When in Doubt, Tax More Progressively: Uncertainty and Progressive Income Taxation*

Minsu Chang† Chunzan Wu‡

January 16, 2022

Abstract

We study the optimal income tax problem when policymakers have only limited information about household preferences and wage process. To this end, we build an incomplete-market life-cycle model of heterogeneous households and measure uncertainty about preference and wage parameters through Bayesian approach using the U.S. data. We find that accounting for such uncertainty leads to more progressive optimal income tax and substantially larger welfare gain from tax reform. The welfare cost of uncertainty through the income tax channel alone is about 0.4% of lifetime consumption. The existence of uncertainty implies that tax reforms are risky, and hence the risk preferences of policymakers also have important tax policy implications.

JEL Codes: E60, H20.
Keywords: Uncertainty, Income Tax, Optimal Taxation.

*We thank Dan Cao, Behzad Diba, Jesús Fernández-Villaverde, Jonathan Heathcote, Mark Huggett, Dirk Krueger, Toshihiko Mukoyama, and Margit Reischer for helpful comments and suggestions.

†Georgetown University. Email: minsu.chang@georgetown.edu.
‡Peking University. Email: chunzan@sas.upenn.edu.

1 Introduction

How should the government tax people's income? This is one of, if not the, most important questions faced by policymakers. Tremendous efforts of the economics profession, both in micro and macroeconomics, through theoretical and quantitative approaches, have been devoted to answering this question; and thanks to that, substantial progress has been made on this front. For example, the previous literature has identified several primitives crucial to the optimal design of income tax such as the elasticity of labor supply and the amount of uninsurable idiosyncratic risk in the economy. However, policymakers only have limited information about these primitives because they are not observable and must be inferred from the data. How does such imperfect information alter our answer to the optimal tax question? What is the welfare cost of these information frictions faced by policymakers?

In this paper, we study *quantitatively* the income tax policy implications of information frictions in the form of policymakers' uncertainty about household preferences and wage process. As a preliminary step towards our quantitative analysis, we first derive general conditions that determine *qualitatively* how parameter uncertainty affects the tax policy. Specifically, we show that accounting for uncertainty leads to larger welfare gains from tax reforms and more progressive optimal income tax *if* welfare gain is convex in the uncertain parameter and such convexity is enhanced by tax progressivity. In a stylized static model with standard preferences, we demonstrate that these conditions are true for both the elasticity of labor supply and the amount of idiosyncratic wage risk.

To quantify the effects of uncertainty, we build an incomplete-market life-cycle model of heterogeneous households with idiosyncratic wage risk and endogenous labor supply. This model serves two goals. First, through the lens of this model, we interpret the data and assess the uncertainty about key economic parameters. Second, it provides us with a laboratory for conducting thought experiments of tax reforms and evaluating their welfare consequences. We extend the standard single-earner household model by incorporating explicitly the endogenous labor supply decisions of the secondary earners (typically females), since they serve as an important private insurance mechanism and have major implications for income tax policy (Blundell et al. 2016; Wu and Krueger 2021).

The severity of information frictions is measured by the degree of policymakers' uncertainty about parameters governing household preferences and wage process (e.g., consumption and labor supply elasticities; persistences and variances of wage shocks), which in turn depends on how much information about these parameters can be extracted from the data. To assess the amount of uncertainty, we estimate these parameters through a limited-information Bayesian approach, using the data from the 1999-2017 Panel Study of Income Dynamics (PSID).[1] The Bayesian inference produces a joint posterior distribution of preference and wage parameters that summarizes policymakers' uncertainty about these aspects of the economy.

To study the tax policy implications of imperfect information, as represented by the posterior distribution of parameters, we conduct thought experiments with our life-cycle model, in which policymakers choose the income tax to maximize the welfare of a newborn cohort subject to a within-cohort budget constraint.[2] Following Bénabou (2002) and Heathcote et al. (2017), the income tax policy is summarized by two parameters of a nonlinear tax function $\widetilde{T}(Y) = Y - (1-\chi)Y^{1-\mu}$, where Y is household income, $\widetilde{T}(Y)$ is tax liability, and μ and χ represent the progressivity and the level of income tax, respectively.

The key trade-off faced by policymakers is the following. On the one hand, a progressive income tax provides valuable social insurance against idiosyncratic wage risk, which is not fully insurable among households because the financial markets are incomplete. On the other hand, a progressive income tax implies a rising marginal tax rate with income, which dampens household labor supply and leads to efficiency losses. Since households can self-insure (albeit imperfectly) against idiosyncratic wage risk through precautionary savings and family labor supply as our model permits, an expansion of public insurance through income tax may also crowd out such private insurance. In the end, the optimal tax progressivity and the welfare consequences of a tax reform depend on the relative strengths of these chan-

[1] The limited-information Bayesian approach can be viewed as the Bayesian version of the generalized method of moments (GMM). It enables us to i) quantify parameter uncertainty as the posterior distribution that the policymakers' objective function is based on, and ii) avoid the forbidding task of constructing the full likelihood function for a complex nonlinear model like ours.

[2] We confine the optimal tax policy problem to a newborn cohort to avoid the redistribution motive of income tax over different generations (and the tricky question of how to weight them).

nels, which are in turn regulated by the underlying preference and wage parameters that policymakers are, unfortunately, uncertain about.

Our first main finding is that taking into account parameter uncertainty in the policymaking process leads to more progressive optimal income tax and substantially larger welfare gain from tax reform. We reach these conclusions by contrasting the results from two scenarios. In the first scenario, as conventionally assumed in the literature, policymakers ignore parameter uncertainty and search for the optimal policy treating the point estimates of parameters as the true values. In the alternative scenario, policymakers reflect parameter uncertainty by maximizing the expected welfare with respect to the posterior distribution.[3] We find that the optimal tax progressivity is about one percentage point higher once parameter uncertainty is accounted for ($\mu = 0.33$ vs. 0.32), and the welfare gain from the optimal tax reform is about 20% larger than that concluded based on the point estimates (1.78% vs. 1.48% of household lifetime consumption).

Two comments are warranted regarding our first main finding. First, the optimal tax policy taking into account parameter uncertainty is quantitatively close to the one deduced from the point estimates. This is somewhat reassuring in the sense that the mistakes we make by ignoring parameter uncertainty are likely small, at least for the choice of the optimal income tax code. Second, however, it would be premature to claim that parameter uncertainty is not important for policy discussions about tax reforms. In the presence of fixed political and economic costs of tax reforms, a vital assessment for any potential tax reform is whether the benefits of the reform outweigh the fixed costs. Ignoring parameter uncertainty biases strongly this comparison in favor of the latter by underestimating the welfare gain from the tax reform. In short, failing to account for parameter uncertainty may falsely deny the worthiness of a tax reform in the first place.

Our second main finding is that the welfare cost of parameter uncertainty is about 0.38% of household lifetime consumption *through the income tax policy channel alone*. To measure the welfare cost of parameter uncertainty, we compare two tax policies: the first-best and the ex ante optimal. The first-best policy is a

[3] We assume that policymakers take parameter uncertainty as exogenous when choosing the tax policy. Allowing policy decisions to potentially influence the information received by policymakers (and thus the uncertainty and policy in the future) is an interesting direction for future research.

parameter-contingent tax plan that maximizes the welfare for each possible combination of parameter values, and it is only feasible when policymakers have perfect information about parameters. In contrast, the ex ante optimal policy is parameter-invariant and maximizes the expected welfare based on the posterior distribution. The differences in welfare between the first-best and the ex ante optimal policies are then the welfare losses due to the imperfect information since they can be eliminated by a signal that reveals the true parameter values to the policymakers.[4] Because in addition to the income tax, there are also other government policy decisions that rely on information about household preferences and wage process, the total welfare cost of uncertainty through distortions to all policy decisions is likely much larger. Our finding thus indicates substantial potential welfare gains from activities that improve our knowledge about these aspects of the economy.

We also show that parameter uncertainty translates to substantial risk associated with tax reforms. Given the optimal tax reform derived from the point estimates as an example, the 95% credible interval of welfare gain ranges from 51% lower to 60% higher than that under the point estimates, and the 95% credible interval of tax revenue is between 12% lower and 24% higher than the original target. Decomposition exercises suggest that uncertainty about wage process explains approximately two thirds of these variations, and the remaining one third is ascribed to uncertainty about household preferences. Similarly, roughly two thirds of these variations are accounted for by uncertainty about parameters related to males, and the rest is due to uncertainty about female and gender-neutral parameters.

Our benchmark analyses assume risk-neutral policymakers with regard to variations in welfare induced by parameter uncertainty, but we also investigate how various risk preferences of policymakers affect our findings about tax reforms. We consider two types of policymakers: i) "career politicians" who maximize the welfare gain relative to the status quo, and ii) "social planners" who minimize the welfare loss relative to the first-best scenario with perfect information.[5] Interest-

[4]An alternative interpretation to the welfare cost measure is thus the amount of welfare the policymakers are willing to pay in exchange for such a signal. The welfare loss depends on the values of parameters, and the reported result is the expected welfare loss based on the posterior distribution.

[5]The two types of policymakers are equivalent when they are risk-neutral, and hence we do not need to differentiate them for the benchmark results.

ingly, we find that career politicians' incentive for a progressive income tax reform is reduced by their risk aversion, whereas the opposite is true for social planners. In particular, for career politicians, the status quo policy is risk-free since the welfare gain is zero independent of parameter values. Therefore, the risk aversion of policymakers pushes the optimal income tax closer to the status quo, which means lower tax progressivity and smaller welfare gain from the optimal tax reform. For social planners, the status quo policy is risky in the sense that the welfare loss of inaction relative to the first-best is uncertain. We find that the optimal income tax in this case is fairly robust with respect to policymakers' risk aversion, and the welfare gain from the optimal tax reform rises when policymakers become more risk-averse.

This paper is most related to the literature that studies the optimal nonlinear income tax in the Ramsey tradition with heterogeneous-agent incomplete-market models. Influential previous studies include Bénabou (2002), Conesa and Krueger (2006), Conesa et al. (2009), and more recently Bakış et al. (2015), Krueger and Ludwig (2016), and Heathcote et al. (2017) among others. A common implicit assumption of these previous studies is that, when choosing the tax policy, policymakers have perfect information about the parameters of their model for the economy. This paper relaxes that assumption and incorporates policymakers' concern about parameter uncertainty into the policymaking process.

The way that parameter uncertainty is analyzed in this paper has several advantages compared to the sensitivity analysis in the literature, in which quantitative exercises are simply repeated for a list of alternative calibrations of parameters. One that is particularly important for optimal policy studies is that sensitivity analysis only produces a parameter-contingent policy plan, which is not implementable when policymakers don't know the true values of parameters.[6] Our approach also differs from the robust control literature as represented by Hansen and Sargent (2008). Their approach aims at finding a policy that performs well across alter-

[6] Additionally, parameter values in sensitivity analysis are often drawn from previous empirical studies, which might be estimated under alternative assumptions that are inconsistent with the model for policy analysis. In contrast, we infer parameter uncertainty from the data through the lens of the same model for policy analysis. The posterior distribution of parameters also carries information about the correlations between parameters, which we show are important for both the qualitative and quantitative effects of uncertainty (Section 2.1 and 5.1.2).

native models without information on the plausibility of each alternative model, and hence the optimal policy becomes the solution to a max-min problem to avoid worst-case scenarios. In contrast, we analyze model uncertainty (as represented by the values of parameters) from a Bayesian perspective and infer the probabilities of alternative models from the data, which are then used to weight different scenarios in the policymaking process à la Brainard (1967).[7]

Earlier research on policy implications of uncertainty focuses primarily on monetary policy. Representative quantitative works include Levin et al. (2005) and Edge et al. (2010), which study the optimal monetary policy under parameter uncertainty in micro-founded New Keynesian models. In comparison, studies of the fiscal policy implications of uncertainty are more limited and mostly theoretical. For example, in a stylized model, Manski (2014) examines the choice of government size by a planner who has partial knowledge about the economy. Lockwood et al. (2020) studies the optimal income tax problem with uncertainty over the elasticity of taxable income in the Mirrleesian tradition. They consider a simple static environment for tractability, whereas our quantitative dynamic model includes additional features of the economy that are important for tax policy evaluations such as household life cycle and endogenous private insurance through precautionary savings and family labor supply. We also quantify the welfare cost of parameter uncertainty and consider the effects of risk preferences of policymakers.

The rest of this paper proceeds as follows. Section 2 derives qualitative theoretical results about uncertainty's effects on tax policy. Section 3 sets up the incomplete-market life-cycle model for our quantitative analysis. Section 4 describes our empirical strategy for estimating model parameters and reports the parameter uncertainty inferred from the data. Section 5 conducts thought experiments of tax reforms and quantifies the tax policy implications of parameter uncertainty. Section 6 concludes.

2 Theoretical Analysis

In this section, we derive general conditions that determine qualitatively how uncertainty affects the welfare evaluation of tax reforms and the optimal tax policy.

[7] We do consider a special case in Section 5.4 that is similar in spirit to the max-min criterion.

We then apply these theoretical results to a static model with standard preferences and show that uncertainty about the elasticity of labor supply and the magnitude of idiosyncratic wage risk leads to larger welfare gains from tax reforms and more progressive optimal income tax.

2.1 General Results

Without loss of generality, suppose that the tax policy is summarized by a single policy parameter μ that represents the progressivity of the tax schedule, and there is a single economic parameter θ that the policymakers are uncertain about. Let $W(\mu, \theta)$ denote the welfare gain from a tax reform that switches from the status quo policy to tax policy μ if the true value of the economic parameter is θ.

Policymakers have partial information about the economic parameter θ, and their posterior about its value is summarized by a probability distribution with the cdf $F_\theta(\theta)$. When uncertainty is ignored, policymakers treat the expected value of θ, $\bar{\theta} = \int \theta dF_\theta(\theta)$, as its true value and choose the tax policy $\bar{\mu}$ to maximize the welfare gain $W(\mu, \bar{\theta})$, i.e.,

$$\bar{\mu} = \arg\max_{\mu} W(\mu, \bar{\theta}).$$

When uncertainty is accounted for, policymakers should choose the optimal policy μ^* that maximizes the expected welfare gain, i.e.,

$$\mu^* = \arg\max_{\mu} \int W(\mu, \theta) dF_\theta(\theta),$$

where policymakers are assumed to be risk-neutral with respect to the welfare risk induced by parameter uncertainty.

Let us first look at the uncertainty's effect on the welfare evaluation of tax reform. Suppose that the uncertainty about θ is small enough such that we can approximate the welfare gain function $W(\mu, \theta)$ by a second-order Taylor expansion at $\bar{\theta}$:

$$W(\mu, \theta) \approx W(\mu, \bar{\theta}) + W_\theta(\mu, \bar{\theta})(\theta - \bar{\theta}) + \frac{1}{2} W_{\theta\theta}(\mu, \bar{\theta})(\theta - \bar{\theta})^2,$$

where W_θ and $W_{\theta\theta}$ represent the first and the second order partial derivatives of W with respect to θ. The actual welfare gain taking into account uncertainty is then

$$\widetilde{W}(\mu) \equiv \int W(\mu, \theta) dF_\theta(\theta) \approx W(\mu, \bar{\theta}) + \frac{1}{2} \sigma_\theta^2 W_{\theta\theta}(\mu, \bar{\theta}),$$

where $\sigma_\theta^2 = \int (\theta - \bar{\theta})^2 dF_\theta(\theta)$ is the variance of θ based on policymakers' posterior distribution. Therefore, the existence of uncertainty implies a larger (smaller) welfare gain from the tax reform if $W_{\theta\theta}(\mu, \bar{\theta}) > 0$ (< 0), i.e., the welfare gain function is strictly convex (concave) in the uncertain parameter.

For the uncertainty's effect on the optimal policy, notice that the slope of the actual welfare gain function $\widetilde{W}$ at $\bar{\mu}$ (the optimal tax progressivity when uncertainty is ignored) is given by

$$\widetilde{W}'(\bar{\mu}) \approx \underbrace{W_\mu(\bar{\mu}, \bar{\theta})}_{=0} + \frac{1}{2}\sigma_\theta^2 W_{\theta\theta\mu}(\bar{\mu}, \bar{\theta}) = \frac{1}{2}\sigma_\theta^2 W_{\theta\theta\mu}(\bar{\mu}, \bar{\theta}),$$

where $W_\mu(\bar{\mu}, \bar{\theta}) = 0$ comes from the fact that $\bar{\mu}$ maximizes $W(\mu, \bar{\theta})$. Suppose $\widetilde{W}(\mu)$ is single-peaked, then the actual optimal policy μ^* must be more (less) progressive than $\bar{\mu}$ if $W_{\theta\theta\mu}(\bar{\mu}, \bar{\theta}) > 0$ (< 0), i.e., the welfare gain function becomes more (less) convex with respect to the uncertainty parameter as tax progressivity increases.

Proposition 1. *Under previous assumptions, accounting for uncertainty leads to*

1. *Larger (smaller) welfare gains from tax reforms if welfare gain is strictly convex (concave) in the uncertain parameter;*
2. *More (less) progressive optimal tax policy if welfare gain becomes more (less) convex in the uncertain parameter as tax progressivity increases.*

Proposition 1 summarizes the effects of uncertainty when policymakers have limited information about a single parameter. What if there are multiple uncertain parameters? Let $\Theta = (\theta_i)_{i \in I}$ denote the vector of uncertain parameters. Then we can approximate the actual welfare gain from a tax reform by

$$\widetilde{W}(\mu) \approx W(\mu, \bar{\Theta}) + \frac{1}{2} \sum_{i,j \in I} W_{\theta_i\theta_j}(\mu, \bar{\Theta}) \sigma_{\theta_i\theta_j},$$

where $\bar{\Theta}$ is the vector of expected values of parameters, $W_{\theta_i\theta_j}$ is the second order derivative of W with respect to θ_i and θ_j, and $\sigma_{\theta_i\theta_j}$ is the covariance of θ_i and θ_j in the joint posterior distribution of Θ. Following the previous logic, we can see that now in addition to the shape of the welfare gain function, the covariances between uncertain parameters $(\sigma_{\theta_i\theta_j})_{i,j \in I}$ may also matter to the signs of the uncertainty's effects.

2.2 Lessons from the Static Model

We now present a static model with standard preferences, for which we can derive the welfare gain function in closed form. We then apply the general conditions from Section 2.1 to determine qualitatively the implications of uncertainty about key economic parameters, namely, the elasticity of labor supply and the magnitude of idiosyncratic wage risk, on progressive income taxation.

Consider a static economy populated by a continuum of measure one households. Households differ in their labor productivity z, which follows a log-normal distribution $LN(-\frac{1}{2}\sigma_z^2, \sigma_z^2)$ in the population. Note that the average productivity is always 1, and σ_z only affects the dispersion of the productivity distribution. Each household chooses labor supply H and consumption C to maximize its utility subject to the household budget constraint:

$$\max_{\{C,H\}} \ln C - \frac{H^{1+\eta^{-1}}}{1+\eta^{-1}}$$

s.t.

$$C = zH - \widetilde{T}(zH),$$

where η is the Frisch elasticity of labor supply, and $\widetilde{T}(\cdot)$ is the income tax function. Following Bénabou (2002) and Heathcote et al. (2017), we set

$$\widetilde{T}(zH) = zH - (1-\chi)(zH)^{1-\mu},$$

where χ controls the tax level, and μ governs the tax progressivity.

When there is no parameter uncertainty, the government budget constraint requires:

$$\int_z \left[zH - (1-\chi)(zH)^{1-\mu}\right] dF_z(z) = 0,$$

where $F_z(z)$ is the cdf for productivity z. When there is parameter uncertainty, for tractability, we assume that policymakers must choose the tax progressivity μ before parameter uncertainty is resolved, whereas the tax level χ is chosen ex post contingent on the true parameter values to balance the ex post government budget. This assumption will be relaxed in our quantitative analysis. The government budget constraint allows us to substitute χ with μ, and therefore, we focus on the choice of tax progressivity μ in the following discussion.

Suppose that the status quo income tax is flat (i.e., $\mu = 0$), and policymakers are considering a tax reform to maximize household welfare. Given tax policy μ, we can derive the indirect utility of productivity z household as

$$U(z,\mu) = \frac{1}{2}\sigma_z^2\mu(1-\mu) + \frac{1}{1+\eta^{-1}}\ln(1-\mu) + (1-\mu)\ln z - \frac{1-\mu}{1+\eta^{-1}}.$$

The social welfare gain from the tax reform is then

$$\Delta(\mu) \equiv \int_z [U(z,\mu) - U(z,0)]dF_z(z) = \frac{1}{2}\sigma_z^2\mu(2-\mu) + \frac{1}{1+\eta^{-1}}[\ln(1-\mu)+\mu].$$

One important issue here is that $\Delta(\mu)$ is in household indirect utility, and hence it may not be comparable across different parameter states, especially when there is uncertainty about household preferences. To address this issue, we measure the welfare gain from a tax reform in each parameter state by the consumption equivalent variation, i.e., the amount of consumption transfers required to generate the same welfare gain as the tax reform. Following the convention, we set the consumption transfers to be proportional to household consumption before the tax reform, and given the log utility in consumption, the welfare gain in consumption equivalent variation is

$$W(\mu) = \int_z (e^{\Delta(\mu)} - 1)\underbrace{C(z,0)}_{=z} dF_z(z) = e^{\Delta(\mu)} - 1,$$

where $C(z,0) = z$ is the consumption of productivity z households under the status quo tax policy. Based on the expression of $\Delta(\mu)$ where $\mu \in (0,1)$, we can see that the welfare gain $W(\mu)$ decreases in the Frisch elasticity η and increases in the magnitude of idiosyncratic risk σ_z.

Suppose that policymakers are uncertain about η and σ_z.[8] Let $\bar{\eta}$ and $\bar{\sigma}_z$ denote the expected values of η and σ_z based on policymakers' posterior belief, and $\bar{\mu}$ denote the policy that maximizes the welfare when uncertainty is ignored. In Appendix A, we show that if $\mu, \bar{\mu} \in (0,1)$, we have i) $W_{\eta\eta}(\mu,\bar{\eta}) > 0$, $W_{\eta\eta\mu}(\bar{\mu},\bar{\eta}) > 0$; and ii) $W_{\sigma_z\sigma_z}(\mu,\bar{\sigma}_z) > 0$, $W_{\sigma_z\sigma_z\mu}(\bar{\mu},\bar{\sigma}_z) > 0$. Applying the results from Section 2.1, and we have:

Corollary 1. *Under the previous assumptions, uncertainty about the elasticity of*

[8] For simplicity, we assume here that there is no correlation between the value of η and σ_z based on policymakers' belief, i.e., $\sigma_{\eta\sigma_z} = 0$. We will allow correlations between uncertain parameters in our quantitative analysis.

labor supply and the magnitude of idiosyncratic wage risk leads to larger welfare gains from tax reforms and more progressive optimal income tax.

3 The Quantitative Model

The sharp theoretical predictions of the static model in Section 2.2 come at the cost of abstracting away from some aspects of reality important for tax policy consideration, for example, household life cycle and endogenous private insurance through precautionary savings and family labor supply. To provide a comprehensive examination of uncertainty's effects on tax policy, a richer model augmented with these additional features is called for.

In this section, we describe the quantitative dynamic model, through which we interpret the data and measure uncertainty about key economic parameters. The model also serves as a laboratory for conducting thought experiments of tax reforms and evaluating their welfare consequences. We first describe the physical environment of our model, and then state the household optimization problem in recursive formulation.

3.1 Environment

We study a partial equilibrium life-cycle model with idiosyncratic wage risk and endogenous labor supply. We follow a cohort of a continuum of measure one households over their life cycle. These households live for T periods, from age 1 to T, work in the first R periods of life, and then are retired from age $R + 1$ onward. Each household consists of two members: a male and a female. For simplicity, we omit the index for different households and denote by $X_{j,t}$ the variable X of member j at age t, with $j = 1$ or 2 indicating the male or the female member.

Households enjoy joint consumption C_t and choose the labor supply of both members $H_{1,t}$ and $H_{2,t}$ that incur disutility. The period utility function is given by $u(C_t, H_{1,t}, H_{2,t})$. An operative extensive margin of female labor supply is included in the model by introducing a fixed per-period utility cost f whenever female hours worked are strictly positive. Households discount the future utility at the constant rate δ, so that $1/(1 + \delta)$ is the household time discount factor.

Members of a household can work at wages $W_{j,t}$ determined by their labor productivity. Log-wages of both household members are stochastic and represent the sum of i) a deterministic life-cycle component $g_{j,t}$ that is common across households, and ii) an idiosyncratic stochastic component $F_{j,t}$:

$$\ln W_{j,t} = g_{j,t} + F_{j,t}. \tag{1}$$

The idiosyncratic component $F_{j,t}$ follows an AR(1) process with persistence ρ_j:

$$F_{j,t} = \rho_j F_{j,t-1} + v_{j,t},$$

where $v_{j,t}$ is the normally distributed random shock to member j's wage at age t, and they may be correlated between two members of the same household, but are independent over time and across different households:

$$\begin{bmatrix} v_{1,t} \\ v_{2,t} \end{bmatrix} \sim N\left(0, \begin{bmatrix} \sigma^2_{v_1} & \sigma_{v_1,v_2} \\ \sigma_{v_1,v_2} & \sigma^2_{v_2} \end{bmatrix}\right).$$

Here $\sigma^2_{v_1}$, $\sigma^2_{v_2}$, and σ_{v_1,v_2} are the variances and covariance of male and female wage shocks. Note that both the deterministic component and the parameters governing the stochastic component are gender-specific.

As is common in standard incomplete-market models, households cannot trade fully state-contingent Arrow securities, but they can save, and potentially borrow, at the risk-free interest rate r subject to a borrowing limit $\underline{A}$.[9] Working-age households need to pay income and payroll taxes in each period, and retired households are eligible for a fixed amount of retirement benefit b in each period in which she is alive.[10]

3.2 Household Optimization Problem

The state variables of a working-age household include the current savings A, the male and female idiosyncratic wage components, F_1 and F_2, and the age of the household t. In each period, members of each household make joint decisions on household consumption C, savings for the next period A', and labor supply of both members H_1 and H_2 to maximize their discounted utility subject to the budget

[9] Households are born with zero savings.

[10] The U.S. social security benefits are piecewise linear functions of average monthly past earnings over the working life. Additional rules govern benefits for spouses. A full representation of the U.S. social security system is costly in terms of computation, since it adds two continuous state variables to the recursive formulation of the problem. Hence we model the U.S. social security benefit formula starkly, by assuming that the benefits per household are independent of past contributions.

constraint. A working-age household's problem is then, in recursive form:

$$
\begin{aligned}
V(A, F_1, F_2, t) = & \max_{\{C, A', H_1, H_2\}} u(C, H_1, H_2) - \mathbf{I}(H_2 > 0) f \\
& + \frac{1}{1+\delta} \mathbb{E}_{(F_1', F_2')} \left[V(A', F_1', F_2', t+1) | F_1, F_2 \right] \\
& \text{s.t.} \quad C + A' = Y - \widetilde{T}(Y) - \tau_{ss} Y + (1+r) A, \\
& \qquad\; Y = W_{1,t} H_1 + W_{2,t} H_2, \\
& \qquad\; C, H_1, H_2 \geq 0, \; A' \geq \underline{A},
\end{aligned}
$$

where $\mathbf{I}(H_2 > 0)$ equals 1 if female hours H_2 is positive. Female hours of $H_2 = 0$ corresponds to non-participation. The term $\widetilde{T}(Y)$ in the budget constraint is the income tax function that determines the tax liability of a household with before-tax income Y, and τ_{ss} is a flat payroll tax representing the Federal Insurance Contribution Act (FICA) taxes. The wage of each member $W_{j,t}$ is determined by household states F_j and t according to (1).

After retirement, labor productivity falls to zero, and hence households optimally do not work in retirement. The state variables of retired households reduce to only current savings and the age of the household. The dynamic programming problem of a retired household is then:

$$
\begin{aligned}
V^R(A, t) = & \max_{\{C, A'\}} u(C, 0, 0) + \frac{1}{1+\delta} V^R(A', t+1) \\
& \text{s.t.} \quad C + A' = b + (1+r) A, \\
& \qquad\; C \geq 0, \; A' \geq \underline{A}.
\end{aligned}
$$

Households are assumed to have an additively separable utility function of the form:

$$
u(C, H_1, H_2) = \frac{C^{1-\sigma}}{1-\sigma} - \psi_1 \frac{H_1^{1+\eta_1^{-1}}}{1+\eta_1^{-1}} - \psi_2 \frac{H_2^{1+\eta_2^{-1}}}{1+\eta_2^{-1}}.
$$

The advantage of using this preference structure is that all the preference parameters are directly interpretable.[11] The parameter σ governs the intertemporal elasticity of substitution for consumption, and its reciprocal is the Frisch elasticity of consumption with respect to its own price. The parameters η_1 and η_2 are the Frisch

[11] In addition, Wu and Krueger (2021) find that a life-cycle model of two-earner households augmented with such preferences matches well the earnings and consumption dynamics in the data.

elasticities of male and female labor supply with respect to their own wages, and ψ_1 and ψ_2 control the levels of disutility from male and female labor supply.

We permit the income tax function $\widetilde{T}(Y)$ to be progressive and use the same two-parameter tax function as in the static model:

$$\widetilde{T}(Y) = Y - (1-\chi)Y^{1-\mu}, \tag{2}$$

where μ and χ are two parameters governing the progressivity and the level of the income tax, respectively. It implies that after-tax income $Y - \widetilde{T}(Y)$ is an increasing and concave function of pre-tax income Y.

4 Measure Uncertainty from the Data

In this section, we take the model in Section 3 to the U.S. data and quantify the degree of uncertainty about parameters governing household preferences and idiosyncratic wage process. We first describe the main data set used for our empirical analysis and how we choose the values of externally calibrated parameters. We then provide a brief summary of the limited-information Bayesian method that we employ to estimate the preference and wage parameters. Finally, we explain our estimation strategy and report the estimation results.

4.1 The Data

Our main data source is the core sample of the 1999-2017 Panel Study of Income Dynamics (PSID), from which we obtain individual and household level information about earnings, hours worked, consumption, net worth, and characteristics such as age, race, education, number of children, and so on.[12] The longitudinal structure of the PSID allows us to follow the same individuals and households over time (i.e., every two years) and observe the comovement in these variables, which is crucial for our estimation strategy. We focus on married households with working male head aged between 30 and 57 because this group of households fit best the specification of our life-cycle model and represent the majority of the U.S. population. To ease the burden of computation, we set each period in our life-cycle model to four years in the data, and hence the biennial PSID data are converted to

[12]We follow closely the procedures in Blundell et al. (2016) when constructing variables from the PSID data and identifying outliers.

four-year frequency. Wages are constructed as earnings divided by hours worked.

All nominal variables are converted to values in year 2000 U.S. dollars based on the consumer price index for all urban consumers (CPI-U) from the U.S. Bureau of Labor Statistics. We set the units of income and labor supply in our model to the average four-year earnings and hours worked by males in our PSID sample, which are \$243,766 and 9,202 hours, respectively.

4.2 External Calibration

Before estimating the preference and wage parameters, we first calibrate the other parameters outside of the life-cycle model, and their values are reported in Table 1. The key difference between these calibrated parameters and the estimated ones is that the former are either directly observable or can be inferred with few to no model assumptions.

4.2.1 Demographic

As mentioned earlier, one period in the model represents four years in the data. We consider the part of household life cycle between age 22 and age 81 with a retirement age of 65. This then translates to a life cycle in the model from age 1 to $T = 15$ with a retirement age of $R = 11$.

4.2.2 Income Tax

There are two parameters in the income tax function given by (2): μ for the income tax progressivity and χ for the income tax level. We estimate these two parameters by taking the natural log of (2) and running the following OLS regression with our PSID sample:

$$\ln(Y - \widetilde{T}(Y)) = \ln(1 - \chi) + (1 - \mu)\ln(Y).$$

Tax liability $\widetilde{T}(Y)$ is defined as federal income tax minus earned income tax credit (EITC) and food stamp benefits. Federal income tax and EITC are calculated based on household income Y and the actual income tax code, and food stamp benefits are obtained directly from the PSID. The estimated income tax parameters are $\mu = 0.128$ and $\chi = 0.126$.

4.2.3 Payroll Tax and Retirement Benefit

The flat payroll tax rate τ_{ss} is set to 7.65%, based on the Federal Insurance Contribution Act (FICA) tax rates on pre-tax income of employees. The retirement benefit b in the model corresponds to the sum of Social Security benefits and benefits from Medicare. We calibrate the Social Security benefits to the sum of average benefits received by males and females aged 66 and older in the 1999-2017 Annual Social and Economic Supplement (ASEC) of the Current Population Survey (CPS), given by $20,975 per year in 2000 dollars. Since the benefits from Medicare are difficult to measure directly, we assume that they are proportional to the social security benefits, based on the ratio of Medicare tax rate to Social Security tax rate. Therefore, the retirement benefit b in the model is calibrated to $\$20,975 \times 7.65\%/6.2\% \times 4 = \$103,522$, i.e., 0.425 in model income units.

4.2.4 Wage Trend and Initial Wage

For the deterministic life-cycle component of log-wage, $g_{j,t}$ in (1), we regress the male or female log-wage on a quadratic polynomial in age, together with a group of controls for the year, education, race, and location, etc. The gender-specific age-profiles of log-wage are then constructed as the predicted values from these regressions at different ages while integrating over the remaining covariates. The resulting wage trends are presented in Figure 7 of Appendix B. For the initial idiosyncratic component of log-wage at the labor market entry, $F_{j,1}$ in (1), we assume that it is normally distributed across households and potentially correlated between the two members of each household. We estimate the covariance matrix of the initial idiosyncratic components from the residuals of the log-wage regressions at model age 1 (data age 22-25). The variances of male and female initial wages are 0.211 and 0.232, respectively, and the initial wages are positively correlated with a correlation coefficient of 0.214.

4.2.5 Interest Rate and Borrowing Limit

We set the annual risk-free interest rate at 2%, which implies a four-year interest rate of $r = 8.24\%$. We exclude non-collateralized debts in the model by imposing a zero borrowing limit (i.e., $\underline{A} = 0$).[13]

4.3 Bayesian Inference

We estimate the remaining parameters about household preferences and idiosyncratic wage process using the PSID data through a limited-information Bayesian

[13] Since we focus on households aged 30-57, who often have significantly positive net worth, the zero-borrowing-limit assumption is inconsequential to our results.

Table 1: Externally Calibrated Parameters

Parameter	Description	Value
T	length of life cycle	15
R	retirement age	11
μ	income tax progressivity	0.128
χ	income tax level	0.126
τ_{ss}	payroll tax rate	7.65%
b	retirement benefit	0.425
$\sigma^2_{F_{1,1}}$	variance of male initial wage	0.211
$\sigma^2_{F_{2,1}}$	variance of female initial wage	0.232
$corr_{F_{1,1},F_{2,1}}$	correlation of male and female initial wage	0.214
r	real interest rate	8.24%
$\underline{A}$	borrowing limit	0

Notes: Each period in the model corresponds to four years in the data.

approach and quantify the uncertainty about these parameters by the posterior distribution.

4.3.1 The Limited-Information Bayesian Method

The limited-information Bayesian method, as described in Kim (2002) and later advocated by Christiano et al. (2010) and Fernández-Villaverde et al. (2016) among others, can be viewed as the Bayesian version of the generalized method of moments (GMM). Similar to GMM, the limited-information Bayesian method only uses a set of moments from the data for parameter inference, and therefore, it does not require strong distributional assumptions about the error terms in the model of the data generating process. There are two main reasons why we adopt the limited-information Bayesian method for our empirical analysis. First, a proper characterization of parameter uncertainty requires the distribution of parameters conditional on the information we already have (e.g., the data we observe), which is conveniently the output of Bayesian inference.[14] Second, the full likelihood function is difficult to construct for a complex nonlinear model like ours, whereas the limited-information Bayesian method allows us to sidestep this obstacle.

Let Θ denote the parameters of interest and $\hat{\mathbf{m}}$ denote the vector of M empirical moments from the data for estimation. Kim (2002) shows that the likelihood of $\hat{\mathbf{m}}$

[14] In contrast, the standard GMM produces estimators for parameters as functions of the data with distributions conditional on the true values of parameters.

conditional on Θ is approximately

$$f(\hat{\mathbf{m}}|\Theta) = (2\pi)^{-\frac{M}{2}}|S|^{-\frac{1}{2}} \exp\left[-\frac{1}{2}\left(\hat{\mathbf{m}} - \mathbf{m}(\Theta)\right)' S^{-1}\left(\hat{\mathbf{m}} - \mathbf{m}(\Theta)\right)\right], \tag{3}$$

where $\mathbf{m}(\Theta)$ is the model's prediction for the moments under parameter Θ, and S is the covariance matrix of $\hat{\mathbf{m}}$. The covariance matrix S is often unknown but can be replaced by a consistent estimator of it, which can be obtained through bootstrap. Bayes' theorem tells us that the posterior density $f(\Theta|\hat{\mathbf{m}})$ is proportional to the product of the likelihood $f(\hat{\mathbf{m}}|\Theta)$ and the prior density $p(\Theta)$:

$$f(\Theta|\hat{\mathbf{m}}) \propto f(\hat{\mathbf{m}}|\Theta)p(\Theta), \tag{4}$$

and we can then apply the standard Markov Chain Monte Carlo (MCMC) techniques such as the Metropolis-Hastings algorithm to obtain a sequence of random samples from the posterior distribution.

4.3.2 Estimation Strategy

Except the Bayesian approach, we follow closely the empirical strategy in Blundell et al. (2016). In particular, using the PSID data, we first regress wage, consumption, and earnings growth on observable characteristics of individuals and households and obtain the residuals $\Delta w_{j,t}$, Δc_t, and $\Delta y_{j,t}$, respectively. Similar to Blundell et al. (2016), we estimate the five wage parameters (ρ_1, ρ_2, σ_{v_1}, σ_{v_2}, $corr_{v_1,v_2}$) using only a group of second-order moments of $\Delta w_{j,t}$, and hence the estimation of wage parameters requires only the statistical model of wage process and is immune to our assumptions about the life-cycle model.[15] Given the estimates of wage parameters, we then combine a group of second-order moments of $\Delta w_{j,t}$, Δc_t, and $\Delta y_{j,t}$ and a group of first-order moments of earnings, hours worked, female participation, and household net worth to estimate jointly the seven preference parameters (σ, η_1, η_2, ψ_1, ψ_2, f, δ).[16] Blundell et al. (2016) show that the second-order moments identify the elasticity parameters (σ, η_1, η_2), and the remaining preference parame-

[15] We estimate the correlation coefficient between male and female wage shocks $corr_{v_1,v_2}$ instead of the covariance σ_{v_1,v_2}.

[16] Note that we do not need to adjust the empirical moments for sample selection due to endogenous female participation decisions since our structural model also features an operative extensive margin of female labor supply that is disciplined by the data.

ters are pinned down by the first-order moments.[17] The list of moment conditions that we employ in our estimation and their values from the data and the model are provided in Table 10 of Appendix B .

When constructing the likelihood in (3) for Bayesian estimation, the estimate of the covariance matrix S is obtained from 5,000 bootstrap samples of the PSID data, and the model predicted moments $\mathbf{m}(\Theta)$ are computed based on a model-simulated panel of 50,000 households.[18] We adopt an uninformative prior $p(\Theta)$ consisting of independent uniform distributions of each parameter. We apply the random-walk Metropolis-Hastings algorithm to simulate draws from the posterior density $f(\Theta|\hat{\mathbf{m}})$ given by (4), and the posterior distribution is characterized by a sequence of 15,000 draws after a burn-in of 5,000 draws.

4.3.3 Estimation Results

Table 2 reports the posterior means and the 95% credible intervals of preference and wage parameters from the Bayesian estimation, together with the uniform priors. For the posterior means, one interesting result is that our estimate of the Frisch elasticity of female labor supply η_2 is somewhat lower than the male counterpart η_1. The reason is partly because we allow gender-specific persistence of wage shocks and find that female shocks are significantly less persistent than male shocks (0.755 vs. 0.903). This difference reduces the income/wealth effect of female wage shocks, and hence a weaker substitution effect (i.e., a lower elasticity of labor supply) for females is enough to explain the variations of female earnings in the data. The 95% credible intervals of the posterior distributions are much narrower than the uniform priors, suggesting that the uncertainty about these parameters is greatly reduced by the information contained in the data.

Figure 1 plots the posterior distributions of each preference and wage parameter. Notice that the posterior distributions are not necessarily normal or symmetric. The distribution of male wage persistence ρ_1 is truncated at one because we restrict the wage persistence parameters ρ_j to be no greater than one in the estimation. As

[17]We verify our empirical strategy with model-simulated data and find that it works well in recovering the true values of parameters.

[18]For consistency, we add measurement errors to model-simulated data as well. Like Blundell et al. (2016), the standard deviation of measurement errors on log-consumption is set to 0.20 based on the data moment $\mathbb{E}(\Delta c_t \Delta c_{t-1})$. For log-earnings and log-hours, the standard deviations of measurement errors are set to 0.15 following the literature.

Table 2: Estimated Parameters

Parameter	Description	Posterior Distribution Mean	Posterior Distribution 95% Interval	Uniform Prior [Min, Max]
σ	inverse of consumption elasticity	2.57	[2.13, 3.02]	[1.00, 6.00]
η_1	male labor supply elasticity	0.329	[0.275, 0.383]	[0.001, 2.000]
η_2	female labor supply elasticity	0.243	[0.184, 0.323]	[0.001, 2.000]
ψ_1	disutility of male labor supply	0.934	[0.841, 1.037]	[0.010, 5.000]
ψ_2	disutility of female labor supply	3.20	[2.09, 4.76]	[0.01, 5.00]
f	fixed utility cost of female participation	0.164	[0.142, 0.186]	[0.001, 0.500]
δ	discount rate of utility	0.0265	[0.0190, 0.0336]	[-0.1000, 0.1000]
ρ_1	persistence of male wage shocks	0.903	[0.781, 0.990]	[0.100, 1.000]
ρ_2	persistence of female wage shocks	0.755	[0.599, 0.923]	[0.100, 1.000]
$\sigma^2_{v_1}$	variance of male wage shocks	0.0744	[0.0665, 0.0823]	[0.0001, 0.3000]
$\sigma^2_{v_2}$	variance of female wage shocks	0.0845	[0.0763, 0.0924]	[0.0001, 0.3000]
$corr_{v_1,v_2}$	correlation of male and female wage shocks	0.064	[0.011,0.115]	[-0.300, 0.500]

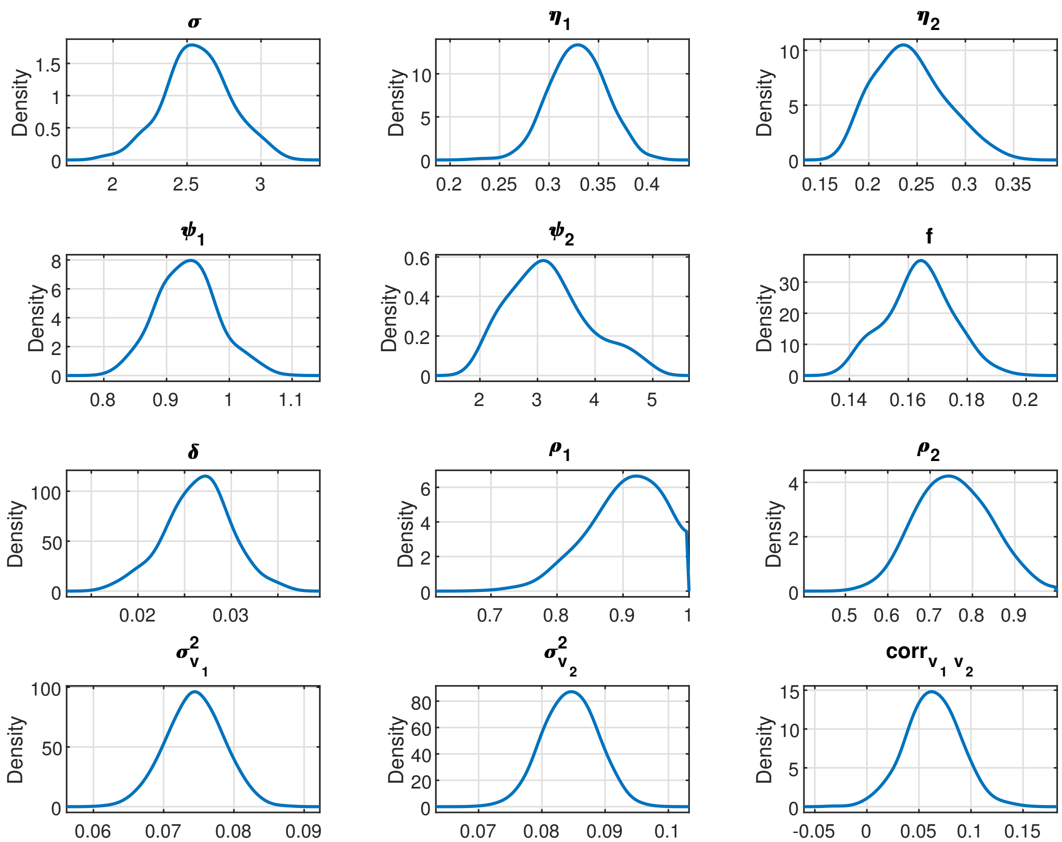

Figure 1: Distributions of Preference and Wage Parameters

Notes: This figure plots the marginal probability density functions of preference and wage parameters based on the joint posterior distribution. The wage persistence parameters ρ_1 and ρ_2 are restricted to be no greater than one.

shown by the graphs, this constraint is binding only for the male but not for the female. Both Table 2 and Figure 1 are about the marginal distributions of parameters, but since these parameters are estimated jointly, they may be correlated in the posterior distribution. Table 11 and 12 of Appendix B report the correlations among preference and wage parameters, respectively. In general, the correlations among preference parameters are stronger than those among the wage parameters.

5 Uncertainty and Income Tax Reform

In this section, we study *quantitatively* how uncertainty about household preferences and wage process affects the optimal design and welfare evaluation of tax reform. Specifically, based on the structural life-cycle model in Section 3 and the estimated posterior distribution of parameters in Section 4, we conduct thought experiments in which policymakers maximize the welfare of a newborn cohort by adjusting the income tax policy subject to a within-cohort budget constraint.[19]

The key trade-off of this policymaking problem is between the welfare benefits of public insurance through progressive income taxation and the efficiency costs of distortions to household labor supply. Since our model allows households to self-insure against idiosyncratic risk through endogenous precautionary savings and family labor supply, an expansion of public insurance may also crowd out such private insurance. In the end, the optimal tax progressivity and the welfare consequences of tax reform depend on the relative strengths of these channels, which are in turn regulated by the underlying preference and wage parameters that policymakers are, unfortunately, uncertain about.

Our quantitative analysis starts with the case when policymakers treat the point estimates of parameters as their true values. We then investigate how taking into account parameter uncertainty in the policymaking process alters our previous conclusions about the optimal tax reform. We quantify the welfare cost of uncertainty by comparing the ex ante optimum to the first-best scenario with perfect informa-

[19] We intentionally choose the welfare measure and government budget constraint for a newborn cohort to avoid the redistribution motive of tax policy over different cohorts and hence the need to take a stand on how to weight different generations. Also, under the partial equilibrium setting, there is no interaction between different cohorts, and the optimal tax policy problem does not require optimizing over the transition path.

tion, in which tax policy can be contingent on the true parameter values. Finally, we explore implications of various risk preferences of policymakers and alternative assumptions about the government budget constraint.[20]

5.1 Tax Reform Ignoring Uncertainty

5.1.1 Optimal Tax Reform Based on Point Estimates

We first consider the case when policymakers ignore the parameter uncertainty when searching for the optimal tax policy, as conventionally done in the previous literature. That is, the policymakers will treat the point estimates of parameters as their true values. We require all potential tax reforms to be revenue-neutral in the sense that the present discounted value (at the fixed interest rate r) of taxes paid by the cohort over its life cycle remains constant. Formally, the optimal tax policy problem is written as follows:

$$(\mu^*(\bar{\Theta}), \chi^*(\bar{\Theta})) = \arg\max_{(\mu,\chi)} V(\mu, \chi; \bar{\Theta}) \tag{5}$$

s.t.

$$\sum_{t=1}^{R} \frac{\text{Tax}_t(\mu, \chi; \bar{\Theta})}{(1+r)^{t-1}} = \bar{G}. \tag{6}$$

Here $V(\mu, \chi; \bar{\Theta})$ is the expected lifetime utility of a newborn cohort, and $\text{Tax}_t(\mu, \chi; \bar{\Theta})$ is the tax revenue collected from this cohort at age t, both under tax policy (μ, χ) and the point estimates (i.e., posterior means) of parameters $\bar{\Theta}$. The fixed amount of government expenditure $\bar{G}$ is set equal to the left-hand side of (6) under the status quo tax policy $(\bar{\mu}, \bar{\chi})$.

Table 3 reports the optimal tax policy and the welfare gain from a tax reform that changes the tax policy from the status quo to the optimal one. The welfare gain is measured by the market value of consumption transfers required to induce the same welfare change as the tax reform.[21] The optimal tax policy is significantly more progressive than the status quo policy (0.321 vs. 0.128) and slightly lower

[20] All quantitative results are based on model-simulated panels of 50,000 households over their life cycles. To keep the computation burden manageable, we approximate the posterior distribution with a sample of 30 independent draws from the joint distribution. Increasing the number of draws further does not affect the results significantly.

[21] Following the convention, we assume that the consumption transfers are proportional to household consumption before the tax reform.

in level (0.111 vs 0.126). The welfare gain from the optimal tax reform is sizable and about 1.48% of household lifetime consumption. These results reflect the fact that based on the point estimates of parameters, there is too little insurance against idiosyncratic wage risk in the status quo economy, and hence there are substantial benefits for the government to step in and provide extra public insurance through progressive income taxation. Moreover, such tax reform would not interfere with the government's budget constraint as it is supposed to be revenue-neutral by design.

Table 3: Optimal Tax Reform Based on Point Estimates

Tax Progressivity (μ)	Tax Level (χ)	Welfare Gain
0.321	0.111	1.48%

Notes: Welfare gain is reported as a percentage of lifetime consumption under the status quo tax policy and the point estimates of parameters.

5.1.2 Unintended Consequences of Tax Reform

However, Table 3 shows only part of the story since the point estimates of parameters may differ, sometimes substantially, from the true values. As a result, if we naively adopt this "optimal" tax reform, we could end up with severe unintended consequences, both in terms of its welfare consequences and its implications on government budget. Based on the parameter uncertainty we estimated from the data (i.e., the posterior distribution), the left panel of Figure 2 plots the probability distribution of the realized welfare gain resulted from the "optimal" tax reform, and the right panel produces a similar graph for tax revenue. The welfare gain and tax revenue at the point estimates are normalized to one. It is evident in these graphs that there are significant chances for the realized welfare gain and tax revenue to be far away from those concluded based on the point estimates. As the benchmark results in Table 4 show, the 95% credible interval of realized welfare gain ranges from 51% lower to 60% higher than the presumed value based on the point estimates. The tax reform would also affect the government budget constraint in a nontrivial way with the 95% credible interval of tax revenue between 12% lower (deficit) and 24% higher (surplus) than the original target.

The counterfactual results in Table 4 demonstrate the importance of properly

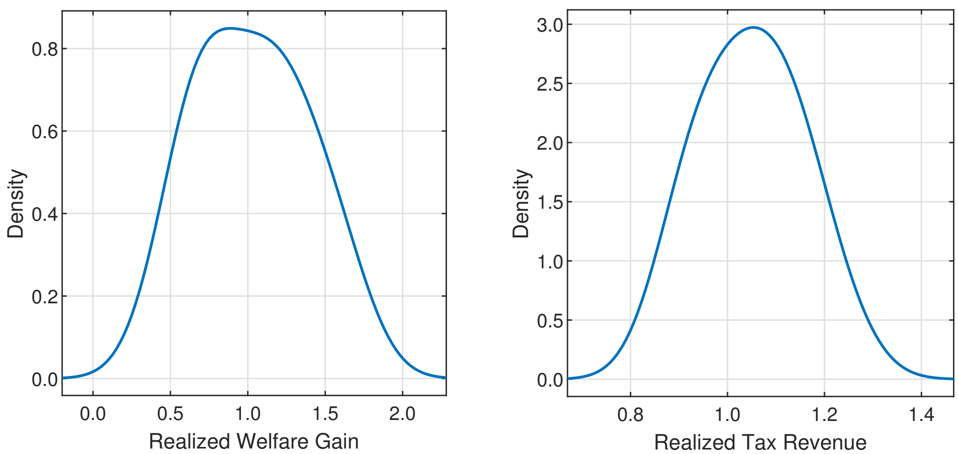

Figure 2: Risky Tax Reform: Distributions of Realized Welfare Gain and Tax Revenue

Notes: This figure shows the distributions of welfare gain (left panel) and tax revenue (right panel) under the optimal tax reform derived from the point estimates of parameters. Welfare gains and tax revenues are converted to fractions of their counterparts under the point estimates of parameters.

Table 4: Risky Tax Reform

	Welfare Gain		Tax Revenue	
	Std.	95% Interval	Std.	95% Interval
Benchmark	0.347	[0.49, 1.60]	0.102	[0.88, 1.24]
Counterfactual				
Reversed Corr.	0.461	[-0.02, 1.93]	0.129	[0.83, 1.27]
Uniform Dist.	0.417	[0.53, 1.85]	0.138	[0.87, 1.38]

Notes: "Reversed Corr." results are based on the posterior distribution with sign-reversed correlations between (ψ_1, ψ_2, f) and other parameters. "Uniform Dist." results are based on the uniform posterior distribution with the same posterior means and variances. Welfare gains and tax revenues are converted to fractions of their counterparts under the point estimates before any statistics are calculated.

capturing the correlations between parameters and the shape of the parameter distribution when evaluating the consequences of parameter uncertainty. This is one of the advantages of our structural Bayesian approach over the standard sensitivity analysis in the literature. From the posterior distribution of parameters, we observe high correlations (> 0.8 in absolution value) between parameters governing the levels of disutility from labor (ψ_1, ψ_2, f) and other preference parameters. In the first counterfactual experiment, we artificially reverse the signs of these correlations while keeping the posterior means and variances of the marginal distributions untouched.[22] This change in the correlation pattern leads to notable differences in the quantitative evaluation of risk: the standard deviations of welfare gain and tax revenue are up by 33% and 26%, respectively; and the 95% interval of welfare gain is much wider and includes negative values (i.e., welfare losses). In the second counterfactual experiment, we modify the shape of the marginal distributions of parameters to uniform distributions with the same means and variances, while maintaining the general correlation pattern between parameters.[23] This change again leads to significantly different evaluations about the magnitude of risk than in the benchmark case, with more remarkable effects on the dispersion of tax revenue.

5.1.3 Risk Decomposition

To understand the importance of uncertainty about different aspects of the economy, we conduct two quantitative decompositions of the risk induced by parameter uncertainty. The results are reported in Table 5. In the first decomposition, we categorize parameters into those related to household preferences and those governing the wage process. We find that uncertainty about the wage process plays a quantitatively more prominent role than uncertainty about household preferences, especially so for tax revenue. In the second decomposition, we divide parameters into three groups based on their relatedness to gender. For example, the male (female) labor supply elasticity η_1 (η_2) and wage persistence ρ_1 (ρ_2) are in the male (female) group, whereas the time discount rate δ and the correlation between male and female wage shocks $corr_{v_1,v_2}$ are considered gender-neutral. Our results suggest that uncertainty about male parameters is the most important source of risk, but the contribution of uncertainty about female parameters is also significant.

5.2 Tax Reform with Uncertainty

We have shown in Section 5.1 that the existence of parameter uncertainty translates into substantial ex post variations in the welfare and government budget implications of tax reform. Therefore, it is essential that we incorporate this aspect

[22] In particular, for ψ_1, ψ_2, and f, we replace each parameter value x with $2\mathbb{E}(x) - x$, where $\mathbb{E}(x)$ is the corresponding posterior mean.

[23] In particular, we replace each parameter value x with $\sqrt{3\text{Var}(x)}(2\text{cdf}(x) - 1) + \mathbb{E}(x)$, where $\mathbb{E}(x)$ and $\text{Var}(x)$ are the corresponding posterior mean and variance, and $\text{cdf}(x)$ is the cumulative distribution function of x. Note that the transformation is strictly increasing, and hence the correlation pattern between parameters is largely preserved, albeit not perfectly.

Table 5: Sources of Risk

	By Category		By Gender		
	Preference	Wage	Male	Female	Neutral
Welfare Gain	34%	66%	73%	17%	10%
Tax Revenue	11%	89%	65%	21%	14%

Notes: The risk decomposition is based on the standard deviations of welfare gain and tax revenue induced by uncertainties about different groups of parameters. The results are normalized such that the total contribution equals 100%.

of reality in policy considerations. Specifically, we study the case in which policymakers correctly account for parameter uncertainty in the policymaking process, and contrast the optimal tax reform and its welfare consequences with those in Section 5.1.

Since the tax revenue depends on the parameters, yet policymakers are uncertain about their values, it is impossible to guarantee a balanced government budget with fixed government expenditure. Therefore, we modify the optimal tax problem to allow endogenous public spending financed by tax revenue.[24] To maintain comparability, we calibrate the welfare benefit of public spending such that when parameter uncertainty is ignored (i.e., treating the point estimates as the true values), the optimal policy and welfare gain are exactly the same as those prescribed by the optimal tax problem in Section 5.1. In that sense, the previous optimal tax problem is a special case of the modified problem with a degenerate parameter distribution at the point estimates (i.e., $\Pr(\Theta = \bar{\Theta}) = 1$). The endogenous public spending in our model can be interpreted as a proxy for the provision of public goods by the government, and it allows our analysis to capture an important source of welfare losses due to parameter uncertainty: the ex post inefficient supply of public goods.

5.2.1 Optimal Tax Reform Based on Posterior Distribution

Specifically, we now assume that the welfare of a newborn cohort under tax policy (μ, χ) and parameter state Θ is given by

$$\widetilde{V}(\mu, \chi; \Theta) \equiv V(\mu, \chi; \Theta) + \gamma \frac{[G(\mu, \chi, \Theta)]^{1-\sigma_G} - 1}{1 - \sigma_G},$$

[24] Another possible fix to this issue is to impose the government budget constraint in expectation. We consider this alternative assumption in Section 5.5.

where

$$G(\mu, \chi, \Theta) = \sum_{t=1}^{R} \frac{\text{Tax}_t(\mu, \chi; \Theta)}{(1+r)^{t-1}}.$$

Here γ is the parameter that controls the importance of utility from public spending relative to private utility, and σ_G governs how fast the marginal utility of public spending declines with its level. We set σ_G to the posterior mean of σ, i.e., σ_G's counterpart for private consumption. For γ, we calibrate its value such that when the policymakers take the posterior means as the true parameter values, the optimal tax policy and welfare gain are the same as those in Table 3.[25]

Since we allow uncertainty about preference parameters, policymakers must trade off welfare gains and losses across states in which household preferences differ. There are two complications here. First, indirect utilities of households with heterogeneous preferences are not directly comparable. We address this issue by converting welfare changes of households induced by tax reform to proportional consumption changes based on their own preferences, and then aggregate the consumption changes based on the market price (i.e., the interest rate). That is, for each parameter state, we measure the welfare change from a tax reform by the market value of consumption transfers required to achieve the same welfare effect without the policy change. Second, aggregation of welfare (i.e., consumption) changes over parameter states requires us to take a stand on the policymakers' risk preferences. For the benchmark analysis, we assume policymakers are risk-neutral. We explore implications of alternative assumptions in Section 5.4.

The optimal tax policy problem is then:

$$(\mu^*, \chi^*) = \arg\max_{(\mu,\chi)} \int W(\mu, \chi; \Theta) d\Pi(\Theta), \tag{7}$$

where $\Pi(\cdot)$ is the probability measure for the posterior distribution of parameters, and $W(\mu, \chi; \Theta)$ denotes the welfare gain from adopting tax policy (μ, χ) under parameter state Θ.[26] By construction, the value of the objective function is zero with the status quo tax policy.

[25] The value of γ calibrated in such way is 5.17.

[26] In particular, $W(\mu, \chi; \Theta)$ is given by

$$W(\mu, \chi; \Theta) \equiv \text{CEV}(\mu, \chi, \Theta) \times \left[\sum_{t=1}^{T} \frac{\text{Consumption}_t(\bar{\mu}, \bar{\chi}; \Theta)}{(1+r)^{t-1}}\right],$$

5.2.2 Tax Policy Implications of Parameter Uncertainty

Table 6 reports the optimal tax policy and the welfare gain from the optimal tax reform based on the posterior distribution of parameters we estimated from the data. Taking into account parameter uncertainty in the policymaking process leads to a more progressive optimal income tax, but only modestly so by 1.1 percentage point (0.332 vs. 0.321). The level of the optimal tax policy also falls slightly by 0.5 percentage point (0.106 vs. 0.111). Overall, the optimal tax policy based on the entire posterior distribution is quantitatively close to the one deduced from the point estimates. This is somewhat reassuring in the sense that ignoring parameter uncertainty might not lead to huge mistakes for the choice of the optimal income tax policy.

Table 6: Optimal Tax Reform Based on Posterior Distribution

Tax Progressivity (μ)	Tax Level (χ)	Welfare Gain
0.332	0.106	1.78%

Notes: Welfare gain is reported as a percentage of lifetime consumption under the status quo tax policy and the point estimates of parameters.

However, it would be incorrect to jump to the conclusion that parameter uncertainty is not important for policy discussions about tax reforms. As shown in Table 6, the welfare gain from the optimal tax reform is substantially larger when parameter uncertainty is properly accounted for. The welfare gain is now 1.78% of household lifetime consumption and 20% higher than that deduced from the point estimates. The difference is 0.30% of household consumption or about 43.6 billion dollars per year based on a back-of-the-envelope calculation using the U.S. personal consumption expenditure in 2019. In practice, tax reforms often require policymakers and the government to commit a significant amount of political and economic resources to push the bill through the congress and to implement the new tax system. A vital assessment of any potential tax reform is thus whether the benefits of the tax reform outweigh these fixed costs, and ignoring parameter uncertainty

where $\text{CEV}(\mu, \chi; \Theta)$ is the percentage change of status quo lifetime consumption required to generate the same welfare change as the tax reform, and $\text{Consumption}_t(\bar{\mu}, \bar{\chi}; \Theta)$ is the cohort's consumption at age t under the status quo policy.

biases strongly this comparison in favor of the latter by underestimating the welfare gain from the tax reform. In short, failing to account for parameter uncertainty may falsely negate the worthiness of a tax reform in the first place.

The left panel of Figure 3 shows how the welfare gain from tax reform changes with tax progressivity based on the posterior distribution of parameters and the point estimates. Welfare gain rises with tax progressivity first since the benefit of extra public insurance provided by a more progressive income tax dominates the efficiency cost of distortions to labor supply. The welfare gain reaches a lower peak slightly earlier based on the point estimates, consistent with our previous findings in Table 3 and 6. When tax progressivity grows too high, the relationship between the benefit of public insurance and the efficiency cost of progressive taxation is reversed, and the welfare gain declines with tax progressivity. The panel also shows that if we increase the tax progressivity from the status quo, household welfare always rises more based on the posterior distribution than based on the point estimates. This implies that not only the maximum welfare gain is higher taking into account parameter uncertainty, but for any potential tax reform that increases tax progressivity, the welfare gain is higher than what we would conclude based on the point estimates.

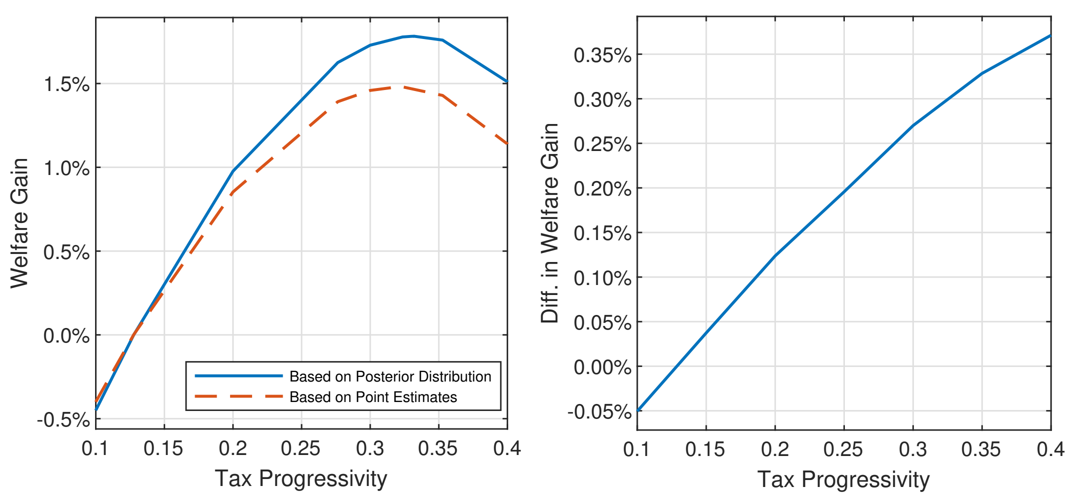

Figure 3: Tax Progressivity and Welfare

Notes: The left panel of this figure shows how the welfare gain relative to the status quo depends on the tax progressivity μ based on the posterior distribution and the point estimates. The right panel plots the difference between the two against tax progressivity.

5.2.3 Understand the Mechanism

Why is the welfare gain from the optimal tax reform larger with parameter uncertainty? Since the optimal policy is similar with or without parameter uncertainty, the majority of the difference in welfare gain is not due to different tax policies. Fixing the tax reform, as indicated by our theoretical analysis in Section 2, the difference in welfare gain is determined by the curvature of welfare gain with respect to the uncertain parameters in a way similar to Jensen's inequality. Because welfare gain is a nonlinear function of the underlying parameters, in general we have

$$\underbrace{\mathbb{E}[W(\mu,\chi;\Theta)]}_{\text{Based on Posterior Distribution}} \neq \underbrace{W(\mu,\chi;\mathbb{E}[\Theta])}_{\text{Based on Point Estimates}} .$$

Whether the left-hand side is greater than the right-hand side (as what we find), however, is a quantitative question. To illustrate the point, in Figure 4, we plot the welfare gain against three parameters. In each plot, we hold the tax policy as in Table 6 and the other parameters at their point estimates. We can see that welfare gain is convex in male wage persistence ρ_1, almost linear in male labor elasticity η_1, and concave in the inverse of consumption elasticity σ. Therefore, the uncertainty about each of the three parameters has positive, zero, or negative contribution to the welfare gain difference. The conclusions are even more obscure when we factor in the interactions and correlations between parameters. For example, although welfare gain appears to be almost linear in male labor elasticity in the center panel, this may no longer be true when other parameters are set at different values, and curvature may still exist along directions of correlations with other parameters.[27] In the end, our quantitative results suggest that welfare gain is overall a "convex" function of the uncertain parameters, and hence the welfare gain from tax reform is higher when parameter uncertainty is considered.

Why is the optimal tax policy more progressive with parameter uncertainty? Our theoretical analysis in Section 2 suggests that it is related to how the curvature of the welfare gain function changes with tax progressivity. The right panel

[27] A simple example is $f(x,y) = xy$. Notice that the function is linear in both x and y individually, but if x and y are perfectly positively correlated, the function is similar to x^2, which is convex. On the other hand, if x and y are perfectly negatively correlated, the function is similar to $-x^2$, which is concave. Again, we see the importance of correlations between parameters for analyses with parameter uncertainty.

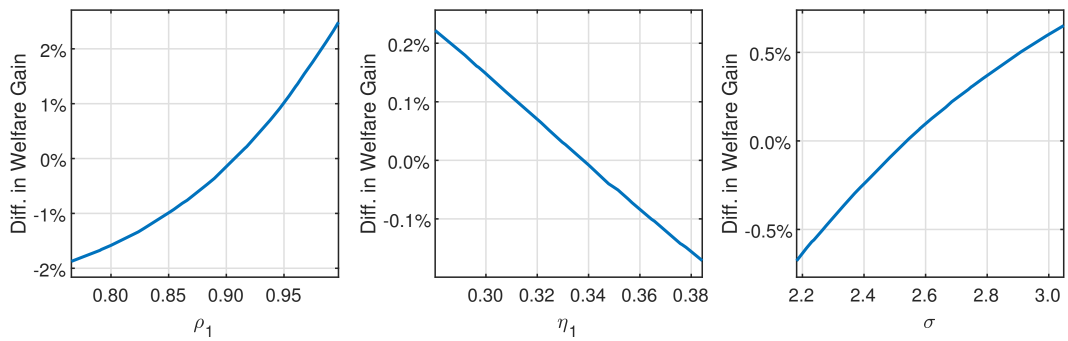

Figure 4: Welfare Gain and Parameters

Notes: This figure plots the welfare gain against the male wage persistence ρ_1 (left panel), male labor elasticitiy η_1 (center panel), and the inverse of consumption elasticity σ (right panel). In each plot, the tax policy is held constant as in Table 6, and the other parameters are set at their point estimates. Welfare gain at the point estimates is normalized to zero in these plots.

of Figure 3 shows that the difference between welfare gain based on the posterior distribution and that based on the point estimates is always increasing in tax progressivity, which implies that the welfare gain function turns overall more "convex" as the tax policy becomes more progressive. Consequently, the welfare gain based on the posterior distribution must reach the peak later as tax progressivity rises, i.e., a more progressive income tax is optimal. Taking the left panel of Figure 4 as an example, note that under the status quo tax policy, the graph would be completely flat since the welfare gain relative to the status quo is always zero no matter what the true value of ρ_1 is. The convexity only appears when we move above the status quo tax progressivity. Again, when there are multiple potentially correlated uncertain parameters, how the overall "convexity" of the welfare gain function changes with tax policy is a quantitative question, and our finding suggests that it is enhanced by tax progressivity.

5.3 Welfare Cost of Uncertainty

When policymakers are uncertain about the true values of key economic parameters, even the ex ante optimal policy such as the one in Section 5.2 is in general not optimal ex post. This leads to welfare losses compared to the first-best scenario in which policymakers have perfect information and can set the tax policy contingent on the state of parameters. We refer to such losses as the welfare cost of uncertainty since they can be eliminated by a signal that reveals the true state of parameters to the policymakers. It is useful to get a sense of how large this welfare cost is since there are potential ways of improving our knowledge about these parameters, yet at various degrees of costs. If the welfare cost of parameter uncertainty turns out to be smaller than the cost of eliminating such uncertainty, it would be rational for the policymakers to remain ignorant. If the opposite is true, more resources should be devoted to activities that can help us achieve better estimates of these parameters.

To quantify the welfare cost of uncertainty, let us first define the parameter-contingent first-best tax policy as

$$(\mu^*(\Theta), \chi^*(\Theta)) = \arg\max_{(\mu,\chi)} W(\mu, \chi; \Theta),$$

where $W(\mu, \chi; \Theta)$ is the same as in Section 5.2, and it represents the welfare gain from a tax reform relative to the status quo. The first-best tax policy maximizes this welfare state-by-state,[28] and hence it is the best the policymakers can do with the perfect information about parameters. The welfare loss from adopting a particular tax policy (μ, χ) relative to the first-best is then

$$\widehat{W}(\mu, \chi; \Theta) \equiv W(\mu^*(\Theta), \chi^*(\Theta); \Theta) - W(\mu, \chi; \Theta). \tag{8}$$

Note that $\widehat{W}(\mu, \chi; \Theta)$ depends on the true parameter state Θ, and it is always non-negative since no policy can induce higher welfare than the first-best policy.

The implementation of the first-best tax policy requires exact knowledge about the true values of parameters, and hence it is infeasible when policymakers face parameter uncertainty. The best the policymakers can do in practice is to choose a non-parameter-contingent tax policy to minimize the welfare loss relative to the first-best. This gives rise to the following optimal tax policy problem:

$$(\mu^*, \chi^*) = \arg\min_{(\mu,\chi)} \int \widehat{W}(\mu, \chi; \Theta) d\Pi(\Theta), \tag{9}$$

where we maintain our previous assumption that policymakers are risk-neutral. It turns out that this problem of minimizing the welfare loss relative to the first-best is equivalent to the problem of maximizing the welfare gain relative to the status quo in Section 5.2. As a result, the ex ante optimal tax policy is still the same as in Table 6.

[28]Here "state" refers to the state of nature that determines the values of parameters.

As Table 7 shows, in the first-best scenario, the expected welfare gain relative to the status quo is 2.17% of household lifetime consumption. The ex ante optimal tax policy can achieve about 80% of this potential welfare gain at 1.78% of lifetime consumption. The remaining 20%, or 0.38% of lifetime consumption, is lost due to the existence of parameter uncertainty. This suggests that the welfare cost of uncertainty is about 55.3 billion dollars per year based on the 2019 U.S. consumption data.

Table 7: Welfare Cost of Parameter Uncertainty

Welfare Gain from		Welfare Cost of
First-Best	Ex Ante Optimal	Parameter Uncertainty
2.17%	1.78%	0.38%

Notes: Welfare is reported as a percentage of lifetime consumption under the status quo tax policy and the point estimates of parameters.

The left panel of Figure 5 plots the distribution of the first-best parameter-contingent tax progressivity. Depending on the parameter state, the first-best tax progressivity varies considerably. However, since policymakers are uncertain about the true parameter values, the ex ante optimal tax progressivity is fixed at $\mu^* = 0.332$, which leads to ex post welfare losses relative to the first-best outcome. The right panel of Figure 5 shows the distribution of this welfare loss, which is right-skewed with a significant chance of exceeding 0.5% of lifetime consumption.

It is worth stressing that the welfare cost of uncertainty we report only captures its distortionary effects on income tax policy. Since there are other government policy decisions that rely on information about household preferences and wage process, the total welfare cost of uncertainty through distortions to all policy decisions is likely much larger. Our finding thus indicates substantial potential welfare gains from activities that improve our knowledge about these aspects of the economy.

5.4 Risk Preferences of Policymakers

In the previous analyses, we have assumed that policymakers are risk-neutral with respect to the variations in welfare induced by parameter uncertainty. In this section, we investigate how our previous findings may be affected by different risk preferences of policymakers.

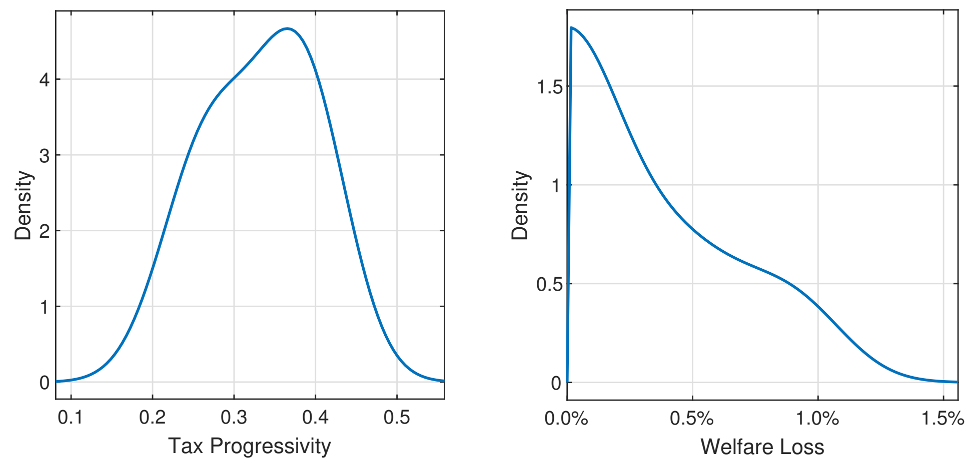

Figure 5: Distributions of the First-Best Policy and Welfare Loss from Uncertainty

Notes: This figure shows the distributions of the first-best tax progressivity (left panel) and the welfare loss relative to the first-best under the ex ante optimal tax policy (right panel).

For this purpose, it is important to differentiate two types of objectives of policymakers: i) maximizing the welfare gain relative to the status quo; and ii) minimizing the welfare loss relative to the first-best. These two types of objectives are equivalent when policymakers are risk-neutral, but it is no longer the case when policymakers are risk-averse. For example, with the first type of objective, maintaining the status quo policy is a risk-free option for policymakers since welfare gain is always zero. However, with the second type of objective, the status quo policy is risky in the sense that the welfare loss of no reform compared to the first-best is uncertain. Put another way, by doing nothing, the policymakers could be giving up very little or a lot in potential welfare gain depending on the true state of parameters. The first type of objective is probably more suitable for "career politicians", whereas the second type is more consistent with the spirit of "social planners", and we consider both specifications in our analysis.

Formally, we introduce the risk-preferences of policymakers by modifying the optimal tax policy problem as follows (with the first type of objective):

$$(\mu^*, \chi^*) = \arg\max_{(\mu,\chi)} \Gamma^{-1}\left(\int \Gamma(W(\mu,\chi;\Theta))d\Pi(\Theta)\right),$$

where

$$\Gamma(X) = \begin{cases} \frac{1-e^{-\alpha X}}{\alpha} & \text{if } \alpha \neq 0, \\ X & \text{if } \alpha = 0. \end{cases}$$

Here $\Gamma(\cdot)$ is the "utility function" of policymakers, which is assumed to take the CARA form. α is the parameter that controls the risk aversion of policymakers. The larger is α, the more the policymakers dislike risk. When $\alpha = 0$, policymakers are risk-neutral, and we are back to the benchmark case. The inverse function $\Gamma^{-1}(\cdot)$ converts the policymakers' "utility" to the certainty equivalent value in consumption. $W(\mu, \chi; \Theta)$ is again the welfare gain from adopting policy (μ, χ) relative to the status quo. For the second type of objective, we only need to replace $W(\mu, \chi; \Theta)$ with the welfare loss relative to the first-best $\widehat{W}(\mu, \chi; \Theta)$ and change the maximization operator to minimization.

Table 8 reports the optimal tax progressivity and the corresponding welfare gain from the optimal tax reform under different degrees of policymakers' risk aversion and for both types of objectives. We consider three degrees of risk aversion corresponding to $\alpha = 0$, 10, and 50. With the first type of objective (maximizing the welfare gain), the risk of tax reform is higher when the new policy is further away from the status quo policy, and hence risk aversion of policymakers pushes the optimal tax policy towards the status quo policy, which means lower progressivity in our model. The welfare gain from the optimal tax reform declines with policymakers' risk aversion because i) the size of the tax reform is smaller, and ii) the "utility" penalty for risk is higher.

Table 8: Optimal Tax Reform and Risk Preferences of Policymakers

	Maximize Welfare Gain		Minimize Welfare Loss	
	Tax Progressivity	Welfare Gain	Tax Progressivity	Welfare Gain
Risk Neutral	0.332	1.78%	0.332	1.78%
Risk Aversion 10	0.291	1.32%	0.333	2.51%
Risk Aversion 50	0.246	0.82%	0.333	3.50%
Tail 10	0.239	0.51%	0.338	3.51%

Notes: "Risk Neutral", "Risk Aversion 10", and "Risk Aversion 50" correspond to the cases with $\alpha = 0$, 10, and 50. "Tail 10" corresponds to the case when the 10th percentile of the welfare distribution is used as the objective. Welfare gain is reported as a percentage of lifetime consumption under the status quo tax policy and the point estimates of parameters.

With the second type of objective (minimizing the welfare loss), the pattern is different. Since the first-best tax policy depends on the true parameter values, a risk-free policy in this case must be parameter-contingent, which is infeasible in the presence of parameter uncertainty. Without a clear risk-free option, it is less straightforward how a higher risk aversion of policymakers would alter the optimal tax policy. Our quantitative results suggest that the optimal tax progressivity is actually quite robust with respect to the degree of policymakers' risk aversion. The welfare gain from the optimal tax reform increases with policymakers' risk aversion mainly because of the decline in welfare under the status quo policy, which is rather risky when measured by the welfare loss relative to the first-best.

Figure 6 shows how welfare measures change with tax progressivity under different degrees of policymakers' risk aversion and for both types of objectives. In addition to reconfirming our findings in Table 8 about the optimal tax policy (i.e., the peaks and troughs of curves), one interesting observation is that with the first type of objective, the welfare gain from any tax reform that increases tax progressivity relative to the status quo is lowered by the policymakers' risk aversion, whereas with the second type of objective, the welfare gain increases with the policymakers' risk aversion. In other words, more risk-averse type-one policymakers have weaker incentives to conduct progressive tax reforms, whereas the opposite is true for type-two policymakers.

We also consider another form of policymakers' risk aversion focusing on the lower tail risk. That is, we assume that the policymakers want to maximize the 10th percentile of the welfare distribution induced by parameter uncertainty.[29] This specification is, in spirit, similar to the max-min criterion from the robust control literature. The corresponding results are reported in Table 8 and Figure 6 under the label "Tail 10". The tax policy implications of this type of risk aversion are quite similar to those of the standard type controlled by α. Compared to the risk-neutral case, with the first type of objective, it reduces the size of the optimal tax reform and the welfare gains from progressive tax reforms. With the second type of objective, the optimal tax progressivity is slightly more progressive, and the welfare gains

[29]For the second type of objective, it would be minimizing the 90th percentile of the welfare loss distribution.

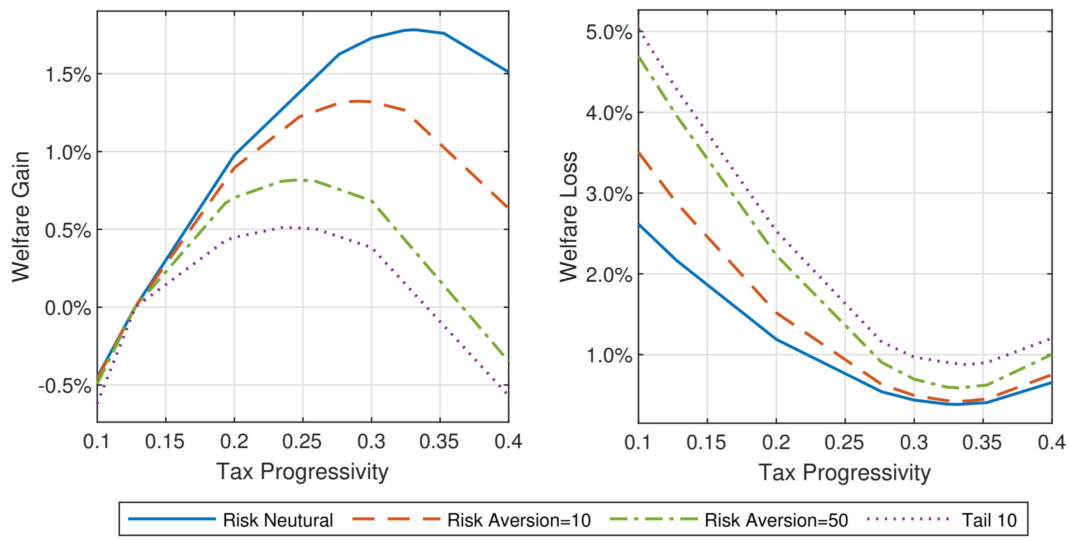

Figure 6: Welfare, Tax Progressivity, and Risk Preferences of Policymakers

Notes: This figure shows how welfare measures change with tax progressivity under different degrees of policymakers' risk aversion and for the first (left panel) and the second (right panel) types of objectives. "Risk Neutral", "Risk Aversion 10", and "Risk Aversion 50" correspond to the cases with $\alpha = 0$, 10, and 50. "Tail 10" corresponds to the case when the 10th percentile of the welfare distribution is used as the objective.

from progressive tax reforms are larger.

5.5 Alternative Government Budget Constraint

As discussed earlier in Section 5.2, when there is parameter uncertainty, no ex ante tax policy can guarantee a balanced government budget with fixed government expenditure for all parameter states. In the benchmark model, we resolve this issue by allowing endogenous public spending that enters utility. An alternative approach is to only impose the government budget constraint in expectation. That is, a tax policy (μ, χ) is considered feasible as long as it satisfies

$$\int \sum_{t=1}^{R} \frac{\text{Tax}_t(\mu, \chi; \Theta)}{(1+r)^{t-1}} d\Pi(\Theta) = \bar{G}. \tag{10}$$

The implicit assumption behind this constraint is that the government can transfer tax revenues across different parameter states and attain full insurance at actuarially fair prices.

We prefer our benchmark model over this alternative specification mainly for two reasons. First, the constraint in (10) implies that government surpluses and deficits induced by parameter uncertainty have no consequences, which is difficult to reconcile with reality. Even if the government can self-insure against such risk through government debts and financial markets, such insurance is hardly perfect or free. In the benchmark model, this type of costs due to fluctuations of government income can be captured by the variations in public spending and the concavity of utility function. Second, with the constraint in (10), the problem for the first-best policy becomes extremely difficult to solve since policymakers also need to optimize over all the possible transfers across parameter states. On the other hand, our benchmark model allows us to conduct all the analyses in a unified framework.

Nevertheless, it is still beneficial to investigate whether our main findings about the implications of uncertainty on tax reforms are robust with respect to this alternative assumption about the government budget constraint. Table 9 reports the results for this alternative specification that are counterparts of those for the benchmark model in Table 6.[30] The optimal tax policy with uncertainty is quite close to that in the benchmark model, and we still find that taking into account uncertainty leads to more progressive optimal income tax and lower tax level than based on the point estimates. There is, however, a notable difference in the welfare gain from the optimal tax reform. With the alternative budget constraint, the welfare gain with parameter uncertainty is even higher (1.92% vs. 1.78% of lifetime consumption in the benchmark model). Since the welfare gain based on the point estimates remains the same with both specifications, this difference only strengthens our previous conclusion that welfare gain from tax reform is underestimated when parameter uncertainty is ignored.

Table 9: Optimal Tax Reform Based on Posterior Distribution
(Balanced Government Budget in Expectation)

Tax Progressivity (μ)	Tax Level (χ)	Welfare Gain
0.338	0.108	1.92%

Notes: Welfare gain is reported as a percentage of lifetime consumption under the status quo tax policy and the point estimates of parameters.

[30] For the alternative specification, we set the fixed government expenditure $\bar{G}$ equal to the tax revenue under the status quo policy, and hence all the feasible tax reforms are revenue-neutral in expectation.

Why is the welfare gain from the optimal tax reform larger than in the benchmark model? Since the optimal tax policy is almost the same, the welfare gain from the tax reform due to changes in private utility must be nearly equal. Therefore, the difference in welfare gain must be due to changes in utility from public spending, which is only present in the benchmark model. Indeed, as tax progressivity rises, the variation of tax revenue across parameter states grows remarkably, and this induces a welfare loss relative to the status quo because the utility function for public spending is concave. Since the alternative specification does not account for this type of costs due to fluctuations in tax revenue, the welfare gain from the tax reform is naturally larger.

6 Conclusions

In this paper, we have examined the income tax policy implications of information frictions in the form of policymakers' uncertainty about household preferences and wage process. We provide general conditions that determine qualitatively the effects of uncertainty on tax policy. To quantify the effects, we employ a limited-information Bayesian approach to measure uncertainty in the form of posterior distribution from the U.S. data through the lens of an incomplete-market life-cycle model of heterogeneous households. Our quantitative analysis suggests that the existence of such uncertainty leads to more progressive optimal income tax and substantially larger welfare gain from tax reform. We have argued that the remarkable difference in the welfare evaluation of tax reform when uncertainty is accounted for could play a vital role in practical policy discussions when fixed political and economic costs of tax reform are present.

The welfare cost of uncertainty through preventing policymakers from implementing the first-best income tax policy alone is about 0.4% of household lifetime consumption. The total welfare cost of uncertainty through distortions to all policy decisions that rely on information about household preferences and wage process is likely much larger. Our finding thus indicates enormous potential welfare gains from activities that can improve our knowledge about these aspects of the economy. We demonstrate that the existence of uncertainty makes tax reform a rather risky decision, and hence the risk preferences of policymakers also have important implications for tax policy.

Our approach of quantifying uncertainty from the data and analyzing its implications on policy design and welfare evaluation is universal and not unique to uncertainty about household preferences and wage process or the income tax policy. Therefore, a natural next step for future research would be to apply the general approach of this paper to uncertainty about other aspects of the economy and related government policies. For computational reasons, we have assumed that policymakers are passive learners in the sense that they take the data they observe as given. It would be interesting to explore what would happen if we introduce active learning such that policymakers can take actions that directly or indirectly affect the information they receive, although the model may have to be simplified substantially along other dimensions to make such extension feasible.

References

Bakış, Ozan, Barış Kaymak, and Markus Poschke (2015). "Transitional Dynamics and the Optimal Progressivity of Income Redistribution". *Review of Economic Dynamics* 18.3, pp. 679–693.

Blundell, Richard, Luigi Pistaferri, and Itay Saporta-Eksten (2016). "Consumption Inequality and Family Labor Supply". *American Economic Review* 106.2, pp. 387–435.

Bénabou, Roland (2002). "Tax and Education Policy in a Heterogeneous-Agent Economy: What Levels of Redistribution Maximize Growth and Efficiency?" *Econometrica* 70.2, pp. 481–517.

Brainard, William C. (1967). "Uncertainty and the Effectiveness of Policy". *The American Economic Review* 57.2, pp. 411–425.

Christiano, Lawrence J., Mathias Trabandt, and Karl Walentin (2010). "Chapter 7 - DSGE Models for Monetary Policy Analysis". Ed. by Benjamin M. Friedman and Michael Woodford. Vol. 3. Handbook of Monetary Economics. Elsevier, pp. 285–367.

Conesa, Juan Carlos, Sagiri Kitao, and Dirk Krueger (2009). "Taxing Capital? Not a Bad Idea after All!" *American Economic Review* 99.1, pp. 25–48.

Conesa, Juan Carlos and Dirk Krueger (2006). "On the Optimal Progressivity of the Income Tax Code". *Journal of Monetary Economics* 53.7, pp. 1425–1450.

Edge, Rochelle M., Thomas Laubach, and John C. Williams (2010). "Welfare-Maximizing Monetary Policy under Parameter Uncertainty". *Journal of Applied Econometrics* 25.1, pp. 129–143.

Fernández-Villaverde, J., J.F. Rubio-Ramírez, and F. Schorfheide (2016). "Chapter 9 - Solution and Estimation Methods for DSGE Models". Ed. by John B. Taylor and Harald Uhlig. Vol. 2. Handbook of Macroeconomics. Elsevier, pp. 527–724.

Hansen, Lars Peter and Thomas J. Sargent (2008). *Robustness*. Princeton University Press.

Heathcote, Jonathan, Kjetil Storesletten, and Giovanni L. Violante (2017). "Optimal Tax Progressivity: An Analytical Framework". *The Quarterly Journal of Economics* 132.4, pp. 1693–1754.

Kim, Jae-Young (2002). "Limited Information Likelihood and Bayesian Analysis". *Journal of Econometrics* 107.1, pp. 175–193.

Krueger, Dirk and Alexander Ludwig (2016). "On the Optimal Provision of Social Insurance: Progressive Taxation versus Education Subsidies in General Equilibrium". *Journal of Monetary Economics* 77, pp. 72–98.

Levin, Andrew T. et al. (2005). "Monetary Policy under Uncertainty in Micro-Founded Macroeconometric Models". *NBER Macroeconomics Annual* 20, pp. 229–287.

Lockwood, Benjamin B., Afras Sial, and Matthew Weinzierl (2020). "Designing, Not Checking, for Policy Robustness: An Example with Optimal Taxation". *Tax Policy and the Economy, Volume 35*. University of Chicago Press, pp. 1–54.

Manski, Charles F. (2014). "Choosing Size of Government Under Ambiguity: Infrastructure Spending and Income Taxation". *The Economic Journal* 124.576, pp. 359–376.

Wu, Chunzan and Dirk Krueger (2021). "Consumption Insurance against Wage Risk: Family Labor Supply and Optimal Progressive Income Taxation". *American Economic Journal: Macroeconomics* 13.1, pp. 79–113.

Online Appendix

When in Doubt, Tax More Progressively: Uncertainty and Progressive Income Taxation

Minsu Chang and Chunzan Wu

A Supplementary Theoretical Results

This section provides additional details about the derivations and claims in Section 2 of the main text.

A.1 The Indirect Utility

The household's problem in Section 2.2 implies that household labor supply is given by

$$H = (1-\mu)^{\frac{1}{1+\eta^{-1}}}.$$

Substitute H in the government budget constraint with the above formula, and we can solve for the tax level χ as a function of tax progressivity μ and parameters (η, σ_z):

$$\chi = 1 - e^{\frac{1}{2}\sigma_z^2\mu(1-\mu)}(1-\mu)^{\frac{\mu}{1+\eta^{-1}}}.$$

Household budget constraint implies $C = (1-\chi)(zH)^{1-\mu}$, and hence household indirect utility is

$$\ln\left((1-\chi)(zH)^{1-\mu}\right) - \frac{H^{1+\eta^{-1}}}{1+\eta^{-1}}.$$

Plug in the formulas for H and χ, and we get the formula for the indirect utility $U(z,\mu)$ in the main text.

A.2 The Signs of Derivatives

For the static model in Section 2.2, when $\mu \in (0,1)$, we have the following signs of derivatives of $\Delta(\mu)$ with respect to the Frisch elasticity η:

$$\Delta_\eta = [\ln(1-\mu)+\mu](1+\eta)^{-2} < 0,$$
$$\Delta_{\eta\mu} = [\frac{-1}{1-\mu}+1](1+\eta)^{-2} < 0,$$
$$\Delta_{\eta\eta} = -2[\ln(1-\mu)+\mu](1+\eta)^{-3} > 0,$$
$$\Delta_{\eta\eta\mu} = -2[\frac{-1}{1-\mu}+1](1+\eta)^{-3} > 0.$$

Since $W(\mu) = e^{\Delta(\mu)} - 1$, we have

$$W_{\eta\eta} = e^{\Delta}[\Delta_{\eta\eta} + (\Delta_\eta)^2],$$
$$W_{\eta\eta\mu} = e^{\Delta}[\Delta_{\eta\eta\mu} + 2\Delta_\eta\Delta_{\eta\mu}] + e^{\Delta}[\Delta_{\eta\eta} + (\Delta_\eta)^2]\Delta_\mu.$$

It is easy to verify that $W_{\eta\eta}(\mu, \bar{\eta}) > 0$. For the sign of $W_{\eta\eta\mu}(\bar{\mu}, \bar{\eta})$, notice that $\Delta_\mu(\bar{\mu}, \bar{\eta}) = 0$ by the definition of $\bar{\mu}$, and hence

$$W_{\eta\eta\mu}(\bar{\mu}, \bar{\eta}) = e^{\Delta}[\Delta_{\eta\eta\mu} + 2\Delta_\eta\Delta_{\eta\mu}] > 0$$

as long as $\bar{\mu} \in (0, 1)$.

Similarly, when $\mu \in (0, 1)$, we have the following signs of derivatives of $\Delta(\mu)$ with respect to the magnitude of idiosyncratic productivity risk σ_z:

$$\Delta_{\sigma_z} = \sigma_z\mu(2 - \mu) > 0,$$
$$\Delta_{\sigma_z\mu} = 2\sigma_z(1 - \mu) > 0,$$
$$\Delta_{\sigma_z\sigma_z} = \mu(2 - \mu) > 0,$$
$$\Delta_{\sigma_z\sigma_z\mu} = 2(1 - \mu) > 0,$$

and

$$W_{\sigma_z\sigma_z} = e^{\Delta}[\Delta_{\sigma_z\sigma_z} + (\Delta_{\sigma_z})^2],$$
$$W_{\sigma_z\sigma_z\mu} = e^{\Delta}[\Delta_{\sigma_z\sigma_z\mu} + 2\Delta_{\sigma_z}\Delta_{\sigma_z\mu}] + e^{\Delta}[\Delta_{\sigma_z\sigma_z} + (\Delta_{\sigma_z})^2]\Delta_\mu.$$

It is easy to verify that $W_{\sigma_z\sigma_z}(\mu, \bar{\sigma}_z) > 0$. For the sign of $W_{\sigma_z\sigma_z\mu}(\bar{\mu}, \bar{\sigma}_z)$, notice again that $\Delta_\mu(\bar{\mu}, \bar{\sigma}_z) = 0$, and hence

$$W_{\sigma_z\sigma_z\mu}(\bar{\mu}, \bar{\sigma}_z) = e^{\Delta}[\Delta_{\sigma_z\sigma_z\mu} + 2\Delta_{\sigma_z}\Delta_{\sigma_z\mu}] > 0$$

as long as $\bar{\mu} \in (0, 1)$.

B Supplementary Quantitative Results

In this section, we provide supplementary results to the quantitative analysis in the main text. Figure 7 presents the estimated male and female log-wage trends over the life cycle. Table 10 reports the list of moment conditions employed in the limited-information Bayesian estimation and their values from the data and from the posterior distribution. Table 11 and 12 report the correlations among preference and wage parameters, respectively.

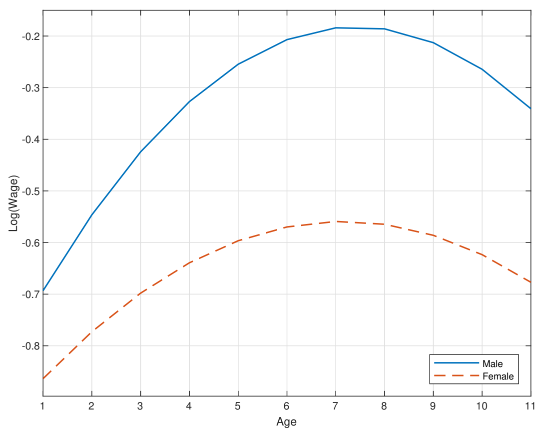

Figure 7: Male and Female Wage Trends

Notes: This figure plots the male (solid line) and female (dashed line) log-wage trends over the life cycle estimated from the data.

Table 10: Moment Conditions

Moment	Data	Model
A. Wage Parameters		
$\mathbb{E}(\Delta w_{1,t}\Delta w_{1,t-1})$	-0.0264	-0.0279
$\mathbb{E}[(\Delta w_{1,t})^2]$	0.1332	0.1337
$\mathbb{E}(\Delta w_{2,t}\Delta w_{2,t-1})$	-0.0352	-0.0351
$\mathbb{E}[(\Delta w_{2,t})^2]$	0.1487	0.1484
$\mathbb{E}(\Delta w_{1,t}\Delta w_{2,t})$	0.0062	0.0061
B. Preference Parameters		
$\mathbb{E}[(\Delta c_t)^2]$	0.0711	0.0471
$\mathbb{E}(\Delta w_{1,t}\Delta y_{1,t})$	0.0988	0.1180
$\mathbb{E}(\Delta w_{2,t}\Delta y_{2,t})$	0.0777	0.1354
$\mathbb{E}[(\Delta y_{1,t})^2]$	0.1275	0.1278
$\mathbb{E}[(\Delta y_{2,t})^2]$	0.1962	0.1558
$\mathbb{E}(\Delta y_{1,t}\Delta y_{1,t-1})$	-0.0264	-0.0180
$\mathbb{E}(\Delta y_{2,t}\Delta y_{2,t-1})$	-0.0211	-0.0298
$\mathbb{E}(\Delta y_{1,t}\Delta c_t)$	0.0078	0.0222
$\mathbb{E}(\Delta y_{2,t}\Delta c_t)$	0.0022	0.0118
$\mathbb{E}(\Delta w_{1,t}\Delta c_t)$	0.0059	0.0222
$\mathbb{E}(\Delta w_{2,t}\Delta c_t)$	-0.0014	0.0136
$\mathbb{E}(\Delta y_{1,t}\Delta y_{2,t})$	-0.0001	-0.0169
$\mathbb{E}(\Delta w_{1,t}\Delta y_{2,t})$	-0.0008	-0.0033
$\mathbb{E}(\Delta w_{2,t}\Delta y_{1,t})$	0.0044	-0.0017
$\mathbb{E}[Y_{1,t}]$	1	0.9276
$\mathbb{E}[Y_{2,t}\|H_{2,t} > 0]$	0.4809	0.4664
$\mathbb{E}[\mathbf{I}(H_{2,t} = 0)]$	0.1539	0.1575
$\mathbb{E}[H_{1,t}]$	1	1.0019
$\mathbb{E}[H_{2,t}\|H_{2,t} > 0]$	0.7058	0.7267
$\mathbb{E}[A_t]$	0.8836	0.8962

Notes: "Model" results are the means of the posterior distributions of moments.

Table 11: Correlations Between Preference Parameters

	σ	η_1	η_2	ψ_1	ψ_2	f	δ
σ	1	0.22	-0.26	0.95	0.47	0.48	-0.51
η_1	0.22	1	-0.21	0.19	0.22	0.22	-0.11
η_2	-0.26	-0.21	1	-0.26	-0.92	-0.88	0.12
ψ_1	0.95	0.19	-0.26	1	0.49	0.52	-0.55
ψ_2	0.47	0.22	-0.92	0.49	1	0.92	-0.24
f	0.48	0.22	-0.88	0.52	0.92	1	-0.29
δ	-0.51	-0.11	0.12	-0.55	-0.24	-0.29	1

Table 12: Correlations Between Wage Parameters

	ρ_1	ρ_2	$\sigma^2_{v_1}$	$\sigma^2_{v_2}$	$corr_{v_1,v_2}$
ρ_1	1	0.04	-0.31	0.00	0.13
ρ_2	0.04	1	-0.01	0.21	-0.03
$\sigma^2_{v_1}$	-0.31	-0.01	1	0.02	0.01
$\sigma^2_{v_2}$	0.00	0.21	0.02	1	-0.08
$corr_{v_1,v_2}$	0.13	-0.03	0.01	-0.08	1