Optimal Taxation with Private Insurance *

Yongsung Chang
Seoul National University

Yena Park
University of Rochester

September 27, 2018

Abstract

We derive a fully *nonlinear* optimal income tax schedule in the presence of private insurance. As in the standard taxation literature without private insurance (e.g., Saez (2001)), the optimal tax formula can still be expressed in terms of sufficient statistics. With private insurance, however, the formula involves additional terms that reflect how the private market interact with public insurance. For example, in our benchmark model–Huggett (1993), the optimal tax formula should also consider pecuniary externalities as well as changes in asset holdings of households. According to our analysis, the difference in optimal tax rates (with and without a private insurance market) can be as large as 10 percentage points.

Keywords: Optimal Taxation, Private Insurance, Pecuniary Externalities, Variational Approach
JEL Classification: E62, H21, D52

*We would like to thank Anmol Bhandari, Narayana Kocherlakota, Dirk Krueger, Aleh Tsyvinski, Louis Kaplow, Yongseok Shin, Kjetil Storesletten, and Maxim Troshkin for helpful comments.

1 Introduction

What is the socially optimal shape of the income tax schedule? This has been one of the classic and central questions in macroeconomics and public finance. Despite significant progress in the literature, surprisingly few studies have investigated the role of private intermediation in the optimal tax system. Understanding the impact of private insurance on the optimal tax is important because, in practice, it is very rare that public insurance can perfectly substitute for a private arrangement. Moreover, even when the government insurance coverage is exactly at the same level that would have been selected by a household from the private market in the absence of government, households may still purchase additional private insurance, if there is moral hazard or a pecuniary externality (see Kaplow (1994)).

In this paper, we study the optimal (fully) nonlinear income tax schedule that highlights the role of the interaction between private and public insurance in determining the optimal tax-and-transfer system. We study a fully nonlinear schedule but focus on a simple class of tax system that is levied on current income only, which allows a direct comparison of our results to those in classic optimal formulas (Saez (2001), Diamond (1998)). The optimal tax formula is derived using a *variational* approach—the tax schedule is optimal, if there is no welfare gain from a small deviation—as in Piketty (1997) and Saez (2001).

The benchmark model we consider for private insurance is incomplete market model with state-noncontingent bond — Huggett (1993). In this economy, consumers can self insure themselves against idiosyncratic income shock through saving and borrowing. The insurance, however, is limited because (i) consumers can only trade non state-contingent bond, and (ii) borrowing is constrained by the exogenous borrowing limit.

We choose Huggett (1993) for several reasons. First, it is one of the most commonly used incomplete market structure in macroeconomic analysis. Second, since its capital market features a pure insurance—households' asset holdings sum to zero in equilibrium, it provides a transparent comparison to those that abstract from private insurance such as Saez (2001). Third, while it assumes a specific (incomplete) market structure, it still allows ready comparisons to the optimal tax formulas from other market structures considered in previous analyses (e.g., Chetty and Saez (2010)). Finally, but not the least, by yielding an analytical expression in the formula, it highlights the effect of pecuniary externalities—emphasized in Dávila, Hong, Krusell, and Ríos-Rull (2012)—on the optimal taxation in an incomplete market economy.

As in Saez (2001), the optimal tax rate can still be expressed in terms of standard statistics—such as the Frisch elasticity of the labor supply and the hazard rate of the income distribution. In the presence of a private insurance market, however, the formula also includes additional terms that reflect the interaction of households' savings with taxes and its welfare effects.

First, the original formula in Saez (2001) needs to be modified to reflect the dispersion of asset holdings. This is likely to lead to a larger inequality in consumption which calls for a stronger redistribution. Second, pecuniary externalities should be considered. As shown in Dávila, Hong, Krusell, and Ríos-Rull (2012), individual's saving decision has externalities because the change in equilibrium interest rate generates an additional redistribution across households. This effect is likely to make the optimal tax schedule less progressive, because providing more progressive tax reform reduces aggregate saving and thus results in a decrease in equilibrium interest rate which makes the asset poor worse off. Third, the formula should also consider the additional welfare effects of some households who are released from the borrowing constraint as a result of tax reform.

Ideally, one would like to express the optimal tax formula in terms of sufficient statistics that can be easily estimated from the data. While we present the generalization of our formula in several directions, we also show that such an attempt is highly challenging for (at least) two reasons. First, the optimal tax formula depends on the welfare effects of the interaction between the private and public insurance. More precisely, the degree to which the envelope theorem can be applied to the response of private intermediation depends on the specifics of the market structure. We illustrate this point in a few well-known market arrangements for private insurance. For example, the optimal formula in Chetty and Saez (2010) is an example where the envelope theorem cannot be applied at all because the savings rate is exogenously (not necessarily at an optimal level) given. On the other hand, the incomplete-markets economy where the interest rate is fixed with no borrowing constraint (Findeisen and Sachs (2017)) is an example where the envelope theorem can still be *fully* applied. Our benchmark model presents an intermediate case where the envelope theorem can be partially applied due to market frictions—which we view as highly common in real world.

Next, we further show that even if the formula can be expressed in terms of sufficient statistic statistics, they are not easy to estimate from the data because because they are not policy invariant. Given these difficulties, we combine the structural and sufficient-statistics methods following the suggestion by Chetty (2009). We obtain the additional statistics from a quantitative general-equilibrium model that is calibrated to resemble some salient features (such a the income and wealth distributions) of the U.S. economy. This allows us to quantify the role of private insurance in determining the optimal tax rate. According to our analysis, the difference in optimal tax rates (with and without a private insurance market) can be as large as more than 10 percentage points. Moreover, these differences in tax rates do not necessarily exhibit the same sign across incomes. For example, the optimal tax rates are higher than those without private markets for the low-income group– mainly because of the increased consumption inequality. The optimal tax rates are lower (than those without a private market) for the middle- to high-income groups—mainly because of pecuniary externalities.

Our paper is most closely related to a literature on optimal labor income taxation using a variational approach, originally pioneered by Piketty (1997) and Saez (2001). In a static model, they express the optimal tax formula in terms of the so-called sufficient statistics (e.g., elasticity of the labor supply and the hazard rate of income), which is obtained by perturbations of a given tax system. This variational approach is a complement to the traditional mechanism-design approach (Mirrlees (1971)) and allows us to understand the key economic forces behind the formula. While this approach has been extended to other contexts such as multi-dimensional screening (Kleven, Kreiner, and Saez (2009)) and dynamic models (Golosov, Tsyvinski, and Werquin (2014), Saez and Stantcheva (2017)), this literature largely abstracts from a private insurance market by assuming that the government is the sole provider of insurance.[1] Chetty and Saez (2010) is an exception that allows for private insurance, but they assume that both private and public insurance are linear, and thus have limited implications for the interactions between the two types

[1]The sufficient statistics approach has been widely used in the taxation literature (e.g., Diamond and Saez (2011), Piketty and Saez (2013a), Piketty, Saez, and Stantcheva (2014), Piketty and Saez (2013b), and Badel and Huggett (2017)).

of insurance.

In the alternative Ramsey approach (Ramsey (1927)), which examines the optimal tax schedule within a class of functional forms, many studies have provided quantitative answers to the optimal amount of redistribution in the presence of self-insurance opportunities (e.g., Aiyagari and McGrattan (1998), Conesa and Krueger (2006), Conesa, Kitao, and Krueger (2009), Heathcote, Storesletten, and Violante (2014), and Bhandari, Evans, Golosov, and Sargent (2016)). However, these studies assume a parametric form for the tax schedule—either affine or log-linear. Moreover, they do not particularly focus on how the introduction of private savings affects the optimal tax schedule. While we allow for a fully nonlinear tax system, our analysis provides a transparent comparison to these papers, as we also compute the optimal tax schedule in a general equilibrium incomplete-markets economy—a workhorse model in macroeconomics. Our quantitative analysis shows that the optimal tax schedule is very different from those commonly assumed—an affine or log-linear tax function—in the literature.[2]

In the *New Dynamic Public Finance* literature, Golosov and Tsyvinski (2007) study optimal taxation in the presence of private insurance under a specific market structure—a competitive insurance industry with private information friction. With this market friction, they also show that internalizing the pecuniary externalities is the role of the government. However, their questions are centered on the welfare gains from government intervention, while we focus on how the optimal tax schedule is affected by the private insurance and understanding of optimal tax formula in general market structure.

Our benchmark analysis is also related to a paper by Findeisen and Sachs (2017), which studies the optimal nonlinear labor income tax and linear capital income tax with self-insurance opportunities. They focus on the interaction between the labor and capital income taxes and as discussed above, their results are an example with full envelope theorem, because they assume fixed interest rate and no borrowing constraint. Our paper has more interaction between public and private insurance by relaxing these assumptions.

Outside the optimal taxation literature, Dávila, Hong, Krusell, and Ríos-Rull (2012)

[2]For example, Heathcote and Tsujiyama (2017) compare three tax systems (affine, log-linear, and Mirrleesian) and find that the optimal tax schedule is close to a log-linear form. Our analysis shows that under a more realistic productivity distribution and private market structure, the optimal tax schedule is highly nonlinear—quite different from log-linear.

analyze the implication of pecuniary externalities in general equilibrium incomplete markets economy. Our paper shows interesting policy implications of the externalities in the context of optimal income taxation. Attanasio and Ríos-Rull (2000) examine the relationship between compulsory public insurance (against aggregate shocks) and private insurance against idiosyncratic shocks. Krueger and Perri (2011) study the crowding-out effect of a progressive income tax on private risk-sharing under limited commitments.

The remainder of the paper is organized as follows. In section 2, we derive the optimal tax formula in Huggett economy. In Section 3, we extend the formula to general private insurance market. Section 4 provides a quantitative analysis. Section 5 studies generalizations in several directions. Section 6 concludes.

2 Optimal Nonlinear Tax Formula with Private Insurance

In this section, we derive optimal nonlinear tax formula in our benchmark economy — Huggett economy. Analyzing optimal tax formula with this special form of market structure is useful itself in understanding the role of private insurance in optimal tax formula, and it will provide implications for an analysis with more general market structure.

In this section, we derive an optimal nonlinear tax formula in our benchmark economy — Huggett-style incomplete market model. As we will show in detail below, the optimal tax formula depends on whether the interaction between private insurance public insurance has welfare effects. Thus, the formula inevitably depends on the structure of private market. As we discuss in the introduction, we chose Huggett (1993) because: (i) it is one of the most commonly used incomplete market structure in macroeconomics, (ii) comparison with Saez (2001) is straightforward due to zero aggregate asset, (iii) it permits a ready comparison with other market structures, (iv) it provides important policy implications of well known pecuniary externalities in incomplete market.

2.1 Restrictions on the Tax System

While we consider a fully nonlinear income tax system without assuming a functional form, we focus on a restrictive class of tax system. The class of tax system we consider is a nonlinear labor income tax with a lump-sum transfer. More precisely, in the benchmark,

(i) we consider a nonlinear labor income tax $T(z)$ where z is current labor income; (ii) the tax is levied on the current period's income only (no history dependency); and (iii) the nonlinear tax function $T(z)$ is age-independent and time invariant. For expositional simplicity, in the benchmark analysis, we assume that there is no capital income taxation, and we relax this assumption later.

We impose these restrictions because they allow for a direct comparison to the static Mirrleesian taxation and Ramsey taxation literature. On one hand, in a static Mirrleesian analysis, the labor income tax depends on income only (not on productivity) because of information frictions. However, in a dynamic environment with stochastic productivity—which we study here—the optimal allocation that solves a mechanism design problem with information frictions (as in the New Dynamic Public Finance literature) will depend on the history of incomes. Moreover, it is well known that a tax system that can implement the constrained-efficient allocation is highly complicated, and thus a direct comparison of tax schedules between a static and a dynamic environment is not straightforward, even without a private market.[3] On the other hand, the Ramsey literature focuses on a tax system with particular functional forms. As in the Ramsey literature, our analysis starts with a simple and implementable tax system, but allows for a fully nonlinear functional form. Thus, our analysis provides a transparent comparison to the theoretical results from Mirrleesian taxation as well as those from Ramsey taxation.

2.2 Economic Environment with Private and Public Insurance

Consider an economy with a continuum of workers with measure one. Workers face uncertainty about their labor productivity in the future. The individual productivity shock x_t follows a Markov process, with transition probability, $f(x_{t+1}|x)$, that has an invariant stationary (cumulative) distribution $F(x)$ whose probability density is $f(x)$.

Individual workers have an identical utility function $\sum_{t=0}^{\infty}\beta^t E_0[U(c_t, l_t)]$, where an instantaneous utility $U(c, l)$ has the following form: $U(c, l) = u(c - v(l))$, where $u(.)$ is concave and increasing in consumption c and $v(.)$ is convex and increasing in labor supply l. We focus on households' preferences that have no wealth effect on the labor supply (the so-called GHH preferences by Greenwood, Hercowitz, and Huffman (1988)). This

[3] Most studies in the new dynamic public finance literature compare the implicit wedge from a dynamic environment to the marginal tax rate from a static one.

assumption is common in the literature because it significantly simplifies the optimal tax formula even without private insurance.[4] The earnings of a worker whose current productivity is x_t are $z(x_t) = x_t l(x_t)$. The cumulative distribution of earnings is denoted by $F_z(z)$ whose density function is $f_z(z)$.

The government provides insurance through a (time-invariant) nonlinear labor income tax and a lump-sum transfer system where the net payment schedule is denoted by $T(z_t)$. The after-tax labor income is $y_t = z_t - T(z_t)$. Workers can also participate in a private market to insure against their income uncertainty. In this benchmark analysis, we consider a Bewley-type incomplete market, where consumers can only self-insure themselves by saving and borrowing via a noncontingent bond (e.g., Huggett (1993)). We also assume that there is an exogenous borrowing limit, $\underline{a}$.

Given prices and government policies, individual consumer solves

$$
\begin{aligned}
V(a_0, x_0) \quad &= \quad \max_{c,l,a} \sum_{t=0}^{\infty} \beta^t \int u(c_t(a_0, x^t) - v(l(x_t))) f(x^t|x_0) dx^t \\
s.t. \qquad & \quad c_t(a_0, x^t) + a_{t+1}(a_0, x^t) = x_t l(x_t) - T(x_t l(x_t)) + (1 + r_t) a_t(a_0, x^{t-1}), \\
& \quad a_{t+1}(a_0, x^t) \geq \underline{a}, \\
& \quad \text{given } a_0, x_0,
\end{aligned}
$$

with solution $\{c_t(a_0, x^t), l(x_t), a_{t+1}(a_0, x^t)\}$. Alternatively, we can represent individual allocation recursively using the individual state, (a_t, x_t), where a_t is current asset holding. Then the allocation will be determined by the policy functions : $h_t^c(a, x), h_t^l(x), h_t^A(a, x)$. We also note that the individual state can be expressed as (a_t, z_t) instead of (a_t, x_t). [5]

The aggregate state of the economy in period t is described by joint measure of asset and productivity, $\mathbf{\Phi}_t(a_t, x_t)$. By abusing the notation, we also denote the distribution of income and asset by $\mathbf{\Phi}_t(a_t, z_t)$, and thus $\mathbf{\Phi}_t(a_t, z(x_t)) = \mathbf{\Phi}(a_t, x_t)$. Let $a_t \in A = [\underline{a}, \bar{a}]$, $x_t \in X = [\underline{x}, \bar{x}]$, and $S = A \times X$. Let $B \in S$ be a Borel set and $\mathscr{M}$ be the set of all finite measures over the measurable space (S, B). An aggregate law of motion of the economy

[4] As we will discuss later, in the presence of private insurance, this assumption is even more crucial for the simplicity of the formula, because we can abstract from the interaction between labor supply response and private insurance.

[5] With no income effects on labor supply, labor income z_t and productivity x_t have a one-to-one relationship and we can use them interchangeably. We also note that even with income effects on labor supply, we can use state variables (a_t, x_t) and (a_t, x_t) interchangeably because $z_t(a_t, x_t) = x_t l_t(a_t, x_t)$ and x_t have a one-to-one relationship given a_t.

is $\mathbf{\Phi}_{t+1} = H_t(\mathbf{\Phi}_t)$, where the function $H_t : \mathscr{M} \to \mathscr{M}$ is defined in the following way. Define a transition function Q by

$$Q(\mathbf{\Phi}_t, a_t, x_t, B; h^A) = \int_{x_{t+1} \in B_x} f(x_{t+1}|x_t) \mathbb{1}_{h^A(a_t, x_t) \in B_a},$$

where $\mathbb{1}$ is the indicator function. Then the distribution of the next period is determined by :

$$\mathbf{\Phi}_{t+1}(B) = \int_S Q(\mathbf{\Phi}_t, a, e, B; h^A) d\mathbf{\Phi}_t.$$

In our benchmark economy (Huggett economy), an equilibrium interest rate, r_t, is determined to clear the asset market:

$$\int a_t(a_t, x_t) d\Phi(a_t, x_t) = 0.$$

That is, in an equilibrium, net asset supplies sum to zero in every period.

In the benchmark, the government evaluates social welfare by:

$$W = \iint G\left(V(a_0, x_0)\right) \phi_0(a_0, x_0)\, da_0\, dx_0,$$

where $G(\cdot)$ is an increasing and concave function that reflects the social preferences for redistribution. The special case is the utilitarian welfare function with $G(V) = V$. Especially, when we compare our optimal tax formula and its simulation with the standard static Saez(2001) formula, we mostly focus on the utilitarian social welfare, since we think this is the most natural comparison between the two economies (with and without private insurance). In the dynamic economy with incomplete market, concave G reflects society's redistribution preferences across the asset as well as across the skills, which makes the comparison with static economy more involved. Later, we will also generalize the social welfare function to consider horizontal equity.

2.3 Deriving an Optimal Formula in a Huggett Economy

In deriving the optimal tax formula, we apply the variational approach (Piketty (1997); Saez (2001)). That is, we consider a perturbation (a small deviation) from a given non-linear tax schedule. If there is no welfare-improving perturbation within the class of tax system, the given tax schedule is optimal. We first derive the tax incidence on individual variables and aggregate variables, then the optimal tax formula will be obtained directly.

2.3.1 Tax Incidence

We start with the tax incidence analysis — the first-order effects of arbitrary tax reforms of a given tax schedule. For a given income tax schedule $T(z)$, the economy we consider converges to a steady state where the distribution of state variables $\Phi(a, x)$ is stationary. We assume that in period 0 the economy starts from that steady state and consider a (revenue-neutral) tax reform in period 0.

Formally, consider an arbitrary tax reform of the initial tax schedule $T(\cdot)$, which can be represented by a continuously differentiable function $\tau(\cdot)$ on $\mathbb{R}_+$. Then, a perturbed tax schedule is $T(\cdot) + \mu\tau(\cdot)$ where $\mu \in \mathbb{R}$ parametrizes the size of the tax reform. As in Golosov, Tsyvinski, and Werquin (2014) and Sachs, Tsyvinski, and Werquin (2016), the first-order effects of this perturbation can be formally represented by the Gateaux derivative in the direction of τ. For example, the incidence on labor supply is

$$dl(x) \equiv \lim_{\mu \to 0} \frac{1}{\mu}[l(x; T + \mu\tau) - l(x; T)],$$

We can define the similar derivatives for the other variables such as indirect utilities of individuals, $V(x_0, a_0)$, government revenue, R_t, and social welfare, W .

From now on, we mostly focus on the *elementary tax reforms*, which can be represented by $\tau(z) = \frac{1}{1-F_z(z^*)}\mathbb{1}\{z \geq z^*\}$ for a given level of income z^*. Under this tax reform, the tax payment of an individual with income above z^* increases by constant amount $\frac{1}{1-F_z(z^*)}$, and the marginal tax rate at the income level z^* is increased by $\frac{1}{1-F_z(z^*)}$ (which is obtained by the marginal perturbation: $\tau'(z) = \frac{1}{1-F_z(z^*)}\delta_{z^*}(z)$). Notice that with tax reform, the increase in government revenue due to mechanical increase in tax payment is equal to \$1. We can focus on this elementary tax reform without loss of generality, because any other perturbations can be expressed as a weighted sum of the elementary tax reforms. See Sachs, Tsyvinski, and Werquin (2016) for further details.[6]

[6] This elementary tax reform is also consistent with the heuristic tax reform in Saez (2001), in which the marginal tax rate $T'(z)$ is increased by $\delta\tau$ on a small income bracket $[z^*, z^*+dz^*]$ and the tax payment $T(z)$ is increased by $\delta\tau \cdot dz^*$ ($= \frac{1}{1-F_z(z^*)}$) in the elementary tax reforms) for the income above z^*.

Incidence of tax reforms on labor supplies

First, we define the elasticity of labor supply with respect to the retention rate $1 - T'(z(x))$. The standard labor supply elasticity with respect to the retention rate along the linear budget constraint is defined as[7]

$$e(x) = \frac{v'(l(x))}{l(x)v''(l(x))},$$

which only takes into account direct effects on labor supply from an exogenous increase in the retention rate. With nonlinear tax system $T(\cdot)$, however, there are additional indirect effects. A change in labor supply $l(x)$ leads to an endogenous change in the marginal tax rate $T'(z(x))$, which in turn results in a further labor supply adjustment. As in Sachs, Tsyvinski, and Werquin (2016), we can define the elasticity of $l(x)$ with respect to retention rate along the nonlinear budget constraint as

$$\epsilon^l_{1-T'}(x) = \frac{dl(x)}{d(1-T')} \cdot \frac{1-T'(xl(x))}{l(x)} = \frac{e(x)}{1+\rho(z(x))e(x)} \tag{1}$$

where $\rho(z(x)) = -\frac{\partial \ln(1-T'(z(x))}{\partial \ln z(x)} = \frac{z(x)T''(z(x))}{1-T'(z(x))}$ denotes the local rate of progressivity of the tax schedule. This elasticity takes into account both direct and indirect effects of change in a retention rate. See appendix for further details.

Using the elasticity along the nonlinear budget, the incidence of a tax reform τ on labor supplies $l(\cdot)$ is represented by

$$dl(x) = -\epsilon^l_{1-T'}(x)\,\frac{\tau'(z(x))}{1-T'(z(x))}\,l(x) = \frac{-\epsilon^l_{1-T'}(x)}{1-F(x^*)} \cdot \frac{\delta_{z*}(z(x))}{1-T'(z(x))}\,l(x).$$

Note that with GHH preferences (with no income effects in labor supply), the increase in the tax level $\tau(\cdot)$ does not have any impact on $dl(x)$. By the definition of the elasticity $\epsilon^l_{1-T'}$, $dl(x)$ represents the change in the labor supply in response to the tax reform, taking into account both exogenous and endogenous (change in $T'(xl(x))$ due to $dl(x)$) changes in the marginal tax rates.

We also remark that $dl(x)$ is constant in all periods. This is the great simplicity we can obtain by assuming no income effects in labor supply. Since the labor supply decision is

[7] With GHH preferences, there is no income effect in labor supply. Thus, the compensated elasticity of labor supply is equal to the uncompensated elasticity of labor supply, and we do not distinguish the notations of the two.

not affected by the wealth level of the individual, $dl(x)$ is time invariant regardless of the private insurance.

Incidence of tax reforms on savings

Individual's saving decision in period t can be represented recursively by the saving policy function $h^A(a_t(a_0, x^{t-1}), x_t)$. Moreover, as long as the mapping $x \mapsto y(x)$ is one to one, we can express the policy function $h^A(a, x)$ as a function of (a, y), so that $h^A(a, x) = h^A(a, y(x))$, where $y(x) = xl(x) - T(xl(x))$ with $l(x)$ that solves $x(1 - T'(xl)) = v'(l)$. We now derive a tax incidence on the saving policy rule $h^A(a, y)$.

Even in the absence of capital income taxation, individual's savings will be changed when the labor income tax schedule is changed.[8] We denote changes in saving policy rule $h^A_t(a, y)$ with respect to current virtual income R_t, state-contingent virtual income in the next period $R_{t+1}(x')$, current interest rate r_t, and future interest rate r_{t+1} by $\epsilon^t_{a',R_t}(a, y) = \frac{\partial h^A_t(a,y)}{\partial R_t}$, $\epsilon^t_{a',R_t}(a, y) = \frac{\partial h^A_t(a,y)}{\partial R_{t+1}(x')}$, $\epsilon^t_{a',r_t}(a, y) = \frac{\partial h^A_t(a,y)}{\partial r_t}$, and $\epsilon^t_{a',r_{t+1}}(a, y) = \frac{\partial h^A_t(a,y)}{\partial r_t+1}$, respectively. In the appendix, we show that the incidence of tax reform on saving policy can be represented by

$$\begin{aligned} dh^A_t(a, y(x)) \quad &= \quad -\epsilon^t_{a',R_t}(a, y) \cdot \tau(z(x)) - \int \epsilon^t_{a',R_{t+1}(x')}(a, y) \cdot \tau(z(x'))dx' \\ &\quad +\epsilon^t_{a',r_t}(a, y)dr_t + \epsilon^t_{a',r_{t+1}}(a, y)dr_{t+1}, \end{aligned} \tag{2}$$

where dr_t and dr_{t+1} represent changes in current interest rate and future interest rate, respectively.[9]

In Huggett economy, aggregate savings sum to zero in equilibrium. Thus, if there is any change in aggregate savings due to tax reform given interest rate, the interest should change to clear the asset market. The incidence on interest rate dr_t is very complicated object, and we do not attempt to solve this analytically. In appendix, we show that the incidence on interest rate dr_t can be expressed in terms of the slope of aggregate supply curve and the incidence on aggregate savings given interest rate.

[8]In a Huggett economy, for example, a change in the labor-income-tax schedule will generate the change in savings through three channels: (i) the change in current income (versus future income), (ii) precautionary savings due to the change in income volatility, and (iii) the general equilibrium effects (the change in equilibrium interest rate).

[9]In an infinite horizon Huggett economy, the responses of saving policy, $\epsilon^t_{a',R_t}(a, y)$, $\epsilon^t_{a',R_t}(a, y)$, $\epsilon^t_{a',r_t}(a, y)$, and $\epsilon^t_{a',r_{t+1}}(a, y)$ are very complicated object because they require solving for an infinite number of unknowns. In Appendix, we show the closed form solution for this in a two-period example.

Incidence of tax reforms on individual welfare

Next, we derive the incidence of a tax reform τ on individual indirect utilities, $V(a_0, x_0)$.

Lemma 1. *The incidence of a tax reform τ of the initial tax schedule T on individual's indirect utilities, $dV(\cdot,\cdot)$, is*

$$\begin{aligned}
dV(a_0,x_0) &= \sum_{t=0}^{\infty}\beta^t \int u'(a_0,x^t)\left[-\tau(z(x_t)) + dr_t\cdot a_t(a_0,x^{t-1})\right] f(x^t|x_0)dx^t \\
&- \sum_{t=0}^{\infty}\beta^t \int \left[u'(a_0,x^t) - \beta(1+r)E[u'(a_0,x^{t+1})|x^t]\right]\cdot dh_t^A(a_t(a_0,x^{t-1}),y(x_t))f(x^t|x_0)dx^t. \quad (3)
\end{aligned}$$

Proof See Appendix. ■

The first term on the right hand side of equation (3), $-\tau(z(x_t))$, is due to a higher tax payment after the tax reform. This decrease in utility is the effect of the standard tax incidence in an economy without private insurance. Note that the welfare effects via $dl(x_t)$ does not show up because of the envelope condition in labor supply.

In the presence of private insurance, however, there are two additional effects on the household's utility. First additional incidence on the utility is the effect from the change in the equilibrium price, dr_t, which arises due to pecuniary externalities: individual households take the market interest rate as given, without considering how their saving decision affects the equilibrium interest rate. The second additional incidence, which is captured by the second integration of equation (3), arises because of the borrowing constraint. If the borrowing constraint is not binding at all, then the Euler equation holds with equality: thus this term is zero. However, for some households who are released from the borrowing constraint as a result of tax reform—i.e., who used to be constrained under the original tax schedule but not any more after the reform, the change in savings, $dh^A(a,y)$, affects welfare.[10]

To understand Equation (3) better, we further decompose the total change in savings of a household with history (a_0, x^t) into:

$$da_{t+1}(a_0,x^t) = dh^A(a_t,y(x_t)) + h_a^A(a_t,y(x_t))\cdot da_t(a_0,x^{t-1}) + h_y^A(a_t,y(x_t))\cdot dy_t(x_t),$$

[10]Technically, the second additional incidence arises because borrowing constrained individual's optimal decision is at the kink of the budget constraint, which does not allow application of the envelope theorem.

where h_a^A and h_y^A are marginal propensity to save out of additional asset holdings and after-tax income, respectively. That is, $dh^A(a,y)$ captures the change in saving policy function for give asset holding, a, and after tax income, y, and this change is the relevant one for the tax incidence. Additional changes in savings due to the change in the state (a,y) for given marginal propensity to save do not have any impact on utility because of the envelope theorem.

We can also express the equation (3) in terms of elasticities, by substituting $dh_t^A(a,y)$ and dr_t with (2) and (15).

Incidence of tax reforms on government revenue and social welfare

The government revenue in period t under the original tax schedule, $R_t = \int T(z(x_t))f(x_t)dx_t$, is constant in the steady state. Thus, the incidence on government revenue, dR_t, directly follows from the change in labor supply $dl(\cdot)$ as:

$$\begin{aligned} dR_t &= \int \tau(z(x))f(x)dx + \int T'(z(x))\Big[-\epsilon^l_{1-T'}(x)) \cdot \frac{\tau'(z(x))}{1-T'(z(x))} z(x)\Big] f(x)\, dx \quad (4) \\ &= \int_{x^*}^{\infty} \frac{f(x)}{1-F(x^*)}\, dx - \frac{T'(z(x^*))}{1-T'(z(x^*))}\epsilon^l_{1-T'}(x^*)\frac{z(x^*)}{z'(x^*)} \cdot \frac{f(x^*)}{1-F(x^*)}, \quad \forall t. \end{aligned}$$

The second equality holds for the elementary tax reform–see Appendix.[11] The change in government revenue $dR_t = dR$ is constant in all periods because the households' labor supply depends on the current productivity only (no wealth effect in labor supply) and the tax system is time invariant.

We now consider the tax incidence on social welfare. Since we consider revenue-neutral tax reforms, any change in government revenue dR will be rebated back to individuals as a lump-sum transfer. Thus, the incidence of a tax reform τ on social welfare dW is:

$$\begin{aligned} dW &= \int G'(V(a_0,x_0)) \sum_{t=0}^{\infty} \beta^t \cdot dR \cdot \Big[\int u'(a_0,x^t) f(x^t \mid x_0) dx^t\Big] \cdot \phi(a_0,x_0) da_0 dx_0 \\ &\quad + \int G'(V(a_0,x_0)) \cdot dV(a_0,x_0)\phi(a_0,x_0)\, da_0\, dx_0. \end{aligned}$$

[11] Note that the step function $1_{z \geq z^*}$ is not differentiable. In appendix, we show that we can nevertheless apply the formula (4) by constructing a sequence of smooth perturbations $\{\tau_n'(z)\}_{n\geq 1}$ which satisfies $\lim_{n\to\infty} \tau_n'(z) = \delta_{z^*(z)}$.

2.3.2 Optimal Tax Formula

The optimal tax schedule maximizes the social welfare subject to the government's budget constraint: $\int T(z(x))f(x)dx = \bar{E}$. Alternatively, given tax schedule, if there is no welfare-improving (revenue-neutral) perturbation within the class of tax system, the given tax schedule is optimal within the class of tax system. By imposing $dW = 0$, we obtain the optimal tax formula.

Proposition 2. *Optimal marginal tax rate at income z^* should satisfy*

$$\frac{T'(z^*)}{1-T'(z^*)} = \frac{1}{\epsilon^l_{1-T'}(z^*)} \cdot \frac{1-F_z(z^*)}{z^* f_z(z^*)} \cdot (1-\beta)\sum_{t=0}^{\infty} \beta^t \left[A_t(z^*) + B_t(z^*) + C_t(z^*)\right], \quad (5)$$

where

$$\begin{aligned}
A_t(z^*) &= \iint_{z^*}^{\infty} (1-g(a,z))\frac{\phi(a,z)}{1-F_z(z^*)}\, dz\, da, \\
B_t(z^*) &= \int g(a,z)\{dr_t \cdot a\}\phi(a,z)\, da\, dz, \\
C_t(z^*) &= -\frac{1}{\lambda}\int \{u'(a,z) - \beta(1+r)E_{z'}\left[u'(a'(a,z),z'))|z\right]\} \cdot dh^A_{t+1}(a,y(z))\phi(a,z)\, da\, dz,
\end{aligned}$$

$\lambda = \int u'(a,z)\phi(a,z)dadz$, *and* $g(a,z) = \frac{u(a,z)}{\lambda}$.

Proof See appendix. ■

Note that the distributions are time invariant because we consider an economy starting from the steady state and the labor supply adjusts instantaneously (no wealth effect). However, private savings may adjust slowly over time, since asset holdings may change slowly. Thus, r_t and $da'_{t+1}(a, y(z))$ can be time varying.

One of the nice features of Saez's (2001) formula is that the optimal tax schedule can be expressed in terms of "sufficient" statistics. According to Saez (2001), the optimal tax rate ($T^{'}$) is decreasing in (i) the Frisch elasticities of the labor supply, e, (ii) the hazard rate of the income distributions, $\frac{z^* f_z(z^*)}{1-F_z(z^*)}$, and (iii) the average social marginal welfare weight of income above z^*, $E[g(a,z)|z \geq z^*]$.[12]

[12] The cost of distortion is proportional to the number of workers ($z^* h(z^*)$) at the margin, while the gain from the tax increase (the increased revenue) is proportional to the fraction of income higher than z^*: $1-F_z(z^*)$. Thus, the optimal tax rate is decreasing in the hazard rate ($\frac{z^* f_z(z^*)}{1-F_z(z^*)}$). The term $1-g(\cdot)$ measures the net benefit of additional lump-sum transfer (lump-sum transfer for all minus extra tax paid by households whose incomes are above z^*) as a result of tax reform. Thus, a larger social welfare weight for households above z^* leads to a lower tax rate.

All three channels remain operative in the new formula (5). However, in the presence of a private insurance market, the standard sufficient statistics are not sufficient to pin down the optimal tax schedule–there are additional terms. The optimal tax schedule also depends on how the private insurance market interacts with public savings, such as dr_t and $da'(a, y(z))$ (which can be expressed in terms of elasticities as shown above).

To simplify the exposition of (5), we rewrite the optimal tax formula (5) with respect to the exogenous productivity distribution by applying change of variables and using the fact that income and productivity densities are related through the equation $f_z(z(x))z'(x) = f(x)$. We also use the following preliminary result.

Lemma 3. *For any regular tax schedule T, the earnings function $z(x)$ is nondecreasing and satisfies:*

$$\frac{z'(x)}{z(x)} = \frac{1+e(x)}{e(x)} \cdot \frac{1}{x} \cdot \epsilon^l_{1-T'}(x)$$

Proof According to Lemma 2 of Saez (2001), $\frac{z'(x)}{z(x)} = \frac{1+e(x)}{x} - \frac{z'(x)}{z(x)}\rho(z(x))e(x)$, where $\rho(z(x)) = \frac{z(x)T''}{1-T'}$. This implies

$$\frac{z'(x)}{z(x)} = \frac{1+e(x)}{x} \cdot \frac{1}{1+\rho(z(x))e(x)} = \frac{1+e(x)}{e(x)} \cdot \frac{\epsilon^l_{1-T'}(x)}{x}.$$

■

Using lemma 3, we express the optimal tax rate in terms of productivity distribution:

$$\begin{aligned} \frac{T'(z(x^*))}{1-T'(z(x^*))} &= \frac{1+e(x)}{e(x)} \cdot \frac{1-F(x^*)}{x^* f(x^*)} \cdot \frac{1}{1-\beta} \\ &\times \sum_{t=0}^{\infty} \beta^t \Big[\iint_{x^*}^{\infty} (1 - \frac{u'(a,x)}{\lambda}) \frac{\phi(a,x)}{1-F(x^*)} \, dx \, da \\ &+ \int \frac{u'(a,x)}{\lambda} \{dr_t \cdot a\} \phi(a,x) \, da \, dx \\ &- \int \big[\frac{u'(a,x)}{\lambda} - \beta(1+r) \int f(x' \mid x) \frac{u'(a'(a,x),x')}{\lambda} dx' \big] \{dh^A_{t+1}(a, y(x))\} \; \phi(a,x) \, da \, dx \Big]. \end{aligned} \tag{6}$$

2.3.3 Role of Incomplete Insurance Market

We now explain the optimal tax formula in detail, along with the comparison to the one without private market in Saez (2001) and Diamond (1998). The optimal tax rate (5) can be decomposed into three terms.

The first term, $A_t(z^*)$, is identical to the original formula in Saez (2001) except that the integration of marginal utility is now over the cross-sectional distribution of assets as well as income. The original Saez effect can be either amplified or mitigated depending on the shape of $\mathbf{\Phi}(z, a)$. Intuitively, incomplete private savings market is likely lead to a larger consumption inequality via more dispersed cross-sectional asset distribution, which in turn implies a larger gain from redistribution–i.e., a higher tax rate is called for. The more incomplete the private insurance markets are, the higher the optimal tax rate is.[13]

The second term, $B_t(z^*)$, reflects whether the tax reform improves pecuniary externalities in an incomplete market. More precisely, this term captures whether the effects of tax reform on saving behavior has positive (or negative) redistribution effects through the change in equilibrium interest rate. As discussed in Dávila, Hong, Krusell, and Ríos-Rull (2012), in an economy with incomplete market where the only available asset is noncontingent bond, a competitive equilibrium is inefficient. Social welfare can be improved by increasing individuals' savings, because a lower interest rate caused by increased savings can improve the welfare of wealth-poor households who has relatively higher marginal utility.[14] The following proposition shows that the sign of the second term in the optimal tax formula (5) is exactly determined by this pecuniary externalities.

Proposition 4. *In the optimal tax formula (5), the sign of the second term $B_t(z^*)$ is determined by*

$$sign(B_t(z^*)) = -sign(dr_t).$$

Proof Since dr_t is constant, $B_t(z^*) = dr_t \cdot \int g(a, z)a\phi(a, z)dadz$. Thus, we only need to show that the sign of the integral in $B_t(z^*)$ is negative. We denote the mean of asset

[13]In Equation (5), a larger consumption inequality increases the dispersion of marginal social welfare weight, $g(a, z)$, without changing the mean $E[g] = 1$, which will in turn decrease $E[1 - g(a, z)|z \geq z^*]$, $\forall z^* > \underline{z}$.

[14]In Dávila, Hong, Krusell, and Ríos-Rull (2012), there is another channel of pecuniary externalities, in which increasing individuals' savings has the opposite welfare implications. Higher wage rate caused by increased savings can generate negative insurance effects by scaling up the stochastic part of the consumer's income. In a Huggett economy, with linear production in labor, wage rate is not affected by the aggregate amount of savings. Thus, we only have redistribution channel of pecuniary externalities.

distribution by $\bar{A}$, and note that in a Huggett economy, $\bar{A} = 0$. We then obtain

$$
\begin{aligned}
& \int g(a,z) a \phi(a,z) da dz \\
= \quad & \int_Z \left[\int_{\underline{a}}^{\bar{A}} \frac{u'(a,z)}{\lambda} [a - \bar{A}] \phi(a|z) da + \int_{\bar{A}}^{\infty} \frac{u'(a,z)}{\lambda} [a - \bar{A}] \phi(a|z) da \right] f_z(z) dz \\
< \quad & \int_Z \left[\frac{u'(\bar{A})}{\lambda} \int_{\underline{a}}^{\bar{A}} [a - \bar{A}] \phi(a|z) da + \int_{\bar{A}}^{\infty} [a - \bar{A}] \phi(a|z) da \right] f_z(z) dz \\
= \quad & \int_Z \frac{u'(\bar{A},x)}{\lambda} \left[E[a|z] - \bar{A} \right] f_z(z) dz \\
< \quad & \frac{u'(\bar{A}, z_m)}{\lambda} \left[\int_{\underline{z}}^{z_m} \left[E[a|z] - \bar{A} \right] f_z(z) dz + \int_{z_m}^{\bar{z}} \left[E[a|z] - \bar{A} \right] f_z(z) dz \right] \\
& \quad \text{(where } z_m \text{ is such that } E[a|z] \geq (<) \bar{A} \text{ for } z \geq (<) z_m) \\
= \quad & \frac{u'(\bar{A}, z_m)}{\lambda} \left[\int E[a|z] f_z(z) dz - \bar{A} \right] = 0.
\end{aligned}
$$

■

Proposition 4 shows that if the elementary tax reform at specific income z^* increases (decreases) interest rate, this has negative (positive) effects on welfare, and thus optimal tax rate at z^* gets lowered (increased). Intuitively, higher interest rate is beneficial to the consumers with positive asset, while it is harmful to the consumers with negative asset. Since the wealth-poor households tend to be consumption poor and have higher marginal utility of consumption, higher interest has negative redistribution effects.

More important, in the context of optimal labor income tax reform, this result implies that pecuniary externalities make the optimal tax system *less progressive*.

A progressive tax reform tends to reduce the precautionary savings motive, which leads to a higher equilibrium interest rate. Higher interest rate has a negative redistribution effects by benefiting the wealth-rich. Thus, the pecuniary externalities makes such a reform less effective in achieving the redistribution of consumption. the optimal tax schedule becomes less progressive to improve pecuniary externalities in an incomplete market.

The third term, $C_t(z^*)$, reflects the change in welfare whose borrowing constraint is no longer binding after the reform. When the tax reform makes some households–who used to be constrained in borrowing—save more (by reducing consumption), this is an additional welfare cost. Next proposition shows that alignment term reduces the optimal tax rate, but in the quantitative analysis, we will show the effect of $C_t(z^*)$ is quantitatively very small because this effect is applied for only small fraction among the borrowing constrained individuals.

Proposition 5. *In the optimal tax formula (5), $C_t(z^*) \leq 0$, for all z^*.*

2.3.4 Alternative Decomposition Comparison with Chetty and Saez (2010)

Chetty and Saez (2010) analyze the optimal tax when both public and private insurance systems are linear–i.e., both tax rate (τ) and saving rate (p) are linear in a static environment.[15] In an economy where private insurance is confined to a linear form, they show the optimal tax rate is: $\frac{\tau}{1-\tau} = -p - \frac{1}{e}(1-\kappa)(1-p)cov(\frac{z}{\bar{z}}, g(z))$, where p is marginal propensity to save, $\kappa = -d\log(1-p)/d\log(1-\tau)$ is crowding elasticity, and $e = d\log(\bar{z})/d\log(1-\tau)$.

To compare our formula to the one in Chetty and Saez (2010) more directly, we can rewrite the formula using two facts: (i) changes in savings sum to zero (aggregate equilibrium condition), and (ii) changes in savings for the history of (a_0, x^t) can be decomposed into: $da_{t+1}(a_0, x^t) = dh^A(a_t, y(x_t)) + h_a^A(a_t, y(x_t)) \cdot da_t(a_0, x^{t-1}) + h_y^A(a_t, y(x_t)) \cdot dy_t(x^t)$, where h_a^A and h_y^A marginal propensity to save out of additional asset holdings and after-tax income, respectively. By adding the changes in aggregate savings, which is equal to zero, we obtain the following formula (see appendix for further detail):

$$
\begin{aligned}
\frac{T'(z^*)}{1-T'(z*)} &= -(1-\beta(1+r)) \int h_y^A(a, y(z))\phi(a \mid z^*)\, da \\
&+ \frac{1}{\epsilon^l_{1-T'}(z^*)} \frac{1-F_z(z^*)}{z^* f_z(z^*)} E\left[1 - g(a,z) | z \geq z^*\right] \\
&+ \frac{1}{\epsilon^l_{1-T'}(z^*)} \frac{1-F_z(z^*)}{z^* f_z(z^*)} (1-\beta) \sum_{t=0}^{\infty} \beta^t \left[\Delta_3 + \Delta_4 + \Delta_5\right],
\end{aligned}
\tag{7}
$$

where

$$
\begin{aligned}
\Delta_{3,t} &= -cov(g(a,z), -dr_t \cdot a) \\
\Delta_{4,t} &= -cov(g(a,z), dh_t^A(a, y(z))) + \beta(1+r)cov(E[g(a'(a,z), z')|z], dh_t^A(a, y(z))) \\
\Delta_{5,t} &= (1-\beta(1+r)) \left[\int h_a^A(a, y(z))\, da d\Phi(a,z) - \int_{z^*}^{\infty} \int h_y^A(a, y(z)) \frac{\phi(a,z)}{1-F_z(z^*)}\, da\, dz\right].
\end{aligned}
$$

[15]More precisely, Chetty and Saez (2010) consider wage compression as a form of private insurance, which is private insurance before paying taxes, but the timing of the private insurance does not change the optimal tax formula as long as tax schedule is committed. That is, in an economy with linear insurance, after tax income is $y = (1-\tau)z + \tau\bar{z}$ and consumption is $c = (1-p)y + p\bar{y}$, where $\bar{z} = E[z]$ and $\bar{y} = E[y]$.

For comparison to our more general formula, the formula in Chetty and Saez (2010) can decomposed into three terms:

$$\frac{\tau}{1-\tau} = -p - \frac{1}{e}cov(\frac{z}{\bar{z}}, g(z)) - \{p + \kappa(1-p)\}\frac{1}{e}cov(\frac{z}{\bar{z}}, g(z)). \tag{8}$$

Each term in our our formula (7) has a counterpart in (8) from Chetty and Saez (2010). The first term in our formula (7) reflects the substitution between private and public insurance, as in the first term, $-p$ in (8), except the adjustment by $(1-\beta(1+r))$. The second term (7) is standard equity-efficiency trade-off term, which is consistent with the second term, $-\frac{1}{e}cov(\frac{z}{\bar{z}}, g(z))$, in (8). The rest of the terms in our formula (7), however, are more involved than the third term in (8), because in Huggett economy, only a part of the change in net savings, $a'(a,x)-(1+r)a$, will have the welfare effects. Despite the optimal response of households, changes in savings might not be optimal from the perspective of government, if there are externalities in savings (shown in Δ_3). The changes in savings by the borrowing-constrained households generate additional welfare effects because the envelope theorem does not hold (show in Δ_4). However, the rest of changes in savings (Δ_5) do not have welfare effects due to the envelope theorem. On the other hand, in Chetty and Saez (2010), exogenous private insurance response is not optimal, thus the total change in private savings $-\{p+r(1-p)\}z$ will have the welfare effects. We discuss this in more detail in section 3.

We learned two important lessens from Chetty and Saez (2010): (i) The formula that ignores the existence of private insurance overstates the optimal tax rate, and (ii) If private insurance does not create moral hazard (in labor supply), the optimal tax formula is identical with and without private insurance. Our analysis shows that these two properties do not necessarily hold in a more general private market.

The optimal tax rate with private savings can be either higher or lower than those without. First, the marginal propensity to savings can be negative, if households are allowed to borrow for consumption smoothing. Second, the presence of private savings can generate a larger consumption inequality, which amplifies the Saez (2001) effects. Third, the pecuniary externalities have either positive or negative sign depending on the change in the equilibrium interest rate as a result of tax reform. All three effects together, we illustrate that the standard optimal formula that ignores the private savings opportunity can either over- or under-state the true optimal tax rate.[16] In fact, our quantitative analysis below shows that there are income regions where the optimal tax rates with private savings are higher than those without.

The appearance of additional terms in the optimal tax formula under private savings market does not necessarily depend on the existence of moral hazard. For example, under an incomplete capital market with self-insurance (like our environment), even when households' labor supply does not depend on wealth, the optimal formula still retains the additional terms that reflect the interaction between private and public insurance. In the next section, we further show that this discrepancy between the formula with and without private insurance crucially depends on whether the envelope theorem can be applied to the response of private intermediation, which in turns depends on the nature of market structure–i.e., frictions in the private insurance market.

3 Implications for Sufficient-Statistics Approach

So far, we have analyzed optimal tax formula with a specific market structure—Huggett (1993). Ideally, one would like to extend the formula to a more general market structure of private insurance market and express the formula in terms of sufficient statistics that can be easily estimated from the data. In this section, we show that such attempt are extremely challenging because of (at least) two reasons. First, the optimal tax rate depends on the welfare effects of the interaction between private and public insurance. Thus, the optimal tax depends on the specific structure of private insurance market. Second, even if we can express the formula in terms of sufficient statistics, they are far more difficult to estimate from the available data compared to the standard ones without private insurance market.

3.1 Optimal Formula with General Incomplete Market

We first analyze how much the optimal tax formula can be generalized to general private insurance market. Our analysis will show that despite a general representation of a wide

[16] In Chetty and Saez (2010) where both the tax rate and private savings are linear, the standard tax formula always overstates the degree of public insurance. This is because of both positive private savings rate and *own* crowding-out effect.

class of private insurance markets, whether the response of private intermediation has welfare effects depends on the specifics (such as incompleteness) of the private insurance market. More precisely, the degree to which the envelope theorem can be applied to the response of private insurance will be different for each structure of the private insurance market. By illustrating how the optimal formula is modified to a few well-known structure of private insurance market, we can better understand the economic insights of previous results in the literature.

We start with a general representation of private insurance markets. Denote the individual state in period t by $(x_t, \mathbf{s}_t)$, where $\mathbf{s}_t = (s_{1,t}, \cdots, s_{M,t}) \in R^M$ is the vector of individual state variables other than individual productivity. For example, if the private insurance market is a Bewley-type incomplete market with non-contingent bond (e.g., Huggett (1993)), we need only one additional state variable: bond holdings a_t: $\mathbf{s}_t = a_t$.

We denote the net payment from private insurance (payment - receipts) by $P_t(x_t, \mathbf{s}_t; T)$. Thus, consumption is $c_t(x_t, \mathbf{s}_t) = z(x_t) - T(z(x_t)) - P_t(x_t, \mathbf{s}_t; T)$. This representation is very general, which can be applied to a wide class of private insurance markets. Note that he private intermediation $P(\cdot; T)$ depends on the government tax/transfer schedule T. From now on, to simplify the notation, we will suppress T in $P(\cdot)$ unless necessary. For expositional simplicity, we assume that the sum of the net payment in the private intermediation is zero:$\int P(\cdot) = 0$,[17] but this can be extended to a more general case. In a Huggett economy with self-insurance only, $P_t(x_t, a_t) = a_{t+1}(x_t, a_t) - (1 + r)a_t$ where r is the rate of return on bond holdings.

We can derive the optimal tax formula using the same perturbation (elementary tax reform). Suppose that in period 0, the economy has converged to a steady state and the steady state distribution is denoted by $\mathbf{\Phi}(x_t, \mathbf{s}_t)$, with its density $\phi(x_t, \mathbf{s}_t)$. The tax reform occurs in period 0. We maintain the assumption of GHH preferences, then the incidence of tax reform on labor supply and government revenue are exactly the same as those in Section 2. However, the incidence of tax reform on private intermediation $P_t(x_t, \mathbf{s}_t)$

[17] That is, we consider a pure insurance market where the aggregate transfer is exactly funded by the aggregate payment in each period.

depends on market structure. The optimal tax formula can be obtained from $dW = 0$:

$$\begin{aligned} \frac{T'(z(x^*))}{1-T'(z(x^*))} &= \left(1+\frac{1}{e(x^*)}\right)\frac{1-F(x^*)}{x^* f(x^*)} \\ &\quad \times (1-\beta)\sum_{t=0}^{\infty}\beta^t \left[\begin{array}{l} \int\int_{x^*}^{\infty}(1-g(x,\mathbf{s}))\frac{\phi(x,\mathbf{s})}{1-F(x^*)}dxd\mathbf{s} \\ -\int\int g(x^t,\mathbf{s}_0)dP_t(x^t,\mathbf{s}_0)f(x^t|x_0)dx^t d\Phi(x_0,\mathbf{s}_0) \end{array}\right], \quad (9) \end{aligned}$$

where $dP_t(x^t, \mathbf{s}_0)$ denotes the incidence of tax reform on private intermediation in period t. Without more information on market structure, we cannot proceed further. From now on, we show that whether the envelope theorem can be applied to the response of private intermediation $dP_t(x^t, \mathbf{s}_0)$ depends on the market structure.

3.1.1 Case1: No Envelope Theorem

If the private intermediation is determined exogenously (not necessarily optimal), the total response of private intermediation will affect individual welfare, and thus none of the second term in the bracket of formula (9) can be ignored when computing optimal tax rate. We first consider an example from Chetty and Saez (2010)—a spot market with a linear payment schedule.[18] With a spot market, we do not need an additional state variable, and the private intermediation can be expressed as follows:

$$P(x) = p \cdot (y(x) - \bar{y}),$$

where p is time-invariant and constant rate of payment to the private intermediaries, $y(x) = xl(x) - T(xl(x))$ is after tax income, and $\bar{y} = E[y(x)]$. As in Chetty and Saez (2010), we consider the case where the rate of payment p does respond to the tax schedule, but it is not necessarily optimal from the perspective of the government (or that of an individual household).

With this private insurance scheme, the incidence of tax reform on the labor supply and the government revenue are exactly as those in Section 2.[19] On the other hand, the

[18] As mentioned above, more precisely, Chetty and Saez (2010) consider a wage compression, but this is essentially identical to a linear-payment spot market.

[19] Individual's first order condition with respect to labor supply is slightly changed: $x(1-T'(xl(x)))(1-p) = v'(l(x))$, but we can easily show that the elasticity of labor supply with respect to retention rate $1-T'$ along the nonlinear budget constraint does not change. See appendix.

incidence on private intermediation yields an analytical expression:

$$\begin{aligned} dP(x) &= -\kappa(x^*)\frac{1-p}{1-T'(z(x^*))}[y(x)-\bar{y}] - p\tau(z(x)) + p(1-T'(z(x)))xdl(x) \\ &\quad + p\epsilon^l_{1-T'}(x^*)\frac{z(x^*)}{z'(x^*)}\frac{f(x^*)}{1-F(x^*)} + p\int \tau(z(x))f(x)dx, \end{aligned}$$

where $\kappa(x) = -\frac{dlog(1-p)}{dlog(1-T'(z(x)))}$ denotes a degree of crowding out.

That is, private intermediation can vary via changes in (i) payment rate, (ii) after tax income, and (iii) transfer from the private intermediaries. Among these changes, only the term associated with the labor supply, $dl(x)$, can be ignored (by applying the envelope theorem) in computing the welfare effects of tax reform. One cannot apply the envelope theorem to all the other terms in $dP(x)$. Then, the second term in the bracket of (9) in this linear spot market is: $-\lambda \cdot p \cdot \epsilon^l_{1-T'}(x^*)\frac{z(x^*)}{z'(x^*)}\frac{f(x^*)}{1-F(x^*)} - \int \frac{u'(x)}{\lambda}\left[-p\left\{\frac{1_{z(x)\geq z^*}}{1-F(x^*)} - 1\right\} - \kappa(x^*)\frac{1-p}{1-T'(z^*)}(y(x)-\bar{y})\right]$. By rearranging the terms, we obtain the optimal formula consistent with that in Chetty and Saez (2010):

$$\frac{T'(z(x^*))}{1-T'(z(x^*))} = -p - \left(1+\frac{1}{e(x^*)}\right)\frac{1-F(x^*)}{x^* f(x^*)}(1-p)\int g(x)\left[\frac{1_{z(x)\geq z^*}}{1-F(x^*)} - 1 - \kappa(x^*)\frac{y(x)-\bar{y}}{1-T'(z(x^*))}\right] f(x)dx.$$

This illustrates that the optimal tax formula in Chetty and Saez (2010) is a special case of private insurance market where the response of private intermediation is highly inefficient from the perspective of the government.

3.1.2 Case 2: Full Envelope Theorem

The other extreme case we consider is the private insurance market where the envelope theorem can be fully applied to the response of private intermediation $dP_t(x^t, \mathbf{s}_0)$. In this case, the second term in the bracket of (9) is zero.

The most straightforward example is the complete market with fully spanned state-contingent assets. Then, the private insurance market can achieve full insurance for any tax schedule. More precisely, with the preferences without income effects on the labor supply, consumption net of the labor supply cost is constant across states under any tax schedule:

$$c(x) - v(l(x)) = \tilde{c}, \quad \forall x, \quad \text{for some constant } \tilde{c}.$$

Thus, $g(x) = \frac{u'(x)}{\lambda} = 1$ for all x, which implies that the response of private intermediation to the tax reform does not have any welfare effects as long as $E[dP(x)] = 0$.[20] With complete market, not only the second term but also the first term in the bracket of optimal tax formula (9) is zero, which implies that optimal tax schedule is zero. That is, if the private market is complete, there is no role for government insurance.

Another example where the envelope theorem can be fully applied is the incomplete market with state noncotingent bond but with exogenously given constant interest rate and natural borrowing limit. For the constant interest rate, we can think of an open economy. In this case, the response of private intermediation $dP(a_0, x^t) = da_{t+1}(a_0, x^t) - (1+r)d_a t(a_0, x^{t-1})$ does not have any welfare effects, because we can apply the envelope theorem to the total change in private intermediation. Recall that in Huggett economy, the incidence of tax reform on private savings had welfare effects through the pecuniary externalities and borrowing constraint. In the absence of both channels, the second term in the bracket of optimal tax formula (9) does not show up. This example shows that even if the private market is incomplete, if the policy tool of the government cannot improve the inefficiency of the incomplete market, we don't need to consider the interaction between the private and public insurance in the optimal tax formula.

Findeisen and Sachs (2017) and Saez and Stantcheva (2017) consider this type of incomplete market with a constant interest rate and natural borrowing limit in their optimal tax analysis. Our analysis illustrate that their optimal tax formula can be viewed as a private insurance market where the envelope theorem is fully applied.

3.1.3 Case 3: Partial Envelope Theorem

The intermediate case is the private insurance market where the envelope theorem is partially applied due to market frictions—which we view as highly common in real world. Individual households' optimal responses (to a tax reform) have no effects on social welfare, if the induced changes in savings neither generate any externalities nor affect the degree of market frictions.

[20] Using the definition of the private intermediation $P(x) = z(x) - T(z(x)) - c(x)$, the private intermediation is represented by $P(x) = z(x) - T(z(x)) - \tilde{c} - v(l(x))$. We can easily show that $E[dP(x)] = 0$.

The benchmark economy we consider in this paper—Huggett (1993)—is a good example of this sort. The private intermediation in Huggett economy is net savings: $P_t(a_0, x^t) = a_{t+1}(a_0, x^t) - (1 + r_t)a_t(a_0, x^{t-1})$. Out of the total change in private intermediation, a change in interest rate has welfare effects, because (i) the change in equilibrium interest rate generates pecuniary externalities and (ii) borrowing constrained households' change in borrowing is not subject to the envelope theorem.

Another example is endogenous incomplete markets with limited commitment (Alvarez and Jermann (2000); Kehoe and Levine (1993)). In this market, households can trade Arrow securities subject to credit lines $\bar{A}_{t+1}(x^t, x_{t+1})$ that are contingent on productivity histories. Consumer's problem is

$$\begin{aligned}
\max_{c_t, a_{t+1}, l_t} \quad & \sum_{t=0}^{\infty} \beta^t \int f(x^t|x_0) u(c_t(a_0, x^t) - v(l_t(x_t))) dx^t \\
s.t. \quad & c_t(a_0, x^t) + \sum_{x_{t+1}} q_t(x^t, x_{t+1}) a_{t+1}(a_0, x^t, x_{t+1}) = x_t l(x_t) - T(x_t l(x_t)) + a + t(a_0, x^t), \quad \forall x^t \\
& a_{t+1}(a_0, x^t, x_{t+1}) \geq \bar{A}_{t+1}(x^t, x_{t+1}), \quad \forall x^t, x_{t+1}.
\end{aligned}$$

The borrowing limits $\{\bar{A}_{t+1}(x^t, x_{t+1})\}$ are endogenously determine to guarantee that individuals have no incentive to default on an allocation at any point in time and any contingency. Following Alvarez and Jermann (2000), the borrowing limits are set as the solvency constraints that are *not too tight*, which satisfies:

$$V_{t+1}(\bar{A}_{t+1}(x^t, x_{t+1}), x^{t+1}) = U_{t+1}^{Aut}(x_{t+1}), \quad \forall (x^t, x_{t+1}),$$

where the continuation utility $V_t(a, x^t)$ denotes the continuation utility of a consumer with history x^t and asset holding a in period t, and the value of autarky is given by

$$\begin{aligned}
U_t^{Aut}(x_t) = \max_{c_s, l_s} \quad & \sum_{s=t}^{\infty} \beta^{s-t} \int f(x^s|x^t) u(c_s(x_s) - v(l_s(x_s))) dx^s \\
s.t. \quad & c_s(x_s) = x_s l_s(x_s) - T(x_s l_s(x_s)).
\end{aligned}$$

We denote the price of risk-free bond by $q_t = \frac{1}{1+r_{t+1}}$ and no arbitrage implies that $q_t(x^t, x_{t+1}) = f(x_{t+1}|x_t t) q_t = \frac{f(x_{t+1}|x_t)}{1+r_{t+1}}$. Private intermediation in this economy is represented by $P_t(a_0, x^t) = q_t \int f(x_{t+1}|x_t) a_{t+1}(a_0, x^t, x_{t+1}) dx_{t+1} - a_t(a_0, x^t)$, and the incidence of tax reform on P_t is $dP_t(a_0, x^t) = dq_t \int f(x_{t+1}|x_t) a_{t+1}(a_0, x^t, x_{t+1}) dx_{t+1} - da_t(a_0, x^t) +$

$q_t \int f(x_{t+1}|x_t) da_{t+1}(a_0, x^t, x_{t+1}) dx_{t+1}$. Then, by rearranging the terms, the second term in the bracket of optimal tax formula (9) writes

$$-dq_t \cdot \int\int f(x^t|x_0) u'(a_0, x^t) \int f(x_{t+1}|x_t) a_{t+1}(a_0, x^t, x_{t+1}) dx_{t+1} dx^t d\Phi(a_0, x_0)$$
$$- \int\int f(x^{t+1}|x_0)[q_t u'(a_0, x^t) - \beta u'(a_0, x^{t+1})] da_{t+1}(a_0, x^t, x_{t+1}) dx^{t+1} d\Phi(a_0, x_0), \qquad (10)$$

where the first term reflects the pecuniary externalities, and the second term represents the welfare effects of change in borrowing for the borrowing constrained consumers.

Although theses terms in the tax formula look similar to those in the Huggett economy, the sign and the source of pecuniary externalities are quite different. With state-contingent assets, the consumption poor in the current period want to borrow from the high productivity state in the future, but this is limited by the endogenous borrowing constraint, and thus the consumption poor's total asset purchase $E[a_{t+1}(a_0, x^t, x_{t+1})]$ is relatively higher which makes sign of the integral positive. In addition, the sign of dq_t is also determined by the degree to which the borrowing constraint is binding in the economy—more binding borrowing constraints lead to lower interest rate and higher q_t. Thus, a progressive tax reform will tighten the endogenous borrowing limit and thus increase price of asset (q_t), which has negative welfare effects in this economy because the poor consumer purchase relatively more asset (higher$E[a_{t+1}(a_0, x^t, x_{t+1})]$). We show this in more detail in the appendix.

3.2 Structural Sufficient-Statistics Approach

A powerful feature of Saez (2001) is that the optimal tax schedule can be expressed in terms of "sufficient" statistics—such as the Frisch elasticity of the labor supply and the cross-sectional distributions of income and marginal utility—which can be estimated or imputed from the data. In principle, we can also express our optimal tax formula in terms of statistics—for example, the marginal propensity to save for $dP_t(\cdot)$. In the presence of a private market, however, it is far more challenging because the formula includes additional statistics that capture the interaction between private and public insurance, which are difficult to obtain from the available data.

Most important, the formula requires the relevant statistics and the distribution of the economy at the *optimal steady state*, which is hard to observe, unless the current tax schedule is already optimal. While the same is true in Saez (2001), given the elasticity of the labor supply, one can still infer the optimal distribution of hours and consumption from an exogenously given distribution of productivity and tax schedule in a static environment. This is no longer the case in a dynamic environment with private savings. We need to know the consumption *rule* and distribution over individual states (e.g., productivity and assets) under the optimal tax. Moreover, these statistics are not policy invariant in general. Thus, it requires out-of-sample predictions. Second, the optimal tax formula involves very detailed micro estimates —e.g., marginal private savings across individual state variables.[21] The formula also requires the elasticity of savings across states, along the transition path of each alternative tax reform.

Faced with these difficulties, we combine the structural and sufficient-statistics methods, following the suggestion by Chetty (2009). We compute the optimal tax schedule using quantitative general equilibrium models calibrated to match some salient features of the U.S. economy. We consider two incomplete markets that are widely used in macroeconomic analysis: Huggett (1993) and Kehoe and Levine (1993).

4 A Quantitative Analysis

4.1 Calibration

We first assume that the individual productivity x can take values from a finite set of N grid points $\{x_1, x_2, \cdots, x_N\}$ and follows a Markov process that has an invariant distribution. We approximate an optimal nonlinear tax and private intermediation with a piecewise-linear over N grid points.[22]

Preferences, Government Expenditure, and Borrowing Constraints

[21]While there are empirical analyses on the marginal propensity to consume (MPC)—e.g., Jappelli and Pistaferri (2014) and Sahm, Shapiro, and Slemrod (2010), these estimates are available for the average or coarsely defined groups of households only.

[22]More precisely, $T(z) = T(0) + \sum_{k=1}^{i-1} T'_k(z_{x_k} - z_{x_{k-1}}) + T'_i(z - z_{x_{i-1}}), \quad z_{x_{i-1}} < z \leq z_{x_i}$, and $\tilde{P}(y) = \tilde{P}(y_{x_0}) + \sum_{k=1}^{i-1} \tilde{P}'_k(y_{x_k} - y_{x_{k-1}}) + \tilde{P}'_i(y - y_{x_{i-1}}), \quad y_{x_{i-1}} < y \leq y_{x_i}$, where $z_{x_0} = 0$ and $y_{x_0} = -T(0)$. Consider a tax reform with an alternative marginal tax rate—suggested by the right-hand side of optimal tax formula (1)—on a grid point T'_i, $i = 1, \cdots, N$. If the tax reform for every grid point no longer improves social welfare—i.e., Equation (1) is satisfied, the optimal tax schedule is found.

The utility function of households is assumed to be a constant relative risk aversion (CRRA):

$$u(c,l) = \frac{(c - v(l))^{1-\sigma}}{1-\sigma}, \quad v(l) = \frac{l^{1+1/e}}{1+1/e},$$

where $\sigma = 1.5$ and the Frisch elasticity of the labor supply (e) is 0.5.[23]

We choose the discount factor (β) so that the rate of return from asset holdings is 4% in the steady state. The government purchase $\bar{E}$ is chosen so that the government expenditure-GDP ratio is 0.188 under the current U.S. income tax schedule (approximated by a log-linear functional form: $T(z) = z - \lambda z^{1-\tau}$) as in Heathcote, Storesletten, and Violante (2014)).[24]

The exogenous borrowing constraint ($\underline{a} = -90.84$) is set so that 10% of households are borrowing constraint. This value is also close to the average annual earnings of households in our model economy under the current U.S. tax schedule. This value is in the range of the credit card limits (between 50% $\sim$ 100% of average annual earnings) in the data. According to Narajabad (2012), based on the 2004 Survey of Consumer Finances data, the mean credit limit of U.S. households is $15,223 measured in 1989 dollars. Finally, we assume that the social welfare function is utilitarian: $G(.)$ is linear. Table 1 summarizes the parameter values in our benchmark case. In Section 4.4 and the appendix below, we perform the sensitivity analysis with respect to different values of σ, e and $\underline{a}$.

Productivity Process

As shown in the optimal tax formula (5), the shape of the income distribution (which is dictated by the stochastic process of a productivity shock and our preferences with no wealth effect in the labor supply) is crucial for the optimal marginal tax schedule. We generate an empirically plausible distribution of productivity as follows. Consider an AR(1) process for log productivity x: $\ln x' = (1-\rho)\mu + \rho \cdot \ln x + \sigma_\epsilon \epsilon'$, where ϵ is distributed normally with mean zero and variance one. The cross-sectional standard deviation of $\ln x$ is $\sigma_x = \frac{\sigma_\epsilon}{\sqrt{1-\rho^2}}$. While this process leads to stationary log-normal distributions of

[23]There is ample evidence of an inter-temporal elasticity of substitution in consumption is smaller than one—e.g., according to the meta analysis by Havránek (2015) based on 169 published articles. The labor supply elasticity ranges between 0.2 and 1—e.g., according to the survey article by Keane and Rogerson (2012).

[24]Given the estimated value for the progressivity, $\tau^{US} = 0.161$ from Heathcote, Storesletten, and Violante (2014), we set λ to match the government expenditure-GDP ratio ($\frac{\bar{E}}{Y}$).

Table 1: Benchmark Parameter Values

Parameter	Description
$\sigma = 1.5$	Relative Risk Aversion
$\beta = 0.9002$	Discount Factor
$e = 0.5$	Frisch Elasticity of Labor Supply
$\underline{a} = -90.84$	Borrowing Constraint
$\frac{\bar{E}}{Y} = 0.181$	Government Expenditure to GDP Ratio under U.S. Tax
$G''(\cdot) = 0$	Utilitarian Social Welfare Function
$\rho_x = 0.92$	Persistence of Log Productivity (before modification)
$\sigma_x = 0.561$	S.D. of Log Productivity
$\frac{xf(x)}{1-F(x)} = 1.6$	Hazard Rate at Top 5% of Wage (Income) Distribution

productivity and earnings, it is well known that the actual distributions of productivity (wages) and earnings have much fatter tails than a log-normal distribution.[25]

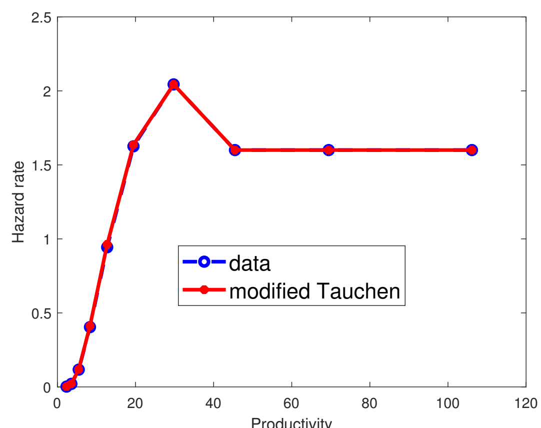
We modify the Markov transition probability matrix to generate a fatter tail as follows. First, we set the persistence of the productivity shock to be $\rho = 0.92$ following Floden and Linde (2001), which is based on PSID wages and largely consistent with other estimates in the literature. We obtain a transition matrix of x in a discrete space using the Tauchen (1986) method, with $N = 10$ states and $(\mu, \sigma_x) = (2.757, 0.5611)$, which are Mankiw, Weinzierl, and Yagan's (2009) estimates from the U.S. wage distribution in 2007. We set the end points of the productivity grid to 3.4 standard deviations of log-normal so that the highest productivity is the top 1% of the productivity distribution in Mankiw, Weinzierl, and Yagan (2009): $(x_1, x_N) = (\exp(\mu - 3.4\sigma_x), \exp(\mu + 3.4\sigma_x))$. Second, in order to generate a fat right tail, we modify the transition matrix of the high productivity grids. More specifically, we increase the transition probability $\pi(x'|x)$ of the highest 3 grids so that the hazard rate of the stationary distribution is $\frac{xf(x)}{1-F(x)} = 1.6$ for the top 5% of productivities.[26] Finally, we also increase the transition probability of the lowest grid,

[25] Saez (2001) and Heathcote, Storesletten, and Violante (2014) estimate the earnings distribution and use tax data to obtain the underlying skill distribution, while Mankiw, Weinzierl, and Yagan (2009) use the wage distribution as a proxy for the productivity distribution.

[26] This hazard rate of 1.6 for the top 5% is slightly smaller than the one reported (which is 2.0) in Mankiw, Weinzierl, and Yagan (2009).

$\pi(x_1|x)$, so that the stationary distribution has a little bit fatter left tail than log normal. This adjustment of the bottom tail of the productivity distribution is designed to take into account disabled workers or those not employed. As Figure 1 shows, the hazard rates of the productivity distribution from our model almost exactly match those in the wage distribution in the data. In Section 4.4, we also study the model economy under a simple log-normal distribution of productivity to examine the impact of fat tails.

Figure 1: Hazard Rates of Wage (Productivity)



Note: The hazard rates are from Mankiw, Weinzierl, and Yagan (2009).

4.2 Indirect Diagnostics

As we described above, difficult-to-estimate statistics in our tax formula call for a numerical simulation of a quantitative model Before we simulate the model economy to compute the optimal tax schedule, we report some key (standard) statistics from our model economy under the *current* U.S. tax schedule because it might still be of interest to compare these statistics to the available estimates in the literature as an indirect diagnostic of our quantitative model.

Distribution of MPC

First, we compare the marginal propensity to consume (MPC) under the current U.S. tax schedule (approximated by the HSV form) to the existing empirical values in the literature. While there are ample empirical studies on the MPC, the MPCs across detailed income and asset levels are not available. Most estimates of MPC are based on the 2001 and 2008 tax rebate policies (e.g., Johnson, Parker, and Souleles (2006) and Sahm, Shapiro, and Slemrod (2010) among others). The estimated MPCs in the literature vary between 0.2 and 0.4. Using a quantile regression method, Misra and Surico (2011) report a wide range of heterogeneity in MPC across households. Jappelli and Pistaferri (2014) also provide detailed MPCs by income and financial assets using the 2010 Italian Survey of Household Income and Wealth.

The average MPC in our benchmark model is 0.88, much higher than the 0.48 reported by Jappelli and Pistaferri (2014) or the 0.33 in Sahm, Shapiro, and Slemrod (2010). This gap is inevitable because the income process is highly persistent in our model—making the MPC close to 1, whereas most empirical estimates are based on idiosyncratic events associated with temporary changes in income, such as tax rebates, which typically imply a small MPC.[27] For this reason, it is not fair to directly compare the levels of MPC between the model and the available estimates. Thus, we rather focus on the *relative* MPCs across different income and asset groups, for which the model is not very far from the data. Table 2 compares the MPCs in our model to those in Jappelli and Pistaferri (2014)—the average and those at the 1st and 5th quintiles (the bottom and top 20%) in the income and asset distributions. The MPCs in the data at the 1st and 5th quintiles are computed using the regression coefficients on dummy variables for the corresponding group (from Table 4 in Jappelli and Pistaferri (2014)). For example, according to Jappelli and Pistaferri (2014), the households in the 1st quintile of the income distribution exhibit MPCs that are 9 to 12% higher than the average MPC of the entire sample, whereas in our model their average MPC is 17% larger than the population average. The households at the 5th quintile show the MPCs that are 11 to 14% smaller than the entire sample average in the data, and they are 14% smaller than the average in our model. Thus, the model generates MPCs that are a little bit more dispersed than those in the data. By assets, the model generates MPCs that are somewhat less dispersed than those in Jappelli and Pistaferri (2014).

<u>Distributions of Income and Assets</u>

Our model is designed to match the income distribution of the U.S. economy fairly well because we calibrate the stochastic process of productivity to mimic the hazard rates

[27]In fact, the average MPC Dupor, Karabarbounis, Kudlyak, and Mehkari (2017)) –which features a Huggett-style incomplete market economy like ours—is 0.22 with respect to a one-time unexpected increase in income, well within the range of empirical estimates.

Table 2: Relative MPC by Income and Assets

	By Income		By Assets	
	Data	Model	Data	Model
Bottom 20%	$+9 \sim +12\%$	$+17\%$	$+22 \sim +25\%$	$+14\%$
Top 20%	$-14 \sim -11\%$	-14%	$-30 \sim -22\%$	-11%

Notes: The numbers represent the average MPC of each group relative to the entire sample mean (0.48 in the data and 0.85 in the model). The data statistics are based on Jappelli and Pistaferri (2014).

of the wage distribution in Mankiw, Weinzierl, and Yagan (2009) (shown in Figure 1). Table 3 shows that the Gini coefficient of earnings in our model is 0.51, not far from those of the U.S. (0.53 – 0.67). The distribution of assets is not necessarily close to that in the data. While the Gini coefficient of wealth in our model is 0.91, even higher than those in the data (0.76-0.86), this comparison is misleading. Given that our model requires zero aggregate savings in equilibrium, there are a large number of households with negative assets. Thus, the Gini is not an appropriate measure and we need a dispersion measure that can accommodate a large fraction of the population with negative values. Instead we report the relative dispersion such as $\frac{a_{80}-a_{20}}{a_{60}-a_{40}}$ where a_{80} is asset holdings at the 80th percentile of the asset distribution. According to Table 3, the model generates an asset distribution whose dispersion is fairly close to that in the data for a wide range of distributions. For example, the relative dispersions in the model are $\frac{a_{90}-a_{10}}{a_{60}-a_{40}} = 4.1$, $\frac{a_{90}-a_{10}}{a_{60}-a_{40}} = 8.9$, and $\frac{a_{95}-a_{05}}{a_{60}-a_{40}} = 18.5$, fairly close to 3.9, 8.6, and 17.3, respectively, in the data. But the dispersion at the tail 2% of the asset distribution, $\frac{a_{99}-a_{01}}{a_{60}-a_{40}}$, is only 39 in the model, much smaller than the 72 in the data. As is well known, this type of incomplete-markets models has difficulty in generating super-rich households. The income and assets are somewhat more strongly correlated in the model (with a correlation coefficient of 0.75) than they are in the data (0.53).

4.3 Optimal Tax Schedule

In our quantitative analysis, for computational convenience, we focus on the optimal tax formula using the so-called "utility-based steady state" approach — optimal tax formulas with the steady-state elasticities. This approach, proposed by Saez and Stantcheva (2018)

Table 3: Distribution of Assets

	Data	Model
Gini (earnings)	0.53-0.67	0.51
Gini (assets)	0.76-0.86	0.91
corr(assets, earnings)	0.53	0.75
$\frac{a_{80}-a_{20}}{a_{60}-a_{40}}$	3.9	4.1
$\frac{a_{90}-a_{10}}{a_{60}-a_{40}}$	8.6	8.9
$\frac{a_{95}-a_{05}}{a_{60}-a_{40}}$	17.3	18.5
$\frac{a_{99}-a_{01}}{a_{60}-a_{40}}$	72.8	39.1

Notes: The data statistics are based on Ríos-Rull and Kuhn (2016) and Chang and Kim (2006). a_{80} denotes asset holdings at the 80th percentile of the asset distribution.

further simplifies the tax formula to:[28]

$$\frac{T'(z^*)}{1-T'(z^*)} = \frac{1}{\epsilon^l_{1-T'}(z^*)} \cdot \frac{1-F_z(z^*)}{z^* f_z(z^*)} \cdot [A(z^*) + B(z^*) + C(z^*)] \tag{11}$$

where

$$\begin{aligned} A(z^*) &= \iint_{z^*}^{\infty} (1-g(a,z)) \frac{\phi(a,z)}{1-F_z(z^*)} \, dz \, da \\ B(z^*) &= \int g(a,z)\{dr \cdot a\} \phi(a,z) \, da \, dz, \\ C(z^*) &= -\frac{1}{\lambda} \int \{u'(a,z) - \beta(1+r)E_{z'}\left[u'(a'(a,z),z'))|z\right]\} \cdot da'(a,y(z))\phi(a,z) \, da \, dz, \end{aligned}$$

$\lambda = \int u'(a,z)\phi(a,z)da\,dz$, and $g(a,z) = \frac{u(a,z)}{\lambda}$.

Using this formula, we compute the optimal tax schedule quantitatively. We start with a given T'. We compute the competitive equilibrium for given T' and the equilibrium with the tax reform to obtain the other statistics—such as dr, $da'(a,y(z))$, $\phi(a,z))$, and $g(a,z)$. We then use the formula to compute the new vector of T'. More precisely, we

[28] There are two possible interpretations for this formula: (i) an approximation of the optimal tax formula where the equilibrium prices and households' asset adjustment during transition are replaced by the change in prices and assets between new and old steady states: $dr_t = dr$ and $a'_{t+1} = da'$ where dr and da' are the change in the equilibrium interest rate and assets holdings, respectively, from the old to new steady state. (ii) this is the formula under the utility-based steady state approach proposed by Saez and Stantcheva (2018). In our context, this is equivalent to steady state welfare maximization but deliberately ignoring the effect of da_0 — change in asset holding in the initial period of new steady state — on individual welfare. Intuitively this means that the government does not consider the change in individual's initial budget in the first period of new steady state, because this change in initial budget is at the cost of individual's past sacrificed consumption in the transition period.

use the formula with respect to the exogenous productivity distribution— formula (6). We repeat the algorithm until the vector converges to a fixed point. This algorithm is a modification of the one used in Brewer, Saez, and Shephard (2010).

Figures 2 and 3 show the optimal marginal tax schedule across productivity and income, respectively, with and without a private insurance market. We normalize the units of quantities in our model so that the average productivity (wage) is \$20 and the average labor income is \$40,000 (comparable to those in 2015 in the U.S.). Without a private insurance market (dotted line), the optimal marginal tax schedule exhibits a well-known U-shape as in the standard Mirrleesian taxation literature (Diamond (1998), Saez (2001)). High marginal tax rates at the very low income levels indicate that net transfers to low-income households should quickly phase out. As seen in Figure 1, the hazard rate of productivity sharply increases, implying that the cost of distorting the labor supply quickly increases (relative to the benefit): the optimal marginal tax rate should start decreasing with income. As income increases, the marginal social welfare weight gradually diminishes—which eventually becomes a dominant factor and results in a higher marginal tax at the high-income group.

Figure 2: Optimal Marginal Tax by Productivity

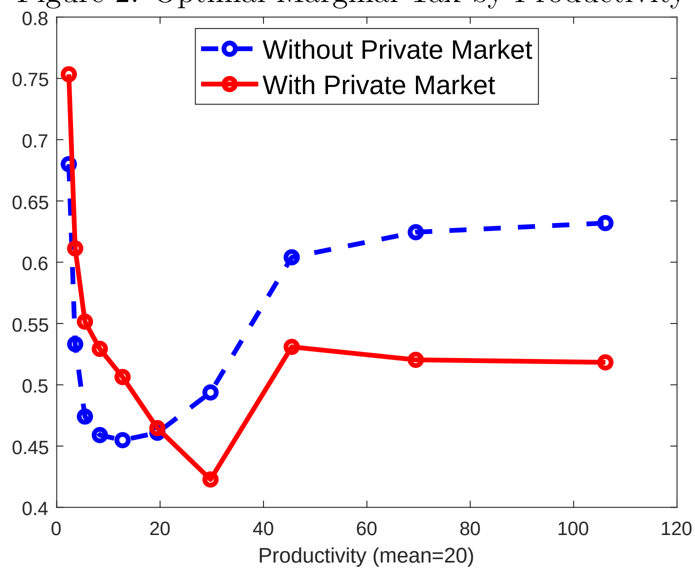
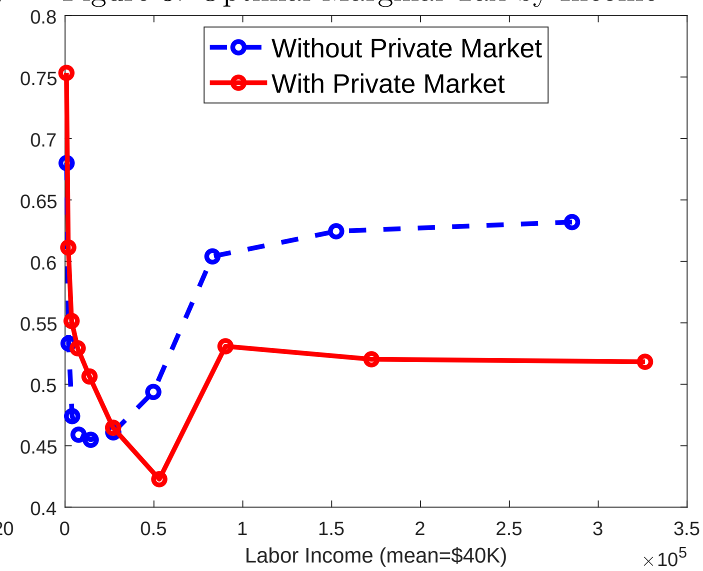


Figure 3: Optimal Marginal Tax by Income



While the same driving forces are operative in an economy with a private insurance market, there are additional factors that make the optimal tax schedule different from that without a private insurance market. Looking at Figures 2 and 3 again, the optimal tax rates in the presence of private insurance (solid line) are higher than those without a private market (dotted line) at the low-income group (wage rates less than \$20). For the middle- and high-income groups (wage rates above \$20), the optimal tax rates are lower than those without private insurance.

We now examine the factors that account for the difference in optimal tax rates with and without private insurance in detail. Comparing our optimal tax formula (11) to that of Saez (2001), the difference between the two formulas consists of three components. The first term in the bracket, $A(z^*)$, is similar to Saez (2001) except that now the distribution of marginal utility of consumption depends on assets as well as income. The second term, $B(z^*)$, reflects the pecuniary externality whose importance in incomplete markets is well-explained by Davilla, Hong, Krusell, and Rios-Rull (2012). That is, a decrease in the equilibrium interest rate (which will make the rich less richer) due to tax reform justifies less progressive tax schedule. The third term, $C(z^*)$, captures the effect of borrowing constraint as explained above. To see the importance of each component, we provide the decomposition of the difference between our optimal tax rate and that of Saez (2001) as:

$$\frac{T'(z^*)}{1-T'(z^*)} - \frac{T'_{Saez}(z^*)}{1-T'_{Saez}(z^*)} = \frac{1}{\epsilon^l_{1-T'}(z^*)} \cdot \frac{1-F_z(z^*)}{z^* f_z(z^*)} \cdot \Big[A(z^*) - A^{Saez}(z^*) + B(z^*) + C(z^*)\Big], \quad (12)$$

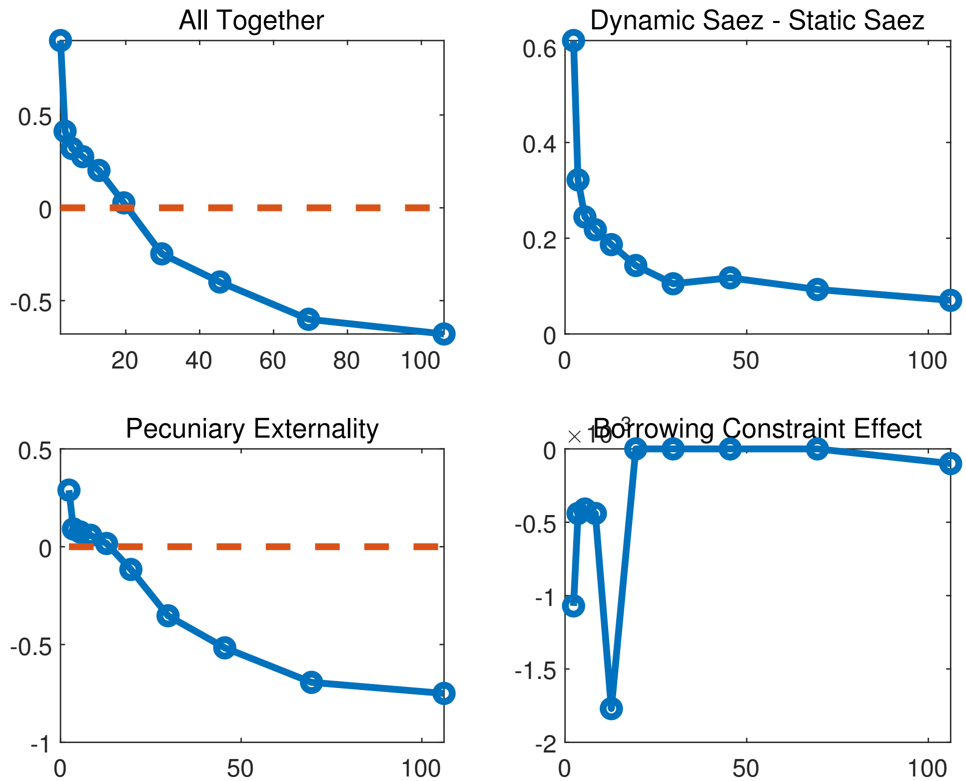
where

$$A(z^*) - A^{Saez}(z^*) = \iint_{z^*}^{\infty} (1-g(a,z)) \frac{\phi(a,z)}{1-F_z(z^*)}\, dz\, da - \int_{z^*}^{\infty} (1-g^{Saez}(z)) \frac{f_z(z)}{1-F_z(z^*)}\, dz.$$

Figure 4 plots each of these three components. The first figure show the difference from all three terms together. The second figure shows $\frac{1}{\epsilon^l_{1-T'}(z^*)} \frac{1-F_z(z^*)}{z^* f_z(z^*)} [A(z^*) - A^{Saez}(z^*)]$, labeled as "Dynamic Saez - Static Saez." The third shows the effect of pecuniary externality, $\frac{1}{\epsilon^l_{1-T'}(z^*)} \frac{1-F_z(z^*)}{z^* f_z(z^*)} B(z^*)$, and the last shows the effect due to the borrowing constraint, $\frac{1}{\epsilon^l_{1-T'}(z^*)} \frac{1-F_z(z^*)}{z^* f_z(z^*)} C(z^*)$.

The distribution of consumption becomes more dispersed in an economy with private savings (due to a skewed asset distribution). Thus, the original Saez formula is amplified and the first term (Dynamic Saez - Static Saez) is always positive, making the optimal tax rate higher. The second term, which represents the effect of pecuniary externality, however, depends on the sign of the equilibrium interest rate change as explained above. This effect is positive for the low-income bracket (less than wage rate of \$15) but becomes negative as income increases, and the negative effect become quite large at the top income.

Figure 4: Decomposition of the Difference in $\frac{T'}{1-T'}$ with and without private insurance



A progressive tax reform of increasing marginal tax rate at the high income bracket reduces the precautionary motive of savings and results in a high equilibrium interest rate (which makes the rich richer) which dampens the redistribution effect of a tax reform. As this effect starts dominating the dynamic Saez - Static Saez term (at wage rates around \$20), the optimal tax rates become lower than those without private insurance market. Finally, the third term, the effect of borrowing constraint, is always negative but quantitatively negligible. All three terms together, the optimal tax rates are higher for the low income group (wages below \$20) but lower for higher income group. In sum, the difference in tax rates with and without a private insurance market is quantitatively important, as the difference between the two can be more than 10 percentage points.

As we have shown in section 2.3.4, an alternative way to decompose the difference between our formula and Saez formula is using the approach in Chetty and Saez (2010), which shows the role of marginal propensity to save more visibly. By comparing the formula (7) with steady state elasticities to that of Saez (2001), the difference between the two formulas consists of three components.

$$\frac{T'(z^*)}{1-T'(z*)} - \frac{T'_{Saez}(z^*)}{1-T'_{Saez}(z^*)} = \Omega_1 + \Omega_2 + \Omega_3, \tag{13}$$

$$\begin{aligned}
&where \\
\Omega_1 = \quad & -(1-\beta(1+r)) \int h_y^A(a, y(z))\phi(a \mid z^*)\, da \\
\Omega_2 = \quad & \frac{1}{\epsilon^l_{1-T'}(z^*)} \frac{1-F_z(z^*)}{z^* f_z(z^*)} \left\{ E\left[1-g(a,z)|z \geq z^*\right] - A^{Saez}(z^*) \right\} \\
\Omega_3 = \quad & \frac{1}{\epsilon^l_{1-T'}(z^*)} \frac{1-F_z(z^*)}{z^* f_z(z^*)} \left[\Delta_3 + \Delta_4 + \Delta_5\right],
\end{aligned}$$

Figure 5: Alternative Decomposition of the Difference in $\frac{T'}{1-T'}$ with and without private insurance

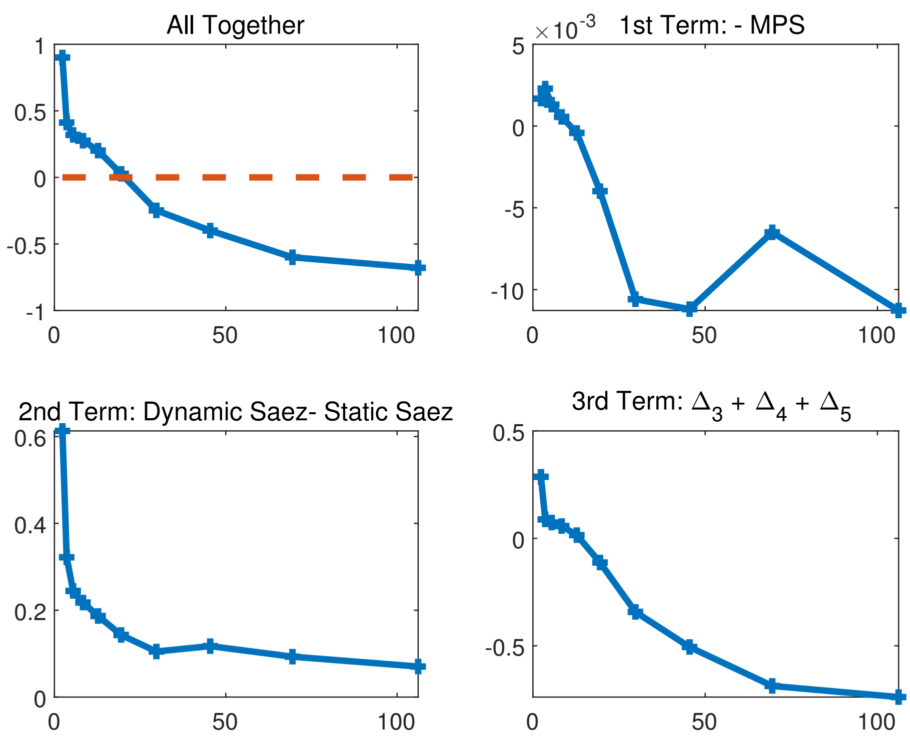
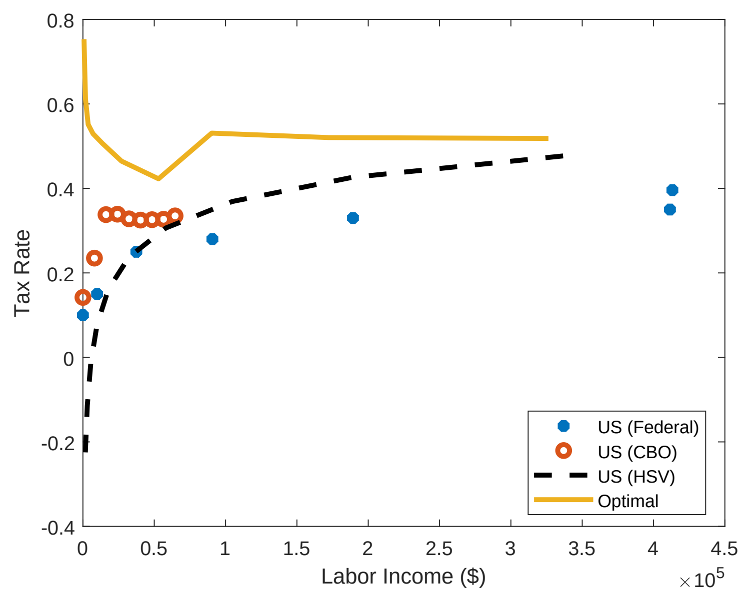


Figure 5 plots each of three components: $\Omega_1, \Omega_2, \Omega_3$. The first term Ω_1 (the second figure) shows the role of marginal propensity to save (MPS), which reflects the substitution between private and public insurance. We can see that MPS has very small effects quantitatively in optimal tax formula, which is because of dynamic adjustment factor $(1-\beta(1+r))$ in Ω_1. Different from spot insurance market in Chetty and Saez (2010), in a dynamic Huggett economy, current saving is asset holding in the future, and thus the direct effect of MPS (substitution effect) is quantitatively small, while MPS can have indirect effect through the distribution of asset.

Finally, we compare our optimal tax schedule to the current U.S. income tax rates. Figure 6 compares the optimal marginal tax rates implied by our model (solid line) to the current U.S. income tax schedule approximated by the HSV functional form (dotted line). First, the optimal marginal tax rates are higher than the current rates in the U.S. for all income groups.[29] However, for the top-income group (i.e., individual income ranges above $250K), the current tax rates are not so far from optimal. This result is very different from those without private insurance seen in Figure 3, where the optimal tax rates are much higher than the current ones.

Figure 6: Marginal Tax Rate: Current vs. Optimal



Note: "US (Federal)" reflects the statutory federal income tax rates for singles in 2015. "US (CBO)" shows the median of effective marginal tax rates for low- and moderate-income workers (single parent with one child) in 2016 published by the CBO.

[29]Heathcote and Tsujiyama (2017) find that the optimal tax schedule is close to a log-linear form. There are at least two important differences between the results of Heathcote and Tsujiyama (2017) and ours. First, we match the exact shape of the hazard rate of the productivity distribution in the data, while they approximate the productivity distribution by an exponentially modified Gaussian. Second, they assume a complete separation between perfectly insurable and noninsurable productivity shocks, whereas we assume a partial insurance market.

4.4 Comparative Statics

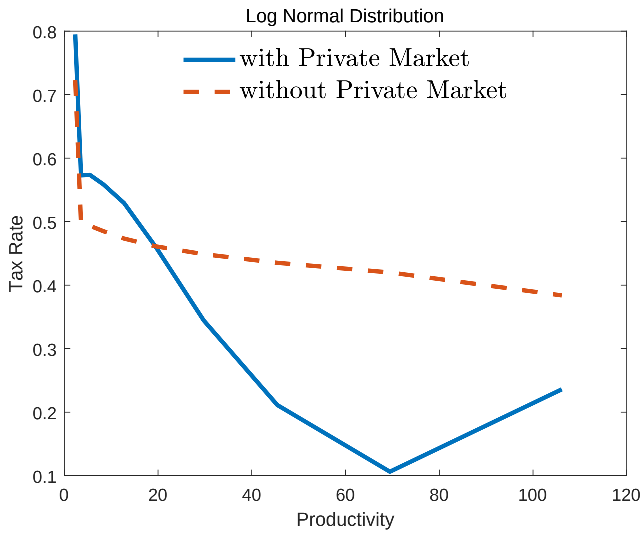
In this section, we investigate how the optimal tax schedule changes with respect to different specifications on (i) the right tail of the income distribution (log-normal rather than Pareto) and (ii) the persistence of productivity shocks. For each alternative specification, we find a new value for the time discount factor β to clear the private insurance market at the given interest rate $r = 4\%$ under the current U.S. tax schedule (approximated by a log-linear form as in HSV). Simultaneously, we recalibrate the exogenous borrowing limit $\underline{a}$ so that about 10% of households are credit-constrained in the steady state. In the appendix, we also carry out the sensitivity analysis with respect to other parameters of the model economy such as relative risk aversion, the Frisch elasticity, and the borrowing constraint. The results are consistent with our economic priors. The optimal tax rates are higher when (i) households are more risk averse, (ii) the labor supply is inelastic, and the role of private insurance is less significant when the borrowing constraint is tighter.

4.4.1 Log-normal Distribution of Income: Effects of Fat Tails

In the benchmark analysis, we have modified the transition probability (from the discretized log-normal distribution) to match the fat tail in the income (and wage) distribution in the data. To examine the role of the fat tail, we compute the optimal tax under a pure log-normal productivity process without modification. The hazard rate $\frac{xf(x)}{1-F(x)}$ of the log-normal distribution monotonically increases. This results in the monotonically decreasing tax rate without a private insurance market in Figure 7. The pattern prevails in the presence of private insurance, suggesting that the fat tail is crucial for the U-shaped optimal marginal tax schedule.

In addition, due to very small fraction of workers at the top income, increasing marginal tax rate at the top has relatively small effects on pecuniary externalities, because relatively few individuals change savings behaviors and thus response of interest rate will be relatively small for the tax reform of increasing marginal tax rate at the top. Thus, the effects of pecuniary externalities becomes weaker at the top.

Figure 7: Log-normal Distribution of Productivity



4.4.2 Persistence of Productivity Shock

Note that the persistence of the productivity shock, ρ, does not appear in the optimal tax formula because we restrict our tax system to be noncontingent on history and the government maximizes steady-state welfare only in the benchmark analysis. However, the persistence of shocks affects households' savings pattern and, as a result, the optimal tax rate in the presence of a private insurance market. We examine the model with $\rho = 0.8$ (lower persistence). We recalibrate the standard deviation to the innovation σ_ϵ to obtain the same standard deviation of log productivity, $\sigma_x = 0.561$, in the benchmark. We also modify the transition probability matrix at both ends of the productivity distribution to match the hazard rates in the data, as we did in our benchmark case.

Figure 8 shows that the optimal tax rates under $\rho = 0.8$ are smaller than those in the benchmark model except for the very low-income (productivity) group. The decomposition of the difference is shown in Figure 9. As the persistence of productivity falls, given the same size of overall income risk, households would like to save more when productivity is high (i.e., a high productivity does not last long). This generates a larger dispersion of asset distribution. This has two effects on the optimal tax rates. (i) It magnifies the Saez term which makes the "Dynamic vs. Static Saez" term larger especially at the low income bracket. (ii) But, at the same time, it leads to a bigger negative pecuniary externality, and this leads to even less progressive optimal tax schedule to improve pecuniary externalities.

Figure 8: Optimal Tax with Private Market

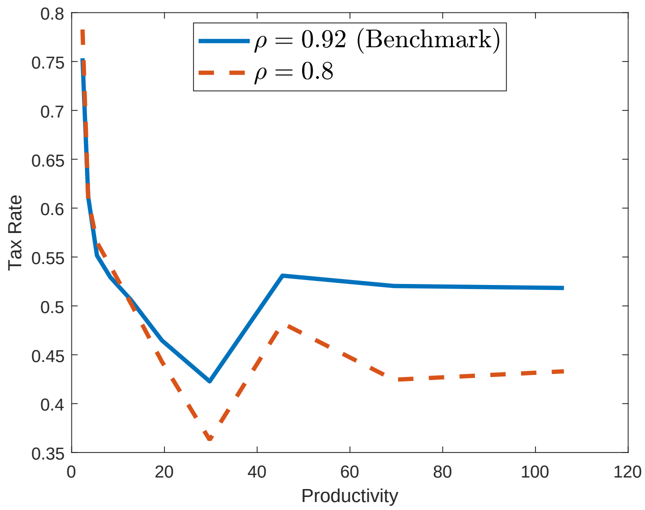
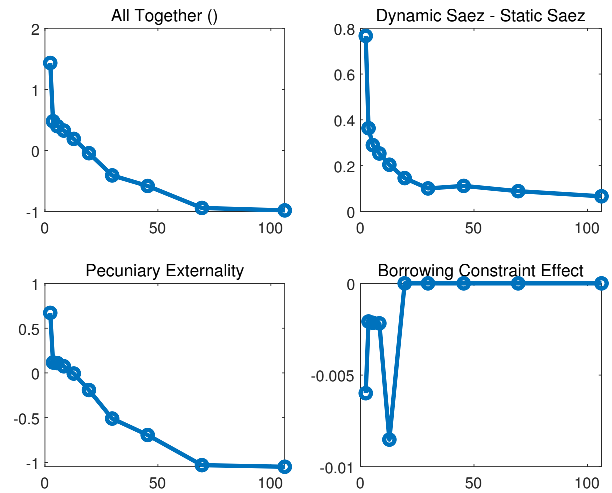


Figure 9: Decomposition ($\rho = 0.8$)



5 Generalizations

In this section we show how the results of Section 2 can be generalized to a more realistic environment. We provide brief descriptions of these extensions only here. The details are gathered in Appendix.

5.1 Aiyagari Economy with Physical Capital

The benchmark endowment economy can be extended to the one with physical production capital (Aiyagari (1994)). In this economy, individual's labor income depends on wage rate w_t, and the budget constraint of the consumer is

$$c(a_0, x^t) + a_{t+1}(a_0, x^t) = w_t x_t l(x_t) - T(w_t x_t l(x_t)) + (1 + r_t) a_t(a_0, x^{t-1}).$$

Production is governed by a constant returns to scale production function $F(K_t, L_t)$, and the aggregate amount of factors of production and their prices are determined by $K_t = \int a_t d\Phi(a_t, x_t)$, $L_t = \int x_t l(x_t) d\Phi(a_t, x_t) = \int x_t l(x_t) f(x_t) dx_t$, $r_t = F_K(K_t, L_t) - \delta$, and $w_t = F_L(K_t, L_t)$, respectively.

In this economy, even with GHH preferences (no income effects), the incidence of tax reform on labor supply is more involved because the responses of labor supply and savings to a tax reform change the wage rate, which in turn further affects the labor supply. The wage rate is affected by the change in all individuals' labor supplies and savings:

$$dw_t = -\alpha_L^w(K_t, L_t) \cdot w_t \int \frac{x_t' dl(x_t')}{L_t} f(x_t') dx_t' + \alpha_K^w(K_t, L_t) \cdot w_t \int \int f(x^{t-1}|x_0) \frac{da_t(a_0, x^{t-1})}{K} dx^{t-1} d\Phi(a_0, x_0),$$

where $\alpha_L^w(K, L) = -\frac{dlogw}{dlogK} = -F_{LL}\frac{L}{F_L}$ denotes the elasticity of wage rate with respect to aggregate labor, and $\alpha_K^w = \frac{dlogw}{dlogK} = F_{LK}\frac{K}{F_L}$ denotes the elasticity of wage rate with respect to aggregate capital. Then, the incidence of tax reform on labor supply should solve the following integral equation:

$$\begin{aligned} dl_t(x_t) &= -\epsilon_{1-T'}^l(x_t) \frac{\tau'(wx_t l(x_t))}{1 - T'(wx_T l(x_t))} l(x_t) \\ &+ \epsilon_w^l(x_t) \left[-\alpha_L^w(K, L) \int \frac{x_t' dl(x_t')}{L_t} f(x_t') dx_t' + \alpha_K^w(K, L) \int \int f(x^{t-1}|x_0) \frac{da_t(a_0, x^{t-1})}{K} dx^{t-1} d\Phi(a_0, x_0) \right] l(x_t), \end{aligned}$$

where $\epsilon_w^l(x) = \frac{dlogl(x)}{dlogw} = \epsilon_{1-T'}^l(x)(1 - \rho(wxl(x)))$ denotes the elasticity of labor supply with respect to wage rate along the nonlinear budget constraint.

Similarly, the interest rate responds to the changes in all individual's labor supply and savings

$$dr_t = -\alpha_K^r(K_t, L_t) \cdot r_t \int \int f(x^{t-1}|x_0) \frac{da_t(a_0, x^{t-1})}{K} dx^{t-1} d\Phi(a_0, x_0) + \alpha_L^r(K_t, L_t) \cdot r_t \int \frac{x_t' dl(x_t')}{L_t} f(x_t') dx_t',$$

where $\alpha_K^r(K, L) = -\frac{dlogr}{dlogK} = -F_{KK}\frac{K}{F_K - \delta}$ and $\alpha_L^r = \frac{dlogr}{dlogL} = F_{KL}\frac{L}{F_K - \delta}$. In turn, this change in interest rate affects saving as in Huggett economy.

We can obtain the optimal tax formula by imposing $dW = 0$.

Proposition 6. *In Aiyagari economy, optimal marginal tax rate at income z^* should satisfy*

$$\frac{T'(z^*)}{1 - T'(z^*)} = \frac{1}{\epsilon_{1-T'}^l(z^*)} \cdot \frac{1 - F_z(z^*)}{z^* f_z(z^*)} \cdot (1 - \beta) \sum_{t=0}^{\infty} \beta^t \left[A_t(z^*) + B_t(z^*) + C_t(z^*) + D_t(z^*) \right], \quad (14)$$

where

$$\begin{aligned} A_t(z^*) &= \iint_{z^*}^{\infty} (1 - g(a, z)) \frac{\phi(a, z)}{1 - F_z(z^*)} \, dz \, da, \\ B_t(z^*) &= \int g(a, z)[dw_t \cdot (1 - T'(wz))z + dr_t \cdot a] \phi(a, z) \, da \, dz, \\ C_t(z^*) &= -\frac{1}{\lambda} \int \{u'(a, z) - \beta(1 + r) E_{z'} \left[u'(a'(a, z), z')) | z \right] \} \cdot da_{t+1}'(a, y(z)) \phi(a, z) \, da \, dz, \\ D_t(z^*) &= dw_t \cdot \int T'(wz(x)) xl(x) f(x) dx. \end{aligned}$$

The formula (14) is different from (5) in our benchmark Huggett economy in two aspects. First, there is an additional term $D_t(z^*)$ which captures additional fiscal externalities due to changes in equilibrium wage rate. Second, the pecuniary externalities term $B_t(z^*)$ captures two channels of pecuniary externalities as:

$$B_t(z^*) = dr_t K \left[\int g(a,z) \left(\frac{a}{K} - 1 \right) d\Phi(a,z) - \int g(a,z) \left(\frac{(1 - T'(wz))z}{L} - 1 \right) d\Phi(a,z) \right].$$

Compared to the term $B_t(z^*)$ in formula (5), there is one more term in the bracket, which captures the insurance channel of pecuniary externalities. If the tax reform decreases the wage rate (and increases the interest rate), it generates positive welfare effects because the stochastic component of household's income–labor earnings— is scaled down. Thus, the sign of the additional pecuniary externalities term is the opposite of that the pecuniary externalities in redistribution of capital.

According to Dávila, Hong, Krusell, and Ríos-Rull (2012), under a realistic calibration, the pecuniary externalities through redistribution dominates the pecuniary externalities through insurance. Thus, the sign of $B_t(z^*)$ term is not likely to change. Moreover, the sign of the additional term in $D_t(z^*)$ is the opposite to that of dr_t, which also makes the optimal tax schedule less progressive.

5.2 Allowing Capital Income Taxes

In the benchmark, we assume that there is no capital income taxation for expositional simplicity. In this section, we show that the formula and intuitions we derived in the benchmark economy carry over to the economy with capital income taxation, as long as the capital income taxation cannot fully complete the market. If the capital income tax function is sophisticated enough (e.g., fully nonlinear history dependent tax), the economy goes back to the complete market case, and there is no need to use labor income taxation to provide insurance. However, with the typical restrictions on the tax system (e.g., history independence), capital income tax cannot complete the market.

If we assume a linear constant capital income tax rate τ_k (either optimally chosen or an arbitrary) in our benchmark, introducing an capital income tax does not change the optimal labor income tax formula (5). We only need to replace r_t into after tax interest rate $r_t(1-\tau_k)$ in the formula. This is, because in a Huggett economy, aggregate asset sum to zero, and thus there is no government revenue from the capital income taxation. The quantitative effects of pecuniary externalities could be dampened with a positive τ_k.

Once we allow for nonlinear capital income taxation, there will be an additional term in the formula because of fiscal externalities caused by capital income tax. Consider a time-invariant capital income tax function $T_k(\cdot)$. The household's budget constraint becomes $c(a_0, x^t) + a_{t+1}(a_0, x^t) = x_t l(x_t) - T(x_t l(x_T)) + (1 + r_t)a_t(a_0, x^{t-1}) - T_k(r_T a_t(a_0, x^{t-1}))$. With this nonlinear asset income taxation, the government revenue from capital income tax is no longer zero. Moreover, the incidence of tax reform on savings will change the government revenue. Thus, the optimal tax formula becomes:

$$\frac{T'(z^*)}{1-T'(z^*)} = \frac{1}{\epsilon^l_{1-T'}(z^*)} \cdot \frac{1-F_z(z^*)}{z^* f_z(z^*)} \cdot (1-\beta) \sum_{t=0}^{\infty} \beta^t \left[A_t(z^*) + B_t(z^*) + C_t(z^*) + D_t(z^*)\right],$$

where $A_t(z^*), B_t(z^*)$, and $C_t(z^*)$ are the same as those in (5) except that r_t is replaced by $r_t(1-T_k'(r_t a_t))$. The additional term $D_t(z^*) = \int\int f(x^{t-1}|x_0) T_k'(r_t a_t(a_0, x^{t-1}))[dr_t a_t(a_0, x^{t-1}) + r_t da_t(a_0, x^{t-1})] dx^{t-1} d\Phi(a_0, x_0)$ represents the fiscal externalities due to capital income tax.

5.3 Income Effects in Labor Supply

[to be filled out]

5.4 Generalized Social Welfare Function

[to be filled out]

6 Conclusion

We study a fully nonlinear optimal income tax schedule in the presence of a private (incomplete) insurance market–Huggett (1993). As in Saez (2001), the optimal tax formula includes standard statistics such as the Frisch elasticity of the labor supply and the income distribution. In the presence of a private market, however, these statistics are no longer sufficient. The optimal tax formula depends on how the private market interact with public insurance and its welfare effects. First, the the optimal tax depends on the shape of joint distribution of assets and income. An economy with incomplete insurance market is likely to lead to a larger inequality in consumption (due to a skewed asset distribution) which typically calls for a stronger redistribution and a more progressive income tax schedule than without private insurance. Second, the optimal tax schedule should consider its pecuniary externalities. As shown in Dávila, Hong, Krusell, and Ríos-Rull (2012), an attempt to reform tax schedule generates changes in equilibrium prices (e.g., interest rate) which have differential welfare impacts across households. This pecuniary externality is likely to deter the optimal tax schedule from being too progressive because a progressive tax is likely to result in an increase in market-clearing interest rate (via reduced precautionary savings), which in turn makes the asset-poor (i.e., borrowers) worse off. Finally, the formula should also consider the additional welfare effects of the households who are released from the borrowing constraint as a result of tax reform.

While we can still express the optimal tax formula in terms of economically meaningful statistics, we argue that it is not practical to adopt a conventional sufficient-statistics approach because it involves additional terms that are hard to estimate from available data. Given these difficulties, we compute the optimal tax schedule based on a structural model—by obtaining those hard-to-measure statistics from a model calibrated to resemble the salient features of U.S. economy (such as the cross-sectional distributions of income and marginal propensity). The presence of a private market is quantitatively important, as the difference in optimal tax rates (with and without private insurance) can be as large as 10 percentage points. For the low-income group, the optimal tax rate is higher in the presence of a private market–mainly due to a more dispersed consumption distribution when savings are allowed. For the middle- and high-income groups, the optimal tax rates are lower in the presence of a private market—as the pecuniary externalities dominates. Unlike previous studies which calls for a very high tax rates at the top income bracket, according to our benchmark simulation, the current U.S. income tax rate for the very rich is not far from optimal.

References

Aiyagari, S., and E. R. McGrattan (1998): “The optimum quantity of debt,” *Journal of Monetary Economics*, 42(3), 447–469.

Aiyagari, S. R. (1994): “Uninsured Idiosyncratic Risk and Aggregate Saving*,” *The Quarterly Journal of Economics*, 109(3), 659–684.

Alvarez, F., and U. J. Jermann (2000): “Efficiency, Equilibrium, and Asset Pricing with Risk of Default,” *Econometrica*, 68(4), 775–797.

Attanasio, O., and J. V. Ríos-Rull (2000): “Consumption smoothing in island economies: Can public insurance reduce welfare?,” *European Economic Review*, 44(7), 1225–1258.

Badel, A., and M. Huggett (2017): “The sufficient statistic approach: Predicting the top of the Laffer curve,” *Journal of Monetary Economics*, 87(Supplement C), 1–12.

Bhandari, A., D. Evans, M. Golosov, and T. J. Sargent (2016): “Fiscal policy and debt management with incomplete markets,” *The Quarterly Journal of Economics*, Volume 132(Issue 2), Pages 617–663.

Brewer, M., E. Saez, and A. Shephard (2010): *Means testing and tax rates on earnings*chap. 2. Oxford University Press.

Chang, Y., and S.-B. Kim (2006): “From Individual to Aggregate Labor Supply: A Quantitative Analysis Based on a Heterogeneous Agent Macroeconomy,” *International Economic Review*, 47(1), 1–27.

Chetty, R. (2009): “Sufficient Statistics for Welfare Analysis: A Bridge Between Structural and Reduced-Form Methods,” *Annual Review of Economics*, 1(1), 451–488.

Chetty, R., and E. Saez (2010): “Optimal Taxation and Social Insurance with Endogenous Private Insurance,” *American Economic Journal: Economic Policy*, 2(2), 85–116.

Conesa, J. C., S. Kitao, and D. Krueger (2009): “Taxing Capital? Not a Bad Idea after All!,” *The American Economic Review*, 99(1), 25–48.

Conesa, J. C., and D. Krueger (2006): “On the optimal progressivity of the income tax code,” *Journal of Monetary Economics*, 53(7), 1425–1450.

Dávila, J., J. H. Hong, P. Krusell, and J.-V. Ríos-Rull (2012): “Constrained efficiency in the neoclassical growth model with uninsurable idiosyncratic shocks,” *Econometrica*, 80(6), 2431–2467.

Diamond, P., and E. Saez (2011): “The Case for a Progressive Tax: From Basic Research to Policy Recommendations,” *The Journal of Economic Perspectives*, 25(4), 165–190.

Diamond, P. A. (1998): “Optimal Income Taxation: An Example with a U-Shaped Pattern of Optimal Marginal Tax Rates,” *The American Economic Review*, 88(1), 83–95.

Dupor, B., M. Karabarbounis, M. Kudlyak, and M. S. Mehkari (2017): "Regional Consumption Responses and the Aggregate Fiscal Multiplier," *Working Paper*.

Findeisen, S., and D. Sachs (2017): "Redistribution and insurance with simple tax instruments," *Journal of Public Economics*, 146(Supplement C), 58–78.

Floden, M., and J. Linde (2001): "Idiosyncratic Risk in the United States and Sweden: Is There a Role for Government Insurance?," *Review of Economic Dynamics*, 4(2), 406–437.

Golosov, M., and A. Tsyvinski (2007): "Optimal Taxation with Endogenous Insurance Markets," *The Quarterly Journal of Economics*, 122(2), 487–534.

Golosov, M., A. Tsyvinski, and N. Werquin (2014): "A Variational Approach to the Analysis of Tax Systems," *National Bureau of Economic Research Working Paper Series*, No. 20780.

Greenwood, J., Z. Hercowitz, and G. W. Huffman (1988): "Investment, Capacity Utilization, and the Real Business Cycle," *The American Economic Review*, 78(3), 402–417.

Havránek, T. (2015): "Measuring Intertemporal Substitution: The Importance of Method Choices and Selective Reporting," *Journal of the European Economic Association*, 13(6), 1180–1204.

Heathcote, J., K. Storesletten, and G. L. Violante (2014): "Optimal Tax Progressivity: An Analytical Framework," *National Bureau of Economic Research Working Paper Series*, No. 19899.

Heathcote, J., and H. Tsujiyama (2017): "Optimal income taxation: Mirrlees meets Ramsey," *Working Paper*.

Huggett, M. (1993): "The risk-free rate in heterogeneous-agent incomplete-insurance economies," *Journal of Economic Dynamics and Control*, 17(5), 953–969.

Jappelli, T., and L. Pistaferri (2014): "Fiscal Policy and MPC Heterogeneity," *American Economic Journal: Macroeconomics*, 6(4), 107–136.

Johnson, D. S., J. A. Parker, and N. S. Souleles (2006): "Household Expenditure and the Income Tax Rebates of 2001," *American Economic Review*, 96(5), 1589–1610.

Kaplow, L. (1994): "A Note on Taxation as Social Insurance for Uncertain Labor Income," *Public Finance*, 49, 244–256.

Keane, M., and R. Rogerson (2012): "Micro and Macro Labor Supply Elasticities: A Reassessment of Conventional Wisdom," *Journal of Economic Literature*, 50(2), 464–476.

Kehoe, T. J., and D. K. Levine (1993): "Debt-Constrained Asset Markets," *The Review of Economic Studies*, 60(4), 865–888.

Kleven, H. J., C. T. Kreiner, and E. Saez (2009): “The Optimal Income Taxation of Couples,” *Econometrica*, 77(2), 537–560.

Krueger, D., and F. Perri (2011): “Public versus private risk sharing,” *Journal of Economic Theory*, 146(3), 920–956.

Mankiw, N. G., M. Weinzierl, and D. Yagan (2009): “Optimal Taxation in Theory and Practice,” *The Journal of Economic Perspectives*, 23(4), 147–174.

Mirrlees, J. A. (1971): “An Exploration in the Theory of Optimum Income Taxation,” *The Review of Economic Studies*, 38(2), 175–208.

Misra, K., and P. Surico (2011): “Heterogeneous Responses and Aggregate Impact of the 2001 Income Tax Rebates,” *CEPR Discussion Papers*, (8306).

Narajabad, B. N. (2012): “Information technology and the rise of household bankruptcy,” *Review of Economic Dynamics*, 15(4), 526–550.

Piketty, T. (1997): “La redistribution fiscale face au chômage,” *Revue française d’économie*, pp. 157–201.

Piketty, T., and E. Saez (2013a): “A Theory of Optimal Inheritance Taxation,” *Econometrica*, 81(5), 1851–1886.

——— (2013b): “A Theory of Optimal Inheritance Taxation,” *Econometrica*, 81(5), 1851–1886.

Piketty, T., E. Saez, and S. Stantcheva (2014): “Optimal Taxation of Top Labor Incomes: A Tale of Three Elasticities,” *American Economic Journal: Economic Policy*, 6(1), 230–271.

Ramsey, F. P. (1927): “A Contribution to the Theory of Taxation,” *The Economic Journal*, 37(145), 47–61.

Ríos-Rull, J.-V., and M. Kuhn (2016): “2013 Update on the U.S. Earnings, Income, and Wealth Distributional Facts: A View from Macroeconomics,” *Quarterly Review*, (April), 1–75.

Sachs, D., A. Tsyvinski, and N. Werquin (2016): “Nonlinear Tax Incidence and Optimal Taxation in General Equilibrium,” *National Bureau of Economic Research Working Paper Series*, No. 22646.

Saez, E. (2001): “Using Elasticities to Derive Optimal Income Tax Rates,” *The Review of Economic Studies*, 68(1), 205–229.

Saez, E., and S. Stantcheva (2017): “A Simpler Theory of Optimal Capital Taxation,” *Journal of Public Economics*.

Sahm, C. R., M. D. Shapiro, and J. Slemrod (2010): "Household Response to the 2008 Tax Rebate: Survey Evidence and Aggregate Implications," *Tax Policy and the Economy*, 24(1), 69–110.

Tauchen, G. (1986): "Finite state Markov-chain approximations to univariate and vector autoregressions," *Economics Letters*, 20(2), 177–181.

Appendix (For Online Publication)

A Proof of the Main Text

A.1 Incidence of tax reforms on labor supplies

Proof of eqautions(1). We derive elasticity of $l(x)$ with respect to retention rate $1 - T'(z(x))$ along the nonlinear budget constraint. The perturbed first order condition when perturbing the retention rate $1 - T'(z(x))$ by $dr(x)$ writes

$$v'(l(x) + dl(x)) = x\{1 - T'[x(l(x) + dl(x))] + dr(x)\}.$$

A first-order Taylor expansion around the initial equilibrium implies:

$$v'(l(x)) + v''(l(x))dl(x) = \{1 - T'(xl(x))\}x - T''(xl(x))x^2 dl(x) + xdr(x),$$

and thus

$$\begin{aligned}\epsilon^l_{1-T'}(x) = \frac{dl(x)}{dr(x)}\frac{1-T'(xl(x))}{l(x)} &= \frac{x}{v''(l(x)) + T''(xl(x))x^2}\frac{1-T'(xl(x))}{l(x)} \\ &= \frac{\frac{v'(l(x))}{v''(l(x))l(x)}}{1 + \frac{T''(xl(x))}{1-T'(xl(x))}xl(x)\frac{v'(l(x))}{v''(l(x))l(x)}} \\ &= \frac{e(x)}{1 + \rho(z(x))e(x)},\end{aligned}$$

where $e(x) = \frac{v'(l(x))}{l(x)v''(l(x))}$ and $\rho(z(x)) = \frac{z(x)T''(z(x))}{1-T'(z(x))}$.

We now derive the incidence of tax reforms on labor supply. We denote the perturbed tax function by $T(z) + \mu\tau(z)$. As we define in the main text, $dl(x)$ denotes the Gateaux derivative of the labor supply of type x in response to this tax reform. The labor supply response $dl(x)$ should solve the perturbed first-order condition:

$$0 = v'(l(x) + \mu dl(x)) - x\{1 - T'[x(l(x) + \mu dl(x))] - \mu\tau'[x(l(x) + \mu dl(x))]\}.$$

A first-order Taylor expansion implies that

$$v'(l(x)) + v''(l(x))\mu dl(x) = x\{1 - T'(xl(x))\} - T''(xl(x))x^2\mu dl(x) - \mu\tau'(xl(x))x.$$

Solving for $dl(x)$,

$$
\begin{aligned}
dl(x) &= \frac{-\tau'(xl(x))x}{v''(l(x)) + T''(xl(x))x^2} \\
&= \frac{-\tau'(xl(x))\frac{v'(l(x))}{v''(l(x))l(x)}}{1 + \frac{T''(xl(x))}{1-T'(xl(x))}xl(x)\frac{v'(l(x))}{v''(l(x))l(x)}} \\
&= \frac{-\tau(xl(x))}{1-T'(xl(x))}\frac{e(x)}{1+\rho(z(x))e(x)}l(x) = -\frac{\epsilon^l_{1-T'}(x)\tau'(z(x))}{1-T'(z(x))}l(x)
\end{aligned}
$$

A.2 Incidence of tax reforms on savings and interest rate

We start with driving the response of saving policy $h_t(a,y)$ to change in (current and future) virtual income and interest rate. Note that the saving policy functions and and interest rate can have the negative value, the following parameters denote the change in savings policy with respect to change in income (or interest rate), instead of percentage change.

The response of saving policy $h_t(a,y)$ to change in current virtual income R given interest rate r can be obtained by taking Taylor expansion of the perturbed Euler equation:

$$
\begin{aligned}
&u'((1+r)a + y(x) + dR - a' - da' - v(l(x))) \\
&= \beta(1+r)E[u'((1+r)(a'+da') + y(x') - h^A(a'+da', y(x')) - dh^A(a', y(x')) - v(l(x')))].
\end{aligned}
$$

Taking the first-order Taylor expansion, we obtain

$$
\begin{aligned}
u'(a,y) - u''(a,y)da' + u''(a,y)dR &= \beta(1+r)E[u'(a',y(x'))] + \beta(1+r)^2E[u''(a',y(x'))]da' \\
&-\beta(1+r)E[u''(a,y(x'))h^A_a(a',y(x'))]da' - \beta(1+r)E[u''(a',y(x'))dh^A_{t+1}(a',y(x'))].
\end{aligned}
$$

Thus,

$$
\epsilon^t_{a',R_t} = \frac{dh^A_t(a,y)}{dR_t} = \frac{u''(a,y) + \beta(1+r)E[u''(a',y(x'))\frac{dh^A_{t+1}(a',y(x'))}{dR_t}]}{u''(a,y) + \beta(1+r)^2E[u''(a',y(x'))] - \beta(1+r)E[u''(a',y(x'))h^A_a(a',y(x'))]}.
$$

Notice that the right hand side of the equation does involve $\frac{dh^A_{t+1}(a',y(x'))}{dR_t}$ and solving for $\frac{dh^A_{t+1}(a',y(x'))}{dR_t}$ requires $\frac{dh^A_{t+w}(a',y(x'))}{dR_t}$ and so on. That is, change in current income affects savings in all future periods, and thus it is very hard to get analytical expression for ϵ^t_{a',R_t}. We can obtain the closed form expression in a simple two period model: $\epsilon^t_{a',R_t} = \frac{u''(a,y)}{u''(a,y)+\beta(1+r)^2E[u''(a',y(x'))]}$.

Next, we can derive the response of saving policy $h_t(a, y)$ to change in state-contingent virtual income in next period $R(x')$ given interest rate r in a similar way. When the virtual income is perturbed by $dR(x') = dR(\hat{x}') \cdot \mathbb{1}_{x'=\hat{x}'}$, the perturbed Euler equation writes

$$u'((1+r)a + y(x) - a' - da' - v(l(x))) = \beta(1+r)E[u'((1+r)(a' + da') + y(x') + dR(x') - h^A(a' + da', y(x') + dR(x')) - dh^A_{t+1}(a', y(x')) - v(l(x')))].$$

Using the first-order Taylor expansion of the perturbed Euler equation, we obtain

$$\epsilon^t_{a',R_{t+1}(\hat{x}')} = \frac{dh^A_t(a,y)}{dR_{t+1}(\hat{x}')} = -\frac{\beta(1+r)\{f(\hat{x}'|x)u'(z, y(\hat{x}'))(1 - h^A_y(a', y(\hat{x}'))) - E[u''(a', y(x'))\frac{dh^A_{t+1}(a',y(x'))}{dR_{t+1}(\hat{x}')}]\}}{u''(a,y) + \beta(1+r)^2 E[u''(a', y(x'))] - \beta(1+r)E[u''(a', y(x'))h^A_a(a', y(x'))]},$$

which requires $\frac{dh^A_{t+1}(a',y(x'))}{dR_{t+1}(\hat{x}')}$ for all x' realizations.

So far, we have derived the response of saving policy given interest rate. The response of saving policy to change in interest rate can be obtained similar way. The perturbed Euler equations with respect to change in current interest rate and future interest rate are

$$\begin{aligned} &u'((1+r+dr_t)a + y(x) - a' - da' - v(l(x))) \\ &= \beta(1+r)E[u'((1+r)(a'+da') + y(x') - h^A(a'+da', y(x')) - dh^A(a', y(x')) - v(l(x')))], \end{aligned}$$

and

$$\begin{aligned} &u'((1+r)a + y(x) - a' - da' - v(l(x))) \\ &= \beta(1+r+dr_{t+1})E[u'((1+r+dr_{t+1})(a'+da') + y(x') - h^A(a'+da', y(x')) - dh^A(a', y(x')) - v(l(x')))], \end{aligned}$$

respectively. The first-order Taylor expansions yield

$$\begin{aligned} \epsilon^t_{a',r_t} &= \frac{dh^A_t(a,y)}{dr_t} = \frac{u''(a,y)a + \beta(1+r)E[u''(a', y(x'))\frac{dh^A_{t+1}(a',y(x'))}{dr_t}]}{u''(a,y) + \beta(1+r)^2E[u''(a', y(x'))] - \beta(1+r)E[u''(a', y(x'))h^A_a(a', y(x'))]}, \\ &\text{and} \\ \epsilon^t_{a',r_{t+1}} &= \frac{dh^A_t(a,y)}{dr_{t+1}} = \frac{-\beta E[u'(a', y(x'))] - \beta(1+r)E[u''(a', y(x'))]a' + \beta(1+r)E[u''(a', y(x'))\frac{dh^A_{t+1}(a',y(x'))}{dr_{t+1}}]}{u''(a,y) + \beta(1+r)^2E[u''(a', y(x'))] - \beta(1+r)E[u''(a', y(x'))h^A_a(a', y(x'))]}. \end{aligned}$$

We now express the incidence of tax reforms on saving policy in terms of the (semi) elasticities we defined above. We denote the Gateaux derivatives of savings in period t of a consumer with current state (a, y) in response of the tax reform by $dh^A_t(a, y)$. The saving response $dh_t^A(a,y)$ should solve the perturbed Euler equation:

$$\begin{aligned}u'((1+r+\mu dr_t)a-\mu\tau(z(x))-a'-\mu da'-v(l(x))) = \beta(1+r+\mu dr_{t+1})E[u'((1+r+\mu dr_{t+1})(a'+\mu da')\\ +y(x')-\mu\tau(z(x'))-h^A(a'+\mu da',y(x')-\mu\tau(z(x'))))-\mu dh^A(a',y(x'))-v(l(x')))].\end{aligned}$$

Solving for $dh_t^A(a,y)$,

$$\begin{aligned}dh_t^A(a,y(x)) &= -\frac{u''(a,y)}{\chi}\tau(z(x))+\frac{\beta(1+r)E[u''(a',y(x'))(1-h_y^A(a',y(x')))\tau(z(x'))]}{\chi}+\frac{u''(a,y(x))a}{\chi}dr_t\\ &\quad -\frac{\beta E[u'(a',y(x'))]+\beta(1+r)E[u''(a',y(x'))]a'}{\chi}dr_{t+1}+\frac{\beta(1+r)E[u''(a',y(x'))dh_{t+1}^A(a',y(x'))]}{\chi}\\ &= -\epsilon^t_{a',R_t}(a,y)\cdot\tau(z(x))-\int\epsilon^t_{a',R_{t+1}(x')}(a,y)\cdot\tau(z(x'))dx'+\epsilon^t_{a',r_t}(a,y)dr_t+\epsilon^t_{a',r_{t+1}}(a,y)dr_{t+1},\end{aligned}$$

where $\chi = u''(a,y)+\beta(1+r)^2E[u''(a',y(x'))]-\beta(1+r)E[u''(a',y(x'))h_a^A(a',y(x'))]$, and the second equality uses
$dh_{t+1}^A(a',y(x')) = -\frac{dh_{t+1}^A(a',y(x'))}{dR}\tau(z(x))-\int\frac{dh_{t+1}^A(a',y(x'))}{dR(\hat{x}')}\tau(z(\hat{x}'))d\hat{x}'+\frac{dh_{t+1}^A(a',y(x'))}{dr_t}dr_t+\frac{dh_{t+1}^A(a',y(x'))}{dr_{t+1}}dr_{t+1}$.

We now briefly discuss how we can express the incidence on the interest rate in terms of the slope of aggregate asset supply curve and the incidence on aggregate savings given interest rate. In Huggett econmy, when there is any change in aggregate asset supply, interest rate should adjust to clear the market (to guarantee that asset sum to zero). Thus, we can define the semi elasticity of interest rate with respect to aggregate asset as $\alpha = \frac{dr}{d\int ad\Phi} = -\frac{1}{A^{s\prime}(r)}$, where $A^s(r) = \int ad\Phi(a,x;r)$ denotes the aggregate asset supply curve, in which $\Phi(a,x;r)$ denotes the steady state distribution associated with the consumer's saving policy function $h^A(a,x;r)$ given interest rate. Then the incidence of tax reforms on interest rate is obtained by

$$dr_t = \alpha\cdot dA_t^s(r), \tag{15}$$

where $dA_t^s(r)$ denotes the incidence of tax reform on aggregate savings given interest rate, $\int\int da_t(x^{t-1};r)f(x^{t-1}|x_0)dx^{t-1}d\Phi(a_0,x_0)$.

A.3 Incidence of tax reforms on individual welfare

Proof of lemma 1

In this section, we derive the incidence of tax reforms on individual welfare (before rebating the change in government revenue as a transfer).

$$
\begin{aligned}
&dV(a_0, x_0) \\
&= d\Big[\sum_{t=0}^{\infty} \beta^t \int f(x^t \mid x_0) u(x_t l(x_t) - T(x_t l(x_t)) - a_{t+1}(a_0, x^t) + (1+r_t) a_t(a_0, x^{t-1}) - v(l(x_t)))\, dx^t\Big] \\
&= \sum_{t=0}^{\infty} \beta^t \int f(x^t \mid x_0) u'(a_0, x^t)\Big[-\tau(z(x_t)) - \, da_{t+1}(a_0, x^t) + (1+r) da_t(a_0, x^{t-1}) + \, dr_t a_t(a_0, x^{t-1})\Big]\, dx^t \\
&= \sum_{t=0}^{\infty} \beta^t \int f(x^t \mid x_0) u'(a_0, x^t)\Big[-\tau(z(x_t)) + dr_t \cdot a_t(a_0, x^{t-1})\Big]\, dx^t \\
&\quad - \sum_{t=0}^{\infty} \beta^t \int f(x^t \mid x_0)\Big[u'(a_0, x^t) - \beta(1+r) \int f(x_{t+1} \mid x_t) u'(a_0, x^{t+1}) dx_{t+1}\Big] da_{t+1}(a_0, x^t) dx^t,
\end{aligned}
$$

where the second equality is due to the Envelope theorem, using the intratemporal first order condition of the consumer.

Note that $a_{t+1}(a_0, x^t)$ can be recursively represented by $h^A(a_t(a_0, x^{t-1}), x_t)$. In addition, as long as $x \to y(x)$ is one-to-one mapping, we can express the saving policy function as a function of (a, y) so that $h^A(a, x) = h^A(a, y(x))$. Thus, $a_{t+1}(a_0, x^t) = h^A(a_t(a_0, x^{t-1}), y(x_t))$

Then the total change in savings, $da_{t+1}(a_0, x^t)$, can be decomposed into

$$
\begin{aligned}
da_{t+1}(a_0, x^t) \quad = \quad & dh^A(a_t(a_0, x^{t-1}), y(x_t)) \\
& + h_a^A(a_t(a_0, x^t), y(x_t)) \cdot da_t(a_0, x^{t-1}) + h_y^A(a_t(a_0, x^t), y(x_t)) \cdot dy_t(x_t),
\end{aligned}
$$

where h_a^A and h_y^A marginal propensity to save out of additional asset holdings and after-tax income, respectively, and h_y^A satisfies $h_x^A(a, x) = h_y^A(a, y(x)) \cdot (1 - T'(xl(x))[l(x) + xl'(x)]$.

When the borrowing constraint is binding, $h_a^A(a, y(x)) = h_y^A(a, y(x)) = 0$. On the other hand, if the borrowing constraint is not binding, $u'(c) - \beta(1+r)E[u'(c')] = 0$. Therefore, $\{u'(c) - \beta(1+r)E[u'(c')]\} \times \{h_a^A(a, y(x))\, da + h_y^A(a, y(x)) dy(x)\} = 0$. Thus,

$$
\begin{aligned}
&dV(a_0, x_0) \\
&= \sum_{t=0}^{\infty} \beta^t \int f(x^t \mid x_0) u'(a_0, x^t)\Big[-\tau(z(x_t)) + dr_t \cdot a_t(a_0, x^{t-1})\Big]\, dx^t \\
&\quad - \sum_{t=0}^{\infty} \beta^t \int f(x^t \mid x_0)\Big[u'(a_0, x^t) - \beta(1+r) E_{x_{t+1}}[u'(a_0, x^{t+1}) | x_t]\Big] dh^A(a_t(a_0, x^{t-1}), y(x_t)) dx^t.
\end{aligned}
$$

A.4 Incidence of tax reforms on government revenue - elementary tax reform

We derive the tax incidence on government revenue for the elementary tax reform. As we discuss in the main text, the elementary tax perturbation $\tau(z) = 1_{z \geq z^*}$ is not differentiable. To apply the formula (4) to this non-differentiable perturbation, we apply the construction technique discussed in Sachs, Tsyvinski, and Werquin (2016). That is, we can construct a sequence of smooth perturbation functions $\kappa_{z^*,\epsilon}(z)$ such that $\lim_{\epsilon \to 0} \kappa_{z^*,\epsilon}(z) = \delta_{z^*}(z)$, in the sense that for all continuous function $h(\cdot)$ with compact support,

$$\lim_{\epsilon \to 0} \int_{\mathbb{R}} \kappa_{z^*,\epsilon}(z) h(z) dz = h(z^*),$$

and by changing variables in the integral, this also implies

$$\lim_{\epsilon \to 0} \int_{X} \kappa_{z^*,\epsilon}(z(x')) \left\{ h(z(x')) \frac{dz(x')}{dx} \right\} dx' = h(z^*).$$

Letting $\tau_{z^*,\epsilon}(\cdot)$ denote the function such that $\tau'_{z^*,\epsilon} = \kappa_{z^*,\epsilon}(\cdot)$, the tax incidence of a tax reform $\tau_{z^*,\epsilon}$ on government revenue $dR(\tau_{z^*,\epsilon})$ is

$$dR(\tau_{z^*,\epsilon}) = \int \tau_{z^*,\epsilon}(z(x)) f(x) dx + \int T'(z(x)) \Big[- \epsilon^{l}_{1-T'}(x) \cdot \frac{\kappa_{z^*,\epsilon}(z(x))}{1 - T'(z(x))} z(x) \Big] f(x) dx.$$

Thus, we can obtain the dR of the elementary tax reform at z^*:

$$\begin{aligned}
\lim_{\epsilon \to 0} dR(\tau_{z^*,\epsilon}) = dR &= \int_{x^*}^{\infty} \frac{f(x)}{1 - F(x^*)} dx - \frac{T'(z(x^*))}{1 - T'(z(x^*))} \epsilon^{l}_{1-T'}(x^*) \frac{z(x^*)}{z'(x^*)} \cdot \frac{f(x^*)}{1 - F(x^*)} \\
&= \int_{z^*}^{\infty} \frac{f(z)}{1 - F_z(z^*)} dz - \frac{T'(z^*)}{1 - T'(z^*)} \cdot \epsilon^{l}_{1-T'}(z^*) \cdot \frac{z^* f_z(z^*)}{1 - F_z(z^*)}.
\end{aligned}$$

A.5 Derivation of Optimal Tax Formula

Proof of proposition 2

We start with deriving the incidence of tax reforms on social welfare. We denote the Gateaux derivative of social welfare in response to the elementary tax reform by dW. Then,

$$dW = \int \sum_{t=0}^{\infty} \beta^t dR \Big[\int f(x^t \mid x_0) u'(a_0, x^t) dx^t \Big] \cdot \phi(a_0, x_0) da_0 dx_0 + \int dV(a_0, x_0) \phi(a_0, x_0) \, da_0 \, dx_0$$

The updating operator for the sequence of distribution densities ϕ_t of savings a and productivity x is $\phi(a_{t+1}, x_{t+1}) = \int f(x_{t+1} \mid x_t) \frac{\phi(a')^{-1}(a_{t+1},x_t),x_t)}{a'_a((a')^{-1}(a_{t+1},x_t),x_t)}\, dx_t$ at any period t. Therefore, for some function $\tilde{h}$ such that $\tilde{h}(a_0, x^t) = h(a_t(x^{t-1}), x_t)$, we can obtain

$$\iint \tilde{h}(a_0, x^*) f(x^* \mid x_0) dx^t \phi(a_0, x_0) da_0 dx_0 = \int h(a_t, x_t)\phi(a_t, x_t)\, da_t\, dx_t,$$

by applying change of variables sequentially.

Thus, we obtain

$$\begin{aligned}
dW &= \int \sum_{t=0}^{\infty} \beta^t dR \cdot \int u'(a_t, x_t)\phi(a_t, x_t)\, da_t dx_t \\
&\quad + \sum_{t=0}^{\infty} \beta^t u'(a_t, x_t)\Big[-\tau(z(x_t)) + dr_t \cdot a_t\Big]\phi(a_t, x_t)\, da_t\, dx_t \\
&\quad - \sum_{t=0}^{\infty} \beta^t \int \Big[u'(a_t, x_t) - \beta(1+r)\int f(x_{t+1} \mid x_t) u'(a_{t+1}(a_t, x_t), x_{t+1}) dx_{t+1}\Big] dh_t^A(a_t, y(x_t))\phi(a_t, x_t)\, da_t\, dx_t \\
&= \frac{\lambda}{1-\beta}\Big[\int_{x^*}^{\infty} \frac{f(x)}{1-F(x^*)}\, dx - \frac{T'(z(x^*))}{1-T'(z(x^*))} \cdot \frac{\epsilon^l_{1-T'}(x^*) z(x^*)}{z'(x^*)} \cdot \frac{f(x^*)}{1-F(x^*)}\Big] \\
&\quad - \frac{1}{1-\beta} \iint_{x^*}^{\infty} u'(a, x) \frac{\phi(a, x)}{1-F(x^*)}\, dx da + \sum_{t=0}^{\infty} \beta^t \int u'(a, x)\{dr_t \cdot a\}\phi(a, x)\, da\, dx \\
&\quad - \sum_{t=0}^{\infty} \beta^t \int \Big[u'(a, x) - \beta(1+r)\int f(x' \mid x) u'(a'(a, x), x')\, dx'\Big]\{dh_t^A(a, y(x))\}\phi(a, x)\, da\, dx,
\end{aligned}$$

where $\lambda = \int u'(a_t, x_t)\phi(a_t, x_t)\, da_t\, dx_t$.

Optimal tax formula can be obtained if no tax reform improves welfare. That is, $dW = 0$ implies

$$\begin{aligned}
\frac{T'(z(x^*))}{1-T'(z(x^*))} =& \frac{1}{\epsilon^l_{1-T'}(x^*)} \cdot \frac{z'(x^*)}{z(x^*)} \cdot \frac{1-F(x^*)}{f(x^*)} \\
& \times \frac{1}{1-\beta} \sum_{t=0}^{\infty} \beta^t \Big[\iint_{x^*}^{\infty} (1 - \frac{u'(a, x)}{\lambda}) \frac{\phi(a, x)}{1-F(x^*)}\, dx\, da - \int \frac{u'(a, x)}{\lambda}\{-dr_t \cdot a\}\phi(a, x)\, da\, dx \\
& - \int \Big[\frac{u'(a, x)}{\lambda} - \beta(1+r)\int f(x' \mid x) \frac{u'(a'(a, x), x')}{\lambda} dx'\Big]\{dh_t^A(a, y(x))\}\ \phi(a, x)\, da\, dx\Big]
\end{aligned}$$

By applying change of variables, we get the formula (5).

B Additional Comparative Statics

B.1 Risk Aversion and Labor Supply Elasticity

We consider the relative risk aversion $\sigma = 1$ (lower than the benchmark risk aversion: $\sigma = 1.5$). Figure 10 shows the optimal tax rates without a private insurance market for $\sigma = 1$ and 1.5. With a smaller risk aversion, the optimal tax rates are lower than those in our benchmark at all productivity levels because there is less need for insurance.[30] This driving force is still present in an economy with a private market.

In the presence of a private insurance market, however, the optimal tax rates at the upper-middle and high-income groups are in fact higher when $\sigma = 1$ (lower risk aversion) in Figure 11. This is because of the general equilibrium effect. With a smaller risk aversion, households save less (a weaker precautionary savings motive). The real interest rate has to increase to clear the private savings market. For example, the real interest rate net of discounting $\beta(1+r)$ increases from 0.9647 (under $\sigma = 1.5$) to 0.9808 (under $\sigma = 1$). This encourages high-income households to save more and makes the marginal private savings (P') schedule steeper, which leads to a larger cross-sectional dispersion in assets and consumption. Thus, the second term and the third term in our formula (amplified Saez effects and aligned private progressivity effects) become larger. For upper-middle and high-income groups (wages above \$35), this general equilibrium effect (larger second and third terms in the formula) dominates a weaker precautionary savings (smaller first term in the formula), resulting in higher marginal tax rates.

Next, we consider a smaller Frisch elasticity of the labor supply ($e = 0.25$). Figure 12 shows that for all income levels the optimal tax rates under $e = 0.25$ are higher than those in our benchmark ($e = 0.5$) because an inelastic labor supply is associated with a smaller cost of distortion.

B.2 Borrowing Constraints

In the benchmark economy, we set the borrowing limit $\underline{a} = -86.55$, which is the average annual earnings in the steady state under the current U.S. tax schedule (approximated by

[30]Under utilitarian social welfare, for example, the social welfare weights at high-income levels ($\frac{u'(c)}{A}$) are relatively high when risk aversion is low.

Figure 10: Optimal Tax without Private Insurance

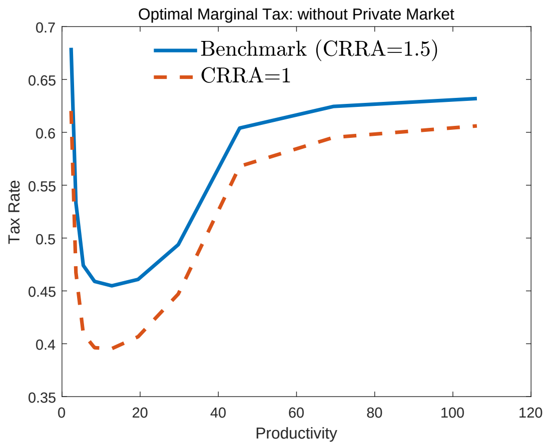


Figure 11: Optimal Tax with Private Insurance

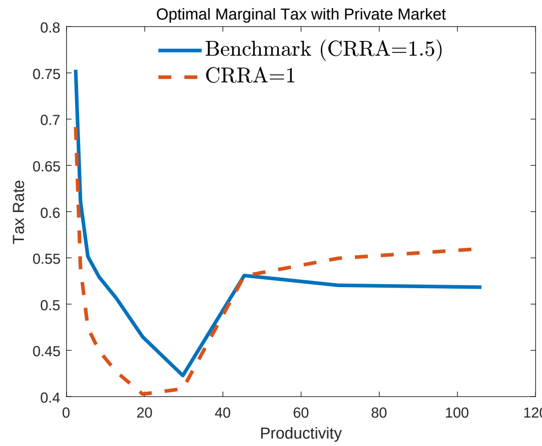
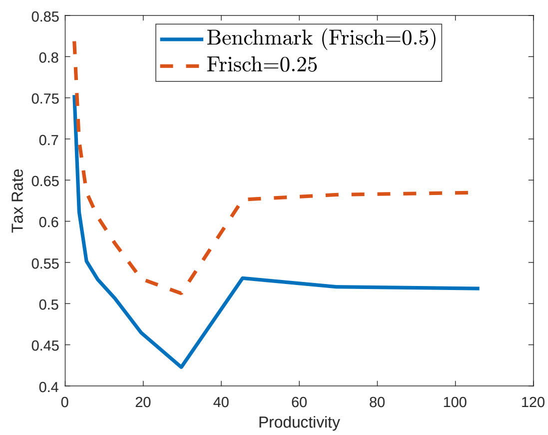


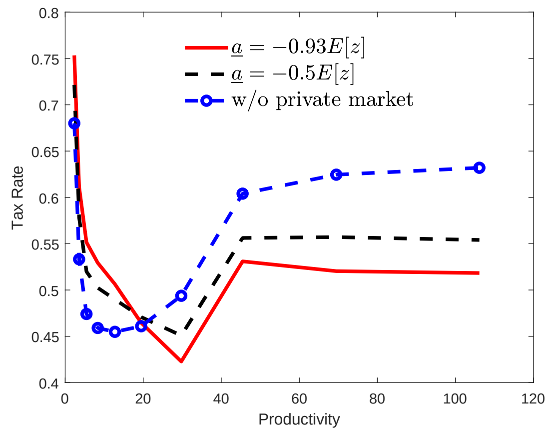
Figure 12: Optimal Tax with Private Market



a log-linear form as in HSV). Under this borrowing limit, 9.7% of households are credit-constrained under the current U.S tax schedule. We consider a tighter borrowing limit, which is half of our benchmark case ($\underline{a} = -43.3$) so that workers can borrow one-half of the average earnings in the economy. With this tighter borrowing limit, about 34% of the population is credit-constrained under the current U.S. tax schedule in our model. Figure 13 shows the optimal tax rate schedules under this tighter borrowing constraint. The optimal tax rates are roughly between those in the benchmark and those without a private insurance market, except for the highest productivity group.

Under the tighter borrowing constraint, households tend to save more due to a stronger precautionary savings motive. To clear the private savings market, the equilibrium interest rate has to decrease. This increases the MPS of the low-income group and decreases the MPS of the high-income group. This will lead to a lower marginal tax for the low-income

Figure 13: Optimal Tax under Tighter Borrowing Constraint



group and a higher marginal tax for the high-income group. In addition, the crowding in/out effects become larger under the tighter borrowing constraint as the tax reform induces a more progressive response of private savings. At the top income group, this effect is strong enough to generate an even higher optimal tax rate.

C Proof of Section 5

[to be filled out]