposite sign when such attributes are considered. While we primarily focus on the canonical model and thus focus on wages, we consider the role of changing job characteristics as a potential mechanism for changes over time in elasticities.

2.3. Evidence on Trends in Labor Supply Elasticities

Several papers have examined changes in the labor supply elasticities of women, with the title of one notable example providing a summary for married women: "The Incredible Shrinking Elasticities: Married Female Labor Supply, 1978-2002" (Heim, 2007). Based on March CPS data from 1979 through 2003, Heim finds that the elasticity declined by about 60 percent from an initial base of 0.36.[1] Blau and Kahn (2007) similarly examine changes for married women, using 1980-2000 March CPS data. They find that the wage elasticity dropped by just over 50% from an initial base of about 0.80. Both papers use a continuous labor supply model with a Heckman selection correction to account for unobserved wages among non-working men. Important for our findings, neither paper finds that the downward trend is diminishing.

[1] Unless otherwise specified, we report results from previous studies for the Marshallian (or uncompensated) elasticity.

Kumar and Liang (2016) similarly examine the labor supply elasticities of married women, but instead use the 1980-2006 PSID.[2] They find large declines in elasticities, but they provide some evidence that the declines level off during the very last years of their sample period.

Bishop et al. (2009) examine the changes in the wage elasticity for single women using the March CPS from 1979 through 2003 and similarly find large declines. Using a continuous choice model, they conclude that wage elasticity of labor supply declined by about 80 percent from an initial base of about 0.20. There is little evidence that the downward trend is diminishing towards the end of their sample period.

Several systematic literature reviews also provide evidence about how the wage elasticity has changed over time. Keane (2011) provides a list of estimates for men from 8 different studies that exhibit little change from 1980 to 2001. Bargain and Peichl (2016) include an exhaustive list of studies from both the US and Europe. While little systematic trend appears to be present for the elasticities for men, Bargain and Peichl (2016) demonstrate that a downward trend exists for married and single women, both in the United States and Europe. Interestingly, Bargain and Piechl (2016) provide evidence that (a) researchers are increasingly using discrete choice estimation methods to estimate wage elasticities, and (b) the downward trends in female elasticities are somewhat flatter when using estimates from these discrete choice studies.

[2] Although the authors provide some estimates that exploit the panel nature of their data, we focus on the estimates obtained from pooled models that are similar to those estimated by Heim (2007) and Blau and Kahn (2007).

3. Empirical Methods

We rely on a static model of labor supply throughout our analysis. The motivation for doing so is that such methods continue to play a prominent role in assessing the responsiveness of hours to wages. For example, Bargain and Peichl (2016) list numerous examples of recent papers that estimate a static model, and two prominent papers that provide a synthesis of our understanding of Hicksian elasticities rely on estimates from static models (Keane, 2011; Chetty et al., 2011). In addition, we are able to use methods that have been applied in numerous other papers and that allow us to focus on utilizing plausibly exogenous variation in wages that stems from between-state and intertemporal variation in marginal tax rates.

Many of the recent papers that estimate the wage elasticity of labor supply do so using a discrete-choice framework pioneered by Van Soest (1995).[3] This approach is particularly well-suited to examining our question. First, it directly incorporates non-participation into the labor supply decision, allowing us to study the extensive and intensive margins in a common framework. Second, it allows us to utilize plausibility exogenous variation in wages caused by variation in the marginal tax rate across states and across time. Third, it allows us to incorporate fixed costs of working as well as other nonconvexities of the wage-hours locus.

3.1. A Static Discrete-Choice Model of Labor Supply

For notational ease, we first discuss the discrete-choice model we use for single individuals. Suppose each individual i chooses from a set 7 different levels of hours per week, denoted $H_{ij} \in \{0, 10, 20, 30, 40, 50, 60\}$. For our baseline analysis, we assume that the same pre-tax

[3] In a recent literature review, Bargain and Peichl (2016) compiles papers that have estimated wage elasticities by time and the method used. Pooling across European and US studies, there is a clear trend towards using the discrete-choice framework.

wage is offered for each of these hour choices. Each choice of weekly work hours delivers a different level of weekly consumption C_{ij} and leisure L_{ij}, where $L_{ij} = 112 - H_{ij}$ (we assume that individuals have 112 total hours per week to devote to work or leisure, with the remaining 56 hours spent on sleep and personal care). Thus, each individual has the choice of 7 different bundles $\{C_{ij}, L_{ij}\}$.

The portion of utility that depends on the hours/consumption bundle and individual taste shifters X_i is denoted by $U_{ij}(C_{ij}, L_{ij}, X_i)$. Total utility is defined to be

(3) $$V_{ij} = U_{ij}(C_{ij}, L_{ij}, X_i) + \epsilon_{ij}$$

where ϵ_{ij} is an i.i.d type 1 extreme-value error term. Based on this error assumption, the probability of individual i choosing choice j, denoted P_{ij}, is given by the closed-form expression

(4) $$P_{ij} = \exp\left(U_{ij}\right) / \sum_{k=1}^{J} \exp(U_{ik}) \,.$$

This expression is the familiar multinomial logit form for choice probabilities.

For our baseline specification, we use the translog specification for U_{ij} proposed in Van Soest (1995):[4]

(5) $$U_{ij} = \beta_{C1} \ln C_{ij} + \beta_{C2}\left(\ln C_{ij}\right)^2 + \pi_i \ln L_{ij} + \beta_{L2}\left(\ln L_{ij}\right)^2 + \beta_{CL} \ln C_{ij} \ln L_{ij} + f(\alpha, H_{ij}).$$

We allow the parameter π_i to vary with several individual characteristics (a quadratic in the logarithm of age, the number of dependents under the age of 3, and the number of dependents under the age of 19), as we discuss in more detail below. The term $f(\alpha, H_{ij})$ allows for fixed utility decrements of working 10, 20, and 30 hours a week; these utility decrements might arise

[4] Other authors who use a translog specification include Flood et al. (2004) and Haan (2006).

due to true fixed costs of working part time, or they might capture employers' choices to not offer these hours packages.

For married households, we amend the basic specification to include a quadratic function of the logarithm of leisure for both spouses, as well as an interaction between the logarithms of leisure of each. The choice set for married couples is then all combinations of the seven hours choices for each spouse, delivering 49 different hours choices for the couple.

At this point, one could estimate the model via a standard multinomial logit, but researchers routinely incorporate two extensions. The first is in recognition that a multinomial logit functional form imposes the property of independence of irrelevant alternatives (IIA), which implies that if one choice is taken away, an individual must substitute towards all other choices proportionally; however, introspection suggests that an individual would be more likely to substitute towards a "nearby" hours choice. To allow for richer substitution patterns, a random term is included in the individual-specific slope on leisure:

(6) $$\pi_i = \delta_0 + \delta_1 \ln age_i + \delta_2 (\ln age_i)^2 + \delta_3 Dep3_i + \delta_4 Dep19_i + u_i,$$

where $Dep3_i$ and $Dep19_i$ are the number of dependents under the ages of 3 and 19, respectively, and $u_i \sim N(0, \sigma_u^2)$. This specification allows otherwise-identical individuals to have different preferences for leisure.

The second extension reflects that wages must be predicted for non-workers (and possibly for workers), which generates measurement error in the wages used in the empirical implementation. Specifically, suppose true wages w_i^* are

(7) $$w_i^* = \widehat{w}(X_i) + v_i,$$

where $\widehat{w}(X_i)$ are predicted wages based on covariates X_i and v_i is the idiosyncratic component to wages. We follow the common practice of "integrating out" the measurement error. In practice, this amounts to pre-specifying the variance of v_i, taking draws from this assumed distribution, and then integrating over the draws when we maximize the simulated likelihood (see van Soest (1995) for a much more rigorous description). We specify the variance of the measurement error to be the estimated variance obtained from a Heckman selection model, as described below.

To accommodate the random slope on leisure and measurement error in wages, we will estimate our models using Maximum Simulated Likelihood (MSL). Define the likelihood contribution for a single individual as

(8) $$L_i = \int \int [P_{ij}(u_i|v_i)^{1[j=1]} p(v_i)] dF(u_i) dF(v_i)$$

where $p(v_i)$ is the density of v_i, the idiosyncratic component of the wage. This likelihood contribution can be approximated by taking *R* independent draws on (u_i, v_i) and averaging over the draws

(9) $$L_i^R = \sum_{r=1}^{R} \left[P_{ij}(u_{ir} \mid v_{ir})\right]^{1[j=1]}.$$

We use these average contributions in the overall likelihood function and choose the parameters accordingly. We use 100 Halton draws in our baseline estimates.[5]

[5] Halton sequences are deterministic sequences that cover domains of integration more evenly than do independent pseudo-random draws. As Train (2000) argues, simulations using Halton draws exhibit greater accuracy than those using pseudo-random draws, particularly in the context of logit choice models.

3.2. Estimating the Elasticities

With parameter estimates in hand, we can directly simulate counterfactuals to allow us compute an elasticity for each individual. We then report the mean or median elasticity across individuals.

For the Marshallian elasticity, we predict the probability of each individual choosing each of the hours choices, both for their (predicted) wage, w^0, and for a wage that is 10 percent higher, w^1, while holding non-labor income constant. We then compute the Marshallian elasticity for each person as

$$(10) \qquad e_M = \frac{\sum_j h_j * \left(p_j(w^1) - p_j(w^0)\right)}{\sum_j h_j * p_j(w^0)} /.1 = \frac{\sum_j h_j * \left(p_j(w^1) - p_j(w^0)\right)}{.1 * \bar{h}(w^0)},$$

where $p_j(w^k)$ denotes the computed probability that the individual will choose choice *j* when faced with a wage of w^k for $k \in \{0,1\}$, and $\bar{h}(w^k)$ is the predicted mean hours worked when an individual is faced with wage w^k, i.e., $\bar{h} = \sum_j h_j * p_j(w^k)$. All of these quantities are evaluated with a pre-tax wage and non-labor income of N^0, which we suppress for now to simplify notation.

To provide further information about the underlying source of the hours change, we decompose the Marshallian elasticity into an extensive (e_M^{Ext}) and intensive (e_M^{Int}) component:

$$(11) \qquad e_M^{Ext} = \frac{\left[\left(1 - p_0(w^1)\right) - \left(1 - p_0(w^0)\right)\right] * \sum_{j \neq 0} h_j * q_j(w^0)}{.1 * \bar{h}(w^0)}$$

$$(12) \qquad e_M^{Int} = \frac{\left[\left(1 - p_0(w^1)\right)\right] * \sum_{j \neq 0} h_j * \left(q_j(w^1) - q_j(w^0)\right)}{.1 * \bar{h}(w^0)},$$

where $q_j(w^k) = p_j(w^k)/(1 - p_0(w^k))$. This decomposition ensures that the underlying components sum to the Marshallian elasticity ($e_M = e_M^{Ext} + e_M^{Int}$), with the first component measuring the change in the probability of not working and the second measuring the change in expected hours conditional on working.

Rather than use the Slutsky decomposition in eq. (3) to calculate the Hicksian elasticity from the Marshallian elasticity, we instead calculate the Hicksian elasticity directly:

(13) $$e_H = \frac{\sum_j h_j * \{p_j(w^1, N^0 - [\widetilde{w}(w^1,N^0) - \widetilde{w}(w^0,N^0)] * \bar{h}(w^0,N^0)) - p_j(w^0,N^0)\}}{.1 * \bar{h}(w^0,N^0)},$$

where $p_j(w^k, N^k)$ denotes the computed probability that the individual will choose choice *j* and $\widetilde{w}(w^k, N^k)$ denotes the expected after-tax wage, both evaluated when an individual faces wage w^k and non-labor income N^k. Expression (13) shows that, in order to calculate the Hicksian elasticity, we must compute the choice probabilities at the higher wage w^1 while reducing non-labor income by the expected increase in after-tax income due to the wage change (reflected by $[\widetilde{w}(w^1, N^0) - \widetilde{w}(w^0, N^0)] * \bar{h}(w^0, N^0)$). This adjustment to non-labor income allows us to abstract from the effect of the wage increase on total income.[6] See Appendix Section A for further details about our elasticity calculations.

4. Data

Our primary data source is the 1979 through 2018 Annual Social and Economic March Supplement of the Current Population Survey (CPS). These surveys provide total earnings and hours worked for the previous year. We list all results by survey year throughout the paper.

The sample is restricted to adults in single-unit households. Single adults must be in the age range of 26 to 55, and couples must have one spouse in the age range 26 to 55 and the other

[6] The fact that the Slutsky equation may not hold in a discrete choice model is well-developed elsewhere, including Dagsvik and Karlstrom (2005) and Dagsvik, Strom, and Locatelli (2014). The formulation we implement is simpler than the approaches laid out in these studies, in part because we use income compensation rather than utility compensation for the worker.

spouse in the age rage 22 to 59. We exclude individuals enrolled in school, disabled, self-employed, or members of the armed forces; earn farm or business income; or work unpaid in a family business. We additionally exclude individuals that have more than seven family members living in their household, or report extreme wages or non-labor income. For couples, if the age-year sample is larger than 10,000 observations, we take a random sample of 10,000.

Hours are discretized into 7 choices: 0, 10, 20, 30, 40, 50 and 60 hours per week. We use usual hours worked per week from the prior year. We compute the hourly wage from wage and salary income in the prior year divided by the usual work hours multiplied by weeks worked. We calculate non-labor income from the sum of dividend income, interest income, rental income and other property income, alimony, child support, and other non-property income. Wages, of course, are unobserved for non-workers. We impute wages for non-workers based on a standard Heckman selection procedure. Appendix Section B provides complete details of the sample construction, basic descriptive statistics, and selected results from the Heckman selection estimation.

As an initial look at the data, Figure 1 shows the labor force participation rate, the average annual hours worked conditional on working, and average wage conditional on working for primary sample period (survey years 1994 through 2018) and for four demographic groups (single and married men and single and married women). The well-documented changes are readily apparent. Labor force participation decreased modestly for men, increased for single women during the 1990s, and remained relatively constant for married women. Among the employed, annual hours worked and wages increased for all four demographic groups. Table 1 shows key descriptive statistics for the four demographic groups for three years: 1994, 2004, and 2014. Except among married women, most individuals work 40 hours and above, with very few working 0-30 hours. Wages increase throughout the sample period, and non-labor income generally increases as well.

We supplement the CPS with two additional data sources. The first is NBER's TAXSIM, which allows us to calculate after-tax income by state and year, utilizing the micro-level information on earnings, state of residence, and household structure. The second is TRIM3, which allows us to calculate social program benefits for which a household qualifies, based on household earnings, state of residence, and household structure. The social benefits that are included are from the Supplemental Nutrition Assistance Program (SNAP), Temporary Assistance to Needy Families (TANF), and Supplemental Security Income (SSI).

See Appendix Section C for further information about how we incorporate TAXSIM and TRIM3.

5. Results

We first present our baseline results for Marshallian own-wage elasticities (OWE) in Section 5.1, and then results based on different assumptions regarding the utility function in Section 5.2 and measurement error in Section 5.3. In Section 5.4, we present results for other elasticities to allow further comparisons back to the literature, including extensive and intensive margin Marshallian elasticities and Hicksian elasticities.

5.1. Baseline Results

For our baseline results, we estimate the model outlined in Section 3.1, except we assume that the variance on the random component of the slope on leisure is zero (i.e., $\sigma_u^2 = 0$); we consider the importance of this assumption in the next subsection. While this assumption implies that we are estimating a typical "conditional logit" as opposed to a "random slope logit" model, we still estimate the model via MSL in order to integrate out the measurement error in wages.

Figure 2 presents mean own-wage elasticities from 1994 through 2018 from the baseline model, shown with a dark line, along with standard error bands (see Appendix Section D for parameter estimates for selected years). These baseline estimates make use of the benefit simulator to capture the outside option of not working, which are only available starting in 1994. For three groups – single men, married men, and married women – we observe small increases in the labor supply elasticity over time. For example, the OWE increases from 0.046 (SE=0.011) to 0.113 (SE=0.011) for single men from 1994 to 2017, from 0.082 (SE=0.004) to 0.108 (SE=0.005) for married men, and from 0.252 (SE=0.009) to 0.279 (SE=0.008) for married women (see Table 2). For the fourth group, single women, the estimates increase after 2000 (from 0.186 (SE=0.009) in 2004 to 0.265 (SE=0.012) in 2014), but given the decline in the elasticity from 1994 to 1999, the elasticity is relatively flat over the full time period. The modest increase in the OWE over the last two decades across various groups represents one of the key findings of our paper.

The finding that the OWE was declining during the initial years is not surprising given previous work (e.g., Heim, 2007). However, our baseline results begin in 1994 because of the data we use to compute benefits, so these results are only partially comparable to previous papers that have examined trends in OWE. We can produce a consistent set of results going back to 1979 by ignoring benefits; we also present these results in Figure 2.[7] For single and married men, the inclusion of benefits has very little effect on the estimates. This result is unsurprising because men tend to qualify for few benefits. The inclusion of the longer time series reveals that the steady upward trend in the mean OWE among men extends back to 1979.

In comparison to men, the estimates for women are more sensitive to the inclusion of benefits. Focusing on the longer time series, the OWE was declining before 1994 for both single and married women; these results are quantitatively similar to previous studies that examined this earlier period.[8]

5.2. Robustness to Alternative Utility Function Assumptions

A simplifying assumption that we made for our baseline results is that the variance of the random component of the slope on leisure is zero (i.e., $\sigma_u^2 = 0$), resulting in a conditional logit specification. Although this assumption is at odds with the assumptions in previous studies, it is not at odds with previous empirical results: numerous papers report results that are consistent with the random slope being unimportant.[9] Regardless of the previous empirical findings, we

[7] Many previous papers that calculate OWE do not include benefits, including Heim (2007), Blau and Kahn (2007) and Bishop et al. (2009). Thus, by excluding benefits, we not only obtain a consistent time series over a longer period, but we also produce OWE that are more comparable to previous estimates.

[8] For single women, our baseline results suggest that the OWE declined by 48% from 1979 to 2000 (from 0.242 (SE=0.013) to 0.125 (SE=0.008)). In Bishop et al. (2009), the authors find about an 80% decline from a base of about 0.20 over a similar period.

[9] Van Soest (1995) estimates random slopes with variances that are not statistically different from zero and very small relative to the mean value. Callan and van Soest (1996) estimate random slopes with variances that are not statistically different from zero, but the variance for women could be practically significant. Van Soest, Das, and Gong (2002) estimate a precise zero for random slope variances for women. Brewer, Duncan, Shephard, and Suarez (2006) include random slopes on income, female hours, FC of work, and welfare participation. For single mothers and couples, the first three variances are precisely estimated zeroes, and the fourth is large but statistically insignificant from zero. Two papers, Haan (2006) and Loffler, Sieglach, and Peichl (2018), examine this issue directly and conclude that labor supply elasticities tend to change very little when allowing for random slopes

directly compare our baseline conditional logit results for 1994-2017 (that use benefits) to results that do not restrict the variance of the random slope to be zero.

Turning to the results in Figure 3, for married men neither the level nor the trend are sensitive to the choice of a baseline conditional logit versus the random slope model. For married women, the overall trends are similar across specifications, but the random slope logit results are systematically larger. For example, in 2018, the estimated mean OWE for married women is 0.327 (SE=0.009) with the conditional logit and 0.441 (SE=0.014) with the random slope logit.

For single women, the point estimates are practically identical across all specifications, but the standard errors for the random slope results become very large. The reason for these results is apparent from the underlying parameter estimates: in 2017, the variance of the random slope is estimated to be small (0.043) and the standard error on the variance is large (0.560). Because the random slope is small, the conditional logit and random slope logit give very similar results, and because the random slope is estimated imprecisely, the estimated elasticities also have large standard errors. Overall, we conclude that the addition of the random slope does not change our findings for single women.

The one group for whom the random slope specification seemingly matters for the trend is single men. However, the steeper trend with random slopes appears to be partly driven by outliers. Because of nonlinear nature of our estimation problem, the mean-zero draws from the underlying distribution of parameters (including the random slope) need not lead to

on just consumption or leisure individually, although both papers find that allowing for more unrestricted unobserved heterogeneity can have some small effects on point estimates.

offsetting labor supply responses. In such circumstances, median elasticities can provide additional insight into the central tendency of the elasticity distribution. As an example, we provide the median elasticities for the baseline and random-slope models in Figure 4. Contrary to the mean elasticities provided in Figure 3, the trends for the median elasticities are very similar for all four demographic groups. This finding supports the interpretation that outliers drive the deviations in mean elasticities for single men in Figure 3; the large standard error bands on the random slope estimates further support this interpretation.

As another robustness check, we also estimated our baseline conditional logit specification by instead specifying a quadratic utility function used in Bargain et al. (2014). We present these results in Appendix E and Figure E.1. The results are very similar to those shown in Figure 3.

5.3. Measurement Error in Wages

To what extent does accounting for measurement error matter for our results? In Figure 5, we provide direct evidence on this question. For our baseline results, the answer is not much at all, particularly if one focuses on trends. The lines for our baseline results that use 100 draws for integrating out measurement error (labeled “Baseline” in the figure) are very similar to those lines that simply use mean predicted earnings for non-workers, labeled “Non, 0 draws”.

One important caveat for this finding is both sets of results only use predicted wages for non-workers. Because most individuals in our samples are working, this procedure treats few individuals as having error-laden wages.[10] Some previous studies use predicted wages for both

[10] In 2014, the percent working is 93% among single men, 88% among single women, 95% among married men, and 74% among married women. See Appendix Tables B.1 and B.2 for results for other years.

workers and non-workers alike, so accounting for measurement error in those circumstances matters more in those cases.[11]

To analyze this possibility, Figure 5 additionally presents results that use predicted wages for everyone. The series labeled "All, 100 Draws" integrates out measurement error using 100 Halton draws, while the series labeled "All, 0 Draws" ignores measurement error. Our results suggest that the decision to predict wages for everyone or for only non-workers matters, in that the elasticities tend to be larger when we predict wages for everyone. However, the estimated trends are insensitive to this choice. In addition, accounting for measurement error once again tends to have very little systematic effect on the results.

Why does accounting for measurement error matter little? For the case at hand, a reasonable representation of the measurement error process is

(14) $w^* = w + \gamma,$

where we denote observed wages as w and actual wages as w^*. We write down the measurement error process in this form rather than the standard classical measurement error form (which reverses the role the role of w and w^*) because we are correcting for measurement error in cases where we utilize predicted wages based on a Heckman selection equation. In other words, we are using predicted wages w in estimation, whereas the actual wages w^* include an individual component γ that is unobserved. While analytic results are not

[11] Studies that predict wages only for non-workers include van Soest (1995), Blundell et al. (1998), and Haan (2006). Studies that predict wages for the full sample include Flood et al. (2004) and Bargain et al. (2014). See Loeffler et al. (2018) for a discussion and analysis of this issue.

available for the non-linear labor supply function we use for estimation, this alternative form of measurement error leads to little or no effects in other settings.[12]

This finding has two important implications. First, to the extent that adjusting for measurement error in wages is not necessary, we can use standard maximum likelihood estimation packages rather than the more computationally costly maximum simulated likelihood methods outlined above. Second, our analysis of measurement error has ignored the fact that, when adjusting for measurement error, we were pre-specifying its variance. By instead relying just on the conditional mean wage, we need to make no such decision about the size of the variance.[13]

5.4. Other Elasticities

To allow further comparisons with previous studies, we next provide evidence on the three other elasticities described in Section 3.2. Given the findings shown above, we present all results for this subsection using a conditional logit specification and use the mean predicted wage for non-workers without integrating out measurement error.

We present results for the mean total, extensive margin, and intensive margin Marshallian own-wage elasticities in Figure 6.[14] Across all four demographic groups, the intensive margin elasticities are small and vary little over time. Thus, changes in the total elasticities are driven

---

[12] For example, it is straightforward to show that the OLS estimate of a regression of y on w is consistent for an OLS regression of y on w^*(assuming that γ is i.i.d. in both cases), which is contrary to the usual attenuation result under classical measurement error from the use of an error-ridden variable.

[13] Changes in the assumption on the size of the variance of the measurement error can appreciably change the results, depending on its relative size compared to other parameters in the model. For example, when predicting wages for everyone (which leads to a large estimated variance), elasticities tend to be more sensitive to the assumption of its size.

[14] In Figure 5, we presented the median of the total Marshallian OWE, whereas in Figure 6 we present the mean. As is readily apparent, the trends are very similar across the two figures.

almost exclusively by changes in the extensive margin elasticities. This finding echoes those from previous studies that have studied trends. For example, Heim (2007) studying married women and Bishop et al. (2009) studying single women find almost identical declines for the total elasticity and the extensive margin elasticities during the 1980s and 1990s. We find that the reversal in the declines in the total elasticities for both groups of women over the last two decades, as well as the increases for men, are similarly driven by changes in the extensive margin elasticity.

6. Conclusion and Discussion

In this paper, we examine trends in labor supply elasticities for four demographic groups: single and married men and single and married women. Our results suggest that the elasticities have steadily increased over the last two decades for all four groups. For single and married women, this finding is a remarkable reversal of the "incredible shrinking elasticities" reported for the 1980s and 1990s in previous studies. These reversals were sufficiently large to unwind all of the previous declines for single women and about half of the previous declines for married women. For all groups, changes in the extensive margin account for nearly all of these trends; the intensive margin contributes little to overall elasticities, regardless of group or time period.

Our results additionally suggest that our major conclusions are robust to some of the common alternative modeling assumptions used in the literature, such as the choice of quadratic utility versus translog utility functions or whether researchers report mean or median elasticities. This insensitivity also encompasses computationally challenging assumptions, such as the use of random-slope models and integrating out measurement error using Maximum

Simulated Likelihood. These latter findings are important because such simplifications can allow for increasing complexity in other dimensions, including non-convexities in the budget set.

Taken together, our results raise an important question: why did labor supply elasticities increase over the last two decades? Because the extensive margin accounted for nearly all of the increases for all four groups, likely candidates include those factors that affect the decision to work. Abraham and Kearney (2020) examine changes in labor force participation rates among young and prime-age adults between 1999 and 2018, assessing the evidence for a host of factors that could explain declining participation rates among men and women younger than 55. While the factors underlying trends in extensive margin elasticities and in participation rates are not necessarily identical, it would be surprising if they were wholly unrelated.

Abraham and Kearney (2020) suggest that one potential driver of the participation trends involves increases in the generosity of safety net assistance, namely through federal disability insurance, SNAP benefits, or publicly provided health insurance. They argue that the latter two factors likely play small roles in participation trends because of timing inconsistencies, and that seems likely to be the case for the trends in elasticities as well. For example, the Affordable Care Act, which went into effect in 2009, cannot explain the increase in estimated elasticities between 1999 and 2008 that we find for each of the four demographic groups. On the other hand, federal disability caseloads grew steadily from 1999 to 2015, and Abraham and Kearney (in addition to Binder and Bound (2019)) suggest that this growth may account for as much as five percent of the decline in participation rates over the same period. Increased access to disability assistance may have also contributed to rising extensive margin elasticities over that period by reducing the relative appeal of working for low wages.

Of course, the increasing trends in (pre-tax) labor supply elasticities may also stem from policies that increase, rather than decrease, in participation rates. The primary candidates involve the expansions of the Earned Income Tax Credit in 1993, 1996, 2001, and 2009. However, our estimated elasticities for both single and married women decline until roughly 2000 (in agreement with previous estimates) before reversing, which is inconsistent with the steady expansions of the EITC. Moreover, the expansions cannot explain the evolution of elasticities for single (and, to some extent, married) men, who are largely ineligible for EITC benefits.

In sum, we find clear evidence that labor supply elasticities have increased in the last twenty years, but convincing explanations for these trends have not emerged. Our estimates suggest that social program benefits influence elasticities for women, but that accounting for intertemporal variation in those benefits does not markedly influence trends. For men, we find that the treatment of benefits has no effect on either the estimated levels or trends in elasticities.

References

Abraham K. and M. Kearney. 2020. Explaining the Decline in the US Employment-to-Population Ratio: A Review of the Evidence. Journal of Economic Literature 58(3): 585-643.

Atrostic B. 1982. The Demand for Leisure and Nonpecuniary Job Characteristics. American Economic Review 72(3): 428-440.

Bargain O, K Orsini, and A Peichl. 2014. Comparing Labor Supply Elasticities in Europe and the United States: New Results. Journal of Human Resources 49(3): 723-838.

Bargain O and A Peichl. 2016. Own-Wage Labor Supply Elasticities: Variation across Time and Estimation Methods. IZA Journal of Labor Economics 2016(5): 1-31.

Binder, A, and J Bound. 2019. The Declining Labor Market Prospects of Less-Educated Men. Journal of Economic Perspectives 33(2): 163-190.

Blundell R, A Duncan, and C Meghir. 1998. Estimating Labor Supply Responses Using Tax Reforms. Econometrica 66(4): 827-861.

Blundell R, A Duncan, J McCrae, and C Meghir. 2000. The Labour Market Impact of the Working Families' Tax Credit. Fiscal Studies 21(1): 75-104.

Bishop, Kelly, Bradley Heim, and Kata Mihaly. 2009. Single Women's Labor Supply Elasticities: Trends and Policy Implications, Industrial and Labor Relations Review, 63 (1): 146-168.

Blau, Francine D., and Lawrence M. Kahn. 2007. Changes in the Labor Supply Behavior of Married Women: 1980-2000, Journal of Labor Economics, 25, 393-438.

Blundell R, E French, and G Tetlow. 2016. Retirement Incentives and Labor Supply. In Handbook of the Economics of Population Aging, Vol 1B (Eds., J Piggott and A Woodland). Amsterdam: North-Holland. 457-566.

Blundell R and T MaCurdy. 1999. Labor Supply: A Review of Alternative Approaches. In Handbook of Labor Economics, Vol 3 (Eds., O Ashenfelter and D Card). Amsterdam: North-Holland. 1559-1695.

Brewer M, A Duncan, A Shephard, and MJ Suarez. 2006. Did Working Families' Tax Credit Work? The Impact of In-Work Support on Labour Supply in Great Britain. Labour Economics 13(6): 699-720.

Card D. 1994. Intertemporal Labor Supply: An Assessment. In Advances in Econometrics Sixth World Congress , Vol 2 (Ed., C Sims). Cambridge University Press. 49-78.

Creedy J and G Kalb. 2005. Discrete Hours Labour Supply Modelling: Specification, Estimation and Simulation. Journal of Economic Surveys 19(5): 697-734.

Currie J and B Madrian. 1999. Health, Health Insurance and the Labor Market. In Handbook of Labor Economics, Vol 3 (Eds., O Ashenfelter and D Card). Amsterdam: North-Holland. 3309-3415.

Eissa N, H J Kleven, and C T Kreiner. 2008. Evaluation of four tax reforms in the United States: Labor supply and welfare effects for single mothers. Journal of Public Economics 92(5): 795-816.

Flood L, J Hansen, and R Wahlberg. 2004. Household Labor Supply and Welfare Participation in Sweden. Journal of Human Resources 39(4):1008-1032.

French E. 2005. The Effects of Health, Wealth, and Wages on Labour Supply and Retirement Behavior. Review of Economic Studies 72(2): 395-427.

Haan P. 2006. Much Ado About Nothing: Conditional Logit vs. Random Coefficient Models for Estimating Labour Supply Elasticities. Applied Economics Letters 13(4):251-256.

Haider S and D Loughran. 2010. Elderly Labor Supply: Work or Play? In Workplace Flexibility: Realigning 20th Century Jobs to 21st Century Workers (Eds., K Christensen and B Schneider). Ithaca, NY: Cornell University Press.

Heim B. 2007. The Incredible Shrinking Elasticities: Married Female Labor Supply, 1978-2002. Journal of Human Resources 42(4): 881-918.

Jaeger, D. 1997. Reconciling educational attainment questions in the CPS and the census. Monthly Labor Review 120: 36.

MaCurdy T. 1981. An Empirical Model of Labor Supply in a Life-Cycle Setting. Journal of Political Economy 89(6): 1059-1085.

Rogerson R and J Wallenius. 2009. Retirement in a Life Cycle Model of Labor Supply With Home Production. University of Michigan, MI Retirement Research Center Working Paper #205.

Rogerson R and J Wallenius. 2013. Nonconvexities, Retirement, and the Elasticity of Labor Supply. American Economic Review 103(4): 1445-1462.

Train, K. 2003. Discrete Choice Models using Simulation. Cambridge, UK: Cambridge University Press.

Van Soest A. 1995. Structural Models of Family Labor Supply: A Discrete Choice Approach. Journal of Human Resources 30(1): 63-88.

Figure 1: Labor Force Participation, Hours, and Wages by Demographic Group

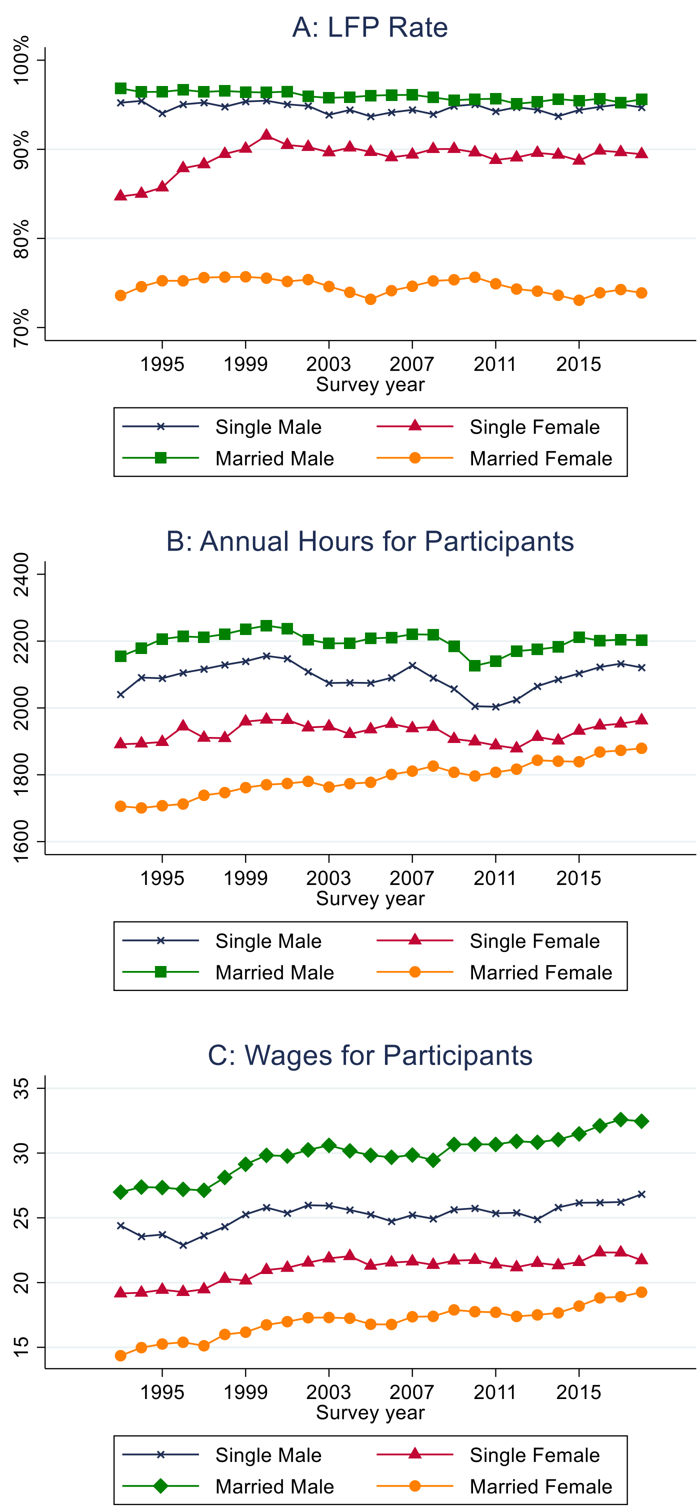

Figure 2: Mean Own-Wage Labor Supply Elasticities, Baseline Estimates (1994-2017) versus Estimates with No Benefits (1979-2017)

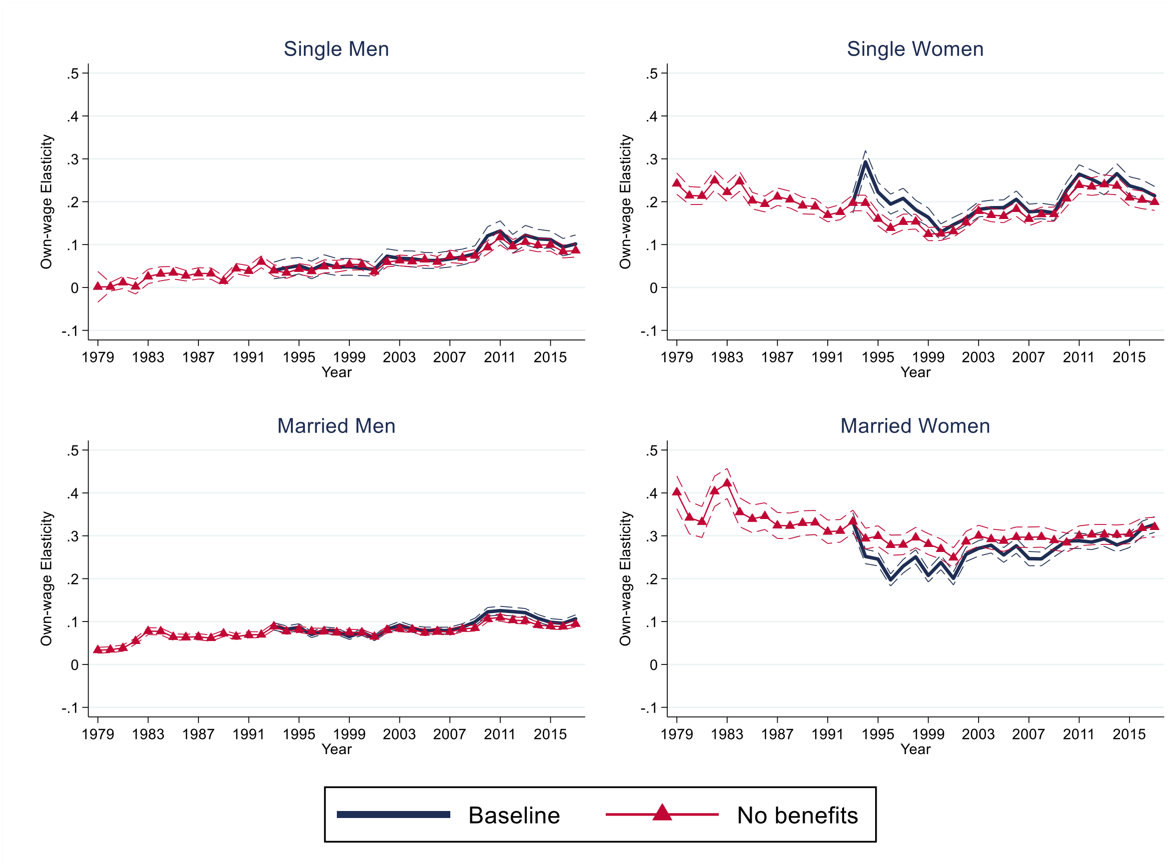

Figure 3: Mean Own-Wage Labor Supply Elasticities, Baseline Estimates versus Estimates from Random-Slope Logit Model

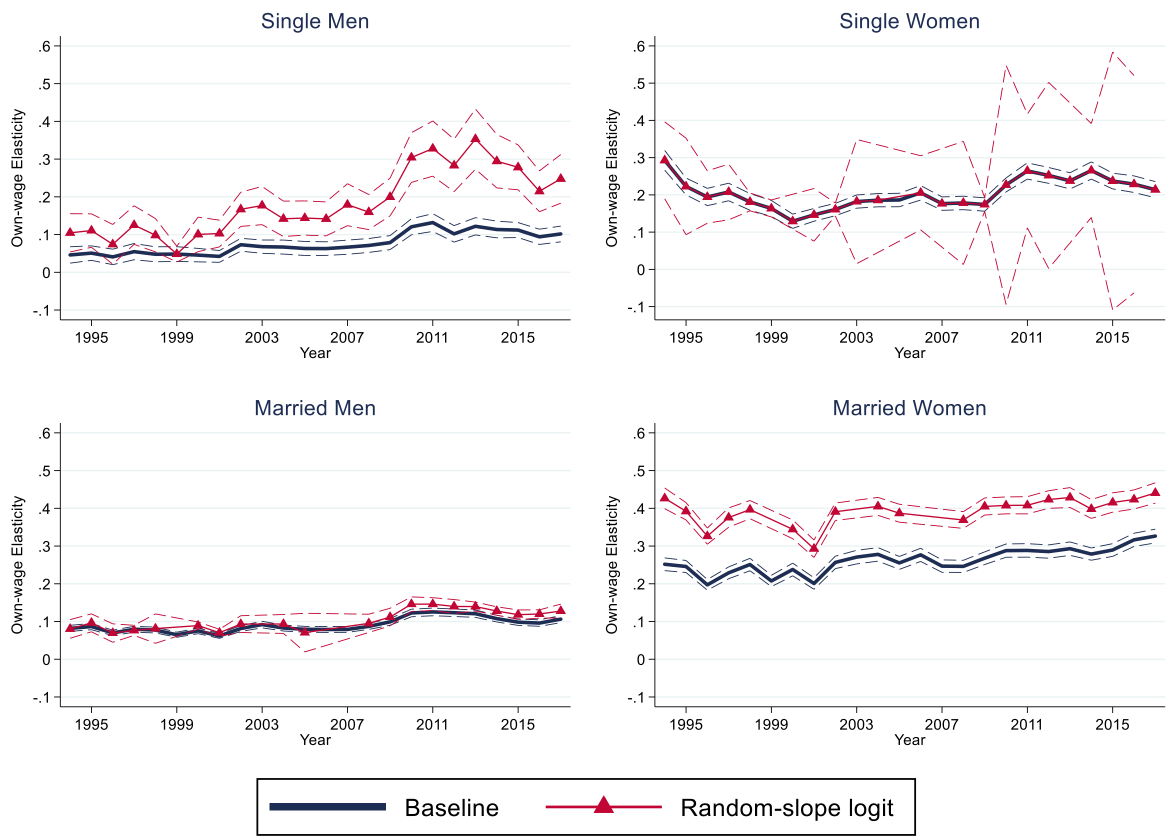

Figure 4: Median Own-Wage Labor Supply Elasticities, Baseline Estimates versus Estimates from Random-Slope Logit Model

Figure 5: Median Own-Wage Labor Supply Elasticities, Baseline Estimates versus Different Treatments of Measurement Error

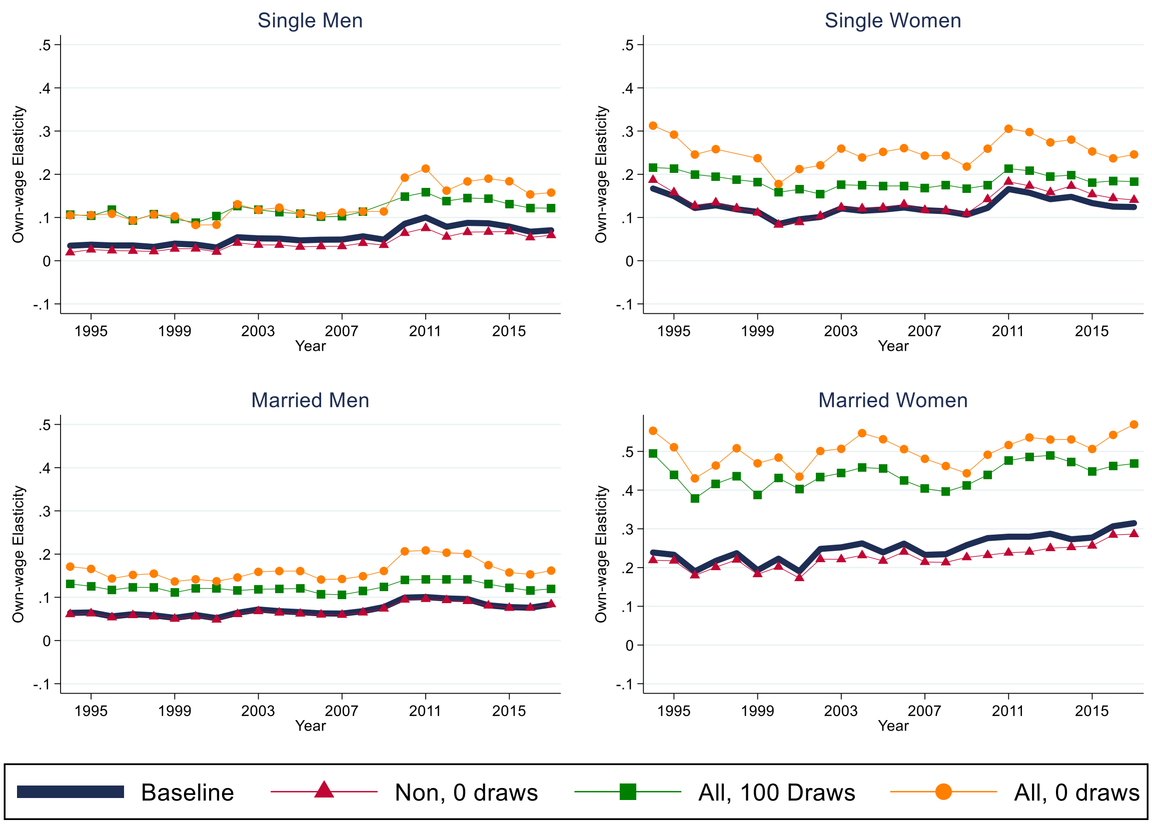

Figure 6: Mean Own-Wage Marshallian Total, Extensive, and Intensive Elasticities

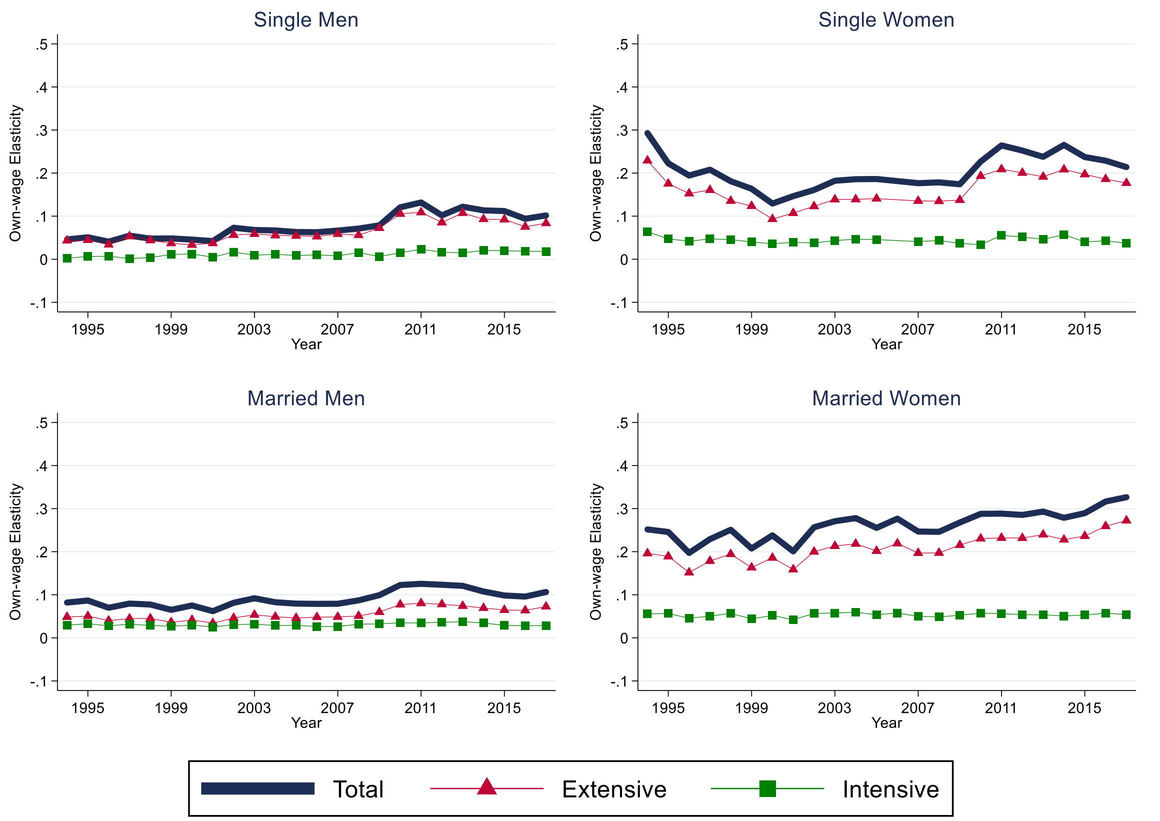

Table 1: Descriptive Statistics

	Single			Married		
	1994	2004	2014	1994	2004	2014
A. Men						
N	3,060	4,274	4,029	14,231	20,634	15,949
Binned Hours						
0	.04	.05	.07	.03	.03	.05
10	.01	.01	.01	.01	.01	.01
20	.02	.02	.02	.01	.01	.01
30	.03	.03	.03	.02	.01	.02
40	.57	.61	.60	.59	.61	.62
50	.24	.19	.18	.23	.22	.20
60	.11	.10	.09	.12	.11	.11
Wages	23.56	25.61	25.80	27.38	30.17	31.05
Non-labor Income	2,342	1,730	3,456	2,786	2,914	3,928
B. Women						
N	5,573	7,674	6,229	15,039	21,223	16,347
Binned Hours						
0	.14	.09	.12	.23	.24	.26
10	.01	.01	.01	.03	.02	.02
20	.04	.04	.05	.08	.07	.06
30	.07	.06	.08	.08	.08	.07
40	.57	.64	.60	.48	.49	.48
50	.11	.12	.10	.08	.08	.08
60	.05	.05	.04	.02	.02	.02
Wages	19.21	22.04	21.34	14.98	17.25	17.66
Non-labor Income	2,597	2,850	2,534	2,786	2,914	3,928

Notes. All dollar values are deflated to 2017 dollars with the CPI-U-RS. Wages are reported for only those who are working. Non-labor income for married women is identical to the non-labor income for married men given our definitions.

Table 2: Mean Marshallian Own-Wage Labor Supply Elasticities, Baseline Estimates for Selected Years

	Total		Extensive		Intensive	
	Single	Married	Single	Married	Single	Married
<u>A. Men</u>						
Bargain et al.	0.20	0.08	0.18	0.08	0.02	0.00
(2004)	(0.01)	(<0.01)	(0.01)	(<0.01)	(<0.01)	(<0.01)
1994	0.046	0.082	0.043	0.052	0.003	0.030
	(0.011)	(0.004)	(0.008)	(0.003)	(0.004)	(0.002)
2004	0.067	0.083	0.055	0.054	0.011	0.029
	(0.009)	(0.004)	(0.007)	(0.003)	(0.003)	(0.002)
2014	0.113	0.108	0.093	0.073	0.020	0.035
	(0.011)	(0.005)	(0.009)	(0.004)	(0.003)	(0.002)
<u>B. Women</u>						
Bargain et al.	0.20	0.08	0.18	0.08	0.02	0.00
(2004)	(0.01)	(<0.01)	(0.01)	(<0.01)	(<0.01)	(<0.01)
1994	0.293	0.252	0.229	0.196	0.064	0.056
	(0.014)	(0.009)	(0.011)	(0.007)	(0.004)	(0.002)
2004	0.186	0.278	0.139	0.218	0.047	0.060
	(0.009)	(0.009)	(0.007)	(0.008)	(0.003)	(0.002)
2014	0.265	0.279	0.208	0.228	0.057	0.051
	(0.012)	(0.008)	(0.010)	(0.007)	(0.003)	(0.002)

Notes: All estimates derived from a conditional logit model. Wages are predicted for non-workers using a Heckman selection model. Measurement error is integrated out based on 100 Halton draws. Standard errors are based on a parametric bootstrap and XXX draws.

Appendix

A. Simulating Elasticities

We estimate each elasticity via simulation. Specifically, for every observation i, choice j, and draw r, we calculate the fitted utility U_{ijr}. Using the closed-form expression for the probability of each choice, we calculate the expected hours worked for each observation and draw. To obtain elasticities, we can compare expected hours to those from alternative wage and non-labor income values. We estimate own-wage elasticities from a 10% increase in the before-tax wage rate, and we estimate income elasticities from a $1,000 increase in after-tax income for observations with positive non-wage income.[15] The percent change in expected hours worked is then used in the numerator of the elasticities.

In implementing the Hicksian elasticity, we use “income compensation” rather than “utility compensation” to adjust income – we adjust income so that the individual can afford the previous bundle, not the utility associated with the previous bundle. As is well-known, income compensation is more than the compensation needed for utility compensation.

We estimate the standard errors on our elasticities using a parametric bootstrap on 100 draws from the joint distribution of parameters.

B. Further Details on the March CPS

The sample is restricted to adults in single-unit households, defined as households with only one sub-family. We include heads of households, their spouses, and other nonrelatives only. Single adults must be in the age range of 26 to 55, and couples must have one spouse in the age range 26 to 55 and the other spouse in the age range 22 to 59.

[15] For those with negative non-wage income, we instead estimate the response to a $1000 reduction in non-labor income to avoid complications around zero.

We use Jaeger's (1997) method to map different education levels into five categories: did not complete high school, completed high school, some college, college degree, advanced degree. We consider any child, stepchild, grandchild, other relative under the age of 18, or sibling under the age of 18 in the household as a dependent. We top-code the number of dependents at 9.

We exclude individuals who (in the prior year) did not work due to illness/disability, attended high school or college full time or part time, earned non-zero farm income or business income, were self-employed or an unpaid family worker, or served in the armed forces. We also exclude individuals with extreme income, defined as a real hourly wage (as calculated by us) less than $4/hour or greater than $150/hour or real non-labor income (as calculated by us) less than -$1,000,000 or greater than $1,000,000.

If either spouse meets our exclusion criteria, we exclude the whole household.

We define hours as usual hours worked per week in the prior year. We top-code hours at 75. When discretized, 0 = {0 to 4}, 10 = {5 to 14}, 20 = {25 to 34}, 30 = {35 to 44}, 40 = {35 to 44}, 50 = {45 to 54}, 60 = {55 to 75}.

We define non-labor income as the sum of household dividend income, interest income, other property income (e.g., rent), and non-property income (e.g., alimony, child support).

We calculate hourly wages as pre-tax wage and salary income for the prior year divided by the product of usual hours worked per week in the prior year and weeks worked for pay in the prior year. We code hourly wage as "missing" if usual hours worked or weeks worked are zero or if wage and salary income is missing.

We use the Consumer Price Index Research Series Using Current Methods (CPI-U-RS) to convert everything to 2017 dollars.

We use observations with imputed values unless otherwise noted.

For our baseline specification, we use the standard Heckman correction to estimate wages for non-workers, following van Soest (1995); because numerous studies use imputed wages for everyone, we examine the sensitivity of our results to this choice below. To do so, we pool individuals in 5 year bins (e.g., 1979-1983, 1984-1988, etc.). We estimate wages separately for each year bin and sex. In the wage equation, we include a third-degree polynomial in the worker's age, indicator variables for their educational attainment, marital status, and number of dependents. We use non-wage income and indicator variables for having a child of various ages (under 2, 3-6 years old, 7-12 years old, 13-17 years old, and 18 years or older) as our excluded variables in the selection equation.

Table B.1 and B.2 present the basic descriptive statistics for all four demographic groups, and Table B.3 presents the results from the Heckman correction models for men and women for the years 2014-2018.

C. Calculating After-Tax Wages and Social Assistance Benefits

To fully exploit variation in the tax system, we would like to calculate after-tax income for each hours choice a household could make (7 for single males, 49 for married couples). Given we are using simulation methods that rely on 100 wage draws for each household, this entails calculating post-tax income 700 times for each single male and 4,900 for each married couple in our sample. This amounts to calculating after-tax income upwards of 319 million times (700*39,360 single men and 4,900*59,405 married couples).

To make this computationally feasible, we rely on a strategy similar to that described by Loffler, Siegloch, and Peichl (2018), estimating a flexible tax function and then using the tax function to simulate after-tax wages. Specifically, for each observation, we use NBER's TAXSIM version 27 to calculate after-tax income given observed wage and non-wage income for all 7 hours choices. For married men, we choose 7 random hours choices for their spouse.[16] Using these data, we estimate a regression model of after-tax income on a quadratic function of pre-tax income for each state-year-marital status group. This regression approximation for TAXSIM works well, in that the R-squared in each regression is at least 0.997.

For social benefits, we use the Urban Institute's TRIM3 and a flexible regression to predict government assistance. TRIM3 provides information on eligibility and receipt of government assistance (i.e., SNAP/Food Stamps, TANF/ADFC, and SSI) for each individual in our sample. We estimated a state-year-marital status-number of dependents-specific kinked regression model to predict benefits given counterfactual labor income bundles. The slope between household income and benefits received is allowed to change after the first $10k and $20k of (nominal) income. Between $30k to $50k, $50k to $70k, and $70k+, we use the average benefits received in that range. This regression approximation for TRIM3 works well in that the R-squared in each regression is at least 0.997.

D. Parameter Estimates

We present our key parameter estimates in Table D.1 and D.2 for key years.

[16] TAXSIM calculates after-tax income based on numerous household characteristics, including year, state, labor income, non-labor income, household demographics, etc. It works through submitting this information through an online request system.

E. Additional Results

We also have estimated models in which we adopt a quadratic utility model, following Bargain et al. (2014). For singles, the portion of utility that depends on observables is

(4) $$U_{ij} = \beta_{ci}C_{ij} + \alpha_c C_{ij}^2 + \beta_{hi}H_{ij} + \alpha_h H_{ij}^2 + \alpha_{ch}C_{ij}H_{ij} - \eta_i * 1\left(H_{ij} > 0\right)$$

where C_{ij} is household income net of taxes and H_{ij} for individual i and choice j. The model includes three parameters that vary across individuals:

$$\beta_{ci} = \beta_c^0 + X_i^c\beta_c^1 + u_i$$

(5) $$\beta_{hi} = \beta_h^0 + X_i^h\beta_h^1$$

$$\eta_i = \eta^0 + X_i^F\eta\ ,$$

where X_i^c and X_i^h include the number of dependents, X_i^F includes the number of children and an indicator for the presence of children aged 0-2 in the household, and u_i is a normally distributed, random preference for consumption. For models that are estimated over all age ranges, the taste modifiers (the X_i terms) also include age.

For couples, we additionally include a quadratic term for spouse hours, spouse hours interacted with indicators for the presence of children aged 0-2, 3-6, 7-12, and 13-17 in the household, and a fixed cost of working for the spouse, interacted with number of children in the household and an indicator for children aged 0-2. The random slope on consumption also includes an interaction with spouse hours.

$$V_{ij} = \beta_{ci}C_{ij} + \beta_{cc}C_{ij}^2 + \beta_{hmi}H_{ij}^m + \beta_{hfi}H_{ij}^f + \beta_{hmm}\left(H_{ij}^m\right)^2 + \beta_{hff}\left(H_{ij}^f\right)^2 + \beta_{chm}C_{ij}H_{ij}^m$$
$$+ \beta_{chf}C_{ij}H_{ij}^f + \beta_{hmhf}H_{ij}^m H_{ij}^f - \eta_i^m * 1\left(H_{ij}^m > 0\right) - \eta_i^f * 1\left(H_{ij}^f > 0\right) + \epsilon_{ij}$$

where $C_{ij}$is household income net of taxes, H_{ij} is the hours worked for males and females, and ϵ_{ij} are i.i.d Type 1 Extreme Value errors capturing unobserved heterogeneity. The model includes five parameters that vary with individual covariates,

$$\beta_{ci} = \beta_c^0 + X_i^c \beta_c + u_i$$

$$\beta_{hmi} = \beta_{hm}^0 + X_i^{hm} \beta_{hm}$$

$$\beta_{hfi} = \beta_{hf}^0 + X_i^{hf} \beta_{hf}$$

$$\eta_{mi} = \eta_m^0 + X_i^m \eta_m$$

$$\eta_{fi} = \eta_f^0 + X_i^f \eta_f$$

where X_i^c and X_i^{hm} include the number of dependents, u_i is a normally distributed individual preference heterogeneity, X_i^{hf} includes indicators for the presence of children aged 0-2, 3-6, 7-12, and 13-17 in the household, and X_i^m and X_i^f include number of children and an indicator for the presence of children aged 0-2 in the household.

In Figure E.1, we compare the quadratic utility model results (without the random slope) to our baseline results. The mean OWE are very similar in both cases.

Table B.1: Key Descriptive Statistics for Single Men and Women

	N	Work?	Hours|>0	Wages	Children	Non-Labor
A. Men						
1979	2,141	0.96	42.7	22.52	0.22	1,161
1984	2,606	0.95	42.0	24.00	0.19	2,007
1989	2,925	0.97	42.9	24.04	0.20	1,585
1994	3,060	0.96	43.3	23.56	0.19	2,342
1999	3,183	0.97	43.4	25.26	0.23	2,869
2004	4,274	0.96	42.3	25.61	0.22	1,730
2009	4,378	0.95	42.4	25.62	0.20	1,749
2014	4,029	0.93	42.5	25.80	0.24	3,456
2018	3,558	0.95	42.7	26.83	0.24	3,175
B. Women						
1979	4,356	0.82	38.4	15.76	1.37	2,686
1984	5,178	0.82	38.5	17.24	1.17	2,717
1989	4,967	0.86	39.6	18.69	1.10	2,697
1994	5,573	0.86	39.7	19.22	1.09	2,597
1999	4,900	0.92	40.3	20.13	1.04	3,011
2004	7,674	0.92	39.7	22.04	0.98	2,850
2009	7,227	0.91	39.3	21.70	1.00	2,515
2014	6,229	0.88	39.0	21.34	1.03	2,534
2018	5,407	0.90	39.8	21.73	1.01	2,780

Table B.2: Key Descriptive Statistics for Married Men and Women

	N	Work?	Hours|>0	Wages	Children	Non-Labor
A. Men						
1979	15,218	0.99	43.5	25.82	1.73	1,668
1984	15,577	0.97	42.7	26.06	1.60	2,140
1989	14,107	0.98	43.7	27.36	1.44	2,547
1994	14,231	0.97	44.1	27.38	1.48	2,786
1999	12,717	0.98	44.4	29.14	1.45	4,009
2004	20,634	0.96	43.8	30.17	1.44	2,915
2009	19,103	0.96	43.7	30.67	1.47	2,677
2014	15,949	0.95	43.5	31.05	1.51	3,929
2018	13,913	0.96	43.5	32.46	1.50	4,901
B. Women						
1979	--	0.64	34.3	9.35	--	--
1984	--	0.68	34.7	10.75	--	--
1989	--	0.75	35.7	13.03	--	--
1994	--	0.78	36.3	14.98	--	--
1999	--	0.79	37.1	17.89	--	--
2004	--	0.74	37.6	17.66	--	--
2009	--	0.77	37.4	17.89	--	--
2014	--	0.74	37.6	17.66	--	--
2018	--	0.75	38.2	19.26	--	--

Table B.3: Heckman Wage Equations for 2014-2018

	Men (N=93,621)		Women (N=103,915)	
	Particip.	Wage	Particip.	Wage
Age	-0.131	0.075	0.165	0.178
	(0.059)	(0.014)	(0.035)	(0.015)
Age square	0.355	-0.114	-0.369	-0.370
	(0.143)	(0.035)	(0.088)	(0.038)
Age cube	-0.033	0.005	0.026	0.026
	(0.011)	(0.003)	(0.007)	(0.006)
Ed: high school only	0.127	0.313	0.477	0.322
	(0.027)	(0.007)	(0.018)	(0.010)
Ed: some college	0.260	0.484	0.805	0.543
	(0.028)	(0.007)	(0.018)	(0.010)
Ed: bachelors	0.395	0.807	0.881	0.887
	(0.030)	(0.008)	(0.018)	(0.010)
Ed: bachelors +	0.554	1.012	1.144	1.162
	(0.035)	(0.008)	(0.021)	(0.011)
In couple	0.045	0.137	-0.537	0.016
	(0.020)	(0.005)	(0.012)	(0.005)
Number of children	0.040	0.012	-0.186	-0.040
	(0.017)	(0.002)	(0.008)	(0.002)
1[Children 0-2]	0.080		-0.198	
	(0.032)		(0.015)	
1[Children 3-7]	0.035		-0.125	
	(0.029)		(0.014)	
1[Children 8-13]	0.066		0.069	
	(0.029)		(0.014)	
1[Children 14-17]	0.087		0.192	
	(0.028)		(0.014)	
1[Children 18+]	0.056		0.205	
	(0.039)		(0.021)	
Relatives in household	0.195		0.005	
	(0.024)		(0.013)	
Non-wage income	-0.002		-0.002	
	(0.001)		(0.001)	
		0.178		0.292
		(0.031)		(0.011)
Constant	3.036	0.894	-0.969	-0.569
	(0.801)	(0.191)	(0.450)	(0.200)

Notes. Standard errors in parentheses.

Table D.1: Utility Parameter Estimates for Single Men

	1994	2004	2014
Log(C)	-0.14	3.43	6.36
	(1.25)	(1.05)	(1.00)
Log(L)	242.53	167.55	131.51
	(37.91)	(32.46)	(30.55)
Log(L) x Log(Age)	-84.85	-33.76	-5.63
	(20.22)	(17.26)	(12.29)
Log(L) x Log(Age)^2	12.10	4.82	1.00
	(2.78)	(2.36)	(2.22)
Log(L) x Dep18	0.10	-0.36	-0.57
	(0.02)	(0.15)	(0.16)
Log(L) x Dep2	0.97	0.66	0.71
	(0.83)	(1.05)	(0.86)
Log(C)^2	0.11	0.10	0.10
	(0.02)	(0.01)	(0.01)
Log(L)^2	-10.40	-11.03	-12.00
	(0.78)	(0.72)	(0.67)
Log(C) x Log (L)	-0.16	-0.88	-1.53
	(0.25)	(0.22)	(0.20)
Hours10	-3.36	-3.45	-3.71
	(0.25)	(0.20)	(0.22)
Hours20	-2.91	-2.97	-3.16
	(0.14)	(0.11)	(0.12)
Hours30	-2.86	-2.91	-2.88
	(0.11)	(0.09)	(0.09)

Notes: These results are for the baseline model in Section 5.1.

Table D.2: Utility Parameter Estimates for Single Women

	1994	2004	2014
Log(C)	12.35	13.56	18.21
	(1.03)	(0.85)	(1.03)
Log(L)	250.63	308.36	295.86
	(27.23)	(24.20)	(26.72)
Log(L) x Log(Age)	-42.87	-61.02	-46.91
	(14.69)	(12.96)	(14.15)
Log(L) x Log(Age)^2	6.06	8.49	6.68
	(2.02)	(1.78)	(1.94)
Log(L) x Dep18	0.47	0.25	0.34
	(0.10)	(0.08)	(0.09)
Log(L) x Dep2	1.04	1.41	0.97
	(0.27)	(0.26)	(0.28)
Log(C)^2	0.19	0.16	0.18
	(0.01)	(0.01)	(0.01)
Log(L)^2	-15.96	-18.42	-18.78
	(0.57)	(0.52)	(0.62)
Log(C) x Log (L)	-2.95	-3.15	-4.16
	(0.20)	(0.18)	(0.20)
Hours10	-3.06	-3.11	-3.28
	(0.12)	(0.11)	(0.12)
Hours20	-2.49	-2.56	-2.68
	(0.07)	(0.06)	(0.07)
Hours30	-2.27	-2.46	-2.28
	(0.06)	(0.05)	(0.05)

Notes: These results are for the baseline model in Section 5.1.

Table D.3: Utility Parameter Estimates for Married Couples

	1994	2004	2014
Log(C)	10.60	14.04	20.37
	(1.18)	(1.14)	(1.00)
Log(Lm)	191.41	278.50	271.76
	(29.52)	(30.88)	(30.59)
Log(Lm) x Log(Age)	-34.86	-87.14	-66.67
	(15.97)	(16.59)	(16.43)
Log(Lm) x Log(Age)^2	5.05	12.32	9.10
	(2.19)	(2.26)	(2.23)
Log(Lm) x Dep18	1.97	12.13	9.87
	(1.24)	(1.28)	(1.18)
Log(Lm) x Dep2	4.94	3.50	1.47
	(3.42)	(3.46)	(3.24)
Log(Lw)	216.93	227.36	254.77
	(24.82)	(25.64)	(24.94)
Log(Lw) x Log(Age)	-31.54	-45.25	-58.11
	(13.72)	(14.01)	(13.54)
Log(Lw) x Log(Age)^2	4.63	6.39	7.84
	(1.92)	(1.94)	(1.86)
Log(Lw) x Dep18	3.06	12.50	10.38
	(1.18)	(1.23)	(1.13)
Log(Lw) x Dep2	5.65	4.79	2.27
	(3.23)	(3.27)	(3.08)
Log(C)^2	0.23	0.22	0.16
	(0.01)	(0.01)	(0.33)
Log(Lm)^2	-13.94	-14.30	-15.43
	(0.33)	(0.34)	(0.32)
Log(Lw)^2	-18.13	-17.50	-16.52
	(0.35)	(0.35)	(0.33)
Log(C) x Log (Lm)	-1.70	-2.11	-3.00
	(0.18)	(0.17)	(0.16)
Log(C) x Log (Lw)	-1.02	-1.29	-1.62
	(0.14)	(0.13)	(0.12)
Log(Lm) x Log (Lw)	10.70	-6.39	-12.90
	(6.26)	(6.47)	(6.30)
Log(Lm) x Log (Lw) x Log(Age)	-4.79	6.87	9.29
	(3.47)	(3.54)	(3.44)
Log(Lm) x Log (Lw) x Log(Age)^2	0.69	-0.96	-1.22
	(0.48)	(0.49)	(0.47)
Log(Lm) x Log(Lw) x Dep18	-0.48	-2.80	-2.31
	(0.28)	(0.29)	(0.27)
Log(Lm) x Log(Lw) x Dep2	-1.07	-0.76	-0.37

	(0.77)	(0.78)	(0.72)
Hours10m	-3.51	-3.88	-4.27
	(0.19)	(0.21)	(0.23)
Hours20m	-3.05	-3.51	-3.52
	(0.10)	(0.12)	(0.11)
Hours30m	-3.11	-3.31	-3.16
	(0.07)	(0.08)	(0.07)
Hours10w	-2.87	-2.88	-3.22
	(0.06)	(0.06)	(0.07)
Hours20w	-2.18	-2.26	-2.49
	(0.04)	(0.04)	(0.04)
Hours30w	-2.22	-2.24	-2.36
	(0.04)	(0.04)	(0.04)

Notes: These results are for the baseline model in Section 5.1.

Figure E.1: Mean Own-Wage Labor Supply Elasticities, Baseline Estimates versus Estimates from Quadratic Utility Model