

2012

2012-4 Heterogeneity and Long-Run Changes in U.S. Hours and the Labor Wedge

Simona E. Cociuba

Alexander Ueberfeldt

Follow this and additional works at: <https://ir.lib.uwo.ca/economicscibc>



Part of the [Economics Commons](#)

Citation of this paper:

Cociuba, Simona E., Alexander Ueberfeldt. "Heterogeneity and Long-Run Changes in U.S. Hours and the Labor Wedge." CIBC Centre for Human Capital and Productivity. CIBC Working Papers, 2012-4. London, ON: Department of Economics, University of Western Ontario (2012).

**Heterogeneity and Long-Run Changes in
U.S. Hours and the Labor Wedge**

by

Simona E. Cociuba and Alexander Ueberfeldt

Working Paper # 2012-4

September 2012

Updated November 2014



CIBC Working Paper Series

Department of Economics
Social Science Centre
Western University
London, Ontario, N6A 5C2
Canada

This working paper is available as a downloadable pdf file on our website
<http://economics.uwo.ca/centres/cibc/>

Heterogeneity and Long-Run Changes in Aggregate Hours and the Labor Wedge*

SIMONA E. COCIUBA[†]

ALEXANDER UEBERFELDT

University of Western Ontario

Bank of Canada

November, 2014

Abstract

From 1961 to 2007, U.S. aggregate hours worked increased and the labor wedge—measured as the discrepancy between a representative household’s marginal rate of substitution and the marginal product of labor—declined substantially. The labor wedge is negatively related to hours and is often attributed to labor income taxes. However, U.S. labor income taxes increased since 1961. We examine a model with gender and marital status heterogeneity which accounts for the trends in the U.S. hours and the labor wedge. Apart from taxes, the model’s labor wedge reflects non-distortionary cross-sectional differences in households’ hours worked and productivity. We provide evidence that household heterogeneity is important for long-run changes in labor wedges and hours in other OECD economies.

Keywords: Labor Wedge, Household Aggregation, Female and Male Labor Supply, Gender Wage Gap, Labor Income Taxation. **JEL-Codes:** E24, H20, H31, J22.

*We thank Jim Dolmas, Lutz Hendricks, Gueorgui Kambovov, Narayana Kocherlakota, Jim MacGee, Ananth Ramanarayanan, Mark Wynne, Carlos Zarazaga, the editor, B. Ravikumar, and an anonymous referee for many useful comments. We thank Wenya Wang for excellent research assistance. [†]Corresponding author. University of Western Ontario, Department of Economics, Social Science Center, Room 4071, London, Ontario, Canada, N6A 5C2. Tel.: +1 519 661 3500; fax: +1 519 661 3666. E-mail address: sociuba@uwo.ca.

1 Introduction

From 1961 to 2007, U.S. average hours worked—defined as total market hours per working-age person—increased by 13 percent. Concurrently, the U.S. labor wedge, measured as the discrepancy between a representative household’s marginal rate of substitution between consumption and leisure (MRS) and the marginal product of labor (MPL), declined by 37 percent. Somewhat surprisingly, despite being much larger than the frequently studied short-run fluctuations, the long-run change in the U.S. labor wedge has received little attention in the literature.¹ Such long-run trends are negatively related to changes in hours and are often attributed to variations in taxes (e.g., Mulligan (2002)) because the labor wedge—denoted here by Δ —resembles a labor income tax (i.e., by definition $1 - \Delta \equiv \frac{MRS}{MPL}$). However, U.S. labor income taxes increased since 1961. As a result, standard representative agent models (e.g., Prescott (2004), Ohanian, Raffo, and Rogerson (2008)) deliver counterfactual predictions for the U.S., as higher taxes imply lower hours and a higher labor wedge.

In this paper, we show that incorporating gender and marital status heterogeneity in an otherwise standard growth model is quantitatively important in accounting for the observed trends in U.S. hours and the labor wedge. Our focus on household heterogeneity is motivated by the large changes in hours and wage rates by gender and marital status since 1961. Married women’s hours more than doubled and men’s hours declined, while gender wage gaps decreased substantially. In our model, shrinking gender wage gaps contribute to an increase in aggregate and women’s hours and deliver a decline in the measured labor wedge, in spite of higher taxes. A key takeaway is that large changes in cross-sectional heterogeneity over time are reflected in long-run changes in the measured labor wedge.

The intuition for why cross-sectional heterogeneity in wages and hours impacts the measured labor wedge is straightforward. In a representative agent model, the labor wedge is

¹Given a high labor supply elasticity (typically used in macro studies), Shimer (2010) documents a decline in the U.S. labor wedge from 1959 to 2007 of about 35 percent, consistent with our calculations (see Figure 1.1 in his book). The U.S. labor wedge decline is about 1.5 times larger for a low labor supply elasticity. In addition, Shimer reports that business cycle fluctuations of the U.S. labor wedge have a standard deviation between 1.8 and 5.5 percent, the larger number corresponding to a low labor supply elasticity.

measured from the intratemporal equilibrium condition which relates the *MRS* to the *MPL* using aggregate data which averages out cross-sectional heterogeneity. In a heterogeneous agent model, the labor wedge is derived from a weighted aggregate of the households' intratemporal equilibrium conditions. These equations are nonlinear relationships between consumption, hours and wages, and, thus, cross-sectional differences in hours and wages do not average out. We formalize this idea in a simple static model where households differ in their labor productivity. We show that, the larger are differences in productivity and hours across households, the larger is the discrepancy between the aggregate *MRS* and the aggregate *MPL*, i.e. the labor wedge. It follows that changes in cross-sectional differences in productivity and hours map into changes in the measured labor wedge.

To quantify the contribution of this mechanism to long-run changes in U.S. hours and the labor wedge, we examine a standard model augmented with three types of households: married couples, single women and single men.² In our model, women receive a lower hourly wage rate compared to men, due to lower productivity and discrimination (as suggested by Goldin (1992)), and the labor income of all households is taxed. We evaluate the impact of taxes and cross-sectional wage heterogeneity on the documented trends in U.S. data. Higher taxes deliver counterfactual predictions for U.S. hours and the labor wedge. However, reductions in gender wage gaps for married couples and singles (reflecting lower discrimination, or higher relative productivity of women, or a combination of the two) generate a long-run increase in aggregate and women's hours and a long-run decline in the aggregate labor wedge.

A calibrated version of our baseline model—with gender wage gaps and taxes measured from U.S. data—accounts for 63 percent of the increase in average hours worked, 86 percent of the increase in married women's hours and about 30 percent of the decline in the labor wedge. To isolate the contribution of gender wage gaps, we consider a variation of our baseline model with constant taxes. This experiment accounts for virtually all of the increase in aggregate and women's hours and about 54 percent the decline in the labor wedge. The model does

²It is possible to extend the model to allow for other types of heterogeneity, e.g. differences in skill levels implied by the rise in the education premium. Such extensions are beyond the scope of this paper.

not account for all of the decline in the labor wedge, since it has difficulty capturing the observed increase in the U.S. consumption to output ratio since the mid 1980s.

We consider two extensions of our analysis which have been suggested as possible explanations for the long-run increase in U.S. hours, and evaluate their impact on the labor wedge. Following Attanasio, Low, and Sánchez-Marcos (2008), we incorporate child care costs in our model. We find that reductions in this cost lead to additional increases in married women’s hours and in aggregate hours, but contribute only a further 6 percentage points to the decline in the labor wedge. Second, we examine whether changes in leisure time which reflect non-market hours (e.g. time spent in home production) can improve predictions for the U.S. labor wedge, as suggested by Ohanian, Raffo, and Rogerson (2008). Our back-of-the-envelope calculations suggest that the increase in U.S. leisure time since mid-1960s—which varies from 2 percent in Ramey and Francis (2009) to a range of 5.4 to 15 percent in Aguiar and Hurst (2007)—can account for 6.5 to 50 percent of the decline in the U.S. labor wedge. We view both extensions as complementary to our analysis.

A natural question is whether the mechanism we analyzed in detail for the U.S. is quantitatively important in other economies. We extend our analysis of long-run changes in hours and labor wedges to Canada and Germany. In Canada, similar to the U.S., the closing of the gender wage gaps and increases in female hours dominate the increase in taxes and lead to a decline in the labor wedge over the last four decades. Germany is especially interesting, since taxes increased by more than the increase in the labor wedge over the last two decades. Reductions in cross-sectional heterogeneity in Germany (captured by a shrinking gender wage gap for married couples and higher married women’s hours) are important as they partly undo the effect of higher taxes, bringing the model’s labor wedge closer to that measured from aggregate data. The improved predictions for the labor wedges lead us to conclude that heterogeneity also helps account for changes in hours in Canada and Germany.

Lack of *long-run* micro survey data prevents us from extending this analysis to a larger number of countries. However, our mechanism is broadly consistent with aggregate data on

hours worked, tax rates and measured labor wedges for a number of other OECD economies where gender wage gaps shrunk. In economies with large changes in hours and the labor wedge, cross-sectional heterogeneity can be quantitatively important in reversing the effects of higher taxes (as observed in Spain, Italy and Belgium), or in accounting for reductions in labor wedges which are larger than reductions in tax rates (as observed in Netherlands, Finland and the U.K.).

The labor wedge is a reduced-form diagnostic tool used extensively in the literature to identify types of distortions that improve a model’s predictions for hours and other aggregate data (Chari, Kehoe, and McGrattan (2007)).³ Much of this literature has focused on accounting for the measured labor wedge via distortions such as taxes, monopoly power, sticky wages or search frictions (e.g., Mulligan (2002), Chari, Kehoe, and McGrattan (2007), Shimer (2009)). Our paper contributes to this literature by highlighting that long-run changes in the measured labor wedge can occur due to non-distortionary cross-sectional differences in households’ labor supplies and productivities.⁴

The idea that aggregation across heterogenous households may lead to a labor wedge was pointed out by Maliar and Maliar (2003) and Chang and Kim (2014) in the context of business cycle models. Similar to Maliar and Maliar (2003), we analytically relate cross-sectional household heterogeneity to our model’s labor wedge, while Chang and Kim (2014) illustrate such a relationship in a quantitative exercise.⁵ Our contributions relative to these works are twofold. First, we quantify the role of household heterogeneity in accounting for the long-run decline in the U.S. labor wedge, despite the observed increase in U.S. labor

³Chari, Kehoe, and McGrattan (2007) show that distortions that manifest themselves as labor wedges are important in understanding the U.S. Great Depression and the 1982 recession. Other studies also show the labor wedge is important in generating predictions in line with macroeconomic aggregates in many countries (e.g., Brinca (2014) and citations therein).

⁴Other papers attribute the labor wedge to shifts in preferences (e.g., Parkin (1988) and Hall (1997)), or firm level financial frictions (e.g., Arellano, Bai, and Kehoe (2012) and Jermann and Quadrini (2012)). Karabarbounis (2014a) also examines the contribution of non-distortionary factors—namely home production—to changes in the labor wedge in an international business cycle model.

⁵In Maliar and Maliar (2003), the representative consumer’s intratemporal condition can be written as $\frac{MRS}{MPL} = \frac{1}{\pi_t}$, where π_t is a composite of individual agents’ characteristics from the equivalent heterogenous agent economy. Maliar and Maliar (2003) refer to π_t as a labor shock from aggregation. In the language of our paper, $1 - \frac{1}{\pi_t}$ represents the labor wedge.

income taxes. Second, we use our model’s analytical relationship between heterogeneity and the labor wedge to extend the analysis to other countries.

Our paper falls into the class of household-based explanations of long-run changes in the labor wedge. Our focus on developing the household side of the standard model is consistent with Karabarbounis (2014b), who shows that fluctuations in the labor wedge for the U.S. and other OECD economies are mostly accounted for by discrepancies between the *MRS* and the real wage, rather than discrepancies between the *MPL* and the real wage. Our contribution is to show that discrepancies between the aggregate *MRS* and the wage rate are also important in a long-run analysis of the labor wedge.

Our paper also relates to the work on taxation and long-run changes in hours. Prescott (2004) and Ohanian, Raffo, and Rogerson (2008) find that taxes account for most of the variations in hours over time and across countries, but identify the U.S. experience as an exception, as both taxes and hours increased since the 1960s.⁶ Our paper shows that incorporating female labor supply and shrinking gender wage gaps allows an otherwise standard growth model to capture most of the observed increase in U.S. hours, despite higher taxes. Moreover, our results connect the labor wedge to the literature analyzing gender wage gaps and women’s hours (e.g., Goldin (1992), Jones, Manuelli, and McGrattan (2003), Bar and Leukhina (2011) and Attanasio, Low, and Sánchez-Marcos (2008)).

The paper is organized as follows. Section 2 documents the U.S. trends in hours and the labor wedge and presents a simple model to illustrate that household heterogeneity can be important in understanding these trends. Section 3 presents our model with gender and marital status heterogeneity and the analytical derivation of the model’s labor wedge. Section 4 presents the quantitative experiments and results. In Section 5, we discuss the importance of changes in child care costs and leisure time for U.S. hours and the labor wedge. Section 6 extends our analysis to other OECD economies. We conclude in Section 7.

⁶Scandinavian countries are also exceptions, see Ragan (2006) and Rogerson (2007). Other countries where taxes and hours increased over time are discussed in Section 6.

2 U.S. Data and a Simple Static Model

In this section, we first document the trends in U.S. hours and the labor wedge. Then, we develop a simple model to illustrate that household heterogeneity in productivity and labor supply can help in understanding these trends.

2.1 Long-Run Changes in U.S. Hours and the Labor Wedge

Throughout the 1960s and the 1970s, the U.S. working-age population worked, on average, 25 hours per week (Figure 1). Since the early 1980s, aggregate hours worked increased steadily to 28 hours per week in 2007. This increase in aggregate hours is driven by women. Married women’s average hours more than doubled from 10 hours a week in 1961 to 23 hours a week in 2007. Single women’s average hours increased slightly, while men’s hours declined over the 47 year period (Appendix A.1 provides details on the data sources and computations).⁷

The increase in U.S. aggregate hours from 1961 to 2007 was accompanied by a decline in the U.S. labor wedge (Figure 1). As is standard in the literature, we measure the labor wedge using U.S. aggregate data and the intratemporal labor equilibrium condition from the neoclassical growth model with a representative household (e.g., Parkin (1988), Hall (1997), Mulligan (2002), Chari, Kehoe, and McGrattan (2007) and Shimer (2009)). Specifically, in this model, the intratemporal condition equates the marginal product of labor (MPL) to the representative household’s marginal rate of substitution between consumption and leisure (MRS). As this relationship does not hold in the data, the aforementioned literature defined the labor wedge at time t —denoted here by Δ_t —as the discrepancy between the MRS and the MPL . Namely, $1 - \Delta_t \equiv MRS_t/MPL_t$.

Many macroeconomic studies (including, but not limited to, the ones cited above) assume a Cobb-Douglas production function. The MPL can then be written as $(1 - \theta) y_t/l_t$, where y_t denotes output per person, l_t denotes aggregate hours worked per person and $1 - \theta$ is the

⁷U.S. aggregate hours worked have declined during the most recent recession, dated by the NBER to last from December 2007 to June 2009. The changes in hours observed since 2007 are interesting in their own right, but are not analyzed in this paper.

labor income share. Time separable log preferences in consumption and leisure—frequently used in macroeconomic studies—give a *MRS* equal to $\alpha (c_t + \phi g_t) / (1 - l_t)$, where c_t denotes private consumption per person, g_t denotes public consumption per person, α is the leisure utility parameter and ϕ measures the marginal rate of substitution between government and private consumption. With these functional forms, the labor wedge, Δ_t , is defined as in (1).

$$1 - \Delta_t \equiv \left(\frac{\alpha (c_t + \phi g_t)}{1 - l_t} \right) / \left(\frac{(1 - \theta) y_t}{l_t} \right) = \frac{\alpha}{1 - \theta} \cdot \frac{(c_t + \phi g_t)}{y_t} \cdot \frac{1}{\frac{1}{l_t} - 1} \quad (1)$$

To measure Δ_t , we use U.S. data on c_t, g_t, y_t, l_t and parameters $\theta = 0.33$, $\alpha = 1.6$ and $\phi = 1$ (as in Prescott (2004) and Ohanian, Raffo, and Rogerson (2008)). As visible in Figure 1, the U.S. measured labor wedge is fairly constant for the period 1961 to 1980, and declines between the early 1980s and 2007. This substantial decline in the labor wedge is also documented by Mulligan (2002) and Shimer (2010), under different functional forms for the marginal value of time (*MRS*).

The labor wedge is used extensively in the literature as a diagnostic tool to help identify types of distortions that improve a model’s predictions for hours worked. Typically, the labor wedge is thought of as measuring labor market distortions. With this interpretation, the increase in U.S. aggregate hours since 1961 could be attributed to a decline in labor market distortions. However, one candidate of such distortions, the effective labor income tax—defined as a combination of consumption and labor income taxes, as in Prescott (2004) and Shimer (2009)—rose over the last five decades (Figure 1). The focus of our paper is to show that changes in the labor wedge are not entirely driven by labor market distortions, such as taxes, but also reflect changes in non-distortionary factors, such as the labor supplies and relative productivities of various subgroups of the population.

2.2 Heterogeneity and the Labor Wedge

We present a simple static model with heterogenous households to build intuition for the idea that cross-sectional differences in productivity and labor supply can be quantitatively important for understanding the long-run decline in the U.S. labor wedge.

The economy consists of J types of households. Each household j has one member who is endowed with one unit of time and has a fixed amount of capital given by k_j . Households supply labor in the market, but differ in their productivity, which is denoted by z_j . The maximization problem solved by household j is:

$$\begin{aligned} & \max_{c_j, l_j} \log(c_j + \phi g) + \alpha \log(1 - l_j) \\ & \text{subject to: } c_j \leq rk_j + (1 - \tau_l) wz_j l_j + \psi_j \end{aligned}$$

where c_j denotes private consumption, g denotes government consumption per person, l_j is the fraction of available time devoted to work and $1 - l_j$ is leisure time. As before, α is the leisure utility parameter and ϕ measures the marginal rate of substitution between government and private consumption. Households receive a wage w per unit of effective labor, $z_j l_j$, and capital income rk_j for renting the capital stock to the firm. We include capital in the static model to be analogous to our model in Section 3. Labor income is taxed at rate τ_l , and ψ_j are lump-sum transfers from the government.

The representative firm uses capital, $K = \sum_j k_j$, and effective labor, $\tilde{L} = \sum_j (z_j l_j N_j)$ where N_j is the number of households of type j , to produce output according to the Cobb-Douglas production function: $Y = AK^\theta \tilde{L}^{1-\theta}$. Here, A denotes the total factor productivity and θ is the capital income share. The wage rate per unit of effective labor is given by $w = (1 - \theta) y / \tilde{l}$, where $y = Y/N$ is the output per person, $N = \sum_j N_j$ is the total population and $\tilde{l} = \tilde{L}/N$ is the aggregate effective labor per person.

In this simple model, it is straightforward to show that there exists a wedge between the aggregate marginal rate of substitution between consumption and leisure (MRS) and the

marginal product of an hour worked (*MPL*). To derive the expression for the labor wedge, we aggregate the optimality conditions which equate household j 's marginal rate of substitution to its after-tax marginal product of labor, i.e. $\alpha(c_j + \phi g)/(1 - l_j) = (1 - \tau_l)wz_j$ for each j . These equations are nonlinear relationships between consumption, hours and wages. As a result, cross-sectional differences between households do not average out, and are mapped into the labor wedge. Aggregation across households yields equation (2).⁸

$$\frac{\alpha(c + \phi g)}{1 - l} = (1 - \tau_l) \left(\sum_j z_j \frac{1 - l_j}{1 - l} \frac{N_j}{N} \right) \frac{l}{\tilde{l}} \cdot (1 - \theta) \frac{y}{l} \quad (2)$$

where $c \equiv \sum_j c_j N_j / N$ is aggregate private consumption per person, $l \equiv \sum_j (l_j N_j) / N$ is aggregate hours worked per person, and where we have used the expression for the wage w . The labor wedge, Δ , defined as in equation (1), can then be written as in (3).

$$1 - \Delta \equiv \left(\frac{\alpha(c + \phi g)}{1 - l} \right) / \left((1 - \theta) \frac{y}{l} \right) = (1 - \tau_l) \left(\sum_j z_j \frac{1 - l_j}{1 - l} \frac{N_j}{N} \right) \frac{l}{\tilde{l}} \quad (3)$$

The labor wedge in this simple model is partly due to distortionary taxes, but also reflects non-distortionary factors such as differences in households' productivities and labor supply decisions. Both of these dimensions of heterogeneity are needed to generate the labor wedge in equation (3). In particular, if households have the same productivity, i.e. $z_j = z$ for all j , or if households work the same number of hours, i.e. $l_j = l$ for all j , the labor wedge in equation (3) reduces to $1 - \tau_l$, the expression encountered in the neoclassical growth model with a representative household and labor income taxes.⁹

To gain further insights from equation (3), we consider a numerical example which illustrates that as differences in households' productivities and hours worked shrink, so does the

⁸We multiply the individual optimality conditions by the fraction of agents of type j in the total population, N_j/N , and sum up to get: $\alpha \sum_j (c_j + \phi g) N_j / N = (1 - \tau_l) \sum_j z_j (1 - l_j) (N_j / N) w$. Next, we substitute in the expression for the wage and divide both sides by $(1 - l)$.

⁹If $z_j = z$ for all j , then $\tilde{l} = z \cdot \frac{\sum_j l_j N_j}{N} = zl$ and $1 - \Delta = (1 - \tau_l) \left(\sum_j \frac{1 - l_j}{1 - l} \frac{N_j}{N} \right) = 1 - \tau_l$. If $l_j = l$ for all j , then $\tilde{l} = l \cdot \sum_j (z_j N_j) / N$ and $1 - \Delta = (1 - \tau_l) \left(\sum_j z_j \frac{N_j}{N} \right) \frac{l}{\tilde{l}} = 1 - \tau_l$.

labor wedge. Consider an economy with two types of households of equal mass, $N_j/N = 0.5$ for $j \in \{1, 2\}$, and no tax distortions, $\tau_l = 0$. Let's examine different scenarios for the households' labor supplies, l_j , and productivities, z_j . First, assume type 2 households are 30 percent less productive and work only a quarter of the time per week compared to type 1 households. Let $z_1 = 1.00$, $z_2 = 0.70$, $l_1 = 0.40$ and $l_2 = 0.10$. These differences generate a wedge between the aggregate *MRS* and the *MPL*—computed from equation (3)—equal to $\Delta = 0.128$. Smaller differences in both hours and productivity ($z_1 = 1.00$, $z_2 = 0.85$, $l_1 = 0.40$ and $l_2 = 0.30$) reduce the labor wedge to about $\Delta = 0.02$. This simple illustration shows that reductions in cross-sectional heterogeneity can result in a substantial decline in the labor wedge. This finding motivates our model in Section 3, where cross-sectional differences between households are reflected in gender wage gaps and hours differences. Section 4 shows that shrinking gender wage gaps and the ensuing increase in women's hours are quantitatively important for understanding the long-run decline in the U.S. labor wedge. Section 6 provides international support for this mechanism using data for other OECD economies.

3 General Model

To quantify the importance of our mechanism to the long-run trends in U.S. data, we consider a neoclassical growth model with three types of households: married couples, single females, and single males. The labor supply decisions of individuals are influenced by several factors, of which the most important are gender wage gaps and effective labor income taxes.

Let N_t be the total population at time t . Let N_{pt} , N_{fst} , and N_{mst} be the total number of married couples, single females and single males, respectively. Males and females in our model differ for two reasons. First, married and single females receive a lower wage than males, consistent with the data. Second, individuals in a married couple have different utility weights. In quantitative experiments, reductions in gender wage gaps generate an increase in female hours worked over time, while the utility weights pin down the relative level of

hours for married men and women in the first period of our model (as detailed in Section 4).

As in Jones, Manuelli, and McGrattan (2003), we assume married couples choose streams of consumption, labor supply and investment to solve their joint decision problem with utility weights given by λ_f and λ_m .

$$\max \sum_{t=0}^{\infty} \beta^t [\lambda_f U_f (c_{fpt} + \phi g_t, 1 - l_{fpt}) + \lambda_m U_m (c_{mpt} + \phi g_t, 1 - l_{mpt})] N_{pt}$$

subject to :

$$c_{fpt} + c_{mpt} + x_{pt} \leq [(1 - \tau_{kt}) r_t + \delta \tau_{kt}] k_{pt} + (1 - \tau_{lt}) w_t [l_{mpt} + (1 - \Gamma_{pt}) l_{fpt}] + \psi_{pt}$$

$$\frac{N_{pt+1}}{N_{pt}} k_{pt+1} \leq x_{pt} + (1 - \delta) k_{pt}$$

Here, subscripts f and m denote female and male, subscript p indicates a married couple or *partnership*, and t is the time subscript. The utility of a married individual of gender $j \in \{f, m\}$ is defined over streams of private consumption, c_{jpt} , average government consumption, g_t , and leisure time, $1 - l_{jpt}$, where available time is normalized to 1 and l_{jpt} is the labor supply expressed as the fraction of available time worked. The discount factor is $\beta \in (0, 1)$. The parameter $\phi \in (0, 1)$ measures the marginal rate of substitution between private and government consumption. The married couple owns capital stock, k_{pt} , which depreciates at rate δ and is augmented by investments, x_{pt} . The capital stock is rented to the firm at interest rate r_t , and capital income net of depreciation is taxed at rate τ_{kt} . The married couple pays taxes on labor income at rate τ_{lt} and receives lump-sum transfers, ψ_{pt} .¹⁰

In our model, married males receive an hourly wage rate of w_t , while married females receive only $w_t (1 - \Gamma_{pt})$ per hour worked. Here, $\Gamma_{pt} \in (0, 1)$ represents the *exogenous* gender wage gap for married couples.¹¹ We assume that productivity differences account for fraction

¹⁰In our model, the effective labor income tax (defined in Section 4.1) is the same for singles and married individuals, as well as for men and women. We have constructed estimates of average income taxes for single men, single women, married men and married women using the methodology in Kryvtsov and Ueberfeldt (2007). We find that while the level of the tax varies slightly, the increase in the income tax between 1961 and 2001 is comparable across groups. Moreover, Bar and Leukhina (2009) find that the U.S. tax reforms of 1980s have a small effect on married females participation. For these reasons, we do not consider different tax rates for the different households in our model.

¹¹A few studies in the literature endogenize the gender wage gap. Erosa, Fuster, and Restuccia (2002,

$\mu \in [0, 1]$ of the gender wage gap, while discrimination accounts for the remainder. In particular, the hourly wage rate received by a married woman can be written as:

$$w_t(1 - \Gamma_{pt}) = w_t(1 - \mu\Gamma_{pt}) - w_t(1 - \mu)\Gamma_{pt} \quad (4)$$

where $w_t(1 - \mu\Gamma_{pt})$ is the wage rate women should receive given their marginal product of labor (i.e. taking into account productivity differences relative to men), while the term $w_t(1 - \mu)\Gamma_{pt}$ represents the portion of the wage rate lost due to discrimination. Our assumption is motivated by Goldin (1992), who shows that some of the U.S. gender gap in earnings for various occupations can be explained by differences in observable attributes between men and women, such as job experience, education. However, a substantial part of the earnings gap remains unexplained and is attributed to discrimination.¹²

Measures of wage discrimination from U.S. data—such as those discussed in Goldin—vary over time. For simplicity, we consider that the fraction of the gender gap accounted for by discrimination is constant over time in the model and is given by $1 - \mu$. In Section 4, we evaluate the importance of this assumption for female hours and the U.S. labor wedge by presenting results under two extreme scenarios: the gender wage gap is due entirely to discrimination or due entirely to productivity differences.

For $\mu \in (0, 1)$, our model is consistent with the view that reductions in the gender gap observed in the U.S. since the early 1960s, were a consequence of improvements in productivity of women and reductions in discrimination. As seen in equation (4), when the gender wage gap, Γ_{pt} , shrinks over time, the marginal product of a married woman’s labor, $w_t(1 - \mu\Gamma_{pt})$, increases, while the wages lost due to discrimination, $w_t(1 - \mu)\Gamma_{pt}$, decline.

2005) endogenize the married women’s gender wage gap, by relating it to the human capital lost after child birth. In Jones, Manuelli, and McGrattan (2003) the gender wage gap is partly endogenous, due to human capital decisions, and partly exogenous, due to direct wage discrimination or to the existence of a “glass ceiling” that keeps women from rising in the hierarchy of organizations.

¹²Goldin (1992) documents that wage discrimination accounted for about 20 percent of the difference in male and female earnings in manufacturing jobs in early 1900, and about 55 percent for office work in 1940.

Single males and females solve the following maximization problem:

$$\begin{aligned} & \max \sum_{t=0}^{\infty} \beta^t U_j (c_{jst} + \phi g_t, 1 - l_{jst}) N_{jst} \\ \text{subject to} \quad & : \\ & c_{jst} + x_{jst} \leq [(1 - \tau_{kt}) r_t + \delta \tau_{kt}] k_{jst} + (1 - \tau_{lt}) w_t (1 - I_j \Gamma_{st}) l_{jst} + \psi_{jst} \\ & \frac{N_{jst+1}}{N_{jst}} k_{jst+1} \leq x_{jst} + (1 - \delta) k_{jst} \end{aligned}$$

where, as before, subscripts $j \in \{f, m\}$ and t denote gender and time, and subscript s indicates a single individual. We use similar notational conventions as in the married couple's problem. The indicator function I_j equals 1 if $j = f$ and zero otherwise and is used to show that single males receive hourly wage rate w_t , while single females receive $(1 - \Gamma_{st}) w_t$. Here, $\Gamma_{st} \in (0, 1)$ represents the *exogenous* gender wage gap for singles. As before, the parameter μ governs the share of the gender wage gap accounted for by productivity differences. In our numerical experiments, the gender wage gap for singles, Γ_{st} , differs from the one for married couples, Γ_{pt} , consistent with U.S. data.

There is a representative firm with a constant returns to scale production function that rents capital, K_t , and pays for effective labor, \tilde{L}_t . The firm's problem is given below.

$$\begin{aligned} & \max F(K_t, \tilde{L}_t) - r_t K_t - w_t \tilde{L}_t \\ \text{subject to: } & F(K_t, \tilde{L}_t) = K_t^\theta (\zeta_t \tilde{L}_t)^{1-\theta} \end{aligned} \tag{5}$$

There is labor augmenting technical progress at a constant yearly rate of $\gamma - 1$, that is, $\zeta_t = \zeta_0 \gamma^t$. The aggregate resource constraints for capital and effective labor are below.

$$\begin{aligned} K_t &= k_{pt} N_{pt} + k_{fst} N_{fst} + k_{mst} N_{mst} \\ \tilde{L}_t &= l_{mpt} N_{pt} + (1 - \mu \Gamma_{pt}) l_{fpt} N_{pt} + (1 - \mu \Gamma_{st}) l_{fst} N_{fst} + l_{mst} N_{mst} \end{aligned}$$

The wage bill in (5) is given by $w_t \tilde{L}_t$. Here, w_t is the wage rate per unit of effective labor and also the wage rate per hour worked by men. In the expression for \tilde{L}_t , the terms $(1 - \mu \Gamma_{it})$ for $i \in \{p, s\}$ measure the productivity of a married or single woman relative to men. Recall that women do not get paid their marginal product of $w_t (1 - \mu \Gamma_{it})$, but receive the lower hourly wage rate of $w_t (1 - \Gamma_{it})$ due to discrimination (as seen in equation (4) for married women). The difference between their marginal product and the wage rate received is equal to $w_t (1 - \mu) \Gamma_{it}$, and is collected by the government as revenue from discrimination.

The resource constraint in the economy is: $F(K_t, \tilde{L}_t) = C_t + X_t + G_t$, where aggregate consumption is $C_t \equiv N_{pt}(c_{mpt} + c_{fpt}) + N_{mst}c_{mst} + N_{fst}c_{fst}$, aggregate investment is $X_t \equiv N_{pt}x_{pt} + N_{mst}x_{mst} + N_{fst}x_{fst}$ and $G_t \equiv N_t g_t$ denotes government spending. In the quantitative analysis, the government consumption is exogenous and varies over time.

The government collects revenues from discrimination and from capital and labor income taxation. Revenues are used for government consumption expenditures and lump-sum rebates to households. The lifetime budget constraint of the government is:

$$\sum_{t=0}^{\infty} \frac{1}{\pi_t} (\Psi_t + G_t) = \sum_{t=0}^{\infty} \frac{1}{\pi_t} \{[\tau_{kt} r_t - \delta \tau_{kt}] K_t + \Upsilon_t\}$$

$$\text{where } \pi_t \equiv \begin{cases} 1 & \text{for } t = 0 \\ \prod_{\zeta=1}^t (1 - \delta + R_{\zeta}) & \text{for } t \geq 0 \end{cases} \quad (6)$$

where, $R_t \equiv (1 - \tau_{kt}) r_t + \delta \tau_{kt}$, aggregate transfers are $\Psi_t \equiv N_{pt} \psi_{pt} + N_{fst} \psi_{fst} + N_{mst} \psi_{mst}$, and aggregate labor revenues, Υ_t , are defined in (7).

$$\begin{aligned} \Upsilon_t \equiv & [\tau_{lt} w_t N_{pt} l_{mpt} + \tau_{lt} w_t N_{mst} l_{mst} + \tau_{lt} (1 - \Gamma_{pt}) w_t N_{pt} l_{fpt} + \tau_{lt} (1 - \Gamma_{st}) w_t N_{fst} l_{fst}] \\ & + [w_t (1 - \mu) \Gamma_{pt} N_{pt} l_{fpt} + w_t (1 - \mu) \Gamma_{st} N_{fst} l_{fst}] \end{aligned} \quad (7)$$

The first four terms in equation (7) represent revenues collected from labor income taxation. In addition, women's labor income is subject to discrimination which raises revenues

equal to $w_t (1 - \mu) \Gamma_{pt} N_{pt} l_{fpt} + w_t (1 - \mu) \Gamma_{st} N_{fst} l_{fst}$.

In our quantitative experiments, we allow the effective labor income taxes, τ_{lt} , the gender wage gaps, Γ_{st} and Γ_{pt} , the government consumption, g_t , and the population fractions, $n_{pt} \equiv N_{pt}/N_t$, $n_{fst} \equiv N_{fst}/N_t$, $n_{mst} \equiv N_{mst}/N_t$, to vary exogenously over time. We allow the population fractions to vary since there has been a large increase in the fraction of singles and a corresponding decline in the fraction of married couples since 1961. These time-varying inputs are measured from U.S. data (see Figure 2) and discussed in Section 4.

3.1 Aggregation and the Labor Wedge

We derive the labor wedge in our model and show that it depends on taxes, as suggested by previous studies, and on gender wage gaps, female labor supplies and aggregate labor supply. In Section 4, we evaluate the quantitative importance of taxes and gender wage gaps in accounting for the changes in U.S. labor wedge and hours worked.

To obtain an expression for the labor wedge we aggregate the model's intratemporal labor equilibrium conditions for married and single men and for married and single women, summarized in equations (8) and (9), respectively.

$$\frac{\alpha (c_{mit} + \phi g_t)}{1 - l_{mit}} = (1 - \tau_{lt}) w_t, \text{ for } i \in \{p, s\} \quad (8)$$

$$\frac{\alpha (c_{fit} + \phi g_t)}{1 - l_{fit}} = (1 - \tau_{lt}) (1 - \Gamma_{it}) w_t, \text{ for } i \in \{p, s\} \quad (9)$$

We multiply each of the intratemporal conditions by the fraction of households of that type (i.e. the fraction of married couples, n_{pt} , and the fractions of singles, n_{fst} and n_{mst}) and sum up to obtain equation (10). A full derivation is provided in Appendix A.2.

$$\frac{\alpha (c_t + \phi g_t)}{1 - l_t} = (1 - \tau_{lt}) \left(1 - \frac{n_{pt} \Gamma_{pt} (1 - l_{fpt}) + n_{fst} \Gamma_{st} (1 - l_{fst})}{1 - l_t} \right) \frac{(1 - \theta) y_t}{\tilde{l}_t} \quad (10)$$

Here, $c_t = C_t/N_t$ denotes aggregate private consumption per person, $g_t = G_t/N_t$ denotes

public consumption per person, $l_t = n_{pt}l_{mpt} + n_{pt}l_{fpt} + n_{mst}l_{mst} + n_{fst}l_{fst}$ denotes aggregate hours worked per person, $\tilde{l}_t = \tilde{L}_t/N_t$ denotes aggregate effective hours per person and $y_t = F(K_t, \tilde{L}_t)/N_t$ denotes output per person.

Combining equation (10) with the definition of the labor wedge given in equation (1), we can rewrite $1 - \Delta_t$ as in (11).

$$1 - \Delta_t = (1 - \tau_{lt}) \left[1 - \frac{n_{pt}\Gamma_{pt}(1 - l_{fpt}) + n_{fst}\Gamma_{st}(1 - l_{fst})}{1 - l_t} \right] \frac{l_t}{\tilde{l}_t} \quad (11)$$

The aggregate labor wedge, Δ_t , depends on endogenous labor supply decisions of the households, as well as time-varying exogenous inputs of the model such as taxes, gender wage gaps and fractions of females in the total population. Notice that μ —the parameter that governs the share of the gender wage gap accounted for by productivity differences—enters equation (11) indirectly through \tilde{l}_t . When $\mu = 1$, the gender gap is due entirely to productivity differences between men and women. Then, changes in the labor wedge reflect changes in distortionary taxes, as well as changes in non-distortionary factors, such as the relative productivity of women, as discussed in the static example in Section 2. When $\mu = 0$, the gender gap is due entirely to discrimination which can be interpreted as another distortion that affects the changes in the labor wedge.

We briefly discuss the model's predictions for the labor wedge under various scenarios. A more detailed analysis is provided in Section 4. First, consider the case when men and women earn the same wage and are equally productive (i.e. $\Gamma_{pt} = \Gamma_{st} = 0$). Moreover, assume that: (i) initial wealth and lifetime transfers are proportional to the lifetime labor income of each household and (ii) the individuals in the married couple have equal utility weights: $\lambda_m = \lambda_f$. Then, the model reduces to a standard growth model with taxes, as in Prescott (2004) and Ohanian, Raffo, and Rogerson (2008). That is, equation (11) simplifies to: $1 - \Delta_t = 1 - \tau_{lt}$. Since taxes, τ_{lt} , increased in U.S. data in the last 50 years, the labor wedge, Δ_t , generated under this scenario increases, contrary to what was observed in U.S. data.

Now, consider the more interesting case in which the gender wage gaps are positive (i.e. $\Gamma_{pt} > 0$, $\Gamma_{st} > 0$). For simplicity, assume that our model has only married couples and no single households, and that the gender wage gap is due entirely to discrimination (i.e. $\mu = 0$). Equation (11) simplifies to: $1 - \Delta_t = (1 - \tau_{lt}) [1 - 0.5 \cdot \Gamma_{pt} (1 - l_{fpt}) / (1 - l_t)]$. Can the model deliver a labor wedge that declines over time as seen in U.S. data? Recall that since the early 1960s, the U.S. gender wage gap shrunk and taxes increased. If the model generates an increase in aggregate hours, l_t , and a larger increase in female hours, l_{fpt} , the term $[1 - 0.5 \cdot \Gamma_{pt} (1 - l_{fpt}) / (1 - l_t)]$ increases over time. In our quantitative analysis, we show that this increase dominates the decline in $(1 - \tau_{lt})$, and the model delivers a decline in the labor wedge, Δ_t , over time (see Section 4.2 for details).

4 Quantitative Analysis

We evaluate the quantitative contribution of taxes, cross-sectional heterogeneity and other exogenous inputs to the long-run changes in U.S. hours and the labor wedge. We compute the equilibrium paths of our model and compare its predictions with U.S. data. In our *baseline* experiment, we treat the effective labor income taxes, the gender wage gaps, the government consumption and population fractions as exogenous, time-varying inputs. We perform other experiments to isolate the quantitative importance of each factor.

4.1 Baseline Calibration

We calibrate the parameters and the exogenous time-varying inputs so that our baseline model matches key statistics of the U.S. economy. We use national accounts and fixed assets data, revenue statistics and survey data for the U.S., as described in detail in Appendix A.1. Unless otherwise noted, we use data for the years 1961 to 2007.

The time-varying exogenous inputs of our model are presented in Figure 2. The effective labor income taxes are defined as in Prescott (2004) and Shimer (2009). Namely,

$\tau_{lt} \equiv 1 - (1 - \tau_{ht}) / (1 + \tau_{ct})$, where τ_{ht} and τ_{ct} are labor income and consumption tax rates constructed following the methodology of Mendoza, Razin, and Tesar (1994). The interpretation of this effective tax is that one additional unit of pre-tax labor income buys $(1 - \tau_{ht}) / (1 + \tau_{ct})$ units of consumption, after labor and consumption taxes are paid for. The government consumption to output ratio is constructed using national accounts data. The gender wage gaps for married and single individuals, Γ_{pt} and Γ_{st} , are measured using microdata from the Current Population Survey (CPS) as detailed in Appendix A.1. Lastly, the population fractions, n_{pt}, n_{fst}, n_{mst} , are also measured from the CPS.

The calibrated parameters are presented in Table 1. We choose η to match the average population growth rate and γ to match the average growth rate of labor augmenting technical change over the 47 year period. We choose θ and δ to match the average capital income share and the average annual depreciation rate, respectively. We set τ_k to the average capital income tax for the U.S. since 1970. The discount factor is chosen to match a steady state after-tax net return $(1 - \tau_k) r^* + \delta \tau_k - \delta$ of 4 percent.

We use the following utility function: $U_f = U_m = U = \frac{1}{1-\sigma} \left\{ [(c + \phi g) \cdot (1 - l)^\alpha]^{1-\sigma} - 1 \right\}$. We follow Prescott (2004) and Ohanian, Raffo, and Rogerson (2008) and set the intertemporal substitution parameter, σ , and the government consumption parameter, ϕ , to 1. In Section 4.2.5, we perform sensitivity analysis with respect to these parameters. The leisure parameter, α , and the utility weight λ_f are calibrated so that the aggregate labor supply and married female labor supply in the initial period in the model are consistent with U.S. data on hours worked in 1961. Labor supply in the model is expressed as a fraction of available time worked. Given 100 hours of available time per week, our calibration ensures that $l_{1961} \cdot 100$ equals 24.6 hours and $l_{fp1961} \cdot 100$ equals 10.3 hours, as observed in U.S. data in 1961. Once α and λ_f are calibrated, the levels of hours for the other individuals for the year 1961 are determined in equilibrium.

The initial wealth of each household, k_{p0}, k_{fs0} and k_{ms0} , is set to be proportional to labor income in 1961. Lifetime transfers are set to be proportional to the total labor income plus

initial capital stock wealth earned by each household. This choice of distributing transfers does not alter the ratios of lifetime income between the three groups of households.¹³

In our baseline calibration, we assume that the gender gap is entirely due to discrimination (i.e. $\mu = 0$). We also consider how our results change when the gender gap is accounted for entirely by productivity differences between males and females (i.e. $\mu = 1$).

4.2 Results

We show that our model is able to replicate the trends in average hours worked for men and women, by marital status, as observed in U.S. data. Moreover, we measure the labor wedge generated in the model and show that it declines over time, consistent with U.S. data. We report results from multiple experiments in order to isolate the relative importance of the different factors considered: taxes, gender wage gaps, government consumption ratio and population fractions. In Section 5, we discuss other factors that may be important for labor supply, such as child care costs, home production and leisure time.

4.2.1 Baseline Experiment: Predictions for Hours and the Labor Wedge

In our baseline model the exogenous inputs—taxes, gender wage gaps, government consumption ratio and population fractions—are set to match their counterparts in U.S. data (see Figure 2). The model delivers an increase in aggregate and women’s hours worked and a decline in the labor wedge (Figure 3). The solid lines in the left side panels of the figure show weekly hours worked by males and females in the U.S. economy between 1961 and 2007. The dashed lines show the baseline model results for hours worked (e.g. for married males, we plot $l_{mpt} \cdot 100$ where l_{mpt} is the fraction of time worked and 100 represents the available

¹³Our assumptions on the initial wealth and lifetime transfers guarantee that, if gender wage gaps are zero and $\lambda_m = \lambda_f$, our model reduces to a standard growth model with taxes. These assumptions are motivated by the fact that the equilibrium level of households’ hours worked depends on the initial wealth and lifetime transfers. To see this, note that the lifetime budget constraints for singles of gender $j \in \{f, m\}$ are: $\sum_{t=0}^{\infty} N_{jst} \frac{1}{\pi_t} c_{jst} \leq \sum_{t=0}^{\infty} \left\{ N_{jst} \frac{1}{\pi_t} \left[(1 - \tau_{lt}) w_t (1 - I_j \Gamma_{st}) l_{jst} + \psi_{jst} \right] \right\} + N_{js0} (1 - \delta + (1 - \tau_{k0}) r_0 + \delta \tau_{k0}) k_{js0}$, where π_t is defined as in equation (6) and $I_j = 1$ if $j = f$ and zero otherwise.

hours per week). The model is successful in matching the level of hours and in accounting for the changes in hours over time. Recall that aggregate hours and married females hours in 1961 are matched through the choice of α and λ_f . The levels of hours worked for married males, single males and single females in 1961 are not pinned down in the calibration, but are determined in equilibrium. While the model does not match these levels exactly, it delivers the same ranking of hours among the different population groups as in the data for the year 1961. For example, in the data, a single male worked about 26 percent more than a single female in year 1961, while the comparable figure in the model is 25 percent.

The upper right panel of Figure 3 plots aggregate weekly hours worked in the data and in the baseline model (variable l_t). The model predicts correctly very little changes in hours between 1960 and 1980, and an increase in hours afterwards. In the data, the overall increase in hours since 1960 was 13.3 percent, while the model delivers an increase of 8.4%. An obvious discrepancy between the model and the data is seen during the 1990s. In the data, aggregate hours worked increase, while the model predicts a decline during this period due to the increase in observed taxes.¹⁴ Lastly, as seen in the lower right panel of Figure 3, the model delivers a decline in the labor wedge since the early 1980s.

4.2.2 Baseline Experiment: Detailed Predictions for Hours

Table 2 presents a detailed comparison of hours worked in the data and the model. We decompose changes in aggregate hours worked per person between 1961 and 2007 as:

$$\frac{l_{2007}}{l_{1961}} = \frac{n_{p2007}l_{mp2007}}{l_{1961}} + \frac{n_{ms2007}l_{ms2007}}{l_{1961}} + \frac{n_{fs2007}l_{fs2007}}{l_{1961}} + \frac{n_{p2007}l_{fp2007}}{l_{1961}} \quad (12)$$

where each term represents the share of hours of a particular group of the population: married males, single males, single females and married females. Each term in (12) can be decomposed further into the change in the group's fraction of the total population between

¹⁴This counterfactual prediction for hours worked during the 1990s is also present in a standard growth model with a representative household. McGrattan and Prescott (2010) show that the U.S. hours boom observed in the 1990s is no longer puzzling after accounting for intangible investment.

1961 and 2007, the group's share in aggregate hours in 1961 and the change in the group's hours between 1961 and 2007. For example, for single females we have:

$$\frac{n_{fs2007}l_{fs2007}}{l_{1961}} = \left(\frac{n_{fs2007}}{n_{fs1961}}\right) \cdot \left(\frac{n_{fs1961} \cdot l_{fs1961}}{l_{1961}}\right) \cdot \left(\frac{l_{fs2007}}{l_{fs1961}}\right).$$

The baseline model matches the decomposition of aggregate hours well, as seen in Table 2. The fractions of married couples and singles in the total population are exogenous inputs into the baseline model, which means that changes in these fractions are matched exactly. Regarding the distribution of hours in U.S. data, in 1961 married men accounted for about 64 percent of hours worked, single men and women accounted for about 10 percent each, and married women for about 17 percent. In the model, the share of hours of each group in the aggregate hours is tightly linked to their predicted level of hours in the initial period. For example, singles contribute slightly more to aggregate hours in 1961 compared to the data, because the model predicts slightly higher hours for them in 1961 (see Figure 3). By the same token, the share of hours of married females in aggregate hours are matched almost exactly. Regarding changes in hours, the model predicts that hours worked by males fall by more than in the data, but hours worked by females increase similarly to what was observed.

4.2.3 Baseline Experiment: Detailed Predictions for the Labor Wedge

Table 3 presents details on the model's labor factor, $1 - \Delta_t$. Although for most of the paper we discuss changes in the labor wedge, Δ_t , Table 3 focuses on the labor factor because its changes over time can be decomposed into several multiplicative components. First, using equation (1) and $\phi = 1$, we can decompose changes in the labor factor into two components: the consumption to output ratio, $(c_t + g_t)/y_t$ and an aggregate labor component (or an aggregate labor to leisure ratio), $l_t/(1 - l_t)$. Notice that changes over time in the labor factor do not depend on the leisure parameter, α , or on the capital income share, θ . Our baseline model predicts an increase of 6.6 percent in the labor factor (which is equivalent to a decline

of about 10.5 percent in the labor wedge). All of the increase in the model’s labor factor is driven by an increase in the aggregate labor component, while the model’s consumption to output ratio declines. When measured using U.S. data, the labor factor increases by more between 1961 and 2007, partly due to an increase in the consumption to output ratio, and partly due to a larger increase in the aggregate labor component. This decomposition underscores one of the counterfactual predictions of the model: the consumption to output ratio declines in the model, while it increased in U.S. data.¹⁵ We will show this result is robust: the model is unable to deliver both an increase in aggregate hours and an increase in the consumption to output ratio, as observed in U.S. data.

A second decomposition of the labor factor from our baseline model makes use of equation (11) and is also presented in Table 3. Changes in the labor factor are now determined by changes in a tax rate component, $1 - \tau_{lt}$, and changes in a female labor component given by $1 - \frac{n_{pt}\Gamma_{pt}(1-l_{fpt})+n_{fst}\Gamma_{st}(1-l_{fst})}{1-l_t}$. Recall that the baseline model attributes all of the gender wage gaps to discrimination (i.e. $\mu = 0$), which means that the labor input equals the effective labor (i.e. $l_t/\tilde{l}_t = 1$). The first lesson from this decomposition is that the increase of 6.6 percent in the labor factor in the baseline model is driven entirely by the female labor component. The exogenous effective tax rate component, $1 - \tau_{lt}$, leads to a decline in the labor factor. The female labor component depends on inputs that are exogenous to the model, such as gender wage gaps and fractions of females in the population, but also on endogenous labor supply decisions of women and on the aggregate labor supply. In the model, the female component increases by about 15 percent over time, which is close to the increase obtained when we evaluate the expression using U.S. data. The takeaway from Table 3 is that a model with changes in gender wage gaps only, and no changes in effective taxes predicts a larger increase in the labor factor (or a larger decline in the labor wedge).

¹⁵We note that the predicted consumption to output ratio declines over time, even if we add time-varying total factor productivity (TFP) to the model. Moreover, allowing for time-varying TFP leaves the predicted decline in the labor wedge essentially unchanged.

4.2.4 Additional Experiments

We perform additional experiments to show that shrinking gender wage gaps are an important driving force for our results. Unless otherwise noted, we use the same parameters in these experiments as given in Table 1. In Figure 4, we plot the results from an experiment in which only gender wage gaps are allowed to vary over time, as measured from U.S. data. All other exogenous inputs shown in Figure 2 are held fixed at their 1961 levels. Overall, the predictions from this experiment for hours of males and females, as well as aggregate hours are closer to U.S. data. The main reason for the improved predictions is that effective income tax rates do not vary over time. As a result, the model predicts a smaller decline in male hours and a slightly larger increase in females hours compared to the baseline model. Moreover, the labor wedge declines by nearly twice as much as in the baseline experiment. The main difference is again due to taxes.

Figure 5 reports results from an experiment in which the gender wage gaps are held fixed at their 1961 levels. All other exogenous inputs—taxes, government consumption ratio and fractions of households—shown in Figure 2 are allowed to vary over time. Without shrinking gender wage gaps, the model fails to generate increases in women’s hours worked. In fact, hours worked for all groups decline marginally over time due to increases in effective labor income taxes. As a result, the model fails to capture the observed increase in aggregate hours worked. Moreover, the increase in the labor wedge is inconsistent with U.S. data. The predictions of this experiment are similar to the predictions of a standard growth model with a representative household and time-varying taxes.

Some additional experiments are summarized in Table 4 and compared with the experiments we already discussed. We present predictions for aggregate hours, l_t , married women’s hours, l_{fpt} , the labor wedge, Δ_t , and the consumption to output ratio, $(c_t + g_t) / y_t$. Neither of the experiments can account for the observed increase in the consumption to output ratio. This result affects negatively the predictions for the labor wedge as discussed earlier (see Table 3). Our baseline model accounts for about 63 percent of the increase in aggregate

hours worked per week, 83 percent of the increase in married women’s hours, while it accounts for only 30 percent of the decline in the labor wedge. Among the experiments with only one time-varying input, the experiment with changes in gender wage gaps performs the best. It accounts for 95 percent of the increase in aggregate hours and about 54 percent of the decline in the labor wedge. The experiment with changes in effective taxes alone has counterfactual predictions for labor supply and for the labor wedge. The experiment in which we allow only the fractions of married couples and singles to vary over time delivers an increase in labor supply, but for the wrong reasons. In this experiment, the hours of all individuals increase slightly over time. The increase in the model’s aggregate hours is then driven mainly by singles, since the fractions of singles increases significantly between 1961 and 2007, as observed in U.S. data.

4.2.5 Sensitivity Analysis

We perform sensitivity analysis with respect to μ , σ and ϕ .

In all experiments discussed so far, we assumed that males and females are equally productive, and the gender wage gaps are due entirely to discrimination, i.e. $\mu = 0$. We perform an experiment in which all exogenous inputs vary over time, but we assume the gender wage gaps are due entirely to productivity differences between females and males, i.e. $\mu = 1$. We recalibrate parameters α and λ_f to match the same targets on hours worked as in the baseline calibration, but keep all other parameters unchanged. We find that a model with $\mu = 1$ implies fairly similar changes in hours worked for males and females (see Table 4). Aggregate hours go up by 7.4 percent compared to 8.4 percent in the baseline model. The decline in the labor wedge is slightly smaller in this experiment compared to the baseline. Recall that using equation (11) the labor factor, $1 - \Delta_t$, can be decomposed into three factors as shown at the bottom of Table 3. When $\mu = 1$, the ratio l_t/\tilde{l}_t is less than one, leading to slightly smaller increases (decreases) in the labor factor (labor wedge).¹⁶

¹⁶Our results are approximately linear in the value of μ . The results from our baseline experiment (which has $\mu = 0$) and from the experiment with $\mu = 1$ provide upper and lower bounds on how successful the

In our baseline calibration, the elasticity of substitution, σ , and the marginal rate of substitution between public and private consumption, ϕ , are equated to 1. We report quantitative results of our baseline model for different values of σ or ϕ in Table 5. In each experiment, we recalibrate parameters α and λ_f to match the same targets on hours worked as in the baseline calibration.

The choice of σ impacts the model’s Frisch elasticities of labor supplies for men and women.¹⁷ With $\sigma = 1$, the aggregate Frisch elasticity is about 3. For $\sigma \in (0, 1)$, the Frisch elasticities are larger and the model appears more successful at predicting the change in aggregate hours, but for the wrong reasons. The predicted hours for males decline due to higher taxes by much more than observed in U.S. data, while the predicted hours for married females have a counterfactually strong increase in response to the shrinking gender wage gap. For $\sigma > 1$, hours worked respond less to changes in taxes and gender gaps. The predictions for the labor wedge do not change as much as hours do in response to changes in σ , since the labor wedge also reflects declines in the consumption to output ratio (see Table 5).

The value of ϕ has little impact on the model’s predictions for hours or the labor wedge. However, ϕ affects the measurement of the labor wedge from the data. A lower value of ϕ corresponds to a larger increase in $\frac{c_t + \phi g_t}{y_t}$ from 1961 to 2007 and a larger decline in the wedge measured from U.S. data. Varying ϕ from 0 to 1, our model accounts for 25 to 30 percent of the observed changes in the U.S. labor wedge (see Table 5).

We conclude that the elasticity of substitution $\sigma = 1$ seems appropriate for our model, as it implies an aggregate labor supply elasticity consistent with other macro studies, such as Prescott (2004) and Ohanian, Raffo, and Rogerson (2008). Moreover, the particular choices of μ —the fraction of the gender wage gaps accounted for by productivity differences—or ϕ —

model is. For example, an experiment in which all exogenous inputs are allowed to vary over time, but the value of μ equals 0.5, predicts that aggregate hours increase by 7.9 percent and the labor wedge declines by 9.4 percent. These values are about midway between the results from the baseline model and the experiment with $\mu = 1$ (see Table 4).

¹⁷For the utility function we consider, $\frac{1}{1-\sigma} \left\{ [(c + \phi g) \cdot (1-l)^\alpha]^{1-\sigma} - 1 \right\}$ with $\sigma > 0$, the Frisch elasticity equals $\frac{1-l}{l} \cdot \frac{1}{1-\alpha \frac{1-\sigma}{\sigma}}$.

the marginal rate of substitution between public and private consumption—do not overturn our conclusion that the gender wage gap is an important driving force behind the long-run changes in U.S. hours and the labor wedge from 1961 to 2007.

5 Other Considerations

We have shown that a model with shrinking gender wage gaps for married couples and singles is successful in delivering an increase in U.S. hours worked and a decline in the U.S. labor wedge, despite the observed increase in U.S. effective labor income taxes. Our baseline model accounts for 63 percent of the increase in aggregate hours, 86 percent of the increase in married women’s hours and 30 percent of the decline in the labor wedge. In this section, we briefly discuss other factors—such as changes in child care costs, leisure time or home production—which may be able to further improve the quantitative predictions of our model for U.S. hours worked and the labor wedge.

5.1 Changes in Child Care Costs

Incorporating reductions in child care costs over time in our model has the potential to deliver a larger increase in the hours of married women, and hence a larger decline in the labor wedge (see equation 11). Attanasio, Low, and Sánchez-Marcos (2008) show that reductions in child care costs along with reductions in the gender wage gap over time help explain the increase in participation rates of females in the U.S. The idea is that, in the past, child care costs were very high and mothers stayed at home after birth to care for their children. Thus, the rise in married women’s labor supply is really a story about their wages increasing, as well as the number of children and the child care costs decreasing.

We extend our baseline model to allow for reductions in child care costs and evaluate the quantitative implications for hours and the labor wedge. We model child care services as a cost paid by the married couple. The married female decides how many hours, l_{fpt} , to

work given the wage rate, $(1 - \Gamma_{pt}) w_t$, she receives and given the hourly cost of child care services, χ_t . The new sequential budget constraints of the married couple are given below.

$$c_{fpt} + c_{mpt} + x_{pt} \leq [(1 - \tau_{kt}) r_t + \delta \tau_{kt}] k_{pt} + (1 - \tau_{lt}) w_t l_{mpt} + [(1 - \tau_{lt}) w_t (1 - \Gamma_{pt}) - \chi_t] l_{fpt} + \psi_{pt}$$

The new resource constraint is: $C_t + X_t + \Xi_t + G_t = F(K_t, \tilde{L}_t)$, where Ξ_t are the total resources spent on child care: $\Xi_t = n_{pt} l_{fpt} \chi_t$.

The intratemporal condition for married females changes to (13), while the intratemporal conditions for married males, single males and single women remain unchanged.

$$\frac{\alpha (c_{fpt} + \phi g_t)}{1 - l_{fpt}} = (1 - \tau_{lt}) (1 - \Gamma_{pt}) w_t - \chi_t \quad (13)$$

We perform an experiment that features all time-varying inputs from our baseline model, plus changes in child care costs. We recalibrate parameters α and λ_f , so that the model with child care costs matches the same targets for hours worked as discussed in the baseline calibration. All other parameters stay unchanged. To determine χ_t , we use data from the U.S. Census Bureau's Survey of Income and Program Participation. According to this survey, the average child care expenditures of families with employed mothers that pay for such services were 15% of the mother's income in the year 2004.¹⁸ Since χ_t are hourly child care costs, we pick χ_{2004} so that $\chi_{2004} / (w_{2004} (1 - \Gamma_{p,2004})) = 0.15$. The change in child care costs from early 1960s to present is harder to measure. Attanasio, Low, and Sánchez-Marcos (2008) consider reductions in child care costs that range between -5% to -20%. In our experiment, we used the midpoint of this range. We pick χ_{1961} such that child care costs as a fraction of a married female's income decline linearly by 12.5% between 1961 and 2004.

The results of this experiment are presented in Table 4. As expected, reductions in child care costs lead to an increase in hours worked by married females. In our baseline model, hours of married females increase by a factor of 2.08 between 1961 and 2007. With reductions

¹⁸Data are available at: <http://www.census.gov>, "Who's Minding the Kids? Child Care Arrangements: Summer 2006", Table 6. The child care expenditures are for families with children 15 years old and younger.

in child care costs, married female hours increase by a factor of 2.2. The hours of all other individuals are comparable across the two experiments. As a result, aggregate hours increase by a bit more and the labor wedge declines by a bit more over the 47 year period.

To summarize, reductions in child care costs lead to additional increases in aggregate and women’s hours, but contribute only a further 6 percentage points to the decline in the labor wedge observed in the U.S. since early 1960s.

5.2 Changes in Time Devoted to Home Production and Leisure

Ohanian, Raffo, and Rogerson (2008) suggest that the counterfactual predictions of the neoclassical growth model for U.S. hours can potentially be reconciled by taking into account changes in the amount of time devoted to home production. In their representative agent model (as well as in the model presented in this paper), hours worked in the market and leisure time are mirror images of each other. In U.S. data, leisure time depends not only on time spent at work, but also on time spent in non-market activities such as home production. Therefore, a model that accounts for the decline in home production observed in the U.S. in the past 50 years has a better chance of matching the increase in U.S. hours worked.

We illustrate that changes in home production and leisure time help improve the predictions of our model for U.S. hours and the labor wedge. A back of the envelope calculation suggests that this mechanism can only account for part of the changes in the labor wedge and is thus complimentary to the mechanism we presented in this paper. To see this, consider the case in which our baseline model reduces to a standard growth model (i.e. all individuals are the same). The labor equilibrium condition is given below:

$$\frac{\alpha (c_t + \phi g_t)}{\text{leisure}_t} = (1 - \tau_{lt}) (1 - \theta) \frac{y_t}{l_t} \tag{14}$$

where leisure is now allowed to include time spent in non-market activities, i.e. $\text{leisure}_t \equiv 1 - l_t - (\text{non-market hours})_t$. What is the change needed in leisure time in order for equation

(14) to be consistent with the data? U.S. data on taxes, private and public consumption, output and hours worked shows that $(1 - \tau_{lt})$ declines by 7.5%, $(c_t + g_t) / y_t$ increases by 6.8% and hours worked increase by 13.3% between 1960 and 2007. Then, leisure time would need to increase by 31%(= $1.068 * 113.3/92.5 - 1$) between 1960 and 2007 in order for equation (14) to hold. That is, a 31 percent increase in leisure time would explain all of the changes in the labor wedge measured from the neoclassical model (with no non-market hours).

The evidence regarding changes in U.S. leisure time is a bit mixed. Ramey and Francis (2009) document that leisure time measured as the difference between time available and time devoted to non-leisure activities (i.e. work, school, home production, commuting and personal care time) changed little for both males and females since 1960 (see Figure 5 in their paper). Males aged 25 to 54 reduced average weekly hours worked and increased the time spent in home production between 1960 and 2005 (see Tables 2 and 4 in their paper). Over the same time period, females in the same age group increased their weekly hours worked, while reducing the time spent in home production. In contrast, Aguiar and Hurst (2007) document that, over a similar time period, leisure time has increased anywhere from four to eight hours per week for males and females of working age (see Table III in their paper). As pointed out by Ramey and Francis, some of the differences in estimates are due to different definitions of leisure. Both studies document that average time spent in home production in the U.S. declined since the 1960s. The larger estimates provided by Aguiar and Hurst (2007) show an increase in leisure time of about 5.4 to 15 percent between 1965 and 2003, while the estimates provided by Ramey and Francis (2009) show an increase of barely 2 percent.

To summarize, a model where leisure time accounts for non-market hours can help improve the predictions of our model for the U.S. labor wedge.¹⁹ We view this mechanism as complementary to the one presented in our paper, since increases in leisure time alone can account for 6.5 to 50 percent of the decline in the U.S. labor wedge.

¹⁹The link between home production and the labor wedge is studied by Karabarbounis (2014a) in an international business cycle model. His estimated model is able to generate a labor wedge volatility and persistence consistent with cross country data.

6 International Evidence

A natural question is whether the mechanism we analyzed in detail for the U.S. is also important in other economies. In this section, we show that shrinking gender wage gaps and labor income taxes are important drivers of the long-run changes in the Canadian and German labor wedges. Our choice of countries is limited by the availability of *long-run* micro survey data which allows us to construct hourly wage rates and hours worked by gender and marital status. However, we show that our mechanism is broadly consistent with aggregate data for other OECD economies.

Our quantitative exercise for Canada and Germany is similar to that performed for the U.S. in Table 3 (see column labeled "Data"). First, we use equation (1) and aggregate data on private and public consumption, output and average hours worked to measure changes in the labor factor, $1 - \Delta_t$. Second, we use tax rates and micro survey data to measure the tax component and the female labor component given in equation (11). The results and sources of data for Canada and Germany are presented in Table 6.

Results for Canada are very similar to those for the U.S. economy. From 1971 to 2012, the labor factor in Canada, $1 - \Delta_t$, increased or, equivalently, the labor wedge, Δ_t , declined. Over the same time period, effective labor income taxes increased, aggregate and women's hours worked increased, men's hours worked were relatively stable and the gender wage gaps for married couples and singles shrunk. A neoclassical growth model with only taxes has counterfactual predictions for Canada. The increase in effective labor income taxes implies a decline in aggregate hours and an increase in the labor wedge. Hence, shrinking gender wage gaps are needed to capture the increase in female and aggregate hours worked. Moreover, the decomposition in Table 6 shows that the female labor component alone, which captures changes in gender wage gaps and female hours worked, can account for about two-thirds of the increase in the labor factor. When considered together, the tax component and the female labor component can account for about a quarter of the increase in the labor factor.

The German example is especially interesting because from 1983 to 2007, the labor factor,

$1 - \Delta_t$, declined or, equivalently, the labor wedge, Δ_t , increased (Table 6). Aggregate hours worked and married women's hours worked increased, while single women's and men's hours worked were relatively stable. The gender wage gap for married couples shrunk, while the gender wage gap for singles was relatively stable. Finally, effective labor income taxes increased. Can the German data be reconciled with our mechanism? Yes. Effective labor income taxes alone predict an increase in the labor wedge, Δ_t , that is larger than observed. In addition, higher taxes imply a counterfactual decline in hours worked. Hence, the closing of the gender wage gap for married females is important in understanding why aggregate and married females' hours have gone up. Moreover, cross-sectional heterogeneity in hours and wages—captured in the female labor component—brings the model labor wedge closer to that measured from German aggregate data.

We extend our analysis to other OECD economies where gender wage gaps shrunk, as documented by Blau and Kahn (2000). While we do not have long-run micro survey data for these economies, Table 7 shows that our mechanism is broadly consistent with aggregate data on hours worked, tax rates and measured labor wedges for a number of other OECD economies. In economies with large changes in aggregate and women's hours and the labor wedge, cross-sectional heterogeneity can be quantitatively important in reversing the effects of higher taxes (as observed in Spain, Italy and Belgium), or in accounting for reductions in labor wedges which are larger than reductions in tax rates (as observed in Netherlands, Finland and the U.K.).

To summarize, for a number of countries—U.S., Canada, Germany, Netherlands, Spain, Finland, Italy, U.K. and Belgium—reductions in cross-sectional heterogeneity in wages and hours can contribute to increases in aggregate hours and reductions in the labor wedge, Δ_t (or equivalently, increases in the labor factor, $1 - \Delta_t$, as shown in Tables 3, 6 and 7).

7 Conclusion

From the early 1960s to 2007, the U.S. labor wedge—measured as the discrepancy between a representative household’s marginal rate of substitution between consumption and leisure and the marginal product of labor—declined substantially. Over the same time period, U.S. aggregate hours worked increased, due to an increase in women’s hours worked. These observations are puzzling from the perspective of a standard neoclassical growth model because they were accompanied by an increase in U.S. effective labor income taxes.

In this paper, we show that incorporating household heterogeneity in productivity and hours worked in an otherwise standard growth model is important in accounting for the long-run trends in U.S. hours and the labor wedge. We show that large cross-sectional differences in productivity and hours worked between households generate a large labor wedge. Consequently, reductions in cross-sectional differences over time contribute to reductions in the measured labor wedge. We focus on a particular split of the population, by gender and marital status, motivated by the increase in women’s hours worked and the decline in the gender wage gaps in the U.S. We show that reductions in gender wage gaps consistent with U.S. data generate an increase in aggregate and women’s hours and a decline in the labor wedge, in spite of higher taxes.

We provide international support for the mechanism we analyzed in detail for the U.S. economy. We show that reductions in cross-sectional heterogeneity in wages and hours worked contribute to reductions in the measured labor wedges in Canada and Germany. In Canada, similar to the U.S., the closing of the gender wage gaps and increases in female hours dominate the increase in taxes and lead to a decline in the labor wedge over the last four decades. Germany is especially interesting, since taxes increased by more than the increase in the labor wedge over the last two decades. Hence, reductions in cross-sectional heterogeneity in Germany (captured by the closing of the gender wage gap for married couples and the increase in married women’s hours) are important as they partly reverse the increase in taxes, bringing the model’s labor wedge closer to that measured from German aggregate

data. The improved predictions for the labor wedges lead us to conclude that household heterogeneity also helps account for the changes in hours worked in Canada and Germany. Moreover, we illustrate that reductions in cross-sectional heterogeneity in wages and hours can be important in accounting for the increase in aggregate hours and the reductions in the measured labor wedges in a broader set of countries.

A Appendix

A.1 U.S. Data

Survey data. We use data from the IPUMS-CPS to construct our measures of average hours worked and the gender wage gaps. The IPUMS-CPS is based on the March Current Population Survey and is available yearly since 1962 at <http://cps.ipums.org/cps/>.

We use the following variables for survey years 1962 – 2008: PERWT (person weight), AGE (person’s age at last birthday), SEX, MARST (current marital status), EMPSTAT (current employment status), HRSWORK (hours worked last week), INCWAGE (wage and salary income last year). We also use WKSWORK1 (weeks worked last year), for 1976 – 2008 and WKSWORK2 (weeks worked last year in interval format), for 1962 – 2008.

We construct total hours worked last year as the product of weeks worked and hours worked per week. Starting 1976, we use the variable WKSWORK1 to obtain weeks worked for each person. Prior to 1976, the survey provides only an interval for the weeks worked for each person (variable WKSWORK2). We replace WKSWORK2 with an average number of weeks worked (given in equation 15) that is calculated based on WKSWORK1 as follows. We take variable WKSWORK1 and group persons according to their number of weeks worked into the same intervals provided in variable WKSWORK2. We then compute the average weeks worked for each of the six intervals from 1976 to 2008. For each interval, the averages obtained vary very little over time. For example, the average number of weeks worked for persons working between 1 and 13 weeks was roughly 8 for all years from 1976 to 2008.

$$\text{weeks worked, 1962} - 1975 = \begin{cases} 8.0 & \text{if WKSWORK2 is 1} - 13 \text{ weeks} \\ 21.7 & \text{if WKSWORK2 is 14} - 26 \text{ weeks} \\ 33.7 & \text{if WKSWORK2 is 27} - 39 \text{ weeks} \\ 42.6 & \text{if WKSWORK2 is 40} - 47 \text{ weeks} \\ 48.3 & \text{if WKSWORK2 is 48} - 49 \text{ weeks} \\ 51.9 & \text{if WKSWORK2 is 50} - 52 \text{ weeks} \end{cases} \quad (15)$$

We use variable INCWAGE to obtain the wage per hours for each person.

We construct population, employment, average hours worked and median wage per hour for the total population, married men, married women, single men, single women. Our measure of married couples includes the following categories from the variable MARST: "married, spouse present", "married, spouse absent" and "separated". We group the categories "divorced", "widowed" and "never married" under our measure of singles. We use population ages 20 to 64. We use the median wage per hour because it is not affected by changes in the top code. To construct employment we take all persons who were employed and at work during the reference week, all persons who were employed but not at work that week, and all persons in the Armed Forces (EMPSTAT = 10, 12 and 13, respectively). We construct hours worked by employed persons using all respondents that report EMPSTAT equal to 10 or 13. The average hours worked per week are then given by: $h_E \cdot \frac{E}{N} \cdot \frac{1}{52}$, where h_E are hours worked during the year by employed people, E is the total number of employed persons, and N is the total population. Our implicit assumption is that persons not at work during the reference week (i.e. people with EMPSTAT equal to 12) work similar yearly hours to those at work during the reference week.

The average hours worked we obtain for the total population are very similar to those reported by Cociuba, Ueberfeldt, and Prescott (2009). Our average hours worked for married and single individuals differ slightly from those reported by McGrattan and Rogerson (2008), who use population 25 – 64.

National accounts and fixed assets data. We obtain these data from the *Bureau of Economic Analysis*. We make a few adjustments to the national accounts. We treat con-

sumer durables as investment. We treat government military investment as government consumption and the remainder of government investment is treated as investment. We also remove sales taxes from the gross domestic product.

Tax rates. We use data from the Organization for Economic Cooperation and Development to construct tax rates following the methodology of Mendoza, Razin, and Tesar (1994). We use Joines (1981) to extend the series of tax rates before 1970.

A.2 Model Aggregate Intra-temporal Condition

Here, we derive the model's aggregate intra-temporal condition. The model has four intra-temporal equations for each type of consumer in the economy. We multiply each intra-temporal condition by the fraction of consumers of that type and sum up. We obtain:

$$\begin{aligned}
& \alpha (n_{pt}c_{mpt} + n_{pt}c_{fpt} + n_{mst}c_{mst} + n_{fst}c_{fst}) + \alpha\phi g_t (2n_{pt} + n_{mst} + n_{fst}) \quad (16) \\
& = (1 - \tau_{lt}) w_t n_{pt} (1 - l_{mpt}) + (1 - \tau_{lt}) w_t n_{pt} (1 - \Gamma_{pt}) (1 - l_{fpt}) \\
& \quad + (1 - \tau_{lt}) w_t n_{mst} (1 - l_{mst}) + (1 - \tau_{lt}) w_t n_{fst} (1 - \Gamma_{st}) (1 - l_{fst})
\end{aligned}$$

Equation (16) can also be written as $\alpha (c_t + \phi g_t) = (1 - \tau_{lt}) w_t \Lambda_t$, where $c_t \equiv \frac{C_t}{N_t}$ denotes aggregate private consumption per person, and where Λ_t is defined as below.

$$\Lambda_t \equiv 2n_{pt} + n_{mst} + n_{fst} - n_{pt}l_{mpt} + n_{pt}(-l_{fpt} - \Gamma_{pt}(1 - l_{fpt})) - n_{mst}l_{mst} + n_{fst}(-l_{fst} - \Gamma_{st}(1 - l_{fst}))$$

$$\Lambda_t = 1 - (n_{pt}l_{mpt} + n_{pt}l_{fpt} + n_{mst}l_{mst} + n_{fst}l_{fst}) - n_{pt}\Gamma_{pt}(1 - l_{fpt}) - n_{fst}\Gamma_{st}(1 - l_{fst})$$

In the last expression we used $2n_{pt} + n_{mst} + n_{fst} = (2N_{pt} + N_{mst} + N_{fst})/N_t = 1$.

Let l_t denote aggregate hours worked per person: $l_t \equiv n_{pt}l_{mpt} + n_{pt}l_{fpt} + n_{mst}l_{mst} + n_{fst}l_{fst}$.

The aggregate intra-temporal equation becomes:

$$\alpha (c_t + \phi g_t) = (1 - \tau_{lt}) w_t [1 - l_t - n_{pt}\Gamma_{pt}(1 - l_{fpt}) - n_{fst}\Gamma_{st}(1 - l_{fst})].$$

We divide both sides by $(1 - l_t)$ and use $w_t = (1 - \theta) y_t / \tilde{l}_t$ to get:

$$\frac{\alpha(c_t + \phi g_t)}{1 - l_t} = (1 - \tau_{lt}) \left(1 - \frac{n_{pt}\Gamma_{pt}(1 - l_{fpt}) + n_{fst}\Gamma_{st}(1 - l_{fst})}{1 - l_t} \right) \frac{(1 - \theta)y_t}{\tilde{l}_t}.$$

References

- AGUIAR, M., AND E. HURST (2007): “Measuring Trends in Leisure: The Allocation of Time Over Five Decades,” *Quarterly Journal of Economics*, 122(3), 969–1006.
- ARELLANO, C., Y. BAI, AND P. J. KEHOE (2012): “Financial Frictions and Fluctuations in Volatility,” *Federal Reserve Bank of Minneapolis Staff Report 466*, pp. 1–56.
- ATTANASIO, O., H. LOW, AND V. SÁNCHEZ-MARCOS (2008): “Explaining Changes in Female Labor Supply in a Life-Cycle Model,” *American Economic Review*, 98(4), 1517–1552.
- BAR, M., AND O. LEUKHINA (2009): “To Work or Not to Work: Did Tax Reforms Affect Labor Force Participation of Married Couples?,” *The B.E. Journal of Macroeconomics*, 9(1), Contributions, Article 28.
- (2011): “On the Time Allocation of Married Couples Since 1960,” *Journal of Macroeconomics*, 33, 491–510.
- BLAU, F. D., AND L. M. KAHN (2000): “Gender Differences in Pay,” *The Journal of Economic Perspectives*, 14(4), 75–99.
- BRINCA, P. (2014): “Distortions in the Neoclassical Growth Model: A cross-country Analysis,” *Journal of Economic Dynamics and Control*, 47, 1–19.
- CHANG, Y., AND S.-B. KIM (2014): “Heterogeneity and Aggregation: Implications for Labor-Market Fluctuations: Reply,” *American Economic Review*, 104(4), 1461–1466.
- CHARI, V., P. J. KEHOE, AND E. R. MCGRATTAN (2007): “Business Cycle Accounting,” *Econometrica*, 75(3), 781–836.
- COCIUBA, S. E., A. UEBERFELDT, AND E. C. PRESCOTT (2009): “U.S. Hours and Productivity Behavior Using CPS Hours Worked Data: 1947-III to 2009-II,” *Mimeo*.

- EROSA, A., L. FUSTER, AND D. RESTUCCIA (2002): “Fertility Decisions and Gender Differences in Labor Turnover, Employment, and Wages,” *Review of Economic Dynamics*, 5(4), 856–891.
- (2005): “A Quantitative Theory of the Gender Gap in Wages,” *University of Toronto, Working Paper*.
- FRICK, J. R., S. P. JENKINS, D. R. LILLARD, O. LIPPS, AND M. WOODEN (2007): “The Cross-National Equivalent File (CNEF) and its Member Country Household Panel Studies,” *Schmollers Jahrbuch (Journal of Applied Social Science Studies)*, 127(4), 627–654.
- GOLDIN, C. (1992): *Understanding the Gender Gap: An Economic History of American Women*, NBER Series on Long-Term Factors in Economic Development. Oxford University Press.
- HALL, R. E. (1997): “Macroeconomic Fluctuations and the Allocation of Time,” *Journal of Labor Economics*, 15(1), S223–S250.
- JERMANN, U., AND V. QUADRINI (2012): “Macroeconomic Effects of Financial Shocks,” *American Economic Review*, 102(1), 238–271.
- JOINES, D. H. (1981): “Estimates of Effective Marginal Tax Rates on Factor Incomes,” *The Journal of Business*, 54(2), 191–226.
- JONES, L. E., R. E. MANUELLI, AND E. R. MCGRATTAN (2003): “Why Are Married Women Working So Much?,” *Federal Reserve Bank of Minneapolis Staff Report 317*.
- KARABARBOUNIS, L. (2014a): “Home production, labor wedges, and international business cycles,” *Journal of Monetary Economics*, 64, 68–84.
- (2014b): “The labor wedge: MRS vs. MPN,” *Review of Economic Dynamics*, 17, 206–223.

- KRYVTSOV, O., AND A. UEBERFELDT (2007): “Schooling, Inequality and Government Policy,” *Bank of Canada Working Paper 2007-12*.
- MALIAR, L., AND S. MALIAR (2003): “The Representative Consumer in the Neoclassical Growth Model with Idiosyncratic Shocks,” *Review of Economic Dynamics*, 6, 362–380.
- MCDANIEL, C. (2007): “Average tax rates on consumption, investment, labor and capital in the OECD 1950-2003,” *Working Paper, Arizona State University*.
- MCGRATTAN, E. R., AND E. C. PRESCOTT (2010): “Unmeasured Investment and the Puzzling U.S. Boom in the 1990s,” *American Economic Journal: Macroeconomics*, 2(4), 88–123.
- MCGRATTAN, E. R., AND R. ROGERSON (2008): “Changes in the Distribution of Family Hours Worked Since 1950,” in *Frontiers of Family Economics*, vol. 1, pp. 115–138. Emerald Group Publishing Limited.
- MENDOZA, E. G., A. RAZIN, AND L. TESAR (1994): “Effective Tax Rates in Macroeconomics: Cross Country Estimates of Tax Rates on Factor Incomes and Consumption,” *Journal of Monetary Economics*, 34, 297–323.
- MULLIGAN, C. B. (2002): “A Century of Labor-Leisure Distortions,” *NBER Working Paper Series*, No. 8774.
- OHANIAN, L., A. RAFFO, AND R. ROGERSON (2008): “Long-term changes in labor supply and taxes: Evidence from OECD countries, 1956-2004,” *Journal of Monetary Economics*, 55(8), 1353–1362.
- PARKIN, M. (1988): “A Method for Determining Whether Parameters in Aggregative Models are Structural,” *Carnegie-Rochester Conference Series on Public Policy*, 29, 215–252.
- PRESCOTT, E. C. (2004): “Why Do Americans Work So Much More Than Europeans?,” *Federal Reserve Bank of Minneapolis Quarterly Review*, 28(1), 2–13.

- RAGAN, K. S. (2006): “Taxes, Transfers, and Time Use: Fiscal Policy in a Household Production Model,” *PhD. dissertation, University of Chicago*.
- RAMEY, V., AND N. FRANCIS (2009): “A Century of Work and Leisure,” *American Economic Journal: Macroeconomics*, 1(2), 189–224.
- ROGERSON, R. (2007): “Taxation and market work: is Scandinavia an outlier?,” *Economic Theory*, 32(1), 59–85.
- SHIMER, R. (2009): “Convergence in Macroeconomics: The Labor Wedge,” *American Economic Journal: Macroeconomics*, 1(1), 280–297.
- (2010): *Labor Markets and Business Cycles*. Princeton University Press.

Figure 1: HOURS WORKED AND THE LABOR WEDGE MEASURED FROM U.S. DATA

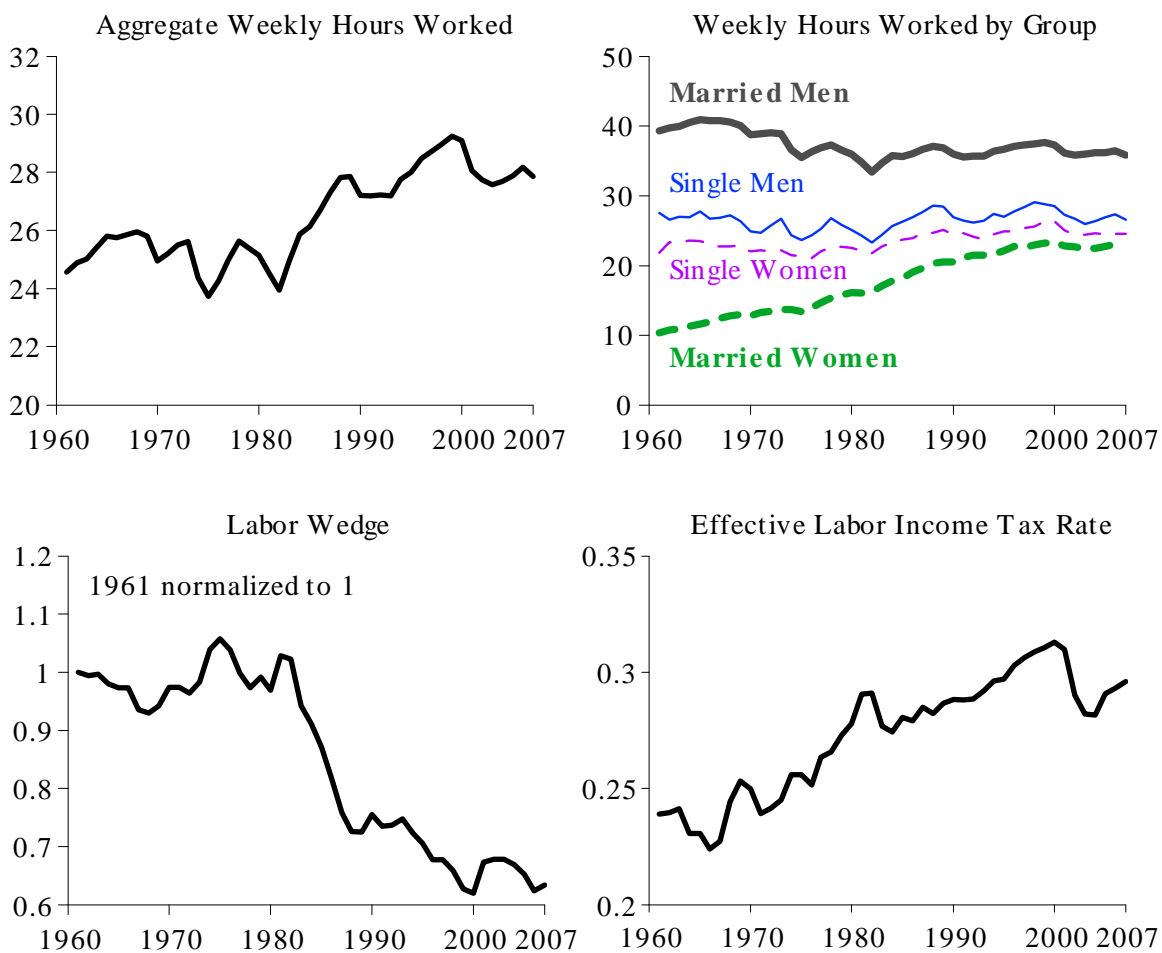


Figure 2: TIME-VARYING MODEL INPUTS MEASURED FROM U.S. DATA

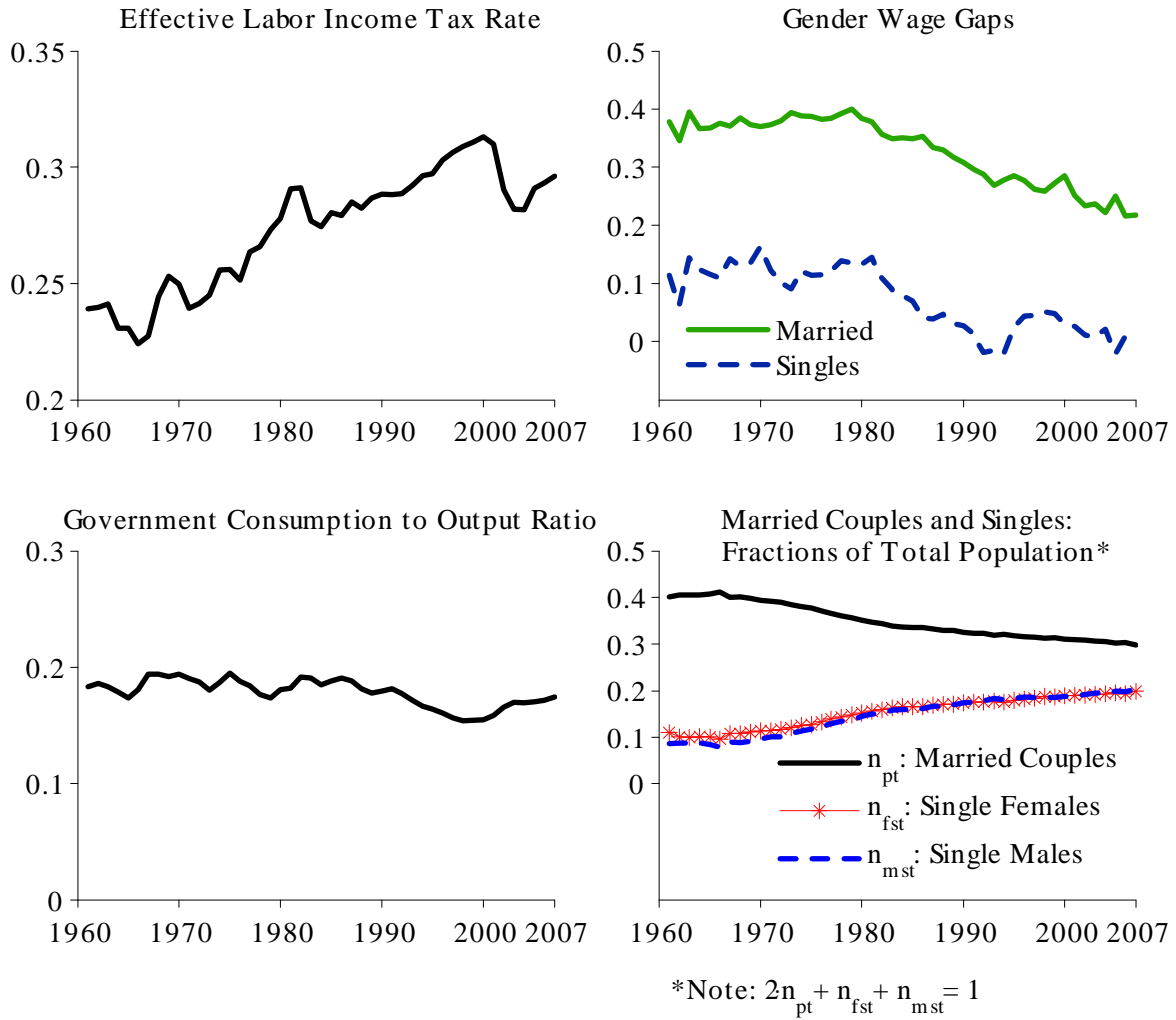


Figure 3: HOURS AND THE LABOR WEDGE IN DATA AND BASELINE MODEL

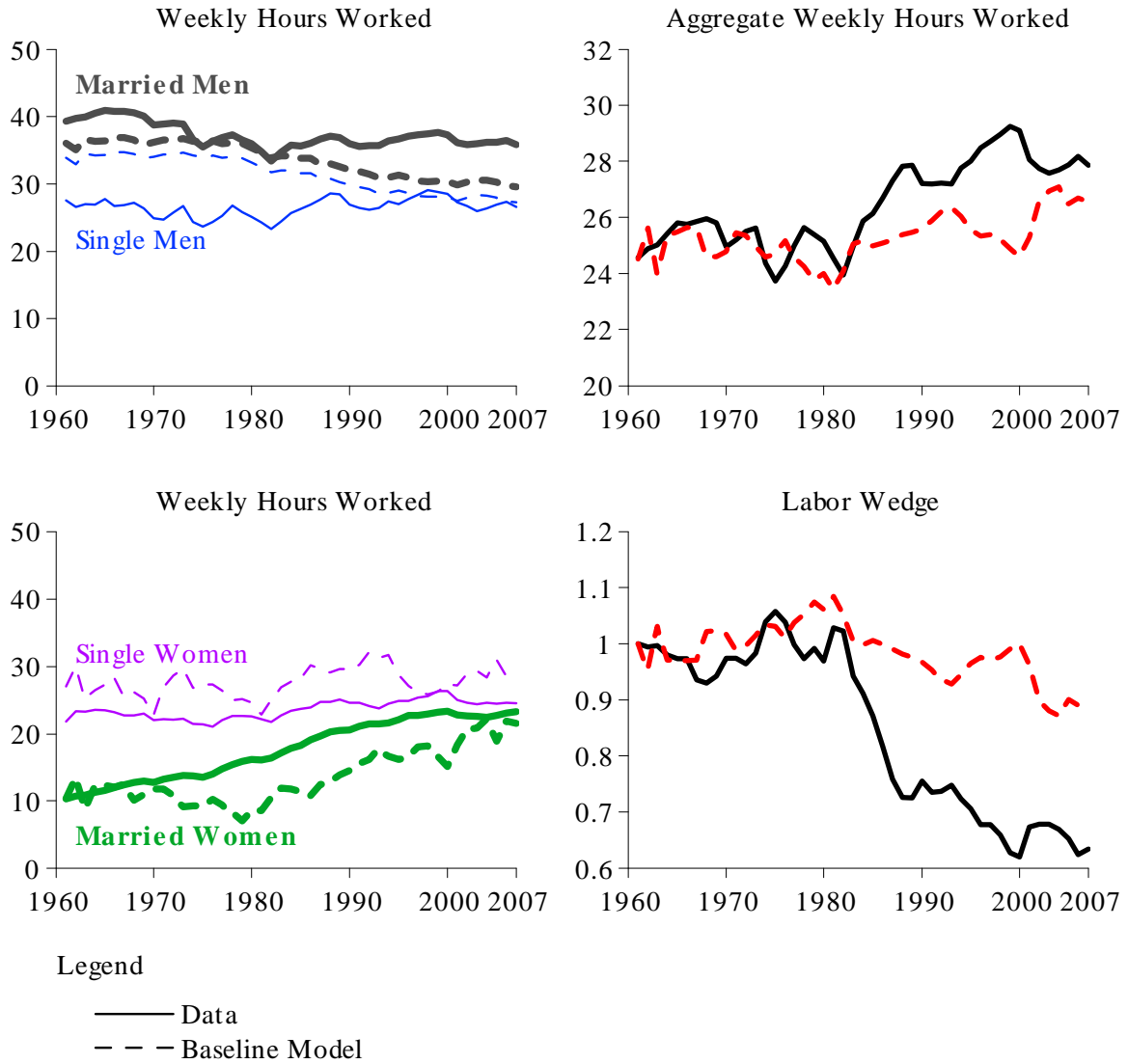


Figure 4: MODEL WITH SHRINKING GENDER WAGE GAPS ONLY

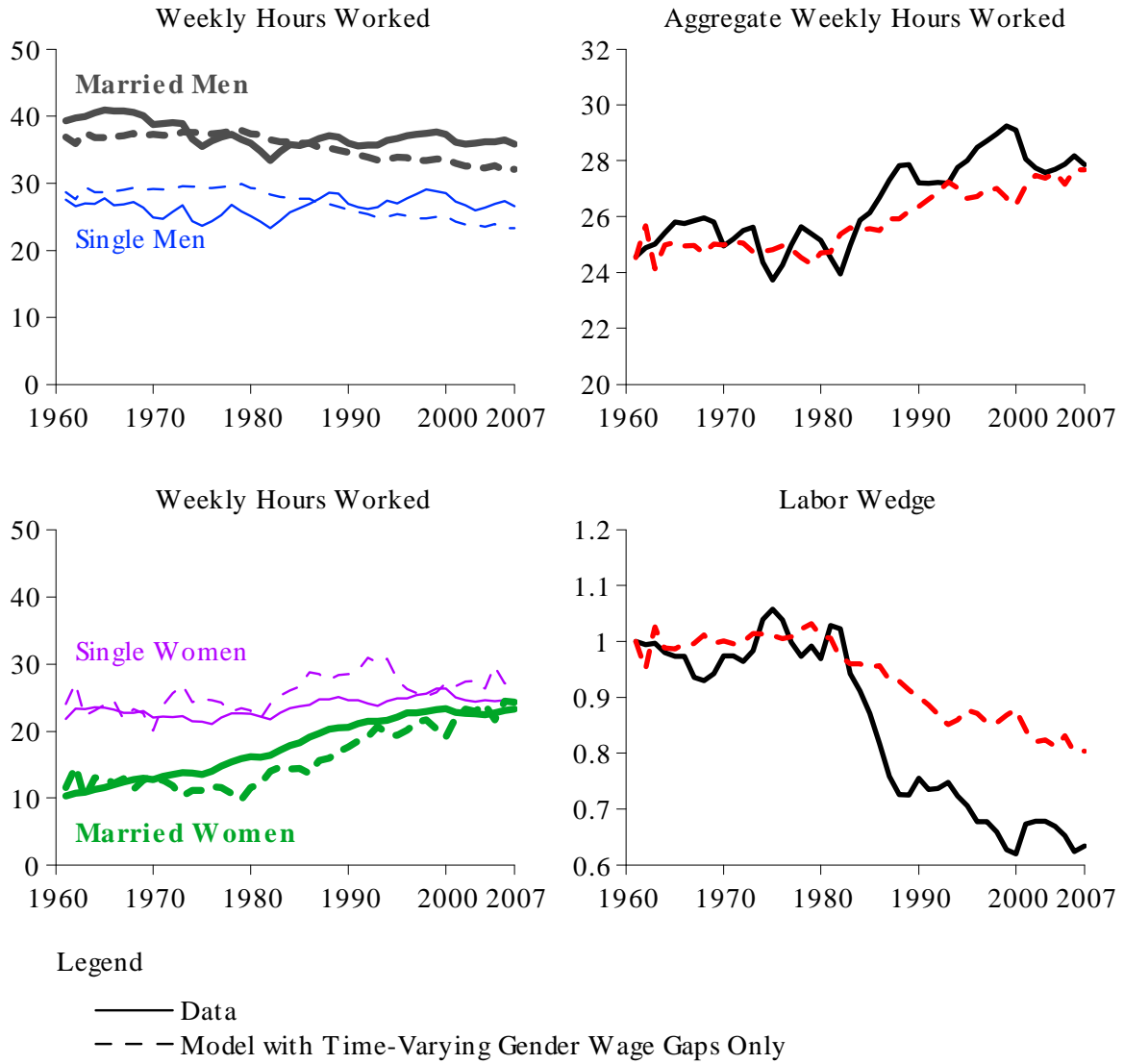


Figure 5: MODEL WITH TIME-VARYING TAXES, GOVERNMENT CONSUMPTION AND FRACTIONS OF HOUSEHOLDS

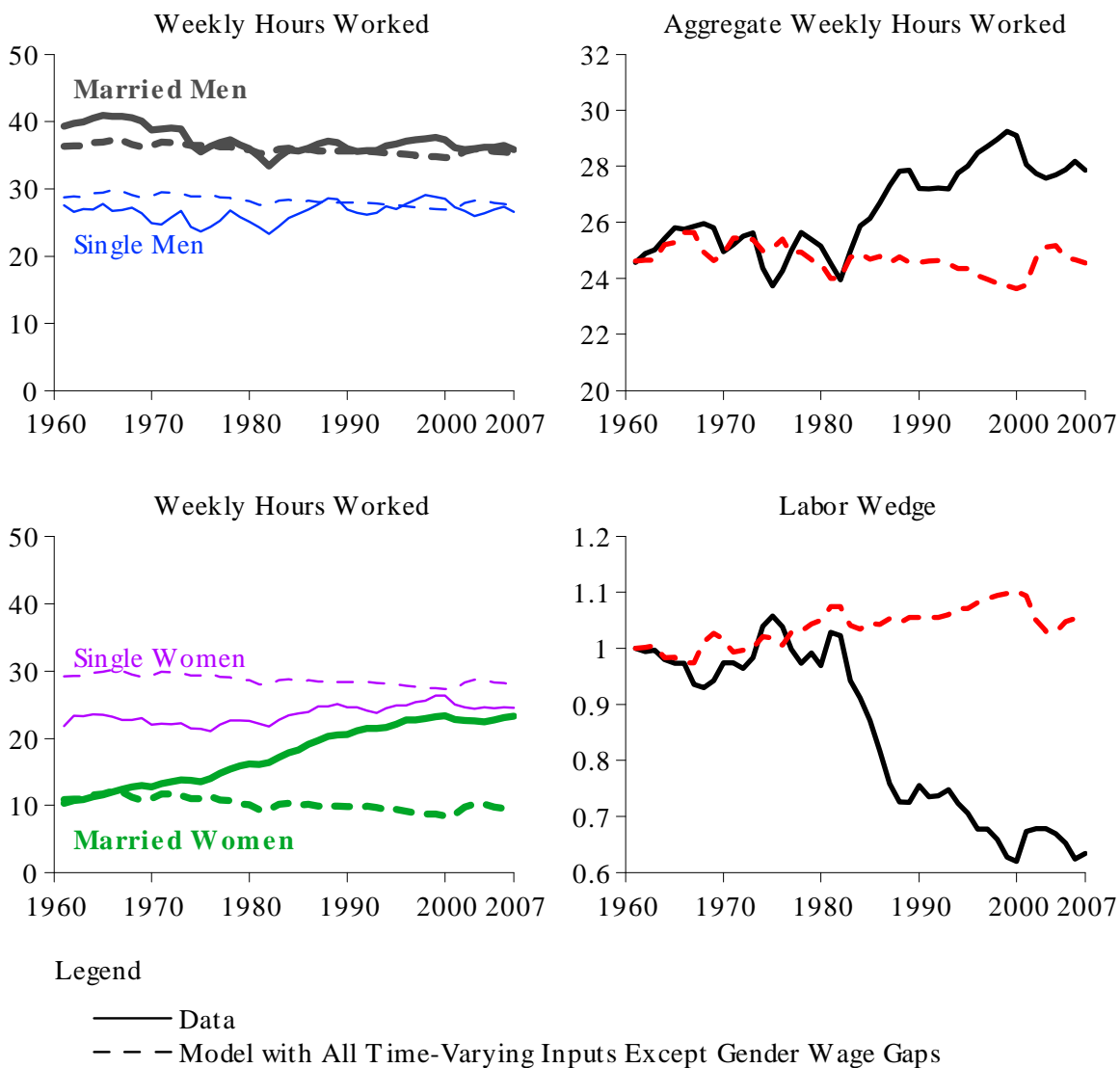


Table 1: BASELINE MODEL PARAMETERS AND TIME-VARYING INPUTS[†]

PARAMETERS	VALUES
Population growth	$\eta = 1.013$
Technology growth	$\gamma = 1.017$
Capital income share	$\theta = 0.33$
Depreciation of capital	$\delta = 0.05$
Capital income tax rate	$\tau_k = 0.40$
Discount factor	$\beta = 0.98$
Intertemporal substitution	$\sigma = 1.00$
Government consumption parameter	$\phi = 1.00$
Leisure parameter	$\alpha = 1.58$
Weights in utility, $\lambda_f U_f + \lambda_m U_m$	$\lambda_f = 0.465, \lambda_m = 1 - \lambda_f$
Share of gender gap due to productivity differences*	$\mu = 0$
TIME-VARYING INPUTS	VALUES (See Figure 2)
Effective labor income tax rate	$\tau_{lt} = \frac{\text{Consumption Tax} + \text{Labor Tax}}{1 + \text{Consumption Tax}}$
Gender wage gaps for married (p) and singles (s)	$\Gamma_{it} = 1 - \frac{\text{hourly wage for females at } t}{\text{hourly wage for males at } t}, i \in \{p, s\}$
Government consumption to output ratio	$\frac{G_t}{Y_t}$
Fractions of married and single households	$n_{pt} = \frac{N_{pt}}{N_t}, n_{fst} = \frac{N_{fst}}{N_t}, n_{mst} = \frac{N_{mst}}{N_t}$

[†]Moments targeted and sources of data are presented in Section 4.1 and Appendix A.1. *In most experiments, we assume $\mu = 0$. We also present results for $\mu = 1$.

Table 2: BASELINE MODEL: CHANGES IN WEEKLY HOURS WORKED, 1961 TO 2007

CHANGES IN AGGREGATE WEEKLY HOURS WORKED[†]:

$$l_{2007}/l_{1961} = \sum_{\nu \in \{mp, ms, fs, fp\}} \left(\frac{n_{\nu 2007}}{n_{\nu 1961}} \cdot \frac{n_{\nu 1961} \cdot l_{\nu 1961}}{l_{1961}} \cdot \frac{l_{\nu 2007}}{l_{\nu 1961}} \right)$$

DATA: $l_{2007}/l_{1961} = \mathbf{1.133}$

CONTRIBUTION OF GROUP ν	Male		Female	
	Married	Single	Single	Married
Change in fraction of population, $\frac{n_{\nu 2007}}{n_{\nu 1961}}$	0.743*	2.364*	1.803*	0.743*
Share of aggregate hours in 1961, $\frac{n_{\nu 1961} \cdot l_{\nu 1961}}{l_{1961}}$	0.641	0.097	0.098	0.169
Change in hours worked, $\frac{l_{\nu 2007}}{l_{\nu 1961}}$	0.913	0.964	1.122	2.258

BASELINE MODEL: $l_{2007}/l_{1961} = \mathbf{1.084}$

CONTRIBUTION OF GROUP ν	Male		Female	
	Married	Single	Single	Married
Change in fraction of population, $\frac{n_{\nu 2007}}{n_{\nu 1961}}$	0.743*	2.364*	1.803*	0.743*
Share of aggregate hours in 1961, $\frac{n_{\nu 1961} \cdot l_{\nu 1961}}{l_{1961}}$	0.589	0.119	0.122	0.169
Change in hours worked, $\frac{l_{\nu 2007}}{l_{\nu 1961}}$	0.822	0.804	1.069	2.083

[†]Changes in aggregate hours are decomposed into the contributions of the different groups in the population: married males (mp), single males (ms), single females (fs) and married females (fp). For ease in writing the formula in this table, we have used notation which is not used in the main text. The number of married males and married females are denoted here by N_{mpt} and N_{fpt} , respectively, whereas the main text just refers to the number of married couples, N_{pt} . Also, here $n_{mpt} \equiv N_{mpt}/N_t$ and $n_{fpt} \equiv N_{fpt}/N_t$. *The fractions of married couples and singles in the total population are exogenous inputs into the baseline model, hence the model matches the changes in these fractions by construction.

Table 3: BASELINE MODEL: CHANGES IN THE LABOR FACTOR, 1961 TO 2007[†]

$$\text{LABOR FACTOR, } 1 - \Delta_t \equiv \frac{\alpha}{(1-\theta)} \frac{(c_t+g_t)}{y_t} \frac{l_t}{1-l_t} = (1 - \tau_{lt}) \left(1 - \frac{n_{pt}\Gamma_{pt}(1-l_{fpt})+n_{fst}\Gamma_{st}(1-l_{fst})}{1-l_t} \right) \frac{l_t}{\tilde{l}_t}$$

CHANGES IN LABOR FACTOR AND COMPONENTS

	Baseline Model ^a	Data ^b
Labor Factor, $1 - \Delta_t$	1.0661	1.2660
Consumption to output ratio, $\frac{(c_t+g_t)}{y_t}$	0.9570	1.0687
Aggregate labor component, $\frac{l_t}{1-l_t}$	1.1141	1.1847
Tax component, $1 - \tau_{lt}$	0.9252	0.9252
Female labor component, $1 - \frac{n_{pt}\Gamma_{pt}(1-l_{fpt})+n_{fst}\Gamma_{st}(1-l_{fst})}{1-l_t}$	1.1523	1.1543
Labor input to effective labor ratio, l_t/\tilde{l}_t	1	

[†]The expressions for the model's labor factor are from Equations (1) and (11). ^aThe column "Baseline Model" reports calculations using data generated from the baseline model. This model attributes all of the gender wage gap to discrimination, so there are no differences between the labor input, l_t , and effective labor, \tilde{l}_t . ^bThe column "Data" shows the changes in the labor factor and components as measured using U.S. data. Detailed data sources are provided in Appendix A.1. To measure the labor factor, we use U.S. data on private consumption, government consumption, output and aggregate hours worked. Notice that α and θ affect the level of the labor factor, but not its changes over time. For other calculations reported in the column "Data", we also use data on female hours worked, taxes, gender wage gaps and fractions of single and married women in the total population.

Table 4: COMPARISON: BASELINE MODEL AND OTHER EXPERIMENTS[†]

	PERCENT CHANGE: 1961 – 2007			
	l_t	l_{fpt}	Δ_t	$\frac{c_t+g_t}{y_t}$
DATA	13.3	125.8	–36.5	6.9
EXPERIMENTS WITH BASELINE CALIBRATION				
<i>Baseline Model*</i>	8.4	108.3	–10.5	–4.3
<i>One Time-Varying Input Only</i>				
Gender wage gaps*	12.7	110.8	–19.6	–4.6
Effective labor income taxes	–3.3	–7.7	12.1	–3.2
Government consumption	2.7	6.4	0.0	–3.5
Fractions of households	5.9	4.2	–6.8	–3.5
<i>All Inputs Except Gender Wage Gaps*</i>	–0.3	–13.0	5.7	–3.2
OTHER EXPERIMENTS: ALL TIME-VARYING INPUTS				
<i>Gender Gaps Due to Productivity ($\mu = 1$)</i>	7.4	108.8	–7.8	–5.6
<i>Reductions in Child Care Costs</i>	11.4	122.0	–12.7	–5.1

[†]We perform a number of experiments under the baseline calibration. In the experiments with only a subset of time-varying inputs, the households’ lump-sum transfers are distributed in the same proportions as in the baseline model. Recalibrating the transfers does not change the results significantly. In the experiment with $\mu = 1$ and the experiment with child care costs, we recalibrate α and λ_f (to match the same targets on hours worked as in the baseline calibration), but leave all other parameters unchanged. *The results from experiments marked with a star are plotted in Figures 3, 4, and 5, respectively.

Table 5: BASELINE MODEL: SENSITIVITY ANALYSIS[†]

PERCENT CHANGE: 1961 – 2007				
	l_t	l_{fpt}	Δ_t	$\frac{c_t + \phi g_t}{y_t}$
VARY σ				
Data, $\phi = 1$	13.3	125.8	-36.5	6.9
Baseline Model				
$\sigma = 0.8$ and $\phi = 1$	10.5	173.4	-11.5	-6.2
$\sigma = 1.0$ and $\phi = 1$	8.4	108.3	-10.5	-4.3
$\sigma = 1.2$ and $\phi = 1$	7.2	85.7	-10.2	-3.1
VARY ϕ				
$\phi = 0$ and $\sigma = 1$				
Data	13.3	125.8	-42.2	10.6
Baseline Model	7.5	107.4	-10.5	-3.6
$\phi = 0.5$ and $\sigma = 1$				
Data	13.3	125.8	-38.9	8.5
Baseline Model	8.0	107.9	-10.5	-3.9

[†]We examine the predictions of our baseline model (in which all exogenous inputs from Figure 2 vary over time) when we change the values of the elasticity of substitution, σ , and the values of the marginal rate of substitution between private and public consumption, ϕ . In each case, we recalibrate α and λ_f to match the same targets on hours worked as in the baseline calibration.

Table 6: INTERNATIONAL EVIDENCE: LONG-RUN CHANGES IN THE LABOR WEDGE IN CANADA (1971-2012) AND GERMANY (1983-2007)[†]

$$\text{LABOR FACTOR, } 1 - \Delta_t \equiv \frac{\alpha}{(1-\theta)} \frac{(c_t+g_t)}{y_t} \frac{l_t}{1-l_t} = (1 - \tau_{lt}) \left(1 - \frac{n_{pt}\Gamma_{pt}(1-l_{fpt})+n_{fst}\Gamma_{st}(1-l_{fst})}{1-l_t} \right) \frac{l_t}{\bar{l}_t}$$

CHANGES IN LABOR FACTOR AND COMPONENTS

	Canada ^a	Germany ^b
Labor Factor, $1 - \Delta_t$	1.2283	0.9505
Consumption to output ratio, $\frac{(c_t+g_t)}{y_t}$	1.0134	0.8989
Aggregate labor component, $\frac{l_t}{1-l_t}$	1.2121	1.0574
Tax component, $1 - \tau_{lt}$	0.9148	0.9183
Female labor component, $1 - \frac{n_{pt}\Gamma_{pt}(1-l_{fpt})+n_{fst}\Gamma_{st}(1-l_{fst})}{1-l_t}$	1.1495	1.0381
Product of tax and female labor components	1.0516	0.9533

[†]The expressions for the model's labor factor are from Equations (1) and (11). We use Canadian or German data to measure changes in the labor factor and components. Note that α and θ affect the level of the labor factor, but not its changes over time, hence we do not need values for these parameters to perform the calculations. Also, note that we have used $\phi=1$ for the marginal rate of substitution between government and private consumption. Choosing $\phi=0$ results in very small changes in the consumption to output ratio, and very small changes to the labor factor. The growth factor for the labor factor in Canada would be 1.2170 instead of 1.2283, while for Germany it would be 0.9684 instead of 0.9505. ^aData sources for Canada: OECD National Accounts and Tax Revenues statistics, National Accounts data from Statistics Canada, and microdata from the Labor Force Survey and the Census. ^bData sources for Germany: OECD National Accounts and Tax Revenues statistics, Regional Accounts VGRdL available at www.vgrdl.de, and microdata available from the German Cross-National Equivalent File, see Frick, Jenkins, Lillard, Lipps, and Wooden (2007) for details.

Table 7: INTERNATIONAL EVIDENCE: OTHER OECD ECONOMIES[†]

	Changes over time			
	$1 - \Delta$	$1 - \tau$	l_{total}	l_{females}
Netherlands (1983-2010)	1.22	1.13	1.20	1.66
Spain (1987-2010)	1.22	0.86	1.15	1.72
Finland (1995-2010)	1.19	1.11	1.08	1.10
Italy (1983-2010)	1.11	0.84	1.04	1.26
UK (1984-2010)	1.08	1.00	1.01	1.25
Belgium (1983-2010)	1.03	0.95	1.07	1.39
Austria (1995-2010)	1.01	1.00	1.02	1.03
Sweden (1995-2010)	1.00	1.03	1.01	1.02
France (1983-2010)	0.97	0.96	0.96	1.12
Australia (1979-2008)	0.93	1.00	0.96	1.11

[†]We use OECD National Accounts and Labor Force Statistics to construct average usual hours worked for all working-age persons, l_{total} , and for females, l_{females} , as well as the labor wedge, Δ . We construct effective labor income taxes, τ , using consumption and labor taxes provided by McDaniel (2007) for 1950-2010, see Excel file available at <http://www.caramcdaniel.com/researchpapers>.