SIER Working Paper Series No. 117

Optimal Taxation with Private Insurance

By

Yongsung Chang and Yena Park

November, 2018

Institute of Economic Research
Seoul National University

Optimal Taxation with Private Insurance *

Yongsung Chang
Seoul National University

Yena Park
University of Rochester

November 1, 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 interacts 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 the 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 11 percentage points.

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

*This is a substantially revised version of previously circulated papers with the same title. We would like to thank Florian Scheuer, Narayana Kocherlakota, Marco Bassetto, Aleh Tsyvinski, Dirk Krueger, Anmol Bhandari, Louis Kaplow, Yongseok Shin, Kjetil Storesletten, and Maxim Troshkin for helpful comments. We also thank Hyun-Tae Kim and Hyun Yeol Kim for their excellent research assistance. This work was supported by the Research Resettlement Fund for the new faculty of Seoul National University. Chang: yohg@snu.ac.kr, Park: yena.park@rochester.edu

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 level that would have been selected by a household in the absence of a 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. We study the optimal schedule within a simple class of tax system that is levied on current income only. While this is a restrictive environment (compared to a more general fully-history-dependent tax system), it allows a direct comparison of our results to those in classic optimal formulas (e.g., Saez (2001) and Diamond (1998)). By using a variational approach within this class of tax system, as in Piketty (1997) and Saez (2001), we can also study the economic insights behind the optimal tax in a more transparent way.

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

We choose Huggett (1993) for several reasons. First, it is one of the most commonly used general equilibrium incomplete market structures in macroeconomic analysis. Second, since its asset 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 least, by yielding an explicit 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: the Frisch elasticity of the labor supply and the hazard rate of the income distribution. In the presence of private insurance, 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 greater inequality in consumption, which calls for a stronger redistribution through tax/transfer. Second, pecuniary externalities should be considered. As shown in Dávila, Hong, Krusell, and Ríos-Rull (2012), individual households' saving decisions have externalities because the change in the equilibrium interest rate generates a redistribution across households. This effect is likely to prevent a tax schedule from becoming too progressive. Providing more insurance through a progressive tax system is likely to reduce the private households' precautionary savings motive and to result in a decrease in the equilibrium interest rate, which in turn makes the asset poor (i.e., borrowers) worse off. Third, the formula should also consider the additional welfare effects of some households that 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 along several dimensions, we also show that such an attempt is highly challenging for (at least) two reasons. First, the optimal tax formula depends on the specific market structure–i.e., various welfare effects from the interaction between private and public insurance. More precisely, the degree to which the envelope theorem can be applied to the response of private intermediation varies across the market. 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 of how 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 and no constraint is imposed on borrowing (Findeisen and Sachs (2017)) is an example of how 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 the real world.

Next, we further show that even if the formula can be expressed in terms of sufficient statistics, it is not easy to estimate them as they are *not* policy invariant. Given this difficulty, we combine the structural and sufficient-statistics methods following Chetty's (2009) suggestion. We obtain the (hard-to-estimate) additional statistics from a general-equilibrium model calibrated to resemble some salient features (such as income and wealth distributions) of the U.S. economy. According to our analysis, the difference in optimal tax rates (with and without a private insurance market) can be as large as 11 percentage points. Moreover, these differences in tax rates do not necessarily exhibit the same sign across income brackets. For example, the optimal tax rates are higher than those without private markets for the low-income group, mainly because of the increased inequality in consumption. The optimal tax rates are lower (than those without a private market) for the middle- and high-income groups, mainly because of pecuniary externalities.

This paper is most closely related to the literature on optimal labor income tax using a variational approach pioneered by Piketty (1997) and Saez (2001). In a static model, they express the optimal tax formula in terms of sufficient statistics (e.g., elasticity of the labor supply and the hazard rate of income), obtained by perturbations of a given tax system. This variational approach complements the traditional mechanism-design approach (Mirrlees (1971)) and helps 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 environments (Golosov, Tsyvinski, and Werquin (2014) and 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 interaction between the two insurances. Scheuer

[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)).

and Werning (2018) study the optimal tax formula in the presence of productivity shocks over the life cycle, but they assume that markets are complete in order to focus on redistribution rather than 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., 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 information friction. With this market structure, they also show that the government can internalize the pecuniary externalities. While their questions are centered on the welfare gains from government intervention, we focus on how the optimal tax formula is affected by private insurance in a more general market structure.

Our benchmark analysis is also related to a recent paper by Findeisen and Sachs (2017), which studies the optimal nonlinear labor income tax and linear capital income tax with self-insurance opportunities and focuses on the interaction between labor and capital income taxes. As discussed above, their results can be viewed as an example of how the envelope theorem can be fully applied, because they assume a fixed interest rate with no

[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.

constraint in borrowing. There are a lot more interactions between public and private insurance in our model, as we relax these assumptions.

Outside the optimal taxation literature, our analysis illustrates interesting policy implications for pecuniary externalities, first analyzed by Dávila, Hong, Krusell, and Ríos-Rull (2012) in a general equilibrium incomplete markets economy. Our analysis is also related to work by (i) Attanasio and Ríos-Rull (2000), who examine the relationship between compulsory public insurance (against aggregate shocks) and private insurance against idiosyncratic shocks, and (ii) Krueger and Perri (2011), who 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 a Huggett economy. In Section 3, we extend the formula to more general private insurance markets. Section 4 provides a quantitative analysis. Section 5 discusses generalizations of the benchmark analysis along several dimensions. Section 6 concludes.

2 Optimal Nonlinear Tax Formula with Private Insurance

In this section, we derive an optimal nonlinear tax formula in our benchmark economy: a Huggett-style incomplete markets model. As we will show in detail below, the optimal tax formula depends on whether the interaction between private insurance and public insurance generates welfare effects. Thus, the formula inevitably depends on a specific structure of the private market. As we discussed in the introduction, we chose Huggett (1993) because: (i) it is one of the most commonly used incomplete markets structure in macroeconomics, (ii) comparison with Saez (2001) is straightforward due to zero aggregate savings, (iii) it permits a ready comparison with other market structures, and (iv) it provides important policy implications of well-known pecuniary externalities in incomplete markets. In Section 3 below, we extend our results along several dimensions with a more general market structure.

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. More precisely, in the benchmark, (i) we consider a nonlinear labor income tax with a lump-sum transfer; (ii) the tax is levied on the current period's income only (no history dependency); and (iii) the tax function is time invariant.[3]

Although these restrictions are less restrictive than the parametric restrictions on the tax function in the Ramsey taxation, they are still subject to the critique by the new dynamic public finance literature. We impose these restrictions for several reasons. First, they allow for a direct comparison to the Saez (2001) tax formula without a private market; thus, we can highlight the role of allowing private insurance in affecting tax schedule. Second, using a variational approach within a simpler tax system allows the tax formula to contain a transparent and intuitive mechanism that demonstrates the effects of taxes, especially the impact of taxes on private insurance. In dynamic optimal taxation constrained only by informational frictions (as in the new dynamic public finance literature), it is much more challenging to obtain transparent insights for at least two reasons: (i) there are many different tax functions that decentralize optimal allocation, and (ii) the implementation requires a complex tax system that uses history dependence as well as joint taxation of different sources of income. Third, under this simpler tax system, it is easier to extend the tax formula to various private insurance markets and it permits ready comparison of the optimal tax formula under different market structures. Fourth, according to our analysis, the presence of private insurance can either improve or undermine social welfare, depending on the degree of completeness of the private market, which we see as a realistic feature. Under the optimal tax system with informational frictions only, allowing private insurance can only reduce welfare: at best, social welfare is unchanged by a complete crowding-out of public insurance.

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, $f(x_{t+1}|x_t)$, that has an invariant stationary (cumulative) distribution $F(x)$, whose probability density is $f(x)$.

[3] For simplicity, the benchmark analysis abstracts from the tax on income from asset holdings. Section 5 extends the optimal tax formula when there is a capital income tax.

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 the labor supply l. We focus on households' preferences that have no wealth effect on the labor supply (the so-called GHH preferences of Greenwood, Hercowitz, and Huffman (1988)) for simplicity, but we relax this assumption later.[4] The earnings of a worker with x_t is $z(x_t) = x_t l(x_t)$. 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 system where the net tax payment (tax − transfer) 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 non-state-contingent bond a_t (e.g., Huggett (1993)). We also assume that there is an exogenous borrowing limit, $\underline{a}$.

Given prices and government policies, the individual consumer solves

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

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

[4] This assumption is common in the literature because it significantly simplifies the optimal tax formula. 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 and private insurance.

[5] Under the preferences that have no wealth effects on the labor supply, labor income z_t and productivity x_t have a one-to-one relationship so that we can use them interchangeably. We also note that even with preferences with wealth effects on the labor supply, we can use state variables (a_t,x_t) and (a_t,z_t)

The aggregate state of the economy in period t is described by a joint measure of assets and productivity, $\mathbf{\Phi}_t(a_t, x_t)$. By abusing the notation, we also denote the distribution of income and assets 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 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 in the next period is:

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

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

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

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

We assume that the government evaluates social welfare according to:

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

which is the utilitarian social welfare function. In Section 5, we extend our analysis to a more general social welfare function.

2.3 Deriving an Optimal Formula in a Huggett Economy

In deriving an optimal tax formula, we apply the variational approach (Piketty (1997); Saez (2001)). Consider a perturbation (a small deviation) from a given nonlinear 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 of individual and aggregate variables, and then the optimal tax formula.

---

interchangeably because x_t and $z_t(a_t, x_t) = x_t l_t(a_t, x_t)$ have a one-to-one relationship given a_t.

2.3.1 Tax Incidence

We start with the tax incidence: 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 will converge to a steady state where the distribution of state variables $\mathbf{\Phi}(a, x)$ is stationary. We assume that in period 0 the economy starts from that steady state.

Consider an arbitrary tax reform of an 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 $\mu \in \mathbb{R}$ parameterizes 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 similar derivatives for other variables such as the 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 a constant amount $\frac{1}{1-F_z(z^*)}$, and the marginal tax rate at income level z^* is increased by $\frac{1}{1-F_z(z^*)}$ (which is obtained by the marginal perturbation: $\tau'(z) = \frac{1}{1-F_z(z^*)}\delta_{z^*}(z)$). Note that with this tax reform, the increased government revenue due to a 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 elementary tax reforms. See Sachs, Tsyvinski, and Werquin (2016) for further details.[6]

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

[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 income above z^*.

the linear budget constraint is defined as[7]

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

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

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

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

From the definition of elasticity $\epsilon^l_{1-T'}$, $dl(x)$ represents the change in the labor supply in response to a tax reform, taking into account both exogenous and endogenous changes in the marginal tax rates–changes in $T'(xl(x))$ due to $dl(x)$.

We also remark that $dl(x)$ is constant in all periods. This is achieved by assuming no wealth effect on the labor supply. Since the labor supply does not depend on wealth, $dl(x)$ is time invariant regardless of private insurance.

Incidence of tax reforms on government revenue

In the absence of capital income tax, the government revenue in period t under the original tax schedule is simply: $R_t = \int T(z(x_t))f(x_t)dx_t$. Then, the incidence on government

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

revenue, dR_t, directly follows from the change in the labor supply $dl(\cdot)$:

$$\begin{aligned} dR_t &= \int \tau(z(x))f(x)dx + \int T'(z(x))\Big[-\epsilon^l_{1-T'}(x))\cdot \frac{\tau'(z(x))}{1-T'(z(x))}z(x)\Big]f(x)\,dx \qquad (2) \\ &= \int_{x^*}^{\infty} \frac{f(x)}{1-F(x^*)}\,dx - \frac{T'(z(x^*))}{1-T'(z(x^*))}\epsilon^l_{1-T'}(x^*)\frac{z(x^*)}{z'(x^*)}\cdot\frac{f(x^*)}{1-F(x^*)}, \quad \forall t. \end{aligned}$$

The second equality holds for the elementary tax reform—see the Appendix.[8] Since we consider revenue-neutral tax reforms, any change in government revenue dR_t will be rebated back to households as a lump-sum transfer. Note that the change in government revenue $dR_t = dR$ is constant in all periods because the household's labor supply depends on current productivity only and the tax system is time invariant.

Incidence of tax reforms on savings

The individual household's savings decision in period t can be represented recursively by the policy function: $h^A(a_t(a_0, x^{t-1}), x_t)$. As long as the mapping $x \mapsto y(x)$ is one to one, we can express $h^A(a,x) = h^A(a, y(x))$, where $y(x) = xl(x) - T(xl(x))$ with $l(x)$ that solves $x(1 - T'(xl)) = v'(l)$.

Even in the absence of capital income taxation, households' saving may adjust when the labor income tax schedule is changed.[9] Deriving the incidence of tax reform on the policy function of savings $h^A(a,y)$ is highly challenging, because saving decisions at different times and histories (of shocks) are inter-linked each other. The following expression of $dh^A(a,y)$—obtained by the Taylor expansion of the perturbed first-order condition—shows this difficulty more explicitly (see the Appendix for the derivation):

$$\begin{aligned} dh^A_{t+1}(a, y(x)) = & \frac{u''(a,x)}{\chi}[-\tau(z(x)) + dR] - \frac{\beta(1+r)E[u''(a',x')(1-h^A_y(a',y(x')))(-\tau(z(x'))+dR)]}{\chi} \qquad (3) \\ & + \frac{u''(a,x)a}{\chi}dr_t - \frac{\beta E[u'(a',x') + (1+r)u''(a',x')a']}{\chi}dr_{t+1} \\ & + \frac{\beta(1+r)E[u''(a',x')dh^A_{t+2}(a',y(x'))]}{\chi}, \end{aligned}$$

[8]Note that the step function $1_{z\geq z^*}$ is not differentiable. In the Appendix, we show that we can nevertheless apply the formula (2) by constructing a sequence of smooth perturbations $\{\tau'_n(z)\}_{n\geq 1}$ which satisfies $\lim_{n\to\infty} \tau'_n(z) = \delta_{z^*(z)}$.

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

where $\chi = u''(a,x) + \beta(1+r)^2 E[u''(a',x')] - \beta(1+r)E[u''(a',x')h_a^A(a',y(x'))]$. The response of savings in period t depends on the response of savings in period $t+1$ with respect to all possible realizations of shocks, $dh_{t+2}^A(a',y(x'))$, which in turn depends on the response of savings in period $t+2$ and so on. Moreover, the changes in equilibrium interest rates, dr_t and dr_{t+1}, also depend on the response of savings with respect to all possible events, which makes Equation (3) a very complicated integral equation.

We do not attempt to solve the incidence on savings and interest rates analytically. Instead, we are interested in how the changes in savings and interest rates appear in the optimal tax formula, which will still illustrate the economic mechanism behind the optimal tax schedule *given* the incidence on savings and interest rates. In principle, we can express $dh_t^A(a,y(x))$ in terms of the (semi-)elasticities of savings with respect to $1-T'(z(x^*))$, if we define the (semi-)elasticity $\epsilon_{1-T'}^{A,t}(x^*,a,x) = \frac{dh_t^A(a,y(x))}{dlog(1-T'(z(x^*)))}$ as the elasticity that reflects the total change in savings, including the effects through changes in future savings and interest rates as well as the income effects from tax rebate dR on $h_t^A(a,y(x))$.[10] That is, $\epsilon_{1-T'}^{A,t}(x^*,a,x)$ measures the causal impact of a change in the tax rate on savings—simply the difference in savings with and without a tax reform.

In a Huggett economy, aggregate savings sum to zero in equilibrium. Thus, if there is any change in aggregate savings due to a tax reform, the equilibrium interest should adjust to clear the market. This implies that the incidence of a tax reform on the equilibrium interest rate depends on the tax incidence on savings. In the Appendix, we show that the incidence on the interest rate dr_t can be expressed in terms of the slope of the aggregate supply of savings.

Incidence of tax reforms on individual welfare

Next, we derive the incidence of a tax reform τ (including the lump-sum rebate) on households' indirect utility, $V(a_0,x_0)$.

Lemma 1. *The incidence of a tax reform τ of an initial tax schedule T on households'*

[10] The elasticity that reflects both substitution and income effects—the effects of the lump-sum rebate—is referred to as a policy elasticity (Hendren (2016)).

indirect utility, $dV(\cdot,\cdot)$, is

$$dV(a_0,x_0) = \sum_{t=0}^{\infty}\beta^t \int u'(a_0,x^t)\left[-\tau(z(x_t)) + dR + dr_t \cdot a_t(a_0,x^{t-1})\right] f(x^t|x_0)dx^t$$
$$-\sum_{t=0}^{\infty}\beta^t \int \left[u'(a_0,x^t) - \beta(1+r)E[u'(a_0,x^{t+1})|x^t]\right] \cdot dh^A_{t+1}(a_t(a_0,x^{t-1}),y(x_t))f(x^t|x_0)dx^t. \quad (4)$$

Proof See the Appendix. ■

The first term on the right-hand side of Equation (4), $-\tau(z(x_t))$, reflects a higher tax payment. This decrease in utility is the effect of the standard tax incidence in an economy without private insurance. The second term, dR, appears because any change in government revenue is rebated as a lump-sum transfer. Note that the welfare effects via $dl(x_t)$ do not show up because of the envelope condition in the labor supply.

In the presence of private insurance, there are two additional effects on households' utility. The first additional incidence on the utility is the effect from the change in the equilibrium interest rate, dr_t, which arises due to pecuniary externalities: individual households take the market interest rate as given, without considering how their savings decision affects the equilibrium interest rate. The second additional incidence, which is captured by the second integration in Equation (4), 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 that are released from the borrowing constraint as a result of tax reform, i.e., that used to be constrained under the original tax schedule but not after the reform, this term shows up. Thus, the increased savings (or less borrowing), $dh^A(a,y)$, implies a decrease in consumption (utility). Technically speaking, this second additional incidence appears because a borrowing-constrained individual's optimal decision is at the kink of the budget constraint where the envelope theorem cannot be applied.

To understand Equation (4) better, we further decompose the total change in savings of a household with history (a_0,x^t) into:

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

where h^A_a and h^A_y are the marginal propensity to save out of additional asset holdings and after-tax income, respectively. That is, $dh^A(a,y)$ captures the change in the policy function of savings for a given asset holding, a, and after-tax income, y. Additional changes in savings due to a change in state (a,y) for a given marginal propensity to save do not have any impact on the households' utility because the envelope theorem applies.

Incidence of tax reforms on social welfare

We now study the tax incidence on social welfare. With the utilitarian social welfare function, the incidence of a tax reform τ on social welfare dW is:

$$dW = \iint dV(a_0, x_0)\phi(a_0, x_0)\, da_0\, dx_0.$$

2.3.2 Optimal Tax Formula

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

Proposition 2. *The 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

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

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

Proof See the Appendix. ■

Note that the incidences of tax reform on the interest rate $dr_t(z^*)$ and the savings policy $dh^A_{t+1}(a, y(z); z^*)$ depend on the level of income z^* at which the elementary tax reform occurs. From now on, for notational simplicity, we drop z^* from these incidences.

We also 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 $dh^A_{t+1}(a, y(z))$ 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^*, $E[g(a,z)|z \geq z^*]$.[11]

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. The optimal tax schedule also depends on how the private insurance market interacts with public savings, such as dr_t and $dh^A(a, y(z))$.

We can rewrite formula (5) in terms of the (exogenous) productivity distribution by applying the change of variables and using $f_z(z(x))z'(x) = f(x)$ along with the following lemma.

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

[11] 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 an 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.

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

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

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 a 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 $\Phi(z,a)$. Intuitively, an incomplete private savings market is likely to lead to greater inequality in consumption via a 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.[12]

The second term, $B_t(z^*)$, reflects the pecuniary externality. More precisely, this term captures whether the changes in savings as a result of tax reform—via the change in the equilibrium interest rate—generate positive (or negative) welfare. As discussed in Dávila, Hong, Krusell, and Ríos-Rull (2012), in an economy with an incomplete market where the only available asset is a non-state-contingent 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

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

households (i.e., borrowers) who have a high marginal utility of consumption.[13] The following proposition shows that the sign of the second term in the optimal tax formula (5) is exactly determined by this pecuniary externality.

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 the asset distribution by $\bar{A}$, and note that in a Huggett economy, $\bar{A} = 0$, which yields:

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

■

Proposition 4 shows that if the elementary tax reform at specific income z^* increases (decreases) the equilibrium interest rate, it has a negative (positive) effect on social welfare, and thus the optimal tax rate at z^* should be lowered (raised). Intuitively, a higher interest rate as a result of progressive tax reform hurts asset-poor households (borrowers). Since these households tend to exhibit a higher marginal utility of consumption, this tax reform is not desirable. Thus, the pecuniary externalities can potentially prevent the optimal tax system from being *overly progressive*.

[13]In Dávila, Hong, Krusell, and Ríos-Rull (2012), there is another channel of pecuniary externalities, where increased households savings generate the opposite welfare implication. A higher wage rate caused by increased savings can generate negative insurance effects by scaling up the stochastic part of the housedhold's income. In a Huggett economy, with a linear production in labor, the wage rate is constant. Thus, this channel is not present.

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 that used to be constrained in borrowing save more (by reducing consumption), this is an additional welfare cost. The next proposition shows that this term is always negative. According to our quantitative analysis below, the magnitude of $C_t(z^*)$ tends to be very small because there are only a small fraction of households to which this effect applies.

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

2.3.4 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 the tax rate (τ) and household's saving rate (p) are linear in a static environment.[14] That is, 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 $\bar{y} = E[y]$. In this economy, they show that the optimal tax rate is: $\frac{\tau}{1-\tau} = -p - \frac{1}{e}(1-\kappa) \cdot cov\left(g(z), (1-p)\frac{z}{\bar{z}}\right)$, where p is the marginal propensity to save, $\kappa = -d\log(1-p)/d\log(1-\tau)$ is the crowding-out elasticity, and $e = d\log(\bar{z})/d\log(1-\tau)$ is the elasticity of earnings with respect to the (after-tax) wage.

We rewrite the formula in Chetty and Saez (2010) for a better comparison to our formula:

$$\frac{\tau}{1-\tau} = -\frac{1}{e}cov\left(g(z), \frac{z}{\bar{z}}\right) - \left[p + \frac{1}{e}cov\left(g(z), \{p + \kappa(1-p)\}\frac{z}{\bar{z}}\right)\right]. \tag{7}$$

The first term in (7) reflects the standard equity-efficiency trade-off, which is essentially the same as the first term ($A(z^*)$) in our formula (5). The second term in (7) reflects the welfare effects of the change in private insurance ($dP(z(x))$), which can be more clearly seen by:

$$p + \frac{1}{e}cov\left(g(z), \{p + \kappa(1-p)\}\frac{z}{\bar{z}}\right) = \frac{1}{e \cdot \delta\tau \cdot \bar{z}} \int g(x) \cdot dP(z(x)) f(x) dx,$$

[14] More precisely, Chetty and Saez (2010) consider a wage compression as a form of private insurance, which is insurance against the before-tax income. The timing of the private insurance — before or after the tax payment —- does not change the optimal tax formula as long as the elasticity of the private insurance with respect to the tax rate is appropriately defined.

where $dP(z(x)) = -\{p + \kappa(1-p)\} \cdot \delta\tau \cdot [z(x) - \bar{z}] - p(1-\tau)d\bar{z}$, and $d\bar{z} = -e\bar{z}\frac{\delta\tau}{1-\tau}$. Thus, the counterparts of the second term in (7) are the second and third terms ($B(z^*)$ and $C(z^*)$) in our formula (5). The key difference of formula (7) compared to our formula (5) is that the envelope theorem does not apply at all to the change in private insurance, which will be discussed in more detail in Section 3.

We learned two important lessons 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 the 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 save can be negative, if households are allowed to borrow for consumption smoothing. Second, the presence of private savings can generate a greater inequality in consumption, which amplifies the Saez (2001) effects. Third, the pecuniary externalities have either a positive or a negative sign depending on the change in the equilibrium interest rate as a result of tax reform. Taking 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.[15] 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 in the presence of a 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 the market structure—i.e., frictions in the private insurance market.

[15] 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 a positive private savings rate and a *own* crowding-out effect.

3 Implications for the Sufficient-Statistics Approach

So far, we have analyzed the optimal tax formula under a specific market structure, Huggett (1993). Ideally, one would like to extend the formula to a more general market structure of private insurance and express the formula in terms of sufficient statistics that can be easily estimated. In this section, we show that such attempts are extremely challenging for (at least) two reasons. First, the formula depends on the welfare effects from the interaction between private and public insurance; thus, its expression depends on the specific structure of the 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 a private insurance market, because they are not likely to be policy invariant.

3.1 Optimal Formula with a General Incomplete Market

We first examine how much the optimal tax formula can be extended to a more general private insurance market. Our analysis will show that despite a general representation of a wide 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. By illustrating how the optimal formula is modified to a few well-known structures of a private insurance market, we can better understand the economic insights of the results in the previous works as well as ours.

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 a noncontingent 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 the 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:$\int P(\cdot) = 0$ but this can be extended to a more general case. In a Huggett economy, $P_t(x_t, a_t) = a_{t+1}(x_t, a_t) - (1+r)a_t$.

We derive the optimal tax formula using the same perturbation (elementary tax reform). Suppose that in period 0, the economy is in a steady state with the distribution 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. If we maintain the assumption of GHH preferences, then the incidences of tax reform on the labor supply and government revenue are exactly the same as those in Section 2. However, the incidence of tax reform on private intermediation $P_t(x_t, \mathbf{s}_t)$ depends on market structure. The optimal tax formula can be obtained by imposing $dW = 0$:

$$\begin{aligned} \frac{T'(z(x^*))}{1-T'(z(x^*))} &= \left(1+\frac{1}{e(x^*)}\right)\frac{1-F(x^*)}{x^* f(x^*)} \\ &\quad \times(1-\beta)\sum_{t=0}^{\infty}\beta^t \left[\begin{array}{l} \int\int_{x^*}^{\infty}(1-g(x,\mathbf{s}))\frac{\phi(x,\mathbf{s})}{1-F(x^*)}dxd\mathbf{s} \\ -\int\int g(x^t,\mathbf{s}_0)dP_t(x^t,\mathbf{s}_0)f(x^t|x_0)dx^t d\mathbf{\Phi}(x_0,\mathbf{s}_0) \end{array}\right] \end{aligned} \tag{8}$$

where $dP_t(x^t, \mathbf{s}_0)$ denotes the incidence of tax reform on private intermediation in period t. Without further information on the market structure of private insurance, we cannot proceed any more. We now discuss three cases (of the market structure) that illustrates how the envelope theorem is applied to $dP_t(x^t, \mathbf{s}_0)$, the response of private intermediation with respect to a tax reform.

3.1.1 Case 1: 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 (8) can be ignored when computing the optimal tax rate. We first consider an example from Chetty and Saez (2010): a spot market with a linear payment schedule.[16] With a spot market, we do not need an additional state variable, and the private intermediation can be expressed as follows:

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

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

where p is a time-invariant 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 incidences of tax reform on the labor supply and the government revenue are exactly the same as those in Section 2.[17] On the other hand, the incidence on private intermediation given income $z(x)$ yields an analytical expression:

$$dP(z(x)) = -\kappa(x^*)\frac{1-p}{1-T'(z(x^*))}[y(x) - \bar{y}] - p\left[\tau(z(x)) - dR\right] - p \cdot d\bar{y},$$

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

That is, private intermediation can vary via changes in (i) the payment rate, (ii) after-tax income, and (iii) the transfer from the private intermediaries. Since one cannot apply the envelope theorem to these exogenous changes in private intermediation, all the changes captured in $dP(z(x))$ have welfare effects.[18] Then, the second term in the bracket of (8) in this linear spot market is:

$$\int \frac{u'(x)}{\lambda}\left[p\tau(z(x)) + \kappa(x^*)\frac{1-p}{1-F(x^*)}\frac{(y(x)-\bar{y})}{1-T'(z^*)}\right]f(x)dx - p\int \tau(z(x))f(x)dx - p\cdot\epsilon^{l}_{1-T'}(x^*)\frac{z(x^*)}{z'(x^*)}\frac{f(x^*)}{1-F(x^*)}.$$

By rearranging the terms, we obtain the optimal formula consistent with that in Chetty and Saez (2010) (see the Appendix for the derivation):

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

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

[17] The individual first-order condition with respect to the 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 the retention rate $1 - T'$ along the nonlinear budget constraint does not change. See the Appendix.

[18] Note that change in private intermediation given productivity x is $dP(x) = dP(z(x)) + p(1 - T'(z(x)))xdl(x)$, and the term associated with $dl(x)$ can be ignored (by applying the envelope theorem) in computing the welfare effects of tax reform. See the Appendix.

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 (8) 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, under the GHH preferences (no wealth effects on the labor supply), consumption net of the utility cost of labor 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 $E[dP(x)] = 0$.[19] With complete markets, not only the second term but also the first term in the bracket of optimal tax formula (8) is zero, which implies that the 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 a non-state-contingent bond but with an exogenous interest rate and a natural borrowing limit. In this case, the response of private intermediation $dP(a_0, x^t) = da_{t+1}(a_0, x^t) - (1+r)d_at(a_0, x^{t-1})$ does not have any welfare effects, because we can apply the envelope theorem to the total change in private intermediation. Recall that in a 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 (8) 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 private and public insurance in the optimal tax formula.

Findeisen and Sachs (2017) and Saez and Stantcheva (2017) consider this type of incomplete market: a constant interest rate with a natural borrowing limit. Our analysis

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

illustrates 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

Individual households' optimal responses (to a tax reform) may not have any effect on social welfare, *if* the induced changes in savings neither generate any externality nor affect the degree of market frictions: the envelope theorem is fully applied. A more realistic case would be an intermediate one where the envelope theorem is partially applied due to various frictions in the market.

The benchmark economy we consider is a good example of this kind. The private intermediation in a 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, the change in savings of households whose borrowing constraint is not binding does not have direct welfare effects (the envelope theorem applies). The change, however, can have an indirect welfare effect through the change in the interest rate, because it generates pecuniary externalities. In addition, the envelope theorem does not apply to the change in savings of borrowing-constrained households.

Another example is an endogenous incomplete market with limited commitment (Alvarez and Jermann (2000); Kehoe and Levine (1993)). In this market, households can trade Arrow-Debreu securities subject to credit lines $\bar{A}_{t+1}(x^t, x_{t+1})$ that are contingent on the history of productivity. The 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 determined 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 $V_t(a, x^t)$ denotes the continuation value of a household with history x^t and asset holding a in period t. The value of the autarky is:

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

We denote the price of a risk-free bond by $q_t = \frac{1}{1+r_{t+1}}$ and the no-arbitrage condition implies $q_t(x^t, x_{t+1}) = f(x_{t+1}|x_t t) q_t = \frac{f(x_{t+1}|x_t)}{1+r_{t+1}}$. 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, the second term in the bracket of optimal tax formula (8) becomes:

$$\begin{aligned} & -dq_t \cdot \int \int f(x^t|x_0) u'(a_0, x^t) \int f(x_{t+1}|x_t) a_{t+1}(a_0, x^t, x_{t+1}) dx_{t+1} dx^t d\Phi(a_0, x_0) \\ & - \int \int f(x^{t+1}|x_0) [q_t u'(a_0, x^t) - \beta u'(a_0, x^{t+1})] da_{t+1}(a_0, x^t, x_{t+1}) dx^{t+1} d\Phi(a_0, x_0), \end{aligned} \tag{9}$$

where the first term reflects the pecuniary externalities, and the second term represents the welfare effects of changes in savings by the borrowing-constrained households.

Although these terms look similar to those in the Huggett economy, the sign and the source of pecuniary externalities are quite different between the two economies. With state-contingent assets, the consumption poor in the current period want to borrow from a high-productivity state in the future. But this is limited due to 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 high, which makes the 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: the more borrowing-constrained households are, the lower the interest rate (the higher q_t) is. Thus, a progressive tax reform will tighten the endogenous borrowing limit and increase the price of assets (q_t), which in turn has negative welfare effects because poor households purchase more assets (higher $E[a_{t+1}(a_0, x^t, x_{t+1})]$).

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. 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 estimate.

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 the 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., the marginal propensity to save across all individual states.[20] 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 in Chetty (2009). We obtain these (hard-to-estimate) statistics using a quantitative general equilibrium model.

4 A Quantitative Analysis

4.1 Calibration

We first assume that 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.

[20]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.

We approximate an optimal nonlinear tax and private intermediation with a piecewise-linear over N grid points.[21]

Preferences, Government Expenditure, and Borrowing Constraints

The households' utility function exhibits 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.[22]

We choose the discount factor (β) so that the equilibrium interest rate 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–which is approximated by a log-linear functional form: $T(z) = z - \lambda z^{1-\tau}$ as in Heathcote, Storesletten, and Violante (2014).[23]

The exogenous borrowing constraint ($\underline{a} = -90.84$) is set so that 10% of households are borrowing-constrained under the current U.S tax schedule. This value of the borrowing limit is close to the average annual earnings of households in our model economy (under the current U.S. tax schedule), which is also in the range of credit card limits (between 50% $\sim$ 100% of average annual earnings) in the data.[24] 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, we perform the sensitivity analysis with respect to different values of σ, e, and $\underline{a}$.

Productivity Process

[21]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.

[22]There is ample evidence that an intertemporal 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 empirically plausible values of labor supply elasticity range between 0.2 and 1, according to the survey article by Keane and Rogerson (2012).

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

[24]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.

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

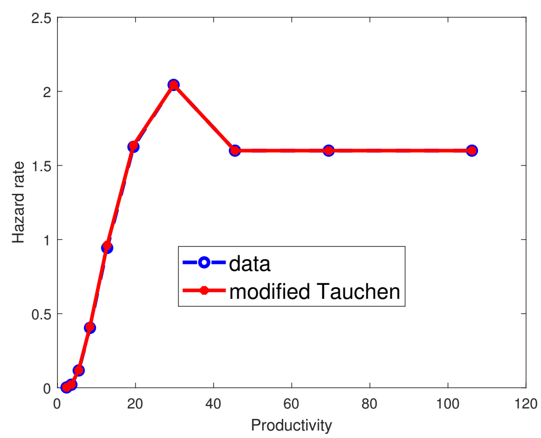
The shape of the income distribution (which is dictated by the stochastic process of productivity given our preferences of no wealth effect on the labor supply) is crucial for the shape of the optimal tax schedule. We generate an empirically plausible distribution of productivity as follows. Consider an AR(1) process for log productivity x: $\ln x' = (1-\rho)\mu + \rho \cdot \ln x + \sigma_\epsilon \cdot \epsilon'$, 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 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

[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.

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, $\pi(x_1|x)$, 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, the hard-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.

[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).

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 very 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 MPCs across households. Jappelli and Pistaferri (2014) provide estimates of MPCs by quintiles of 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. For this reason, it would not be fair to directly compare the levels of MPC between the model and the available estimates.[27]

Thus, we rather focus on the comparison of the *relative* MPCs between the model and data. Table 2 reports the relative MPCs (in the 1st and 5th quintiles of the income and asset distributions relative to the mean) in our model to those in Jappelli and Pistaferri (2014). The relative MPCs in the data at the 1st and 5th quintiles are computed using the estimates of 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 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

[27] In fact, previous studies show that the average MPC in a Huggett-style incomplete markets model is 0.22 with respect to a one-time unexpected increase in income (Dupor, Karabarbounis, Kudlyak, and Mehkari (2017)), which is comparable to the 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).

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).

Distributions of Income and Assets

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 of the wage distribution (shown in Figure 1). Table 3 shows that the Gini coefficient of earnings in our model is 0.51, not far from those in the U.S. (0.53 - 0.67). The distribution of assets is not that 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 (because the negative assets are imputed as zeros) 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 between tails (e.g., between the top and bottom 1%) of the asset distribution, $\frac{a_{99}-a_{01}}{a_{60}-a_{40}}$, is only 39 in the model, much smaller than the 73 in the data.[28] The

[28] As is well known, this type of incomplete markets model has difficulty generating super-rich households.

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).

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.

4.3 Optimal Tax Schedule

In our quantitative analysis, for computational convenience, we focus on the optimal tax formula under the so-called "utility-based steady-state" approach. This method, proposed by Saez and Stantcheva (2018), further simplifies the tax formula to (by replacing some values during the transition with the steady-state values):[29]

$$\frac{T'(z^*)}{1-T'(z^*)} = \frac{1}{\epsilon^l_{1-T'}(z^*)} \cdot \frac{1-F_z(z^*)}{z^* f_z(z^*)} \cdot [A(z^*) + B(z^*) + C(z^*)] \tag{10}$$

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

where

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

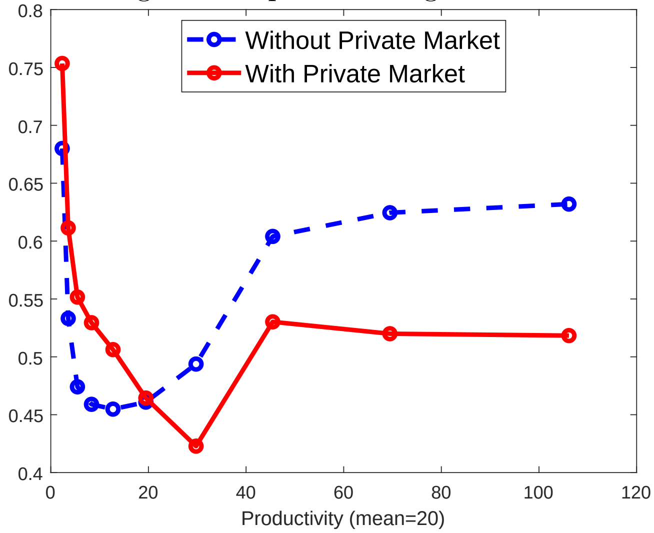
$\lambda = \int u'(a,z)\phi(a,z)da\,dz$, and $g(a,z)=\frac{u(a,z)}{\lambda}$.

The algorithm to find the optimal marginal tax rates is a modification of Brewer, Saez, and Shephard (2010). Starting with a given vector of T', we compute the competitive equilibrium and necessary statistics, including the hard-to-estimate ones such as dr, $dh^A(a,y(z))$, $\phi(a,z))$, and $g(a,z)$, for a tax reform at each of 10 income brackets. We then use the formula to compute the new vector of T'. More precisely, we use the formula with respect to the exogenous productivity distribution—formula (6). We repeat the algorithm until the vector T' converges to a fixed point.

Figure 2 shows the optimal marginal tax schedule across productivity with and without a private insurance market. We normalize the units of quantities in our model so that the average productivity (wage) is \$20. 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 (e.g., Diamond (1998) and Saez (2001)). High marginal tax rates at the very low-income levels indicate that net transfers to low-income households should quickly phase out. (See also Figure 8 in the Appendix for the average tax rates.) 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.

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 Figure 2 , 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

Figure 2: Optimal Marginal Tax



those without private insurance. The difference is as large as 11 percentage points: 64% (without private market) vs. 53% (with private market) at the top income bracket.

We now examine the factors that account for the difference in optimal tax rates with and without private insurance quantitatively. Comparing our optimal tax formula (10) 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 the distribution of the marginal utility of consumption now depends on assets as well as income. The second term, $B(z^*)$, reflects the effect of pecuniary externalities (Dávila, Hong, Krusell, and Ríos-Rull (2012)). For example, an increase (or decrease) in the equilibrium interest rate as a result of tax reform makes the poor worse off (better off), undermining (improving) social welfare. The third term, $C(z^*)$, captures the effect of changes in savings of borrowing-constrained households 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:

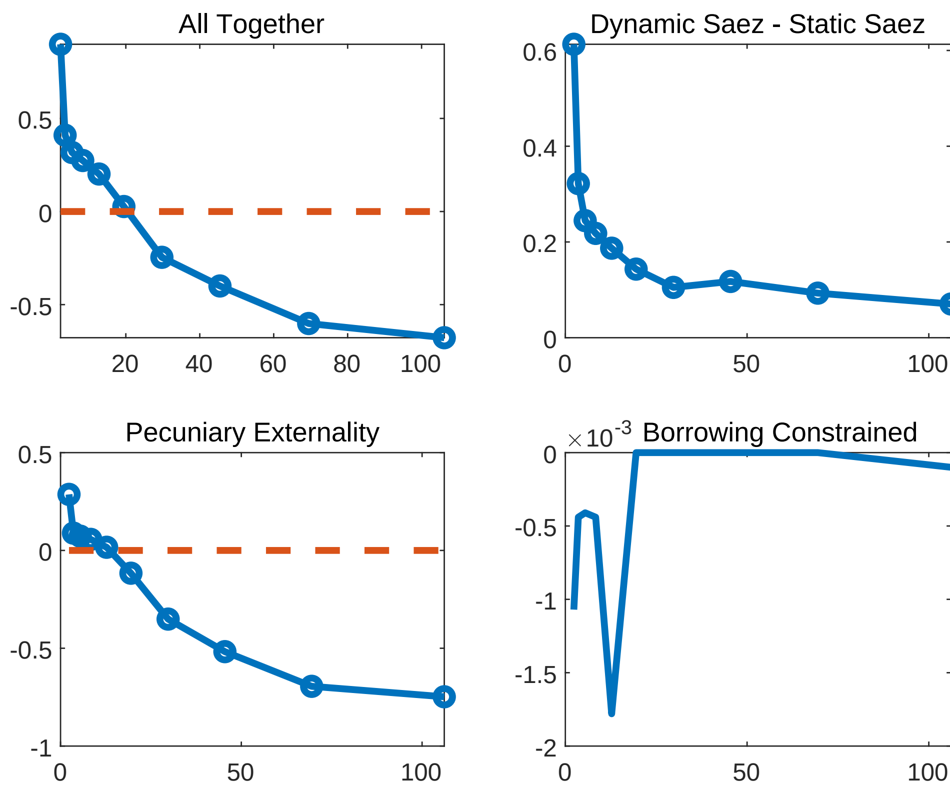
$$\frac{T'(z^*)}{1-T'(z^*)} - \frac{T'_{Saez}(z^*)}{1-T'_{Saez}(z^*)} = \frac{1}{\epsilon^l_{1-T'}(z^*)} \cdot \frac{1-F_z(z^*)}{z^* f_z(z^*)} \cdot \left[A(z^*) - A^{Saez}(z^*) + B(z^*) + C(z^*)\right], \quad (11)$$

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 3 plots each of these three components. The first figure shows 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 the 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-constrained households, $\frac{1}{\epsilon^{l}_{1-T'}(z^*)}\frac{1-F_z(z^*)}{z^* f_z(z^*)}C(z^*)$.

Figure 3: Decomposition of the Difference in $\frac{T'}{1-T'}$ with and without private insurance



The distribution of consumption becomes more dispersed in an economy with private savings (due to a skewed distribution of wealth). 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 the pecuniary externality, however, depends on the sign of the equilibrium interest rate movement as a result of tax reform. This effect is positive for an increase in marginal tax at the low-income bracket (a wage rate of less than \$15). But it becomes negative for an increase in the marginal tax rate for the middle- and high-income brackets. The magnitude of this effect is quite large at the top income bracket. A progressive tax reform of increasing the marginal tax rate at high-income brackets reduces the precautionary motive of savings and results in an increase in the equilibrium interest rate. This makes the poor (i.e., borrowers) worse off, which makes a progressive tax reform less effective in achieving redistribution. As this effect starts dominating the "Dynamic Saez - Static Saez" term (at wage rates above \$20), the optimal tax rate becomes lower than that without a private insurance market. Finally, the third term, the effect of borrowing-constrained households, is always negative (as we showed above) but quantitatively negligible. Taking all three terms together, the optimal tax rates are higher for the low-income group (wages below \$20) but lower for the 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.

We now compare social welfare under three tax schedules: the optimal tax with a private market (based on our formula), the optimal tax without a private market (Saez (2001)), and the current U.S. tax system (approximated by a log-linear form as in HSV). Social welfare is compared based on a constant-compensating differential in (steady-state) consumption. The welfare cost (or gain) of a tax system T relative to our optimal tax formula (T^*) is Δ that satisfies:

$$\begin{aligned} \int V(x,a;\Delta)\, d\Phi(x,a;T) &= SWF(T^*) \\ V(x,a;\Delta) &= u((1+\Delta)c(x,a) - v(l(x,a))) + \beta E[V(x',a'(x,a);\Delta)|x] \end{aligned}$$

where $\Phi(a,x,T)$ is the steady-state distribution under tax system T and $SWF(T^*)$ is the steady-state social welfare under our optimal tax (with a private market) schedule T^*. According to this measure, the welfare cost of the U.S. tax system compared to our optimal tax schedule is 12.1%. As we discussed above, it may generate higher welfare if the government simply shuts down the private insurance market (if this is feasible at no cost) and adopts the optimal tax without a private market (i.e., the original Saez (2001) formula). Indeed, social welfare under the optimal tax without a private market is higher than that with a private market by 7.1% (i.e., $\Delta = -7.1\%$).

This result is not actually surprising, given that the market structure of private intermediation assumed in our quantitative analysis is rather primitive—the only asset available for households is a noncontingent bond for self-insurance. If the private market has richer tools for intermediation and faces fewer frictions than the government does, the welfare results can be quite different: social welfare under the optimal tax in the presence of a private market could be higher than that under the optimal tax without a private market.

Finally, we compare our optimal tax schedule to the current U.S. income tax rates. Figure 4 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).[30] We also plot the median of the effective marginal tax rate of low- and moderate-income workers (single parents with one child) in 2016 published by the Congressional Budget Office (in red circles) as well as the marginal income tax rates based on the Federal Statutory Income Tax Schedule (in blue dots).

The optimal marginal tax rates are higher than the current ones for all income groups.[31] However, for the top income group (i.e., individual income ranges above \$250K), the current tax rates are not so far from optimal (54%). This result is very different from those without private insurance, where the optimal tax rates (65%) are much higher than the current ones.

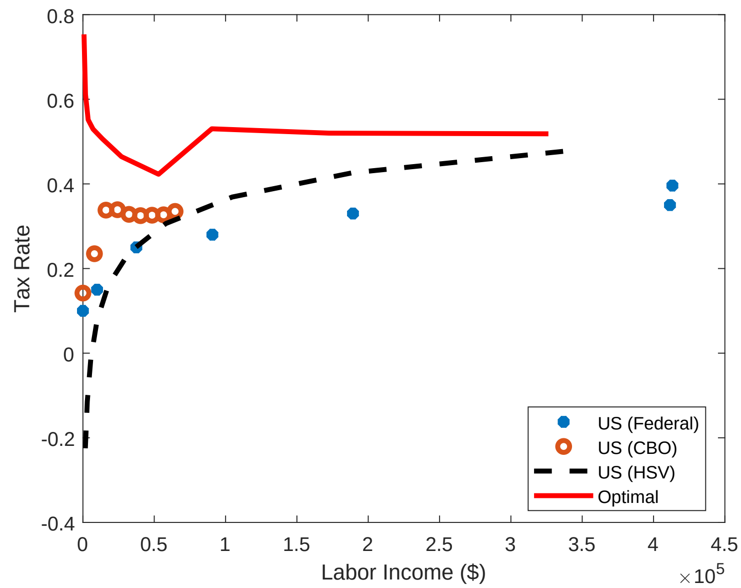
4.4 Comparative Statics

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. Although the exact tax schedule depends on these specifications, the pattern

[30]The HSV functional form does not allow any lumpsum transfer to the zero income. Thus, the HSV approximation cannot capture the transfer schedule at the bottom very well.

[31]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.

Figure 4: Marginal Tax Rate: Current vs. Optimal



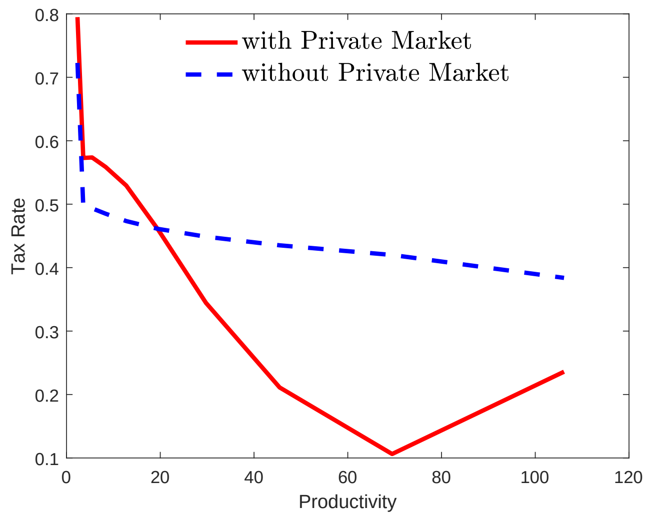
Note: "US (Federal)" reflects the statutory federal income tax rates for singles in 2015. "US (CBO)" shows the median of effective marginal tax rates for low- and moderate-income workers (single parent with one child) in 2016 published by the CBO.

of change in the optimal tax schedule by the introduction of private insurance is the same across all sensitivity analyses: the optimal tax rates are higher (than those without a private insurance market) at the low-income level mainly because of greater inequality in consumption, whereas they are lower at the middle- and high-income levels due to the dominating effects from pecuniary externalities.

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 5. The same 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, under the log-normal distribution, there are only a small fraction of workers at the top income bracket. Thus, the effect of

Figure 5: Log-normal Distribution of Productivity



pecuniary externalities associated with the tax reform at the top is relatively small, which results in a slightly increasing marginal tax rate schedule at the top.

4.4.2 Persistence of the 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 not being history dependent. 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 re-calibrate 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. Thus, the change in persistence does not have any impact on the optimal tax schedule without a private market.

Figure 6 shows that the introduction of private insurance changes the tax schedule in a similar way as in the benchmark. It increases the optimal tax rates for the very low-income group and decreases them for the middle- and high-income groups. But the difference (of the optimal tax rates with and without a private market) is more pronounced when productivity shocks are less persistent ($\rho = 0.8$). The decomposition of the difference in

Figure 6: Optimal Tax with Private Market

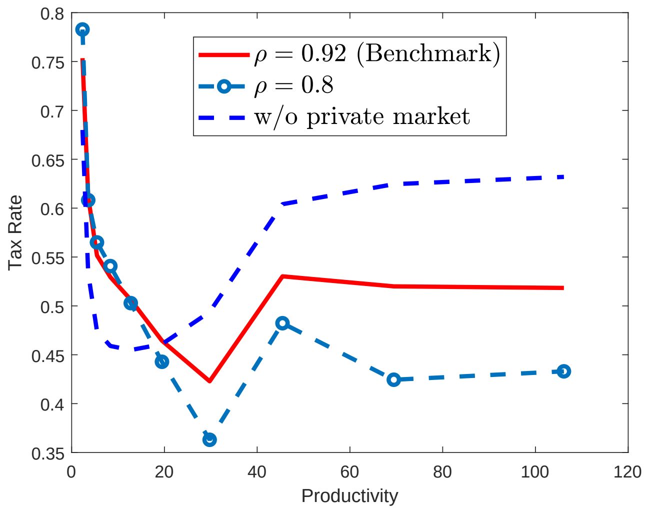


Figure 7: Decomposition ($\rho = 0.8$)

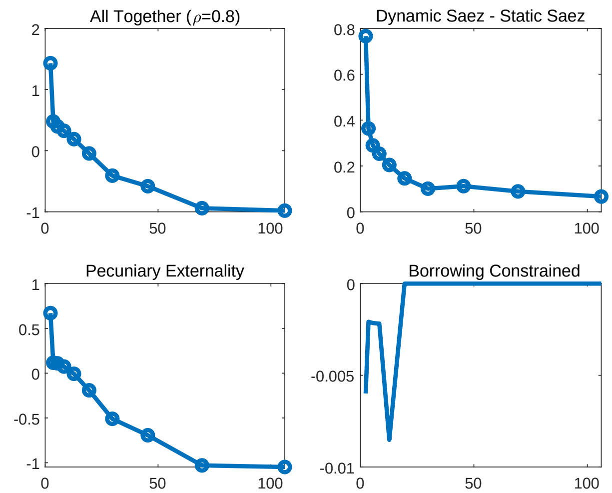


Figure 7 shows that the effects of both Saez and the pecuniary externalities are stronger under less persistent shocks.

A less persistent stochastic productivity (given the same size of overall income risk), generates a larger dispersion in the wealth distribution because the precautionary savings motive is stronger—high productivity does not last long. This has two effects on the optimal tax rates. (i) It magnifies the Saez effect, which makes the "Dynamic vs. Static Saez" term particularly large at the low-income bracket. For example, the size of this term is 0.8 for the lowest productivity level when $\rho = 0.8$, whereas it was 0.6 under the benchmark. (ii) At the same time, it leads to bigger pecuniary externalities for higher-income brackets, making the optimal tax rate lower than that without. For example, at a wage rate of \$70, the size of this term is close to -1, whereas it was -0.7 under the benchmark specification.

5 Generalizations

In this section we discuss how to generalize the results of Section 2 when we augment our model economy to accommodate (i) productive physical capital, (ii) capital income tax, (iii) generalized social welfare, and (iv) wealth effects on the labor supply. We provide brief descriptions of these extensions here. The details are relegated to the Appendix.

5.1 Aiyagari Economy with Physical Capital

The main result from the benchmark (endowment) economy can be extended to a production economy; e.g., Aiyagari (1994). Suppose that the aggregate production function exhibits a constant returns to scale in aggregate capital K_t and labor L_t: $F(K_t, L_t)$ where $K_t = \int a_t \, d\Phi(a_t, x_t)$ and $L_t = \int x_t l(x_t) \, d\Phi(a_t, x_t) = \int x_t l(x_t) f(x_t) \, dx_t$.[32] We assume that both goods and factor markets are perfectly competitive so that the equilibrium factor prices are marginal products: $r_t = F_K(K_t, L_t) - \delta$ (where δ is the depreciation rate of capital), and the wage rate (for an efficiency unit of labor) $w_t = F_L(K_t, L_t)$. The household's budget constraint becomes: $c(a_0, x^t) + a_{t+1}(a_0, x^t) = w_t x_t l(x_t) - T(w_t x_t l(x_t)) + (1 + r_t) a_t(a_0, x^{t-1})$.

In this economy, even with GHH preferences (no wealth effects on the labor supply), the incidence of tax reform on labor supply is complicated 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 change in the wage rate is:

$$\begin{aligned} dw_t &= -\alpha_L^w(K_t, L_t) \cdot w_t \int \frac{x_t' dl(x_t')}{L_t} f(x_t') dx_t' \\ &\quad + \alpha_K^w(K_t, L_t) \cdot w_t \iint f(x^{t-1}|x_0) \frac{da_t(a_0, x^{t-1})}{K} dx^{t-1} d\Phi(a_0, x_0), \end{aligned} \tag{12}$$

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

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

where $\epsilon_w^l(x) = \frac{d \log l(x)}{d \log w} = \epsilon_{1-T'}^l(x)(1 - \rho(w x l(x)))$ denotes the elasticity of the labor supply with respect to the wage rate along the nonlinear budget constraint. The response of savings $da_t(a_0, x^{t-1})$ measures the causal impact of tax reform on savings, including the income effects of a tax rebate, which in turn depends on $dl_t(x_t)$. Thus, obtaining a

[32] We implicitly assume that all workers are perfect substitutes for each other. This greatly simplifies the labor market equilibrium because the equilibrium wage rate depends on the aggregate effective units of labor only (not the distribution of productivity across workers).

closed-form solution of the tax incidence on labor supply becomes highly challenging. We derive the optimal tax formula *given* the incidence on the wage rate dw_t. This will still illustrate the economic mechanism behind the formula.

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

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

where $\alpha_K^r(K, L) = -\frac{d\log r}{d\log K} = -F_{KK}\frac{K}{F_K-\delta}$ and $\alpha_L^r = \frac{d\log r}{d\log L} = F_{KL}\frac{L}{F_K-\delta}$. In turn, this change in the interest rate affects saving (as in the benchmark economy without production capital). We obtain the optimal tax formula by imposing $dW = 0$.

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

$$\frac{T'(z^*)}{1-T'(z^*)} = \frac{1}{\epsilon_{1-T'}^l(z^*)}\cdot\frac{1-F_z(z^*)}{z^* f_z(z^*)}\cdot(1-\beta)\sum_{t=0}^{\infty}\beta^t\left[A_t(z^*)+B_t(z^*)+C_t(z^*)+D_t(z^*)\right], \quad (13)$$

where

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

Formula (13) is different from (5) in our benchmark Huggett economy in two ways. First, there is an additional term $D_t(z^*)$, which captures additional fiscal externalities due to changes in the equilibrium wage rate. Second, the pecuniary externalities term $B_t(z^*)$ captures two channels of pecuniary externalities as:

$$B_t(z^*) = dr_t K\Bigg[\underbrace{\int g(a,z)\left(\frac{a}{K}-1\right)d\Phi(a,z)}_{\equiv\Delta_K} - \underbrace{\int g(a,z)\left(\frac{(1-T'(wz))z}{L}-1\right)d\Phi(a,z)}_{\Delta_L}\Bigg]. \quad (14)$$

Compared to $B_t(z^*)$ in the benchmark formula (5), there is one more term in the bracket (Δ_L), 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 households' income–labor earnings, which makes up a large portion of the total income of poor households, is scaled down. Thus, the sign of the additional pecuniary externalities term (Δ_L) is opposite to that of Δ_K (pecuniary externalities through redistribution).

According to Dávila, Hong, Krusell, and Ríos-Rull (2012), under a realistic calibration, the pecuniary externalities through redistribution (Δ_K) dominate the pecuniary externalities through insurance (Δ_L). Thus, the sign of the $B_t(z^*)$ term is not likely to change. Moreover, the sign of the additional term in $D_t(z^*)$ is opposite to that of dr_t, which also makes the optimal tax schedule less progressive (than the one without a private market).

5.2 Capital Income Taxes

For simplicity, we abstract from capital income tax in our benchmark analysis. We show that the same formula and intuition carry over to the economy with a capital income tax (more exactly, tax on income from asset holdings in a Hugget economy), 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 and history dependent), the economy goes back to the complete market case, and there is no need to use labor income taxation to provide insurance. However, with typical restrictions (e.g., history independence), the capital income tax cannot complete the market.

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

However, if the capital tax is levied on positive capital income only, the government revenue from the capital tax is not zero. Under the nonlinear capital income taxation, there will be an additional term in the formula because of fiscal externalities caused by the 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 the 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 the capital income tax, where dr_t and $da_t(a_0,x^{t-1})$ measure the causal effects considering all chains of responses.

5.3 Generalized Social Welfare Function

We derive the optimal tax formula under the utilitarian social welfare function. We show that the main result from the benchmark can be extended to a more general social welfare criterion. Consider a government that evaluates social welfare according to:

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

where $G(\cdot)$ is an increasing and concave function that reflects the social preferences for redistribution. The utilitarian social welfare function is a special case where $G(V) = V$.

Note that in a dynamic economy with incomplete markets, a concave $G(\cdot)$ reflects society's preference for redistribution across assets as well as productivity. Under a more general social welfare function, the optimal tax formula becomes:

$$\begin{aligned} \frac{T'(z(x^*))}{1-T'(z(x^*))} &= \frac{1+e(x^*)}{e(x^*)} \frac{1-F(x^*)}{x^* f(x^*)} (1-\beta) \\ &\times \sum_t \beta^t \Big[\iint \Big[1-g(a_0,x^t)\Big] \mathbb{1}_{\{x_t \geq x^*\}} f(x^t|x_0) dx^t d\Phi(a_0,x_0) \\ &+ dr_t \iint g(a_0,x^t) a_t(a_0,x^{t-1}) f(x^t|x_0) d\Phi(a_0,x_0) \\ &- \iint g(a_0,x^t) \left(1-\beta(1+r)E\left\{\frac{u'(a_0,x^{t+1})}{u'(a_0,x^t)}\middle| x^t\right\}\right) dh^A(a_t(x^{t-1}), y(x_t)) f(x^t|x_0) dx^t d\Phi(a_0,x_0)\Big], \end{aligned} \tag{15}$$

where $g(a_0, x^t) = \frac{G'(V(a_0,x^t))u'(a_0,x^t)}{\lambda}$ and $\lambda = (1-\beta)\sum_t \beta^t \iint G'(V)u'(a_0, x^t)f(x^t|x_0)dx^t d\Phi(a_0, x_0)$. This is almost identical to the benchmark except that the marginal social welfare weight $g(a_0, x^t)$ reflects the concavity of $G(\cdot)$. Since the marginal social welfare depends on the initial state (a_0, x_0), we should evaluate the welfare effects using the distribution of the history,

Next, we consider a government that is concerned about horizontal equity. More specifically, consider the following social welfare function:

$$W = \sum_{t=0}^{\infty} \beta^t \int G(u(a_t, x_t))d\Phi(a_t, x_t),$$

where social utility $G(\cdot)$ is a function of flow utility. Compared to the benchmark formula (5), there are two differences. First, as in the case of the general social welfare function (just described above), the marginal social welfare weight is $g(a_t, x_t) = \frac{G'(u(a_t,x_t))u'(a_t,x_t)}{\lambda}$, where $\lambda = \int G'(u(a_t, x_t))u'(a_t, x_t)d\Phi(a_t, x_t)$. Second, and more important, the change in asset holdings will have welfare effects even for households that are not borrowing constrained, i.e., the envelope theorem is not applied. An individual Euler equation is not optimal from the perspective of social welfare, because the ratio of the marginal utilities of consumption between the present and the future is not exactly aligned with that of society. The third term in formula (6) will be replaced by:

$$\iint [g(a_t(a_0, x^{t-1}), x_t) - \beta(1+r)E[g(a_{t+1}(a_0, x^t), x_{t+1})|x^t]]da_{t+1}(a_0, x^t)f(x^t|x_0)dx^t d\Phi(a_0, x_0).$$

5.4 Labor Supply with Wealth Effects

When there are wealth effects on the labor supply, the tax incidence (on labor supply) should consider both substitution and income effects, as in Saez (2001). Moreover, in the presence of a private insurance market, we also need to consider the income effects through the change in savings and interest rates. Given the causal effect of a tax reform on savings and interest rate, (da_{t+1} and dr_t), which reflect all possible responses, the incidence on the labor supply can be expressed by

$$\begin{aligned} dl_t(a_0, x^t) \quad &= \quad -\epsilon^c_{1-T',l}(a_t(a_0, x^{t-1}), x_t)\frac{\tau'(x_t l_t)}{1 - T'(x_t l_t)} l_t \qquad (16) \\ &\quad + \frac{\epsilon_{R,l}(a_t(a_0, x^{t-1}), x_t)}{x_t(1 - T'(x_t l_t))}[-\tau(x_t l_t) - da_{t+1}(a_0, x^t) + (1 + r)da_t(a_0, x^{t-1}) + dr_t a_t(a_0, x^{t-1})], \end{aligned}$$

where $\epsilon^c_{1-T',l}(a_t, x_t)$ denotes the compensated elasticity of the labor supply with respect to the retention rate along the nonlinear budget constraint, and $\epsilon_{R,l}(a_t, x_t)$ denotes the income effect along the nonlinear budget constraint (see the Appendix for details).

With wealth effects, the labor supply decision depends on asset holdings as well as current productivity. Thus, labor income $z(a_t, x_t)$ and productivity x_t do not have a one-to-one relationship, and the elementary tax reform at income level z^* will change the marginal tax rate of the worker with state $(a, x^*(a) = z^{-1}(a, z^*))$, where $z^{-1}(a, z^*)$ denotes the value of the inverse of $z(a, \cdot)$ at z^* given a. We denote the average compensated elasticity of labor by $\epsilon^c_{1-T',l}(z^*) \equiv \int \epsilon^c_{1-T',l}(a, x^*(a))\phi(a|x^*(a))da$.

Incorporating the additional income effects on government revenue, the optimal tax formula will be

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

where $A_t(z^*), B_t(z^*)$, and $C_t(z^*)$ are the same as those in (5), and the additional $D_t(z(x^*))$ reflects the income effects given the total causal effects on savings and interest rates:[33]

$$\begin{aligned} D_t(z^*) \quad &= \quad \iint \epsilon_{R,l}(a_t(a_0, x^{t-1}), x_t) \frac{T'(x_t l_t)}{1-T'(x_t l_t)} \\ &\times \left[\frac{\mathbb{1}_{\{z(a_t(a_0,x^{t-1}),x_t)\geq z^*\}}}{1-F_z(z^*)} - da_{t+1} + (1+r)da_t + dr_t a_t\right] f(x^t|x_0)dx^t d\Phi(a_0, x_0). \end{aligned}$$

6 Conclusion

We study a fully nonlinear optimal income tax schedule in the presence of a private (incomplete) insurance market. As in Saez (2001), the optimal tax formula includes standard statistics, such as the Frisch elasticity of the labor supply and the cross-sectional distribution of income. In the presence of a private market, however, these statistics are no longer sufficient. The optimal tax formula depends on how the private market interacts with public insurance and its welfare effects. First, the optimal tax depends on the shape of the joint distribution of assets and income. An economy with an incomplete insurance

[33]The marginal social welfare weight in this formula will be $\frac{u'(a_0,x^t)}{\lambda}$ where λ is the marginal value of the public fund considering the wealth effect on the labor supply.

market is likely to lead to a greater inequality in consumption, which typically calls for a stronger redistribution. Second, the optimal tax schedule should consider its pecuniary externalities. Reforming a tax schedule may change equilibrium prices (e.g., the interest rate), which have differential welfare impacts across households, as shown in Dávila, Hong, Krusell, and Ríos-Rull (2012). This pecuniary externality is likely to prevent the optimal tax schedule from being overly progressive because an overly progressive tax is likely to result in an increase in the market-clearing interest rate (via reduced precautionary savings), which in turn makes poor households (i.e., borrowers) worse off. Finally, the formula should also consider the additional welfare effects of households that are released from the borrowing constraint as a result of tax reform (to whom the envelope theorem does not apply).

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 calibrated to resemble the salient features (such as the cross-sectional distributions of income and the marginal propensity to save) of the U.S. economy . According to our analysis, 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 11 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 (because of highly uneven cross-sectional asset holdings). For the middle- and high-income groups, the optimal tax rates are lower in the presence of a private market due to pecuniary externalities.

References

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*.

Hendren, N. (2016): "The Policy Elasticity," *Tax Policy and the Economy*, 30(1), 51–89.

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.

Scheuer, F., and I. Werning (2018): "Mirrlees meets Diamond-Mirrlees: Simplifying Nonlinear Income Taxation," *Working Paper*.

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 Equations

A.1 Incidence of tax reforms on labor supply

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

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

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

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

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 $\rho(z(x)) = \frac{z(x)T''(z(x))}{1-T'(z(x))}$.

Now let's consider the incidence of tax reforms on the labor supply. We denote the perturbed tax function by $T(z) + \mu\tau(z)$. As 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's 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:

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

This yields the solution for $dl(x)$ as:

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

A.2 Incidence of tax reforms on government revenue

We derive the tax incidence on government revenue for the elementary tax reform. As we discussed in the main text, the elementary tax perturbation $\tau(z) = \frac{\mathbb{1}_{z\geq z^*}}{1-F(z^*)}$ is not differentiable. To apply formula (2) 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 functions $h(\cdot)$ with a 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 implies:

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

Let $\tau_{z^*,\epsilon}(\cdot)$ denote the function such that $\tau'_{z^*,\epsilon} = \frac{\kappa_{z^*,\epsilon}(\cdot)}{1-F_z(z^*)}$, the tax incidence of a tax reform $\tau_{z^*,\epsilon}$ on government revenue $dR(\tau_{z^*,\epsilon})$ is:

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

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.3 Incidence of tax reforms on savings and interest rate

As we discuss in the main text, deriving the analytical expression for the incidence of tax reform on savings and interest rates is a highly challenging task. Here, we derive an integral equation for the incidence on savings and the interest rate instead of attempting to find a closed-form solution.

Denote the Gateaux derivatives of the savings function in period t of a consumer with current state (a, y) in response to a tax reform by $dh_t^A(a, y)$. The savings policy response $dh^A(a, y(x))$ is given by the solution to the perturbed first-order condition:

$$
\begin{aligned}
&u'((1+r+\mu dr_t)a + y(x) - \mu\tau(z(x)) + \mu dR - a' - \mu dh_{t+1}^A(a, y(x)) - v(l(x)))) \\
&= \beta(1+r+\mu dr_{t+1})E\left[\begin{array}{l} u'((1+r+\mu dr_{t+1})(a' + \mu dh_{t+1}^A(a, y(x))) + y(x') - \mu\tau(z(x')) + \mu dR \\ -h^A(a' + \mu dh_{t+1}^A(a, y(x)), y(x')) - \mu\tau(z(x')) - \mu dh_{t+2}^A(a', y(x')) - v(l(x'))) \end{array}\right]
\end{aligned}
$$

A first-order Taylor expansion of the perturbed first-order condition around the baseline allocation yields:

$$
\begin{aligned}
&u'(a, y(x)) + u''(a, y(x))\mu dr_t a + u''(a, y(x))[-\mu\tau(z(x)) + \mu dR] - u''(a, y(x))\mu dh_{t+1}^A(a, y(x)) \\
&= \beta(1+r)E[u'(a', y(x'))] + \beta\mu dr_{t+1}E[u'(a', y(x'))] + \beta(1+r)E[u''(a', y(x'))]\mu dr_{t+1}a' \\
&\quad +\beta(1+r)^2E[u''(a', y(x'))]\mu dh_{t+1}^A(a, y(x)) + \beta(1+r)E[u''(a', y(x'))\{-\mu\tau(z(x'))] + \mu dR\}] \\
&\quad -\beta(1+r)E[u''(a', y(x'))h_a^A(a', y(x'))]\mu dh_{t+1}^A(a, y(x)) + \beta(1+r)E[u''(a', y(x'))h_y^A(a', y(x'))\mu\tau(z(x'))] \\
&\quad -\beta(1+r)E[u''(a', y(x'))\mu dh_{t+2}^A(a', y(x'))].
\end{aligned}
$$

Solving for $dh_{t+1}^A(a, y)$ yields:

$$
\begin{aligned}
dh_{t+1}^A(a, y(x)) \quad = \quad & \frac{u''(a, y)}{\chi}[-\tau(z(x)) + dR] - \frac{\beta(1+r)E[u''(a', y(x'))(1 - h_y^A(a', y(x')))\{-\tau(z(x')) + dR\}]}{\chi} \\
& + \frac{u''(a, y(x))a}{\chi}dr_t - \frac{\beta E[u'(a', y(x'))] + \beta(1+r)E[u''(a', y(x'))]a'}{\chi}dr_{t+1} \\
& + \frac{\beta(1+r)E[u''(a', y(x'))dh_{t+2}^A(a', y(x'))]}{\chi}
\end{aligned}
$$

where $\chi = u''(a, y) + \beta(1+r)^2E[u''(a', y(x'))] - \beta(1+r)E[u''(a', y(x'))h_a^A(a', y(x'))]$.

We now briefly discuss how to express the incidence on the interest rate in terms of the slope of the aggregate asset supply curve and the incidence on aggregate savings given the interest rate. In a Huggett economy, when there is any change in the aggregate asset supply, the interest rate should adjust to clear the market (to guarantee that asset holdings sum to zero). Thus, we can define the semi-elasticity of the interest rate with respect to aggregate assets 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 savings function $h^A(a,x;r)$ given the interest rate. Then, the incidence of tax reforms on the interest rate is:

$$dr_t = \alpha \cdot dA_t^s(r), \tag{17}$$

where $dA_t^s(r) = \iint da_t^s(a_0, x^{t-1}; r) f(x^{t-1}|x_0) dx^{t-1} d\Phi(a_0, x_0)$ is the incidence of tax reform on aggregate savings given the interest rate, including the responses to the lum-psum rebate dR.

A.4 Incidence of tax reforms on individual welfare

Proof of lemma 1: The incidence of tax reforms on individual welfare is:

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

where the second equality is obtained thanks to the envelope theorem of the intratemporal first-order condition.

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 policy function for savings 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

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

where h_a^A and h_y^A are the 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, $h_a^A(a, y(x)) = h_y^A(a, y(x)) = 0$. 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,

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

A.5 Derivation of Optimal Tax Formula

Proof of Proposition 2

We begin by deriving the incidence of tax reforms on social welfare. Denote the Gateaux derivative of social welfare in response to the elementary tax reform by dW:

$$dW = \iint dV(a_0, x_0) \phi(a_0, x_0)\, da_0\, dx_0$$

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, where $(a')^{-1}(\cdot, x)$ is the inverse of $a'(\cdot, x)$ given x. Therefore, for some function $\tilde{h}$ such that $\tilde{h}(a_0, x^t) = h(a_t(x^{t-1}), x_t)$, by applying the change of variables sequentially:

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

This yields:

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

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

By imposing no further improvement in social welfare ($dW = 0$) and with the change of variables, we obtain the optimal tax formula (5):

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

A.6 Derivation of Formula with Exogenous Private Insurance (Case 1)

In Section 3.1.1, we consider an exogenous private insurance, considered in Chetty and Saez (2010), as an example of how the envelope theorem is not applied at all to the changes in private intermediation. Here, we derive the optimal tax formula for this case. With this exogenous spot market, we can write the individual's problem as a repeated static optimization:

$$\max_{l(x)} u\left(xl(x) - T(xl(x)) - p\left(xl(x) - T(xl(x)) - \bar{y}\right) - v(l(x))\right),$$

whose first-order condition is $x(1 - T'(xl(x)))(1-p) = v'(l(x))$.

It is easy to show that the elasticity of the labor supply with respect to the retention rate is equivalent to the one in the benchmark model. The first-order condition when perturbing the retention rate $1-T'(z(x))$ by $dr(x)$ (along the nonlinear budget constraint) is: $v'(l(x)+dl(x)) = x\left[1-T'(xl(x)+dl(x))+dr(x)\right](1-p)$. A first-order Taylor expansion implies:

$$v'(l(x))+v''(l(x))dl(x) = x(1-T'(xl(x)))(1-p)-T''(xl(x))x^2(1-p)dl(x)+x(1-p)dr(x).$$

It is straightforward to show $\epsilon^{l}_{1-T'}(x) = \frac{e(x)}{1+\rho(z(x))e(x)}$, where $e(x) = \frac{v'(l(x))}{l(x)v''(l(x))}$. Using this definition, we obtain the incidence of tax reform on the labor supply as: $dl(x) = -\epsilon^{l}_{1-T'}(x)l(x)\frac{\delta_{z^*}(z(x^*))}{1-T'(z(x))}$, which is equivalent to the one in the benchmark. Then, the incidence of the tax reform on government revenue dR is exactly the same as the one in Appendix A.2.

The private intermediation in this spot insurance market can be expressed as $P(x) = p\cdot(y(x)-\bar{y})$ where $y(x) = xl(x)-T(xl(x))$. The incidence of the tax reform on private intermediation $P(x)$ is

$$dP(x;x^*) = dp(x^*)\cdot(y(x)-\bar{y})+p\cdot(dy(x)-d\bar{y}),$$

where $dp(x^*) = -\kappa(x^*)(1-p)\frac{\tau'(z(x^*))}{1-T'(z(x^*))}$ denotes the change in the payment rate and $\kappa(x^*) = -\frac{dlog(1-p)}{dlog(1-T'(z(x^*)))}$. The change in after-tax income including the lump-sum rebate is $dy(x) = x(1-T'(z(x)))dl(x)-\tau(z(x))+dR$ and the change in $\bar{y}$ is $d\bar{y} = d\bar{z} = \int dz(x)f(x)dx = -\epsilon^{l}_{1-T'}(x^*)\frac{z(x^*)}{z'(x^*)}\frac{f(x^*)}{1-F(x^*)}$ because we consider revenue-neutral tax reform. Since the household's consumption is $c(x) = y(x)-P(x) = y(x)-p\cdot(y(x)-\bar{y})$, the incidence on consumption is $dc(x) = -dp\cdot(y(x)-\bar{y})+(1-p)dy(x)+pd\bar{y}$, and the change in consumption net of the disutility of labor is $d[c(x)-v(l(x))] = -dp\cdot(y(x)-\bar{y})+(1-p)[-\tau(z(x))+dR]+p\bar{y}$, where the terms associated with $dl(x)$ disappear by the envelope theorem.

Note that the allocations and the individual utility $(u(x))$ depend on current productivity x only. Thus, the social welfare is simply $W = \frac{1}{1-\beta}\int u(x)f(x)dx$ and the incidence on social welfare is:

$$
\begin{aligned}
dW &= \frac{1}{1-\beta}\int u'(x)\left[-dp\cdot(y(x)-\bar{y})+(1-p)\left(-\tau(z(x))+dR\right)+p\bar{y}\right]f(x)dx \\
&= \frac{1}{1-\beta}\left[-dp\int u'(x)(y(x)-\bar{y})f(x)dx-(1-p)\int u'(x)\tau(z(x))f(x)dx+dR\cdot\lambda+p(d\bar{y}-dR)\lambda\right],
\end{aligned}
$$

where $\lambda = \int u'(x)f(x)dx$. By replacing dp, dR, and $d\bar{y}$ using the expressions above, $dW = 0$ implies:

$$
\begin{aligned}
\frac{T'(z(x^*))}{1-T'(z(x^*))} &= \frac{1}{\epsilon^l_{1-T'}(x^*)}\frac{z'(x^*)}{z(x^*)}\frac{1-F(x^*)}{f(x^*)} \\
&\quad\times\left[\begin{array}{l}\int(1-\frac{u'(x)}{\lambda})\tau(z(x))f(x)dx-p\epsilon^l_{1-T'}(x^*)\frac{z(x^*)}{z'(x^*)}\frac{f(x^*)}{1-F(x^*)} \\ +\int\frac{u'(x)}{\lambda}\left[p\cdot(\tau(z(x))-\int\tau(z(x))f(x)dx)+\kappa(x^*)\frac{1-p}{1-F(x^*)}\frac{y(x)-\bar{y}}{1-T'(z(x^*))}\right]f(x)dx\end{array}\right] \\
&= -p+\frac{1}{\epsilon^l_{1-T'}(x^*)}\frac{z'(x^*)}{z(x^*)}\frac{1-F(x^*)}{f(x^*)}\left[\begin{array}{l}-\int\frac{u'(x)}{\lambda}(\tau(z(x))-1)f(x)dx \\ +\int\frac{u'(x)}{\lambda}\left[\begin{array}{l}p\cdot(\tau(z(x))-1) \\ +\kappa(x^*)\frac{1-p}{1-F(x^*)}\frac{y(x)-\bar{y}}{1-T'(z(x^*))}\end{array}\right]f(x)dx\end{array}\right] \\
&= -p-\frac{1}{\epsilon^l_{1-T'}(x^*)}\frac{z'(x^*)}{z(x^*)}\frac{1-F(x^*)}{f(x^*)}(1-p)\int\frac{u'(x)}{\lambda}\left[\begin{array}{l}\tau(z(x))-1 \\ +\frac{\kappa(x^*)}{1-F(x^*)}\frac{y(x)-\bar{y}}{1-T'(z(x^*))}\end{array}\right]f(x)dx.
\end{aligned}
$$

B Additional Figures and Comparative Statics

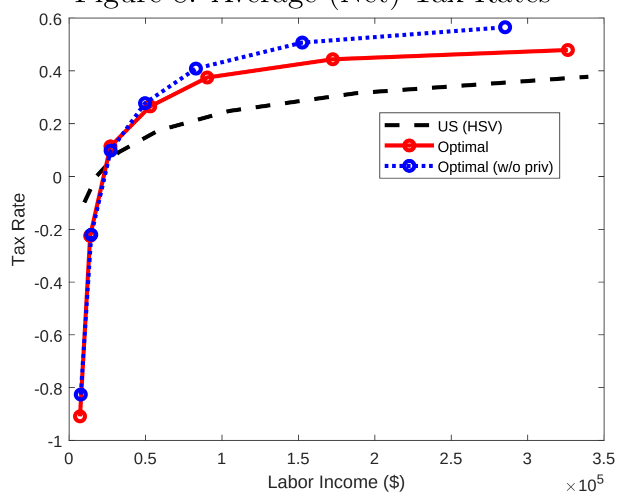
B.1 Average Tax Rates under Optimal Tax Schedules

Figure 8 shows the average (net) tax rates from the model (with and without a private insurance market) and the U.S. economy (when the tax schedule is approximated by the log-linear HSV form). Here, the unit of the model is normalized so that the average labor income in the benchmark model with a private market is $40,000, which is comparable to that in 2015 in the U.S. We dropped the first three productivity grid points for better readability of the plot, and show average tax rates from the 4^{th} grid. Due to the large lump-sum transfers under the optimal tax scheme, the average tax rates for the first three grids are big negative numbers.[34]

[34]The average tax rates for the first three grids are -1,302%, -513% and -225% with a private market and -1,090%, -449%, -200% without a private market.

As we discussed in the main text, high optimal (marginal) tax rates at the low productivity level imply a rapid phase-out of the transfer. The optimal tax schedule with private insurance exhibits an even larger transfer with a faster phase-out, compared to that without private insurance.

Figure 8: Average (Net) Tax Rates



B.2 Risk Aversion and Labor Supply Elasticity

We consider the relative risk aversion of $\sigma = 1$ (smaller than the benchmark value of 1.5). Figure 9 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.

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 10. This is due to weaker pecuniary externalities. With a smaller risk aversion, the pecuniary externality effect, a dominant force to reduce tax rates for the high-income group, is now smaller (see the third panel of Figure 11.)

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 distorting the labor supply.

Figure 9: Optimal Tax w/o Private Insurance

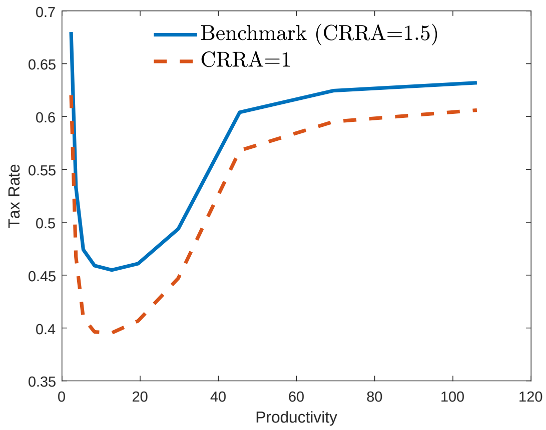


Figure 10: Optimal Tax with Private Insurance

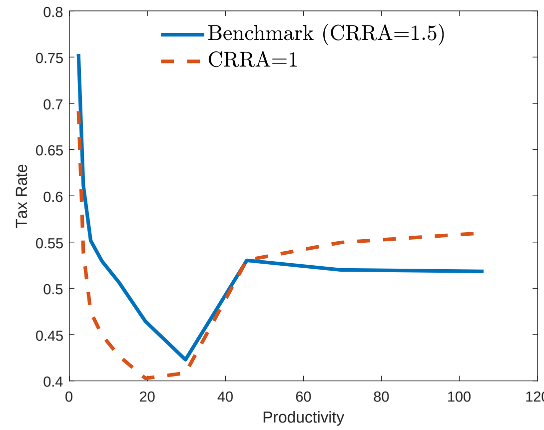


Figure 11: Decompostion: CRRA=1

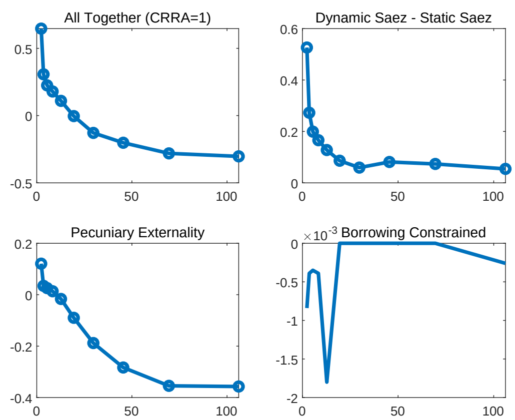
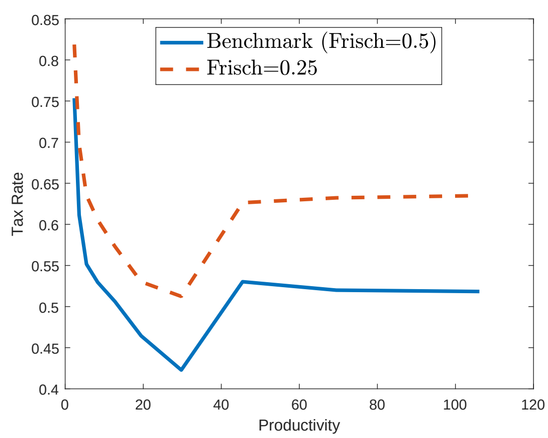


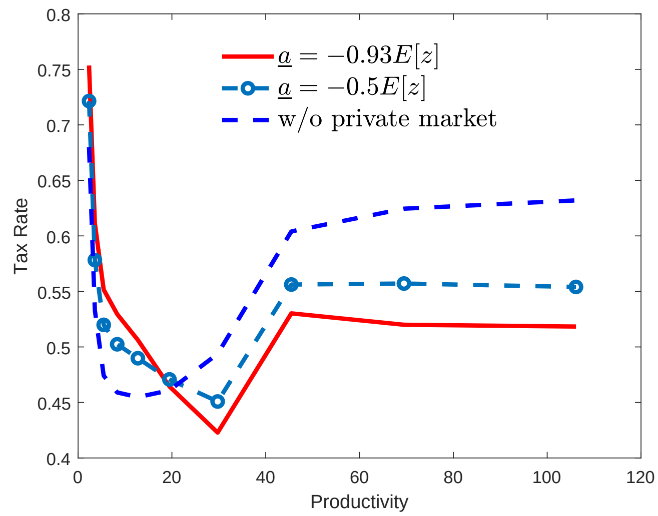
Figure 12: Optimal Tax with Private Market



B.3 Borrowing Constraints

In the benchmark economy, we set the borrowing limit to $\underline{a} = -90.84$, which is about 93% of the average annual earnings in the steady state under the current U.S. tax sched-

Figure 13: Optimal Tax under Tighter Borrowing Constraint



ule (approximated by a log-linear form as in HSV). Under this borrowing limit, 10% of households are borrowing constrained, again under the current U.S tax schedule. We consider a tighter borrowing limit, which is about half of our benchmark value ($\underline{a} = -48.84$): workers can borrow up to one-half of the average earnings in the economy. With this tighter borrowing limit, 33% of the population is credit constrained under the current U.S. tax schedule in our model. Figure 13 shows that the optimal tax rates under this tighter borrowing constraint are roughly between those in the benchmark case and those without a private insurance market.

C Proofs in Section 5

C.1 Aiyagari Economy with Physical Capital

As we discuss in the main text, in an Aiayagari economy, even under the GHH preferences, the incidence of tax reform on the labor supply incorporates the effects of changes in the wage rate as well as the effects of changes in the retention rate. The definition of the elasticity of the labor supply with respect to the retention rate $1 - T'(z(x))$ along the nonlinear budget constraint is the same as in the Huggett economy: $\epsilon^l_{1-T'}(x) = \frac{e(x)}{1+\rho(wz(x))e(x)}$, where $\rho(wz(x)) = \frac{T''(wz)wz}{1-T'(wz)}$.

We now define the elasticity of the labor supply with respect to the wage rate along the nonlinear budget constraint. The first-order condition when perturbing the wage rate w by dw is:

$$v'(l(x)+dl(x)) = [1-T'((w+dw)x(l(x)+dl(x)))](w+dw)x.$$

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

$$v'(l(x))+v''(l(x))dl(x) = [1-T'(wxl(x))]wx-T''(wxl(x))w^2x^2dl(x)-T''(wxl(x))wx^2l(x)dw+(1-T'(wxl(x))$$

Thus,

$$\begin{aligned}
\epsilon_w^l(x) = \frac{dl(x)}{dw}\frac{w}{l(x)} &= \frac{(1-T'(wxl(x))x-T''(wxl(x))wx^2l(x)}{v''(l(x))+T''(wxl(x))w^2x^2}\frac{w}{l(x)} \\
&= \frac{\frac{v'(l(x))}{l(x)v''(l(x))}-\frac{v'(l(x))}{l(x)v''(l(x))}\frac{T''(wxl(x))wxl(x)}{1-T'(wxl(x))}}{1+\frac{v'(l(x))}{l(x)v''(l(x))}\frac{T''(wxl(x))wxl(x)}{1-T'(wxl(x))}} \\
&= \frac{e(x)(1-\rho(wz(x)))}{1+\rho(wz(x))e(x)} = \epsilon_{1-T'}^l(x)(1-\rho(wz(x))).
\end{aligned}$$

The incidence of tax reform on the labor supply can be obtained from the following first-order condition under the perturbed tax function $T(z)+\mu\tau(z)$:

$$v'(l(x)+\mu dl(x)) = [1-T'((w+\mu dw_t)x(l(x)+\mu dl(x)))-\mu\tau'((w+\mu dw_t)+x(l(x)+\mu dl(x)))]x(w+\mu dw_t).$$

A first-order Taylor expansion around the equilibrium implies:

$$\begin{aligned}
v'(l(x))+v''(l(x))\mu dl(x) &= [1-T'(wxl(x))]wx-T''(wxl(x))[xl(x)\mu dw_t+wx\mu dl(x)]xw \\
&\quad -\tau'(wxl(x))xw+[1-T'(wxl(x))]xdw_t.
\end{aligned}$$

By solving for $dl(x)$, we obtain:

$$\begin{aligned}
dl(x) &= \frac{-\tau'(wxl(x))+[1-T'(wxl(x))-T''(wxl(x))]\frac{dw_t}{w}}{\left[\frac{l(x)v''(l(x))}{v'(l(x))}+\frac{T''(wxl(x))}{1-T'(wxl(x))}xwl(x)\right]\frac{1-T'(wxl(x))}{l(x)}} \\
&= -\frac{e(x)}{1+\rho(wz(x))e(x)}\frac{\tau'(wxl(x))}{1-T'(wxl(x))}l(x)+\frac{e(x)(1-\rho(wz(x)))}{1+\rho(wz(x))e(x)}\frac{dw_t}{w} \\
&= -\epsilon_{1-T'}^l(x)\frac{\tau'(wxl(x))}{1-T'(wxl(x))}l(x)+\epsilon_w^l(x)\frac{dw_t}{w}.
\end{aligned}$$

Replacing dw_t with (12) yields the expression of $dl(x)$ in the main text. Thus, the tax incidence on government revenue dR_t should also consider the effect of the change in the wage.

$$\begin{aligned} dR_t &= \int \tau(wz(x))f(x)dx + \int T'(wz(x))[dw_t xl(x) + wxdl(x)]f(x)dx \\ &= \int_{x^*}^{\infty} \frac{f(x)}{1-F(x^*)}dx + dw_t \int T'(wz(x))xl(x)f(x)dx \\ &\quad + \int T'(wz(x)) \left[-\frac{\epsilon^l_{1-T'}(x)}{1-F(x^*)} \frac{\delta_{z^*}(z(x))}{1-T'(z(x))} wz(x) + dw_t x \epsilon^l_w(x) l(x) \right] f(x)dx \\ &= \int_{x^*}^{\infty} \frac{f(x)}{1-F(x^*)}dx - \frac{T'(wz(x^*))}{1-T'(wz(x^*))}\epsilon^l_{1-T'}(x^*)\frac{wz(x^*)}{z'(x^*)}\frac{f(x^*)}{1-F(x^*)} + dw_t \int T'(wz(x))(1+\epsilon^l_w(x))z(x)f(x)dx, \end{aligned}$$

which shows additional fiscal externalities due to changes in the equilibrium wage. The tax incidence on individual welfare also shows additional welfare effects due to the change in the wage:

$$dV(a_0,x_0) = \sum_{t=0}^{\infty} \beta^t \int f(x^t|x_0)u'(a_t(x^{t-1}),x_t)[-\tau(w_t x_t l(x_t)) + (1-T'(w_t x_t l(x_t)))x_t l(x_t) dw_t + dr_t a_t(x^{t-1})]dx^t$$
$$-\sum_{t=0}^{\infty} \beta^t \int f(x^t|x_0)[u'(a_t(a_0,x^{t-1}),x_t) - \beta(1+r)E\{u'(a_{t+1}(a_0,x^t),x_{t+1})|x^t\}dh^A(a_t(a_0,x^{t-1}),y(x_t))dx^t.$$

Using these incidences, deriving the incidence on social welfare is straightforward. By imposing $dW = 0$, we obtain the formula (13).

We now rewrite the pecuniary externalities term $B_t(z^*)$ in (13) to show the two channels of pecuniary externalities. First, we can rewrite dr_t and dw_t using the constant returns to scale in production technology ($F_{KK}K + F_{KL}L = 0$, and $F_{LL}L + F_{LK}K = 0$):

$$\begin{aligned} dr_t &= F_{KK}dK_t + F_{KL}dL_t = F_{KL}(dL_t - \frac{L_t}{K_t}dK_t) = F_{KL}L_t(\frac{dL_t}{L_t} - \frac{dK_t}{K_t}), \\ dw_t &= F_{LL}dL_t + F_{LK}dK_t = F_{KL}(dK_t - \frac{K_t}{L_t}dL_t) = F_{KL}K_t(\frac{dK_t}{K_t} - \frac{dL_t}{L_t}). \end{aligned}$$

Thus, we can rewrite the bracket in the integrand of $B_t(z^*)$ by replacing dr_t and dw_t:

$$\begin{aligned} & (1-T'(wz(x_t))x_t l(x_t)dw_t + a_t dr_t \\ =\ & F_{KL}K_t L_t \left(\frac{dK_t}{K_t} - \frac{dL_t}{L_t} \right) \left[\frac{(1-T'(wz(x_t)))x_t l(x_t)}{L_t} - \frac{a_t}{K_t} \right] \\ =\ & dr_t K_t \left[\frac{a_t}{K_t} - \frac{(1-T'(wz(x_t)))x_t l(x_t)}{L_t} \right] \\ =\ & dr_t K_t \left[\left(\frac{a_t}{K_t} - 1 \right) - \left(\frac{(1-T'(wz(x_t)))x_t l(x_t)}{L_t} - 1 \right) \right], \end{aligned}$$

which implies (14) in the main text.

C.2 Generalized Social Welfare Function

With the generalized social welfare function $G(V(a_0, x_0))$, the incidence of tax reform on social welfare is:

$$\begin{aligned}
dW &= \left[\int G'(V(a_0,x_0))\sum_{t=0}^{\infty}\beta^t \int f(x^t|x_0)u'(a_0,x^t)dx^t d\Phi(a_0,x_0)\right]\cdot dR \\
&\quad + \int G'(V(a_0,x_0)) \left[\begin{array}{c} \sum_{t=0}^{\infty}\beta^t \int f(x^t|x_0)u'(a_0,x^t)[-\tau(z(x_t)) + dr_t a_t(x^{t-1})]dx^t \\ -\sum_{t=0}^{\infty}\beta^t \int \{u'(a_0,x^t) - \beta(1+r)E[u'(a_0,x^{t+1})|x^t]\} \\ \times dh^A(a_t(x^{t-1}), y(x_t))f(x^t|x_0)dx^t \end{array}\right] d\Phi(a_0,x_0) \\
&= \frac{\lambda}{1-\beta}\int \tau(z(x))f(x)dx - \frac{\lambda}{1-\beta}\frac{T'(z(x^*))}{1-T'(z^*)}\epsilon^l_{1-T'}(x^*)\frac{z(x^*)}{z'(x^*)}\frac{f(x^*)}{1-F(x^*)} \\
&\quad + \sum_{t=0}^{\infty}\beta^t \iint G'(V(a_0,x_0))u'(a_0,x^t)[-\tau(z(x_t)) + dr_t a_t(x^{t-1})]f(x^t|x_0)dx^t d\Phi(a_0,x_0) \\
&\quad - \sum_{t=0}^{\infty}\beta^t \iint G'(V(a_0,x_0))u'(a_0,x^t)\left\{1-\beta(1+r)E\left[\frac{u'(a_0,x^{t+1})}{u'(a_0,x^t)}\middle| x^t\right]\right\} f(x^t|x_0)dx^t d\Phi(a_0,x_0),
\end{aligned}$$

where $\lambda = \sum_{t=0}^{\infty}\beta^t \iint G'(V(a_0,x_0))u'(a_0,x^t)f(x^t|x_0)dx^t d\Phi(a_0,x_0)$. By imposing $dW = 0$, we can obtain the optimal tax formula (15).

C.3 Labor Supply with Income Effects

We consider the utility that has the general form $u(c, l)$, preferences with wealth effects on the labor supply. Define the standard compensated elasticity of the labor supply with respect to the retention rate $1 - T'$ along the linear budget constraint as:

$$e^c_{1-T',l}(a_t,x_t) = \frac{dl(a_t,x_t)}{d(1-T'(x_t l_t))}\frac{1-T'(x_t l_t)}{l(a_t,x_t)}\bigg|_u = \frac{u_l(a_t,x_t)/l(a_t,x_t)}{u_{ll}(a_t,x_t) + \left(\frac{u_l(a_t,x_t)}{u_c(a_t,x_t)}\right)^2 u_{cc}(a_t,x_t) - 2\left(\frac{u_l(a_t,x_t)}{u_c(a_t,x_t)}\right)u_{cl}(a_t,x_t)},$$

and the wealth effect on the labor supply along the linear budget constraint as:

$$e_{R,l}(a_t,x_t) = \frac{dl(a_t,x_t)}{dR}x_t(1-T'(x_t l_t)) = \frac{-\left(\frac{u_l(a_t,x_t)}{u_c(a_t,x_t)}\right)^2 u_{cc}(a_t,x_t) + \left(\frac{u_l(a_t,x_t)}{u_c(a_t,x_t)}\right)u_{cl}(a_t,x_t)}{u_{ll}(a_t,x_t) + \left(\frac{u_l(a_t,x_t)}{u_c(a_t,x_t)}\right)^2 u_{cc}(a_t,x_t) - 2\left(\frac{u_l(a_t,x_t)}{u_c(a_t,x_t)}\right)u_{cl}(a_t,x_t)}.$$

As we do in the benchmark analysis without wealth effects on the labor supply, define the elasticities along the nonlinear budget constraint as:

$$\begin{aligned}
\epsilon^c_{1-T',l}(a_t,x_t) &= \frac{e^c_{1-T',l}(a_t,x_t)}{1+\rho(z(a_t,x_t))e^c_{1-T',l}(a_t,x_t)}, \\
\epsilon_{R,l}(a_t,x_t) &= \frac{e_{R,l}(a_t,x_t)}{1+\rho(z(a_t,x_t))e^c_{1-T',l}(a_t,x_t)}.
\end{aligned}$$

We can write the incidence of a tax reform τ on the labor supply in a similar way. A Taylor expansion of the perturbed first-order condition with respect to the labor supply can be solved for $dl(a_t,x_t)$, given the causal effect of a tax reform on savings and the interest rate (da_{t+1} and dr_t) reflecting all possible channels of responses. This leads to Equation (16).

Then, the tax incidence on government revenue is:

$$\begin{aligned}
dR_t &= \int \tau(z_t)d\Phi(a_t,z_t) + \iint T'(z(a_t(a_0,x^{t-1}),x_t))x_t dl(a_0,x^t)f(x^t|x_0)dx^t d\Phi(a_0,x_0) \\
&= \iint_{z^*}^{\infty} \frac{\phi(a,z)}{1-F_z(z^*)}dzda \\
&\quad - \frac{T'(z^*)}{1-T'(z^*)}\int \epsilon^c_{1-T',l}(a,z^*)\phi(a|z^*)da\frac{z^* f_z(z^*)}{1-F_z(z^*)} \\
&\quad + \iint \epsilon_{R,l}(a_t(a_0,x^{t-1}),x_t)\frac{T'(x_t l_t(a_t(a_0,x^{t-1}),x_t))}{1-T'(x_t l_t(a_t(a_0,x^{t-1}),x_t))} \\
&\quad \times \left[\frac{\mathbb{1}_{\{z(a_t(a_0,x^{t-1}),x_t)\geq z^*\}}}{1-F_z(z^*)} - da_{t+1}(a_0,x^t) + (1+r)da_t(a_0,x^{t-1}) + dr_t a_t(a_0,x^{t-1})\right] f(x^t|x_0)dx^t d\Phi(a_0,x_0).
\end{aligned}$$

The change in government revenue dR_t is rebated back to households as a lump-sum transfer, which increases social welfare by:

$$\sum_{t=0}^{\infty}\beta^t \cdot \lambda \cdot dR_t,$$

where λ denotes the marginal value of public funds, which measures the value of distributing an additional unit of revenue uniformly to the entire population, considering the income effect of the lump-sum rebate on the labor supply. The rest of the change in social welfare is the same as that in the benchmark. By imposing $dW = 0$, we obtain the optimal tax formula in the text.