# Taxing the Rich[1]

Augustin Landier[2]
Toulouse School of Economics

Guillaume Plantin[3]
Toulouse School of Economics and CEPR

September 12, 2013

[1]We thank Daron Acemoglu, Bruno Biais, Patrick Bolton, Philippe De Donder, Christian Hellwig, Thomas Mariotti, Thomas Piketty, Sven Rady, Sylvia Serfaty, Konrad Stahl, and seminar participants at Gerzensee (ESSET), Ecole Polytechnique, London School of Economics, Stockholm School of Economics, and Toulouse School of Economics for very helpful comments. Plantin benefitted from a European Research Council Starting Grant (N°263673 – RIFIFI).

[2]E-mail: augustin.landier@tse-fr.eu

[3]E-mail: guillaume.plantin@tse-fr.eu

**Abstract**

Affluent households can respond to taxation with means that are not economically viable for the rest of the population, such as sophisticated tax plans and international tax arbitrage. This paper studies an economy in which an inequality-averse social planner faces agents who have access to a tax-avoidance technology with increasing returns to scale, and who can shape the risk profile of their income as they see fit. Scale economies in avoidance imply that optimal taxation is regressive at the top. This in turn may trigger excessive risk taking.

# Introduction

The taxation of affluent households periodically comes to the forefront of the public debate. The view that the rich should pay more taxes than they currently do recently gained influence in the U.S. and in Europe. Likely reasons include the need for fiscal consolidation resulting from the financial and economic crisis that erupted in 2008, and long-term trends of increasing income inequality and decreasing top marginal tax rates, in particular in the U.S. The "Buffett rule" proposed by the Obama administration responded in particular to the spread of the sentiment that effective tax rates have become overly regressive.[1] The view is that low effective tax rates for the rich result not only from low nominal tax rates, but also from increased tax avoidance by the most affluent households.

This paper develops a new theoretical framework to study the taxation of the rich. Our motivation is twofold. First, the taxation of the rich has first-order implications for public finances, simply because affluent households collect a significant fraction of aggregate income. In 2011, the top quintile and percentile of the U.S. income distribution respectively collected 36% and 20% of aggregate income (Piketty and Saez, 2013). Second, taxing the rich raises issues that are, in our view, quite different from that raised by the taxation of the rest of the population.

Taxing affluent households raises specific issues because the rich can respond to taxation with means that are by and large unavailable to the rest of the population. This paper studies the taxation of a population of agents that can avail themselves of two such means: tax avoidance - the minimization of one's tax liabilities by legal or quasi-legal means, and income-risk shifting.

Consider first risk shifting. Households that are "rich" in the sense that they are on top of the income distribution are also typically "rich" in the sense that they are on top of the wealth distribution (Diaz-Gimenez, Glover, and Rios Rull, 2011). High wealth gives them access to a large set of sophisticated financial instruments with few risk-taking restrictions (e.g., investments in hedge funds), and enables them to hire the expertise required to handle them. Rich households therefore have a free hand at shaping the risk profile of their capital income, which is a sizeable fraction of their total income. By contrast, for the rest of the population, labor income, whose risk profile can be altered only through occupational choices, is a larger fraction

[1] This rule applies a minimum tax rate of 30 percent on individuals making more than a million dollars a year.

of income. Investment is limited to a more narrow and more regulated set of vehicles. Gambling options are available but typically come at a high expected loss.

Second, regarding tax avoidance, the saying that "the poor evade and the rich avoid" epitomizes that avoidance is concentrated at the top of the income distribution. We believe that this is so because common tax-avoidance techniques typically display scale economies. They are therefore profitable only when spread over sufficiently large pre-tax resources. We actually view this as the key economic distinction between tax avoidance and tax evasion. Evasion is an outright breach of tax law, which should naturally be thought of as displaying diseconomies of scale. It seems evident that concealing larger amounts from the tax authority, and converting them into secret consumption comes at a higher unit cost. Obvious reasons include the difficulty of settling large transactions with cash, and the lack of discretion entailed by an affluent lifestyle.

By contrast, tax-avoidance techniques entail costs that are not very sensitive to the income base to which they are applied. There are two main forms of tax avoidance. First, a major source of tax avoidance consists in tax plans that shape the timing, nature, and amount of taxable income so as to minimize taxes. Typical schemes consist in relabelling labor income as capital income, or in borrowing against capital gains instead of realizing them to consume. The ability of private equity and hedge fund managers to structure their pay as carried interest, which is taxed as dividends instead of labor income, is a simple example of such avoidance. Sophisticated tax planning involves significant fixed costs associated with the setup of complex legal structures and the remuneration of tax planners' human capital. Shackelford (2000) describes several widespread tax-avoidance plans, and notes that "these plans are restricted to the wealthiest of taxpayers because the implementation fees are so large that the income or transfer taxes saved must be enormous to justify purchasing the tax plan." That wealth managers and sophisticated tax planners impose high minimum accounts is consistent with such significant fixed costs. Also consistent with this, Lang, Norhass, and Stahl (1997) use detailed consumer survey data in Germany to show that the difference between legislated and effective tax rates increases with respect to income, and that a sizeable fraction of it is due to the exploitation of legal tax write-offs. In the U.S., the Congressional Budget Office estimates that more than 90 percent of the benefits of reduced tax rates on capital gains and dividends will accrue to households in the highest income quintile in 2013, with almost 70 percent going to households in the top percentile.

A second important form of tax avoidance consists in international tax arbitrage, by locating assets or establishing fiscal residence in low-tax countries.[2] This form of tax avoidance also involves legal and transportation costs that are not very sensitive to income. Consistent with this, using data on the geographic mobility of soccer players, Kleven, Landais and Saez (2013) document that it is only at the top of the income distribution that location choice is highly elastic to taxes. Studying the impact of the Danish preferential tax scheme for high-earning immigrants, Kleven, Landais, Saez, and Schultz (2013) also document a high elasticity of migration of top earners. Another international tax arbitrage technique consists in keeping one's fiscal residence unchanged while making undeclared bank deposits in countries with strong bank secrecy. Strictly speaking, this pertains to evasion rather than avoidance, as it is illegal. But it it is a virtually undetectable fraud (holding international treaties fixed), and involves the same type of fixed legal and administrative costs as avoidance. Thus we consider it to be part of the tax-optimization techniques that we seek to model in this paper. Exploiting inconsistencies in international accounts, Zucman (2013) estimates that 8% of total household financial wealth is held in tax havens. Recent attempts by the G20 at cracking down on this type of evasion may reduce its magnitude in the near future. Yet, Johannesen and Zucman (2013) offer suggestive evidence that rather than repatriating funds in response, evaders tend to relocate them to alternative less compliant havens.

We develop a model that simply formalizes tax avoidance as follows. We study the situation of a social planner who seeks to implement inequality-averse views in an endowment economy. The planner faces an informational friction. Agents privately observe their endowments, and can convert the fraction that they do not report to the planner into secret consumption at some cost. In line with the above discussion, we capture avoidance by assuming that this conversion has increasing scale returns. The optimal redistribution scheme implemented by the social planner in the presence of such tax avoidance is simple. Net income has a fixed component and a variable one that increases with respect to the reported pre-tax income. This scheme is such that agents report their entire income: there is no avoidance in equilibrium. The fixed component equally splits among agents the total tax capacity of the planner, defined as the total resources that he could extract from the population if he was not redistributing any of it. The variable component makes every agent indifferent between reporting his entire

[2] An extreme form of tax avoidance is that of "perpetual travelers" - individuals who spend sufficiently little time in any given country that they have no identified fiscal residence.

income or reporting the lowest income level in the population. This simple scheme has several interesting properties. First, taxation cannot be progressive above some income threshold. This is consistent with Landais, Piketty, and Saez (2011), who show that in France, income taxes have become regressive above the 5% top income quantile. Second, pre-tax inequality may reduce the tax capacity of the social planner. Tax capacity is lower when pre-tax income distribution is riskier in the sense of second-order stochastic dominance. As a result, more pre-tax inequality among very top earners affects net consumption relatively more severely at the other end of the distribution - low pre-tax income levels.

We then add a risk-shifting friction to the tax-avoidance one. We introduce an initial stage during which agents can add any fair lottery to their endowments. A key consideration then is whether the social planner can credibly commit to a scheme that he announces before agents take their gambles, or if he lacks commitment, and redistributes instead in an *ex post* optimal fashion after gambles are realized. Absent commitment power, agents do gamble in equilibrium because they expect their utility to be non concave over pre-tax income given the regressivity of *ex post* optimal taxation. Thus in this case, tax avoidance does not occur in equilibrium but yet creates more pre-tax inequality. In the case of a fixed cost of avoidance, agents whose endowment is below that cost favor long-shot gambles that yield a large positive return with a small probability. Those that earn more than this cost prefer bets that generate frequent small gains at the cost of large rare losses. We thus predict that the demand for this latter risk profile, deemed "fake alpha" in the finance literature, should increase as the right tail of income distribution becomes fatter. If the planner can commit, then he prefers to concavify himself agents' utility, rather than let agents do it themselves with gambles that increase inequality. Thus agents do not shift risk in equilibrium in this case. The planner uses an *ex post* inefficient tax scheme that is less regressive and grants a lower minimum net consumption level.

The paper also studies a dynamic version of this economy. If the planner cannot commit, then agents no longer report their entire income initially because they expect to be then taxed in a confiscatory fashion in the future: there is tax avoidance in equilibrium. Further, not only do agents take excessive risk on their investments, but they also distort their savings behaviour so that their current consumption overreacts to life-income shocks.

Finally, we discuss a version of the model in which agents do not take idiosyncratic bets, but rather continuously trade a single risky asset. We find that the distribution of the now risky taxes collected by the planner reflects that of the income distribution rather than the distribution of asset returns. We also find that agents trade with positive feedback: The aggregate demand for the risky asset increases with respect to its realized return.

Overall, our analysis suggests that, absent radical changes in international coordination and drastic simplifications of the major tax codes, any attempt at suddenly increasing tax pressure on the rich would yield limited additional tax revenues, and spur excessive financial risk taking.

The paper is organized as follows. Section 1 studies optimal taxation in the presence of a tax-avoidance friction. Section 2 adds the risk-shifting friction and studies optimal taxation with and without commitment. Section 3 studies a multi-period extension of the model. Section 4 develops a version of the model in which agents gamble by continuously trading a unique risky asset. Section 5 discusses the relation of the paper to various strands of existing research. Proofs are relegated to an appendix.

# 1 Economies of scale in tax avoidance and optimal taxation

## 1.1 Baseline model

Consider a one-date economy populated by a continuum of agents with unit mass. There is a single consumption good. The agents have identical preferences represented by a utility over consumption $u$ that is increasing and strictly concave. Agents differ only with respect to their endowments of the consumption good - their "incomes". The cumulative income distribution $F$ has support $[0, +\infty)$ and a finite mean:

$$\int w dF(w) < +\infty.$$

We study the problem of a social planner who redistributes income in order to maximize the utilitarian welfare of the population. Because agents have identical concave preferences, the first-best policy clearly consists in ensuring that each agent consumes the same amount, and that aggregate consumption is equal to the aggregate endowment:

**Proposition 1**

*In the first-best, each agent consumes* $\int w dF(w)$.

**Proof.** Jensen's inequality.■

We depart from this first-best, and assume that the planner faces the following informational friction. Each agent privately observes his income. An agent with income $x$ may report any amount $y \in [0, x]$ to the planner, and conceal the residual $x - y$. This concealed income $x - y$ can be converted into $g(x - y)$ units of secret consumption, where $g$ is a continuous function that satisfies:

$$0 \leq g(z) \leq z. \tag{1}$$

This secret consumption adds up to the public one, which is the net transfer that the agent receives after the social planner redistributes aggregate reported income.

Under this general formulation, the friction facing the social planner could be interpreted either as tax evasion or as tax avoidance. As we explained in the introduction, we believe, however, that evasion and avoidance correspond to very distinct properties of the avoidance technology represented by $g$. In the case of evasion, which is an outright breach of tax law, the technology $g$ should be thought of as displaying diseconomies of scale. In the case of tax avoidance, the technology $g$ should feature economies of scale, at least within some range, given in particular important fixed costs of avoidance. Accordingly, we introduce economies of scale by positing that:

**Assumption 1.** *The function $g$ is superadditive. For all $w$, $w' \geq 0$*

$$g\left(w + w'\right) \geq g(w) + g(w'). \tag{2}$$

Notice that superadditivity and $g \geq 0$ imply that $g$ is increasing. One interpretation of (2) is that a single affluent agent avoids more efficiently than a group of agents with the same aggregate tax base, which seems natural. The following lemma exhibits simple cases in which $g$ is superadditive.

**Lemma 2**

*i) Suppose that $g$ is convex. Then $g$ is superadditive.*

*ii) Suppose that converting non-reported income into secret consumption requires spending a fixed cost $k > 0$. After this cost is sunk, $x$ units of non-reported income translate into $f(x)$ units of secret consumption, where $f$ is an increasing concave function that satisfies (1). Suppose that for all $w \geq 0$,*

$$2f\left(\frac{w}{2}\right) - f(w) \leq k. \tag{3}$$

*Then $g(w) = \max\{f(w) - k; 0\}$ is superadditive.*

**Proof.** See the appendix.■

In point *ii)* of Lemma 2, the cost of avoidance corresponds to a textbook cost function, with increasing marginal costs topping up a fixed setup cost. Agents with an income below the fixed cost cannot profitably convert their secret income into secret consumption. Condition (3) shows that superadditivity holds in this case provided the fixed cost of avoidance $k$ is sufficiently important and the marginal cost of avoidance does not increase too sharply with the tax base ($f$ is not too concave).

We now solve the planner's problem in the presence of this tax-avoidance friction. In application of the revelation principle, one can write down the planner's problem using only direct mechanisms. A direct mechanism is a pair of functions $(r(.), v(.))$ such that an individual with endowment $w$ has the incentive to report $r(w) \in [0, w]$, and receives a net transfer $v(r(w))$ from the social planner after doing so. Notice that we rule out random mechanisms. Proposition 7 establishes a sufficient condition on $u$ for this to be without loss of generality.

The social planner solves the program $(\wp)$ :

$$\max_{r,v} \int_0^{+\infty} u\left(v(r(w)) + g\left(w - r(w)\right)\right) dF(w) \tag{4}$$

$$s.t. \left\{ \begin{array}{c} \int_0^{+\infty} v(r(w))dF(w) \leq \int_0^{+\infty} r(w)dF(w), \\ \forall w, w' \geq 0 \ s.t. \ r(w') \leq w, \\ v(r(w)) + g\left(w - r(w)\right) \geq v(r(w')) + g\left(w - r(w')\right). \end{array} \right. \tag{5}$$

The first inequality in (5) is the resource constraint of the planner. The other inequalities are incentive-compatibility constraints, ensuring that individuals report according to their types (which of course does not necessarily imply that they report their entire income). We show that the solution to this program $(\wp)$ is very simple when tax avoidance displays scale economies.

**Proposition 3**

*Under Assumption* 1, *the solution to* $(\wp)$ *is attained with* $(r^*, v^*)$ *defined as*

$$\left\{ \begin{array}{c} r^*(w) = w, \\ v^*(w) = g\left(w\right) + \int_0^{+\infty} \left(t - g(t)\right) dF(t) \end{array} \right. . \tag{6}$$

**Proof.** See the appendix.■

Proposition 3 states that there is no tax avoidance in equilibrium: agents report their entire income. This result was first established in Grochulski (2007). This is a direct consequence from the superadditivity of $g$. Any incentive-compatible tax scheme that implies some avoidance can be replaced with a more efficient one that does not entail any. To see this, suppose that a mechanism $(r, v)$ implies $\int r(w)dF(w) < \int wdF(w)$. Then a scheme whereby an individual with income $w$ reports $w$ and receives $v(r(w)) + g(w - r(w)) + \varepsilon$ satisfies the resource constraint for $\varepsilon > 0$ sufficiently small. Further, it is incentive-compatible:

$$\begin{aligned} v(r(w)) + g(w - r(w)) &\geq v(r(w')) + g(w - r(w')), \\ &\geq v(r(w')) + g(w' - r(w')) + g(w - w'). \end{aligned}$$

The first inequality stems from the incentive-compatibility of $(r, v)$, the second one from the superadditivity of $g$. This second inequality means that this new mechanism is also incentive-compatible. It is strictly preferable to $(r, v)$ because the income destruction induced by tax avoidance disappears.

Proposition 3 exhibits the most redistributive scheme among all "avoidance-free" ones. It simply consists in making every agent indifferent between reporting his entire income or none of it. Scale economies imply that an agent who is indifferent between reporting everything and reporting nothing also prefers a full report to any partial report.

The broad message from Proposition 3 is that in the presence of scale economies in tax avoidance, the overriding priority of the social planner is to make sure that agents report any income at all. Scale economies leave him with no other option but making agents indifferent between reporting their entire income and none of it. We now show that this is a robust result, in the sense that this overriding constraint also shapes the optimal tax scheme under alternative modelling assumptions.

## 1.2 Alternative settings

This section studies the optimal tax scheme designed by the social planner in modifications of the above baseline model. A reader interested in getting directly at our main results may skip this Section 1.2, and move on directly to Section 2.

### 1.2.1 Alternative planner's objectives

The optimal tax scheme in Proposition 3 does not depend on the utility function $u$, and the proof of the proposition only uses that $u$ is increasing and strictly concave. Thus the scheme (6) is optimal given any concave objective:

**Corollary 4** *Replace the planner's objective (4) in* ($\wp$) *with*

$$\max_{r,v} \int_0^{+\infty} \Omega \left[ u \left( v(r(w)) + g \left( w - r(w) \right) \right) \right] dF(w),$$

*where* $\Omega \circ u$ *is increasing, strictly concave. Then the optimal tax scheme is given by (6).*

**Proof.** See the appendix.■

This invariance result implies that the tax scheme given by (6) also maximizes a Rawlsian criterion:

**Corollary 5** *Replace the planner's objective (4) in* ($\wp$) *with*

$$\max_{r,v} \inf_{w} \left\{ v(r(w)) + g \left( w - r(w) \right) \right\}.$$

*Then the optimal tax scheme is given by (6).*

**Proof.** See the appendix.■

Notice that if $g$ is convex, the function $t \rightarrow t - g(t)$ is concave. This implies that if, other things equal, the income distribution $F$ is riskier in the sense of second-order stochastic dominance, then the constant in $v^*$, $\int_0^{+\infty} (t - g(t))\, dF(t)$, is smaller. Thus, more pre-tax inequality makes the poorest relatively worse off in the presence of the tax-avoidance friction.

**Tax capacity.** The reason tax scheme (6) maximizes any concave objective is basically that the constant term in $v^*$, $\int_0^{+\infty} (t - g(t))\, dF(t)$, is also the *tax capacity* of the social planner. Namely, it is the maximal revenue that a planner who does not redistribute can extract from the population. To see this, consider the following program ($\wp'$) :

$$\max_{r,\tau} \int_0^{+\infty} \tau \left( r(w) \right) dF(w) \tag{7}$$
$$s.t. \begin{cases} \forall w \geq 0, \;\; \tau \left( r(w) \right) \leq r(w), \\ \forall w, w' \geq 0 \; s.t. \; r(w') \leq w, \\ r(w) - \tau \left( r(w) \right) + g \left( w - r(w) \right) \geq r(w') - \tau \left( r(w') \right) + g \left( w - r(w') \right). \end{cases}$$

This program formalizes (applying the revelation principle) the situation in which a planner seeks to extract as many resources as possible from the population for purposes that are outside the model. The function $r(w)$ describes the report of an agent with income $w$, while $\tau(r(w))$ describes by how much he is taxed. Notice that we impose that no agent be taxed beyond his reported income, which is plausible and ensures that the program ($\wp'$) has a finite solution. We have

**Corollary 6** *The solution to* $(\wp')$ *is* $\int_0^{+\infty}(t-g(t))\,dF(t)$*, and is attained with* $(r^*,\tau^*)$ *defined as*

$$\left\{\begin{array}{r} r^*(w)=w, \\ \tau^*(w)=w-g\left(w\right) \end{array}\right. . \tag{8}$$

*Therefore, optimal taxes are not strictly progressive.*

**Proof.** See the appendix.■

The average tax rate

$$\frac{\tau^*(w)}{w}=1-\frac{g\left(w\right)}{w}$$

can of course be constant if $g$ is linear, but cannot be strictly increasing. This would imply $g$ strictly concave and thus strictly subadditive. If $g$ is strictly convex over some interval, then taxation is strictly regressive over this income range. If

$$g(w)=\max\left\{f(w)-k;0\right\}$$

as in Lemma 2 ii), then taxation is progressive on each side of $k$, but the average tax rate jumps downwards at $k$.

Overall, these three corollaries suggest that in the presence of scale economies in avoidance, the same tax scheme given by (6) addresses both the question of the *optimal* taxation - provided the objective is inequality-averse, and that of the *maximal feasible* taxation. This tax scheme depends only on pre-tax income distribution and on the avoidance technology. Pre-tax income distribution affects only $v^*(0)$, but not $v^*\left(w\right)-v^*\left(0\right)$. This contrasts with solutions to the standard Mirrlees problem, for which the optimal tax scheme is typically more sensitive to the fine details of the model.

### 1.2.2 Random Taxation

It is well-known that a utilitarian planner may find it optimal to randomize taxes (see Stiglitz, 1981, 1987). Such a randomization of tax schemes can take place *ex ante* and/or *ex post. Ex ante* randomization means that the planner randomly splits the population in several groups before they report any income, and then subjects them to different tax schemes. This is a concavification device that may be desirable when welfare is non-concave in the total income raised by the planner. *Ex post* randomization means that upon reporting $r(w)$, an agent receives from the planner a random net income. In general, *ex post* randomization may make it easier to satisfy incentive compatibility because an agent may report a higher income so as to incur a lower disutility from this source of risk. The next proposition exhibits a sufficient condition under which randomization does not improve upon the deterministic mechanisms studied thus far.

**Proposition 7**

*Suppose that $u'$ exists and is weakly more concave than $u$. When ex ante and ex post randomization are allowed, the tax scheme that maximizes utilitarian welfare is still given by (6).*

**Proof.** See the appendix.■

That $u'$ is more concave than $u$ means that $u' \circ u^{-1}$ is concave. This condition is satisfied for example for utility functions that exhibit constant or increasing absolute risk aversion. It is not satisfied when the agents have strictly decreasing absolute risk aversion. These results mirror that in Stiglitz (1981). In an environment with unobservable labor decisions, he shows that randomization of taxes may be desirable in particular if the marginal utility over consumption is sufficiently convex.[3] The rest of the paper will focus on the case of CARA preferences, in which the restriction to deterministic mechanisms is therefore without any loss of generality.

### 1.2.3 Alternative supports of income distribution

We assume that income distribution has support $[0, +\infty)$ in the baseline model in order to simplify the notations and the exposition of some results. This comes with no loss of generality, the only substantial part of this assumption is that the support be connected. For example, if the income distribution has support over $[\underline{w}, \overline{w}]$, it is easy to check that the optimal tax scheme in Proposition 3 simply becomes

$$\begin{aligned} r^*(w) &= w, \\ v^*(w) &= g\left(w - \underline{w}\right) + \int_{[\underline{w},\overline{w}]} \left(t - g(t - \underline{w})\right) dF(t). \end{aligned} \tag{9}$$

Notice that we do not impose any restriction on the c.d.f. $F$ besides a connected support and a finite mean.

[3] in the sense that $\frac{\partial^3 U}{\partial C^3}$ be sufficiently large.

### 1.2.4 Exogenous expenditures

Another straightforward extension is the addition of some exogenously given expenditures $S$ that the social planner must cover with taxes. The optimal tax scheme in Proposition 3 simply becomes

$$\begin{gathered} r^{*}(w) = w, \\ v^{*}(w) = g\left(w\right) + \int_{0}^{+\infty}\left(t - g(t)\right)dF(t) - S. \end{gathered}$$

Overall, this Section 1.2 suggests that the optimal tax scheme with scale economies in tax avoidance presented in the baseline model is robust to a number of alternative specifications. We now add to this setup a second friction - secret risk shifting.

# 2 Economies of scale in tax avoidance, secret risk shifting, and optimal taxation

We argued in the introduction that a distinctive feature of affluent individuals is their access to a rich set of financial instruments that gives them a free hand at selecting the risk profile of their pre-tax income. This section introduces a second friction in our baseline model that formalizes this feature, and studies its interplay with scale economies in tax avoidance.

## 2.1 Setup

As in the baseline model of the previous section, a social planner seeks to maximize the utilitarian welfare of a continuum of agents with unit mass. The economy now has two dates, 0 and 1. Preferences, endowments, and the information structure are as follows.

*Preferences.* Agents value only date-1 consumption, over which they have CARA utility $u$.

*Endowments.* Agents receive their entire endowment at date 0. The cumulative income distribution $F_0$ has support $[0, +\infty)$ and a finite mean. Agents need to store their income from date 0 to date 1 in order to consume. A risk-free storage technology with unit return is available. Agents may also enter into risk shifting in the following sense. Each agent may add to this risk-free return an idiosyncratic risky return. In this case, he has a free hand at choosing the unit-mean distribution of this risky return. Formally, an agent with initial income $w_0$ can choose a date-1 income with any cumulative distribution function with mean $w_0$ and support included in $[0, +\infty)$.

We assume that for each date-0 income level, an arbitrarily small measure of agents only has access to the risk-free storage and cannot gamble this way. This is a technical assumption meant to ensure that date-1 pre-tax income distribution has full support over $[0, +\infty)$.[4]

*Information.* Agents privately observe their date-0 income. The social planner does not observe their investment decisions, nor their resulting date-1 individual incomes. As in the previous section, agents can convert $x$ units of concealed income into $g(x)$ units of secret consumption, where $g$ is a continuous superadditive function that satisfies (1). The (exogenous) date-0 income distribution, that we denote $F_0$, is publicly observed. So is the (endogenous) date-1 income distribution, that we denote $F_1$.[5]

Thus, the social planner now faces two informational frictions. First, agents can divert income and secretly consume as in the previous section. Second, they can also secretly shift income risk. We model this risk-shifting ability as the possibility to add fair lotteries with arbitrary distribution to their income. This modelling choice has two advantages. First, excessive risk taking is simply and clearly characterized in our model as the addition of non-rewarded risk to a safe endowment by a risk-averse agent. Second, this delivers sharp insights into the type of risk distributions that households willing to shift risk demand. A number of other restrictive assumptions are made in this setup. We will explain their roles and relax them in subsequent sections. More precisely, Section 2.4.2 relaxes the assumption of a CARA utility. Section 3 tackles the case in which agents value consumption at several dates. Section 4 studies the situation in which agents take correlated risks.

Before tackling such extensions, we solve for the social planner's problem in this baseline setup. Section 2.2 first tackles the case in which the social planner cannot commit to a tax scheme, but instead redistributes *ex post* optimally at date 1, after date-1 incomes are realized. Section 2.3 then tackles the case in which the planner can commit to a scheme. It may seem unusual to first solve the model with the additional friction of limited commitment, and to then remove it. It will be apparent that we do so because the optimal tax scheme under full commitment actually derives from that without commitment.

[4] Alternatively, one could assume that agents only have access to distributions with full support over $[0, +\infty)$ when shifting risk. An agent interested in a discrete distribution could approximate it arbitrarily well with such a continuous one.

[5] We only need that the date-0 income distribution be common knowledge. Assuming that so is $F_1$ slightly simplifies the exposition.

## 2.2 Optimal taxation without commitment

In this section, we suppose that the social planner redistributes in an *ex post* optimal fashion at date 1, after agents have made their investment decision and received their date-1 income. More precisely, the timeline of this economy is as follows. At date 0, each of the agents decides on the distribution of the mean-preserving spread that he wishes to add to his date-0 endowment. He may of course also prefer to store at the risk-free rate. At date 1, agents receive their date-1 incomes. The resulting *ex post* income distribution $F_1$ is publicly observed. At this stage, the social planner announces a taxation mechanism $(r, v)$ so as to maximize utilitarian welfare.

An equilibrium in this economy consists in an *ex post* income distribution $F_1$, a date-1 tax scheme $(r^{**}, v^{**})$, and date-0 investment decisions such that:

- The tax scheme $(r^{**}, v^{**})$ solves the program $(\wp)$ described by (4) and (5) for an income distribution $F = F_1$.
- Each agent makes a date-0 investment decision that maximizes his date-1 expected utility given his initial income and his beliefs about $F_1$ and $(r^{**}, v^{**})$.
- The distribution $F_1$ correctly aggregates the impact of these individual investment decisions on the initial income distribution $F_0$.
- Agents have correct beliefs about $F_1$ and $(r^{**}, v^{**})$.

We now characterize such equilibria. Notice first that since the planner maximizes *ex post* social welfare, Proposition 3 applies at date 1, and

**Lemma 8** *Upon observing $F_1$, the social planner sets*

$$\left\{ \begin{array}{c} r^{**}(w) = w, \\ v^{**}(w) = g\left(w\right) + \int_0^{+\infty} \left(t - g(t)\right) dF_1(t) \end{array} \right. . \tag{10}$$

**Proof.** Discussion above.■

When facing his investment decision at date 0 and forming beliefs about $F_1$, an agent with initial income $w_0$ expects his date-1 consumption to be the random variable $v^{**}\left(\widetilde{w}_1\right)$, where $\widetilde{w}_1$ are his possibly random proceeds from investment and $v^{**}$ is defined in (10). Thus he optimally chooses the distribution of $\widetilde{w}_1$ that maximizes his expected utility subject to the constraint that he expects to earn his initial endowment $w_0$ before taxes. Formally, this agent solves the following problem:

$$\begin{aligned} V(w_0, F_1) &= \max_{G\in\Gamma} \int_0^{+\infty} u\left(g(w) + \int_0^{+\infty} (t - g(t))\, dF_1(t)\right) dG(w), \quad (11) \\ & s.t. \quad \int_0^{+\infty} w dG(w) = w_0. \end{aligned}$$

where $\Gamma$ is the set of cumulative distribution functions with support included in $[0, +\infty)$. CARA preferences imply that the value function $V(w_0, F_1)$ of this program satisfies

$$V(w_0, F_1) = \exp\left(-\alpha \int_0^{+\infty} (t - g(t))\, dF_1(t)\right) W(w_0), \tag{12}$$

where we adopt the convention $u(x) = -e^{-\alpha x}$, and $W(w_0)$ is the value function of the program below that depends only on $w_0$:

$$\begin{aligned} W(w_0) &= \max_{G\in\Gamma} \int_0^{+\infty} u(g(w)) dG(w), \quad (13) \\ & s.t. \quad \int_0^{+\infty} w dG(w) = w_0. \end{aligned}$$

Thus, CARA preferences imply that the agent's beliefs about $F_1$ do not affect his investment choice. The reason is simply that $F_1$ only affects the constant term in $v^{**}$. The value of this constant has no impact on the attitude of a CARA agent towards the riskiness of his date-1 consumption. All that is left to complete this characterization is to solve for (13) for all $w_0 \geq 0$. To do so, we introduce the concavification of the function $u \circ g$, and denote it $\overline{u \circ g}$. That is, the function $\overline{u \circ g}$ is the smallest concave function such that for all $w \geq 0$,

$$\overline{u \circ g}(x) \geq u\left(g\left(x\right)\right).$$

The function $\overline{u \circ g}$ exists and is unique (see, e.g., Aumann and Perles, 1965).

**Lemma 9**

*We have*

$$W(w_0) = \overline{u \circ g}\left(w_0\right).$$

*If $\overline{u \circ g}\left(w_0\right) = u\left(g\left(w_0\right)\right)$, then the agent stores at the risk-free rate.*

*If $\overline{u \circ g}\left(w_0\right) > u\left(g\left(w_0\right)\right)$, then the agent takes additional risk. He may be indifferent among several distributions. In this case, the least risky one* *in the sense of second-order stochastic dominance is the binary distribution with support* $\{\underline{w}(w_0); \overline{w}(w_0)\}$, *where*

$$\begin{cases} \underline{w}(w_0) = \sup\{w \leq w_0 \;\; s.t. \overline{u \circ g}(w) = u(g(w)) \;\;\} \\ \overline{w}(w_0) = \inf\{w \geq w_0 \;\; s.t. \overline{u \circ g}(w) = u(g(w)) \;\;\} \end{cases}. \tag{14}$$

**Proof.** See the appendix.■

Lemma 9 formalizes that agents use lotteries in order to concavify their date-1 utility when the tax scheme given by (10) implies that their utility may be non concave in their pre-tax income.

### A simple example

It is useful to illustrate this result in the simple particular case in which the tax-avoidance technology $g$ is piecewise linear with a convex kink. Suppose that

$$g(x) = (1-\lambda)x + \mathbf{1}_{\{x \geq c\}} \Delta\lambda (x-c). \tag{15}$$

This corresponds to the case in which two tax-avoidance technologies are available. The first one dissipates a fixed fraction $\lambda \in (0,1)$ of each diverted unit of income. The second one wastes only $\lambda - \Delta\lambda \in (0, \lambda)$ out of each diverted income unit, but comes at a fixed cost $c\Delta\lambda > 0$. An agent chooses the latter if and only if his date-1 income is larger than $c$. Figure 1 displays in this particular case the functions $u \circ g$ and $\overline{u \circ g}$.

Figure 1 here.

The functions $u \circ g$ and its concavification coincide over two intervals $[0, \underline{w}]$ and $[\overline{w}, +\infty)$, where $\underline{w} < c < \overline{w}$.[6] The concavification $\overline{u \circ g}$ is strictly above $u \circ g$ over $(\underline{w}, \overline{w})$, where it is equal to the chord linking the points $(\underline{w}, u(g(\underline{w}))$ and $(\overline{w}, u(g(\overline{w})))$. This chord is tangent to $u \circ g$ at these two points. From Lemma 9, any agent with an initial income $w_0 \in [0, \underline{w}] \cup [\overline{w}, +\infty)$ stores at the risk-free rate. If $w_0 \in (\underline{w}, \overline{w})$, then the agent shifts risk and invests with a binary risky return so as to obtain a date-1 income equal to $\underline{w}$ with probability $\frac{\overline{w}-w_0}{\overline{w}-\underline{w}}$ or $\overline{w}$ with probability $\frac{w_0-\underline{w}}{\overline{w}-\underline{w}}$. The date-1 income distribution $F_1$ is thus riskier than $F_0$ in the sense of second-order stochastic dominance because the mass of $F_0$ between $\underline{w}$ and $\overline{w}$ is split into two atoms of $F_1$, in $\underline{w}$ (with mass $\int_{(w,\overline{w})} \frac{\overline{w}-w}{\overline{w}-\underline{w}} dF_0(w)$) and $\overline{w}$ (with mass $\int_{(\underline{w},\overline{w})} \frac{w-\underline{w}}{\overline{w}-\underline{w}} dF_0(w)$). This fully characterizes the equilibrium in this particular case.

As is obvious in this simple example, the distribution of the risk taken by risk-shifting agents depends on their initial income. When they are relatively

[6] It is possible that $\underline{w} = 0$.

"poor", so that $w_0$ is on the right neighborhood of $\underline{w}$, they purchase payoffs that are negative and small in absolute value with a large probability, and large and positive with a small probability - like a lottery ticket. Conversely, when income is in the left neighborhood of $\overline{w}$, the investors favor trades that pay off a small excess return most of the time and generate rare, large losses. These risk-profiles, labelled by Rajan (2010) as "fake alpha" strategies, are produced by collecting a fair premium for exposure to a large disaster risk. The profile of $F_0$ thus determines aggregate risk taking. As the tail of $F_0$ becomes fatter, the demand for fake-alpha strategies increases.

**Equilibrium in the general case**

With a general function $g$, there are two additional technical difficulties. Figure 2 illustrates them.

Figure 2 here.

First, the set of income levels for which $u \circ g < \overline{u \circ g}$ is not necessarily a single interval. Thus, unlike in the simple example, the support $\{\underline{w}(w_0); \overline{w}(w_0)\}$ of the lottery associated to an initial income level $w_0$ may depend on $w_0$, as stated in Lemma 9.

Second, it is possible to construct cases in which, unlike in the simple example above, the lottery that solves (13) for a given income level $w_0$ is no longer necessarily unique. Thus $F_1$ is not uniquely defined, and nor is the equilibrium. The lottery that Lemma 9 singles out in this case is the least risky in the sense of second-order stochastic dominance among all those that solve (13). In what follows, we assume that when indifferent among lotteries, agents pick this least risky lottery. We believe that this is a reasonable selection criterion.

The next proposition summarizes the equilibrium characterization above. In the remainder of the paper, we denote $\widetilde{\rho}(w)$ the equilibrium lottery for an agent with initial income $w$ defined in Lemma 9. That is, $\widetilde{\rho}(w) = w$ if the agent decides to store at the risk-free rate, and $\widetilde{\rho}(w)$ is the least risky lottery defined in Lemma 9 otherwise. By definition, for all $w \geq 0$,

$$E\left[u \circ g\left(\widetilde{\rho}(w)\right)\right] = \overline{u \circ g}(w).$$

**Proposition 10**

*Suppose that risk-shifting agents choose the least risky lottery $\widetilde{\rho}(w)$ among those that deliver the same expected utility. Then the equilibrium is such that an agent who receives an initial income $w$ obtains a date-1 pre-tax income* *equal to $\widetilde{\rho}(w)$. His date-1 net income is*

$$v^{**}\left(\widetilde{\rho}(w)\right) = g(\widetilde{\rho}(w)) + v^{**}\left(0\right), \tag{16}$$

*with*

$$v^{**}\left(0\right) = \int_0^{+\infty} \left(t - g(t)\right) dF_0(t) - \int_0^{+\infty} \left(E\left[g\left(\widetilde{\rho}(t)\right)\right] - g(t)\right) dF_0(t). \tag{17}$$

**Proof.** See above.■

Proposition 10 establishes a theoretical link between tax avoidance, excessive risk taking, and the rise of inequality at the top. Absent any commitment power of the tax authority, increasing returns to tax avoidance create a demand for excessive risk taking, which in turn makes the *ex post* pre-tax income distribution more diffuse. It is interesting to study how the addition of the risk-shifting friction to the avoidance problem affects utilitarian welfare.

**Corollary 11**

*i) Utilitarian welfare is strictly lower in the presence of risk shifting than without it if and only if $\overline{u \circ g} \neq u \circ g$.*

*ii) If $\overline{u \circ g} \neq u \circ g$, it may be that some agents are better off in the presence of risk shifting than without it. It may also be that every single agent is worse off in the presence of risk shifting.*

**Proof.** See the appendix.■

The first point states that risk shifting matters if and only if $u \circ g$ is not concave (otherwise $\widetilde{\rho}(w) = w$ everywhere), and reduces utilitarian welfare in this case. It relates to Proposition 7: if risk shifting was desirable, the planner would implement himself the *ex post* randomization $\widetilde{\rho}(w)$ in the optimal scheme absent the risk-shifting friction.

The second point may be understood as follows. Each given agent benefits from risk shifting since the variable component of his net income is larger than absent risk taking. On the other hand, risk shifting by the other agents creates negative externalities for him since the constant part of the tax scheme $v^{**}\left(0\right)$ is smaller than absent risk shifting. The difference in constant terms is $-\int_0^{+\infty} \left(E\left[g\left(\widetilde{\rho}(t)\right)\right] - g(t)\right) dF_0(t)$.[7] Point ii) in Corollary 11 states that the negative impact of risk shifting on the constant term may or may not overcome the positive one on the variable term for all agents depending on parameter values. Essentially, if the mass of agents who do

[7] Inequality (18) shows that this difference is negative.

shift risk in equilibrium is sufficiently small, then these agents are better off than absent the risk shifting friction. Every agent is worse off in the presence of risk shifting in the polar case in which a large fraction of the population takes excessive risk in equilibrium.

This second point has interesting potential political-economy implications, as it suggests that for some parameter values, a fraction of the population would strongly oppose a ban of risk shifting, which one could broadly interpret as a tighter prudential regulation of financial institutions.

### 2.3 Optimal taxation with commitment

We now consider the situation in which the planner can announce a tax scheme at date 0 and commit to it at date 1. More precisely, the timing is as follows. The planner first announces a date-1 tax scheme $(r^{***}, v^{***})$. Then agents make their investment decisions. After date-1 incomes are realized, the tax scheme $(r^{***}, v^{***})$ is enforced: agents make their reports and the planner redistributes as announced. We have

**Proposition 12**

*The planner sets the tax scheme*

$$\begin{cases} r^{***}(w) = w \\ v^{***}(w) = u^{-1} \circ \overline{u \circ g}(w) + v^{***}(0) \end{cases},$$

*with*

$$v^{***}(0) = \int_0^{+\infty} (t - g(t))\, dF_0(t) - \int_0^{+\infty} \left(u^{-1} \circ \overline{u \circ g}(t) - g(t)\right) dF_0(t).$$

*There is no risk shifting in equilibrium, so that $F_1 = F_0$.*

*If $\overline{u \circ g} \neq u \circ g$, utilitarian welfare is strictly smaller than absent risk shifting and strictly larger than absent commitment.*

**Proof.** See the appendix.■

This optimal scheme simply consists in having the planner committing to a tax scheme that generates concave expected utility over pre-tax income, instead of letting agents concavify themselves by means of risk shifting. This requires commitment power as soon as $\overline{u \circ g} \neq u \circ g$ because this requires the enforcement of an *ex post* inefficient tax scheme.

The gains from not having agents concavifying their utility themselves is that the date-1 income distribution does not get more diffuse as a result, which raises tax capacity, thereby generating a higher constant term in $v^{***}$ than in $v^{**}$. We have indeed for all $w \geq 0$:

$$\frac{u \circ v^{***}(w)}{E\left[u(v^{**}(\widetilde{\rho}(w)))\right]} = e^{\alpha(v^{**}(0)-v^{***}(0))} < 1$$

if $\overline{u \circ g} \neq u \circ g$, because Jensen inequality implies

$$\int_0^{+\infty} E\left[g\left(\widetilde{\rho}(t)\right)\right] dF_0\left(t\right) > \int_0^{+\infty} u^{-1}\left(E\left[u \circ g\left(\widetilde{\rho}(t)\right)\right]\right) dF_0\left(t\right) > \int_0^{+\infty} g(t) dF_0(t). \tag{18}$$

Even though risk shifting does not occur in equilibrium, commitment does not fully eliminate the costs from the risk-shifting friction, however. The planner must still leave rents to agents for whom the risk-shifting option is valuable. Utilitarian welfare under $v^{***}$ is still strictly lower than absent the risk-shifting friction.

## 2.4 Extensions

### 2.4.1 Date-0 income reports

In order to make the situations with and without commitment easier to compare, Section 2.3 studies mechanisms whereby income reports take place at date 1 as in the case without commitment. If the planner can commit, however, it must be weakly dominant that he asks for income reports at date 0, collects the reported incomes, and commits to date-1 net transfers based on such reports. With such date-0 reports, agents must first decide on how much income to conceal and then on how to gamble with it. With date-1 reports, agents first gamble with their entire income, and then decide on how much to conceal conditionnally on observing the outcome of the gamble. This latter situation obviously gives agents more room for opportunistic behaviour than the one with date-0 reports.[8] The following lemma shows that the introduction of such date-0 income reports does actually not strictly improve welfare.

**Lemma 13**

*Suppose that the planner asks for reports at date* 0. *The optimal scheme is still* $(r^{***}, v^{***})$ *and there is still no risk shifting in equilibrium.*

[8] To see this, notice that an agent can always gamble with a fraction $w - w'$ of his income $w$, store $w'$ at the risk-free rate and report $w'$ at date 1, thereby replicating a date-0 report of $w'$.

**Proof.** See the Appendix.■

This result is essentially driven by the superadditivity of $g$. The binding constraint at any income level is that agents report any income at all, in which case the ordering of the gambling and reporting decisions is irrelevant.

### 2.4.2 Non CARA utility

The CARA assumption greatly simplifies equilibrium characterization because, as formalized in (12), the attitude of an agent towards risk is not affected by his beliefs about the date-1 income distribution $F_1$. In other words, one's investment decision does not depend on others' investment decisions. This is no longer the case in the presence of varying absolute risk-aversion. Equilibrium characterization in this case requires that one characterizes optimal individual risk-taking decisions for given beliefs about $F_1$, then derives how a given collection of individual risk-taking decisions aggregate into a date-1 income distribution $F_1$, and finally finds a fixed point for this system. We solve this problem here in the particular example studied above in which $g$ is the piecewise linear function given by (15). Within this example, we solve for the equilibrium without commitment, only relaxing the assumption that $u$ is CARA, and assuming instead that it is an increasing and strictly concave function such that

$$\frac{u(y)}{y} \underset{y \to +\infty}{\to} 0. \tag{19}$$

In this case, equation (12) no longer holds, and (11) does not simplify into (13). Yet, for any given *ex post* distribution $F_1$, it is still the case that the tax scheme is given by (10). Thus it must be that $u \circ v^{**}$ has the generic shape depicted in Figure 2: piecewise concave with a unique convex kink. Thus, for a given arbitrary *ex post* income distribution $F_1$, there exists two thresholds $\underline{w}^{F_1} \leq c \leq \overline{w}^{F_1}$ such that agents store at the risk-free rate when their initial income $w_0 \notin \left[\underline{w}^{F_1}, \overline{w}^{F_1}\right]$, and invest in mean-preserving binary lotteries with support $\left\{\underline{w}^{F_1}; \overline{w}^{F_1}\right\}$ otherwise.

But this means in turn that possible date-1 income distributions are characterized by two numbers $\underline{W} \leq c \leq \overline{W}$ such that:

- $F_1$ coincides with $F_0$ over $[0, \underline{W})$ and $\left(\overline{W}, +\infty\right)$,

- The mass of $F_0$ between $\underline{W}$ and $\overline{W}$ is split into two atoms of $F_1$, in $\underline{W}$ and $\overline{W}$, while preserving the mean.

We denote $F_1^{\left(\underline{W}, \overline{W}\right)}$ such a date-1 distribution built this way from $F_0$ and a given pair $\left(\underline{W}, \overline{W}\right)$. Thus we have introduced two mappings. One maps a pair $\left(\underline{W}, \overline{W}\right)$ into the transformation of $F_0$ denoted $F_1^{\left(\underline{W}, \overline{W}\right)}$. The other maps a distribution $F_1$ into a pair $\left(\underline{w}^{F_1}, \overline{w}^{F_1}\right)$. Any fixed point of this system defines an equilibrium. In other words, any pair $(x, y)$ such that

$$(x, y) = \left(\underline{w}^{F_1^{(x,y)}}, \overline{w}^{F_1^{(x,y)}}\right) \tag{20}$$

defines an equilibrium. Standard compacity and continuity arguments ensure that the fixed-point problem (20) has a solution, so that there exists at least one equilibrium. Equilibrium uniqueness is an open question.

### 2.4.3 Alternative planner's objectives

It is easy to see that it is still the case in the presence of risk shifting that the constants $v^{**}(0)$ and $v^{***}(0)$ do correspond to the maximal tax capacity of the planner. That is, solving the program $(\wp')$ described in (7) in the presence of risk shifting leads to an *ex post* tax capacity $v^{**}(0)$ and an *ex ante* tax capacity (assuming commitment power) $v^{***}(0)$. More generally, the tax schemes $v^{**}$ and $v^{***}$ are the optimal ones under the alternative planner's objectives studied in Section 1.2.1 in the presence of risk shifting, in the respective cases of limited and full commitment.

# 3 Dynamics

The assumption that consumption takes place at a single date is meant to focus on risk-taking decisions, and to abstract from consumption smoothing. This section relaxes this assumption. More precisely, we leave the setup described in section 2.1 unchanged, except for agents' preferences. We suppose in this section that they are of the form

$$E\left[u(c_0) + u(c_1)\right],$$

where $u$ is CARA, and $c_0$ and $c_1$ are respective consumptions at dates 0 and 1. We let $u(x) = -e^{-\alpha x}$. We suppose that agents privately learn the value of their entire life income $w$ at date 0. This life income has c.d.f. $F$ within the population. Agents can secretly borrow and lend at the zero risk-free rate. They can also take arbitrary positions in assets with zero expected return and arbitrary risk profiles. Secret consumption is subject to the same technology as before: at each date, concealed income $x$ can be converted into $g(x)$ units of consumption.

In sum, we only add the ingredient that agents value date-0 consumption. We adopt again the standpoint of a social planner seeking to maximize utilitarian welfare. We suppose that the planner can borrow and store at the risk-free rate. He thus faces only an intertemporal budget constraint.

Note that we focus on the case in which agents face no income uncertainty. This way, we abstract from insurance concerns. Consistent with the rest of the paper, this enables us to study the impact of redistribution with frictions on an economy that is trivially Pareto efficient at the outset.

We do three things in this section. First, we study the optimal consumption and saving choices of an agent who hides all of his income. Second, we solve for the problem facing a social planner who can commit to a long-term redistribution tax scheme. Last, we study the case in which the planner lacks commitment.

To start, we compute the utility that an agent who secretly consumes his entire endowment can derive. Suppose that an agent with an endowment $w$ seeks to consume it entirely secretly, and that he does not receive any transfer from the social planner. This agent now faces two decisions. First, he decides on the fraction of his total income $w$ that he saves for date-1 consumption. Second he sets the risk profile of his saving vehicle. From the previous sections, we know that if he saves $w_1$ for future consumption, the agent will optimally invest so as to generate the binary lottery $\widetilde{\rho}\,(w_1)$ with mean $w_1$ and realizations $\{\underline{w}\,(w_1)\,;\overline{w}\,(w_1)\}$ such that

$$E\left[u \circ g\left(\widetilde{\rho}(w_1)\right)\right] = \overline{u \circ g}(w_1).$$

This means that the agent's decision boils down to

$$\max_{w_1} u \circ g\left(w - w_1\right) + \overline{u \circ g}\left(w_1\right). \tag{21}$$

This is a non-convex program which yet admits a simple solution:

**Lemma 14** *The agent reaches utility*

$$2\overline{u \circ g}\left(\frac{w}{2}\right)$$

*as follows:*

*If* $u \circ g\left(\frac{w}{2}\right) = \overline{u \circ g}\left(\frac{w}{2}\right)$*, then the agent invests* $\frac{w}{2}$ *at the risk-free rate, and consumes the rest at date 0.*

*Otherwise, if* $\underline{w}\left(\frac{w}{2}\right) \leq \frac{w}{2} \leq \frac{\underline{w}\left(\frac{w}{2}\right)+\overline{w}\left(\frac{w}{2}\right)}{2}$*, then the agent consumes* $\underline{w}\left(\frac{w}{2}\right)$ *at date 0, and invests the residual* $w - \underline{w}\left(\frac{w}{2}\right)$ *in a fair lottery with support* $\left\{\underline{w}\left(\frac{w}{2}\right);\overline{w}\left(\frac{w}{2}\right)\right\}$ *. If* $\frac{\underline{w}\left(\frac{w}{2}\right)+\overline{w}\left(\frac{w}{2}\right)}{2} < \frac{w}{2} \leq \overline{w}\left(\frac{w}{2}\right)$*, then the agent consumes* $\overline{w}\left(\frac{w}{2}\right)$ *at date 0, and invests the residual* $w - \overline{w}\left(\frac{w}{2}\right)$ *in a fair lottery with support* $\left\{\underline{w}\left(\frac{w}{2}\right);\overline{w}\left(\frac{w}{2}\right)\right\}$*.*

**Proof.** See the appendix.■

It is instructive to discuss Lemma 14 in the particular case in which $g$ is piecewise linear, given by (15). In this case, $\overline{u \circ g}$ coincides with $u \circ g$ outside the interval $[\underline{w}, \overline{w}]$ over which it is a straight line. Lemma 14 shows that an agent such that $w \in (2\underline{w}, 2\overline{w})$ uses both risk shifting and the shifting of consumption over time in order to concavify his utility. As in the previous sections, risk shifting serves to concavify utility over date-1 consumption. In addition, shifting income over time in a distorted fashion serves to equate marginal utilities over pre-tax income at dates 0 and 1. If $w \in (2\underline{w}, \underline{w} + \overline{w})$, the agent consumes only $\underline{w} < \frac{w}{2}$ at date 0, and thus over-saves compared with a standard situation of concave preferences. If $w \in (\underline{w} + \overline{w}, 2\overline{w})$, then conversely the agent consumes $\overline{w} > \frac{w}{2}$ at date 0, and thus under-saves in comparison with a standard situation with concave preferences. Risk shifting makes date-1 pre-tax income distribution more diffuse - with atoms in $\underline{w}$ and $\overline{w}$, while the distorted saving behaviour also generates atoms in $\underline{w}$ and $\overline{w}$ in the date-0 pre-tax income distribution. Notice that this cross-sectional savings patterns can be interpreted as an overreaction of current consumption to life-income shocks, since agents switch from over-saving to under-saving as their life-income increases and crosses $\underline{w} + \overline{w}$.

Lemma 14 is useful in the analysis for two reasons: First, it generates the agent's outside option from secret consumption $2\overline{u \circ g}\left(\frac{w}{2}\right)$ which plays a key role in solving the case in which the planner can commit. Second, we will show that absent commitment, agents do not report their entire income and do secretly consume in equilibrium. The insight from Lemma 14 will help derive their consumption and savings profile in this case. We first study the case in which the planner can fully commit to a long-term scheme.

### 3.1 The planner can commit

Suppose the planner can commit to date-1 actions. In this case, a direct mechanism can be described by the functions $(r, v_0, v_1, \widetilde{\sigma}, \sigma)$ s.t. an agent with life income $w$ reports $r(w) \in [0, w]$ at date 0, and receives net consumption $v_i(r(w))$ from the planner at date $i \in \{0; 1\}$. The function $\sigma(w)$ describes the share of concealed income $w - r(w)$ that the agent stores for secret date-1 consumption, using a fair lottery $\widetilde{\sigma}(w)$ with mean $\sigma(w)$. The residual $w - r(w) - \sigma(w)$ may be secretly consumed at date 0. Formally, the planner solves the program $\left(\wp^D\right)$ :

$$\max_{(r,v_0,v_1,\widetilde{\sigma},\sigma)} \int_0^{+\infty} \left[ \begin{array}{c} u\left(v_0(r(w)) + g\left(w - r(w) - \sigma(w)\right)\right) \\ +E\left[u\left(v_1(r(w)) + g\left(\widetilde{\sigma}(w)\right)\right)\right] \end{array} \right] dF(w) \tag{22}$$

$$s.t. \left\{ \begin{array}{c} \int_0^{+\infty} \left(v_0(r(w)) + v_1(r(w))\right) dF(w) \leq \int_0^{+\infty} r(w) dF(w), \\ \left[ \begin{array}{c} u\left(v_0(r(w)) + g\left(w - r(w) - \sigma(w)\right)\right) \\ +E\left[u\left(v_1(r(w)) + g\left(\widetilde{\sigma}(w)\right)\right)\right] \end{array} \right] \\ = \max_{w',s,\widetilde{s}} \left\{ \begin{array}{c} u\left(v_0(r(w')) + g\left(w - r(w') - s\right)\right) \\ +E\left[u\left(v_1(r(w')) + g\left(\widetilde{s}\right)\right)\right] \end{array} \right\} \\ s.t.\ 0 \leq r(w') \leq w, \\ 0 \leq s \leq w - r(w'), \\ E\left[\widetilde{s}\right] = s. \end{array} \right. \tag{23}$$

We denote $\widetilde{s}$ a generic random variable with support included in $[0, +\infty)$. The program $\left(\wp^D\right)$ simply states that the mechanism must satisfy a budget constraint (this is the first constraint), and that it must be incentive-compatible (second constraint) given that agents now have three degrees of freedom: they can choose a report, decide on how much of their concealed endowment to save, and on the risk profile of their savings.[9]

Notice that with commitment, the mechanisms studied here whereby agents report their entire income at date 0 must be weakly better than more realistic mechanisms whereby agents report their disposable income at date 0 (the income that they do not store) and then their disposable income at date 1. The reasoning is the same as that in Section 2.4.1.[10] We have:

**Proposition 15**

*The solution to program $\left(\wp^D\right)$ is attained with $\left(r^D, v^D, v^D\right)$ defined as*

$$\left\{ \begin{array}{c} r^D(w) = w, \\ v^D(w) = v^D(0) + u^{-1} \circ \overline{u \circ g}\left(\frac{w}{2}\right) \end{array} \right. ,$$

*where*

$$v^D(0) = \int_0^{+\infty} \left( \frac{t}{2} - u^{-1} \circ \overline{u \circ g}\left(\frac{t}{2}\right)\right) dF(t).$$

*Agents report their entire life income. This delivers an agent with life-income $w$ an expected utility*

$$-2e^{-\alpha\left(\int_0^{+\infty}\left(\frac{t}{2} - u^{-1}\circ\overline{u\circ g}\left(\frac{t}{2}\right)\right)dF(t)\right)}\overline{u \circ g}\left(\frac{w}{2}\right).$$

[9] Notice that we consider only deterministic mechanisms. It is for brevity, and without loss of generality for exactly the same reasons as in the static case.

[10] It is possible to show that, as in Lemma 13, a full date-0 report does not strictly improve upon mechanisms with reports at dates 0 and 1.

**Proof.** See the appendix.■

Proposition 15 essentially shows that the insights from the static model extend to the dynamic case when the planner can commit. First, the optimal redistribution consists in making every agent indifferent between reporting his entire income or none of it. Second, the planner finds it preferable to concavify utilities over pre-tax income rather than let agents do it themselves through risk shifting and distorted savings in the way described in Lemma 14. By concavifying himself, not only does the planner reduce the dispersion of pre-tax income in the cross-section at date 1 (by eliminating excessive risk taking), but he also eliminates under and over savings at date 0, and thus smoothes income over time compared with the case in which agents secretly consume.

## 3.2 The planner cannot commit

Suppose now that the planner cannot commit at date 0 to date-1 actions. At date 1, the planner maximizes date-1 utilitarian welfare given his information about each agent's endowment and his budget constraint. Interestingly, the no-avoidance result that held in all the cases considered thus far breaks down in this case: there must be secret consumption in equilibrium. In this section, we make the following mild technical assumption:

**Assumption 2.** *The function $g$ has left and right derivatives that are uniformly bounded over* $[0, +\infty)$ .

The functions described in Lemma 2 satisfy this condition.

**Proposition 16** *Suppose $g \neq 0$. An optimal mechanism must be such that date-0 income reports have an upper bound, and that secret consumption at date* 0 *grows without bound and tends to $\frac{w}{2}$ as $w \to +\infty$.*

**Proof.** See the appendix.■

This result is essentially a form of the ratchet effect in dynamic agency problems without commitment (see, e.g., Laffont and Tirole, 1988). The intuition is the following. It is very costly for an agent to report a sizeable fraction of his total income at date 0 because this raises the planner's beliefs about this agent's pre-tax income at date 1. The planner taxes in an *ex post* optimal fashion at date 1. Therefore, if he infers from the date-0 report how much an agent has saved for date 1, he cannot commit not to extract most of it (net of a date-1 minimum payment). This implies that date-1 net consumption has an upper bound when date-0 reports are unbounded. This cannot be incentive-compatible, as a sufficiently wealthy agent will always prefer to report no income at all initially, so as to be able to freely smooth consumption across dates.

Interestingly, the risk-shifting friction can ease the revelation of income at date 0. Suppose that $g$ is such that in equilibrium, an agent with income $w_0$ gambles between dates 0 and 1 and that his low payoff is 0. Then, revealing his entire income $w_0$ at date 0 can be an equilibrium strategy because it comes at no cost to the agent at date 1. The planner cannot rule out that the agent has no income at date 1 given that 0 is in the support of his gamble, and thus he cannot exploit the date-0 report at date 1. But this potentially positive role of gambling is irrelevant at high income levels at which gambles have very small supports because $u'$ decreases exponentially while $g$ has bounded derivatives.

Notice that absent commitment, since there is secret consumption, the behaviour described in Lemma 14, featuring distorted savings and excessive risk taking, must occur in equilibrium if agents with a sufficiently large income $w$ have a nonconcave utility in $\frac{w}{2}$.

The no-avoidance property was instrumental in generating simple optimal schemes in the cases previously studied. Absent this property here, we are unable to characterize the optimal scheme in general. The next result studies an interesting limiting case in which agents report no income at all at date 0. In this case, the optimal scheme can be characterized. We derive the optimal scheme in the limiting case in which agents are infinitely risk-averse. This corresponds for example to letting the index of absolute risk aversion tend to infinity in the CARA case. Infinite risk-aversion is the limiting case in which agents care only about the worst outcome across dates and states. In this case, a utilitarian planner maximizes the lowest consumption level across agents and dates.

**Lemma 17** *Suppose that agents are infinitely risk-averse and that the left and right derivatives of $g$ are larger than $\frac{1}{2}$. Then the optimal scheme is such that all agents report a zero-income at date 0. They secretly consume half of their income at date 0 and store the other half at the risk-free rate. At date 1, they entirely report the proceeds. The net transfers at dates 0 and 1, $v_0$ and $v_1$, are*

$$\begin{aligned} v_0(w) &= K, \\ v_1(w_1) &= K + g(w_1), \end{aligned}$$

*where $w_1$ is the date-1 report and $K = \frac{1}{2}\int \left(\frac{w}{2} - g(\frac{w}{2})\right) dF(w)$.*

**Proof.** See the appendix.■

In other words, the planner can only collect taxes at date 1 on half the aggregate income. He splits the proceeds into two equal constant terms at dates 0 and 1. The intuition for this result is the following. Recall the intuition behind Proposition 16 that agents are reluctant to reveal how affluent they are at date 0 because the planner cannot commit not to tax them in a confiscatory fashion at date 1. One way the report of any income at all at date 0 can be made incentive-compatible is if, unlike in the full commitment case, the minimum net consumption level at date 1 is larger than that at date 0. In this case, agents do not need to save in order to smooth consumption provided their total income is not too large, and thus are not worried about future confiscatory taxes. But this solution is not desirable when risk-aversion is infinite and the avoidance technology is sufficiently efficient (i.e., the derivatives of $g$ are larger than $\frac{1}{2}$) because the required difference in fixed payments at dates 0 and 1 must be large in this case. This makes the date-0 minimum consumption level - the one that matters to the social planner - too low.

Finally, the reader may wonder whether it would conversely be preferable to give a lower minimum consumption level at date 1 than at date 0 when there are no initial reports. This would induce agents to save more, and thus would increase the date-1 tax collection on realized savings. This is actually not the case when the derivatives of $g$ are larger than $\frac{1}{2}$. If avoidance is too efficient, the net effect of an increasing gap between minimum levels and of an increase in their sum due to higher tax collection is a decrease in the date-1 minimum level - which is the one the planner cares for since it is lower than the date-0 minimum consumption. With less efficient avoidance technologies, we conjecture that the planner may find it worthwhile to induce high savings with unequal minimum consumption across dates.

# 4 Systematic risk taking

This section develops a version of the model in which agents do not take idiosyncratic risks. It studies the polar case in which all agents share access to a unique source of risk. We modify the model set up in Section 2.1 in two ways.

First, agents store their income from date 0 to date 1 by continuously trading two assets over the time interval $[0, 1]$. The first asset has a constant unit price at all date $t \in [0, 1]$. The second one is a risky asset whose price process $(P_t)_{t\in[0,1]}$ obeys

$$\left\{ \begin{array}{c} P_0 = 1 \\ \frac{dP_t}{P_t} = \sigma W_t \end{array} \right. ,$$

where $\sigma > 0$ and $W_t$ is a standard Wiener process. This asset pays no dividend, and its price fluctuations are the only source of risk in this economy. These available assets correspond to a version of the Black-Scholes (1973) model with a zero risk-free rate and a zero-return on the risky asset.

Second, we do not assume that the social planner seeks to maximize utilitarian welfare. We suppose instead that he seeks to maximize the taxes that he collects from the population, as described in Section 1.2.1. We motivate this assumption below.

For expositional brevity only, we assume that $g$ is piecewise linear, given by (15). We first solve the model in the case in which the planner taxes in an *ex post* optimal fashion at date 1. Then we derive the full commitment case. Before proceeding, we collect some elementary properties of binary options in the Black-Scholes model that will prove useful.

## 4.1 Replicating binary options

A cash-or-nothing call of maturity 1 with strike $K$ is a derivative security that pays off 1 if $P_1$, the price of risky asset at date 1, is above $K$, and pays off 0 otherwise. Any date-1 binary lottery can be viewed as the payoff of a portfolio invested in the risk-free asset and in a cash-or-nothing call with the appropriate strike. Denoting $\Phi$ and $\varphi$ the c.d.f. and p.d.f. of a standard normal variable, we have

**Lemma 18**

*The payoff of a cash-or-nothing call of maturity* 1 *with strike* $K$ *can be replicated with a dynamic self-financed strategy that consists in continuously trading the risky and risk-free assets between* $t = 0$ *and* $t = 1$*. The initial total amount that needs to be invested to replicate the call is:*

$$\Phi\left(-\frac{\ln K}{\sigma} - \frac{\sigma}{2}\right),$$

*which is also the probability that the call pays off* 1 *at date* 1*. The number of units of the risky security that must be held at date* $t \in [0, 1]$ *as part of the replicating strategy is*

$$q(K, P_t, t) = \frac{1}{\sigma\sqrt{1-t}P_t}\varphi\left(\frac{1}{\sigma\sqrt{1-t}}\ln\frac{P_t}{K} - \frac{\sigma}{2}\sqrt{1-t}\right). \tag{24}$$

**Proof.** See the appendix.■

It is easy to see from (24) that the replicating strategy consists in holding a number of shares $q(K, P_t, t)$ that is non-monotonic in $P_t$. These textbook results on option replication are interesting in our context. They show that the abstract binary lotteries considered in the rest of the paper could be viewed more concretely as the resulting payoffs from continuous trading strategies with time-varying risk exposures on assets with idiosyncratic risk.

## 4.2 Tax capacity without commitment

Suppose that the social planner solves at date 1 the program $(\wp')$ described in (7) after the date-1 income distribution is realized. From Corollary 6, the optimal tax scheme does not depend on this date-1 distribution. It is simply such that each agent fully reports his date-income $w_1$ and is left with $g(w_1)$. Thus, each agent trades between date 0 and date 1 so that the distribution of his date-1 income viewed from date 0 solves the program (13). We know from Lemma 9 that there exists $\underline{w} < c < \overline{w}$ such that an agent with a date-0 income $w \in [0, \underline{w}] \cup [\overline{w}, +\infty)$ finds it optimal to store at the risk-free rate. If the initial income $w \in (\underline{w}, \overline{w})$, then the agent would like to receive a fair binary payoff with support $\{\underline{w}; \overline{w}\}$ at date 1. This agent would like to invest $\underline{w}$ in the risk-free asset, and the residual $w - \underline{w}$ in $\overline{w} - \underline{w}$ units of an asset that pays off 1 with probability $\frac{w-\underline{w}}{\overline{w}-\underline{w}}$ and 0 otherwise. Lemma 18 shows that this agent can readily replicate this asset. This yields the following outcome.

**Proposition 19**

*An agent with initial income $w \notin (\underline{w}, \overline{w})$ invests at the risk-free rate. An agent with initial income $w \in (\underline{w}, \overline{w})$ invests $\underline{w}$ in the risk-free asset, and with the residual replicates a cash-or-nothing call with strike*

$$K(w) = e^{-\sigma \Phi^{-1}\left(\frac{w-\underline{w}}{\overline{w}-\underline{w}}\right) - \frac{\sigma^2}{2}}. \tag{25}$$

*As a result, the aggregate demand for the risky asset at date $t$ is*

$$\frac{(\overline{w} - \underline{w})\, e^{\frac{\sigma^2 t}{2} - \sigma W_t}}{\sigma\sqrt{1-t}} \int_{\underline{w}}^{\overline{w}} \varphi\left(\frac{W_t + \Phi^{-1}\left(\frac{w-\underline{w}}{\overline{w}-\underline{w}}\right)}{\sqrt{1-t}}\right) dF_0(w). \tag{26}$$

*The date-1 taxes collected by the planner are of the form*

$$T - \tau F_0\left(\overline{w} - (\overline{w} - \underline{w})\, \Phi(W_1)\right), \tag{27}$$

*where* $\Phi(W_1)$ *has by construction a uniform distribution over* $(0,1)$ *viewed from date* 0, *and*

$$T = \int_{[0,\underline{w}]\cup[\overline{w},+\infty)} (t - g(t))\, dF_0(t) + (\overline{w} - g(\overline{w}))\, F_0(\overline{w}) - (\underline{w} - g(\underline{w}))\, F_0(\underline{w}),$$
$$\tau = (\overline{w} - \underline{w} + g(\underline{w}) - g(\overline{w})).$$

**Proof.** See the appendix.■

Proposition 19 yields two interesting insights. First, the date-1 revenues of the planner are not surprisingly random since agents gain exposure to a systematic source of risk. More interestingly, the consequence from dynamic replication is that the distribution of these revenues is very different from that of the return on the risky asset. It reflects instead the date-0 distribution of pre-tax income: Expression (27) shows that the random component of taxes is of the form $-F_0\left(-\widetilde{X}\right)$, with $\widetilde{X}$ uniformly distributed over $(\underline{w}, \overline{w})$. Thus the Gaussian nature of asset risk is not reflected in (27). This connection between inequality among top incomes and the risk profile of their taxes is a novel prediction to our knowledge. More generally, the prediction that endogenous portfolio rebalancing creates a nonlinear relationship between realized asset returns and collected taxes is novel to our knowledge.

Second, expression (26) shows how the demand for the risky asset dynamically evolves. It is easy to analyze in the case in which $t$ is very close to 1. It is transparent from (26) that for $t$ sufficiently close to 1, the exposure on the risky asset is concentrated on agents with an initial income $w$ such that:

$$w = \underline{w} + (\overline{w} - \underline{w})\, \Phi(-W_t).$$

Thus for $t$ close to 1, the demand for the risky asset is driven by agents with initial income close to $\overline{w}$ after bad realizations of the asset return ($W_t$ small), and by agents with initial income close to $\underline{w}$ after positive realizations of the asset return. It seems reasonable to assume that income distribution $F_0$ has a decreasing density over $[\underline{w}, \overline{w}]$ because avoidance is is concentrated at the right tail of income distribution, where density is decreasing. In this case, total demand for the risky asset should chase the realized return. It increases in particular after positive realized returns.

### 4.3 Tax capacity with commitment

We now consider the case in which the planner still seeks to maximize his expected date-1 revenues, but can credibly commit to a tax scheme that he announces at date 0. In this case,

**Proposition 20**

*The planner announces the tax scheme*

$$\begin{aligned} r(w) &= w, \\ \tau(w) &= w - u^{-1} \circ \overline{u \circ g}(w). \end{aligned}$$

*Agents invest at the risk-free rate and the tax capacity is*

$$\int_0^{+\infty} \left(w - u^{-1} \circ \overline{u \circ g}(w)\right) dF_0(w).$$

**Proof.** See the Appendix.■

Commitment power is particularly valuable in the case of systematic risk taking. Not only does it raise the planner's tax capacity, but it also severs the link between tax avoidance and financial instability discussed above. The planner commits in this case to a tax scheme that is less regressive than the *ex post* optimal one. This eliminates risk-taking incentives and thus yields to a stable distribution of taxable income, but comes at the cost of a smaller tax capacity given this *ex post* distribution.

## 4.4 Discussion

### General avoidance function

It is easy to see that the results are qualitatively similar for a general avoidance function $g$. In this case each agent with initial income $w$ uses dynamic trading in order to replicate the binary option $\widetilde{\rho}(w)$ corresponding to his optimal lottery.

### Welfare-maximizing planner

Maximizing tax capacity no longer coincides with utilitarian welfare maximization when risk shifting involves exposure to systematic risk. Consider the case of taxation without commitment. In this case, the constant term $v(0)$ of the welfare-maximizing tax scheme is also contingent on $W_1$, and positively correlated with the variable term $v(w) - v(0)$. Thus a hedging motive arises for each agent and his optimal trading strategy is more complex. We have been unable to characterize it simply.

The situation in which maximization of tax capacity and that of utilitarian welfare coincide is that in which the planner must announce a fixed amount of expenditures at date 0, and then collects taxes at date 1. In this case he breaks even on average but bears systematic risk because of revenue uncertainty. This may be a realistic description of public finances in the short to medium run, where expenditures are much less sensitive to the evolution of the economy than collected taxes.

## 5 Related literature

This paper studies how two frictions - avoidance and risk shifting - affect the tax and redistribution capacities of a social planner. As such, it blends ingredients that have been studied in distinct literatures, and from different angles. This section discusses how this paper relates to the respective existing literatures on tax avoidance, on risk shifting, and on the effect of frictions on inequality.

First, there is a surprising contrast between the large evidence that taxpayers do take advantage of available legal methods of reducing their fiscal obligations, and the relatively sparse theoretical literature on this topic. The literature on tax avoidance is by and large descriptive (see, e.g., Stiglitz, 1985). Slemrod and Kopczuk (2002) and Piketty, Staez, and Stancheva (2013) capture avoidance in a reduced form, as an exogenous elasticity of taxable income to the tax rate. Like us, Casamatta (2013) and Grochulski (2007) adopt the alternative approach of modelling tax avoidance as a primitive informational friction, and then deriving optimal fiscal policy as an optimal mechanism. We share with these contributions the modelling of avoidance as an *ex post* moral-hazard problem of costly diversion. Grochulski (2007) establishes the result that increasing returns to avoidance imply the optimality of avoidance-free schemes. Casamatta (2013) shows that this no longer need be the case when the function $g$ is concave.

Second, the risk-shifting friction is a form of *ex ante* moral hazard that has been thoroughly studied in financial economics. In their seminal paper, Jensen and Meckling (1976) show that overly leveraged firms may undertake value-destroying projects provided these are sufficiently risky. A large asset-pricing literature studies how nonconcavities stemming from compensation schemes or career concerns create risk-shifting incentives for fund managers. Contributions include Basak, Pavlova, and Shapiro (2007), Carpenter (2000), Ross (2004), and Makarov and Plantin (2013). We borrow our formal modelling of risk shifting as a choice among arbitrary distributions from the latter.

Our focus on how risk taking shapes the wealth distribution relates to a number of paper that study economies in which agents care not only for consumption but also for their status (see, e.g., Becker, Murphy, and Werning, 2003, Ray and Robson, 2012, or Robson, 1992). In these contributions, status may induce non concavities in utility over endowment, so that only wealth distribution that are sufficiently inequal discourage agents from gambling. By contrast, agents care only about consumption in our economy, and only scale economies in avoidance and lack of commitment generate endogenous pre-tax inequality.

While risk shifting is largely absent from the public-finance literature to our knowledge, it relates to various forms of secret side-trading studied by this literature. Contributions include Cole and Kocherlakota (2001), Golosov and Tchivisty (2007), or Ales and Maziero (2013). In Cole and Kocherlakota (2001), agents can secretly save at an exogenously given rate. Golosov and Tsyvinski (2007) endogenize the price of the assets that agents secretly trade. Ales and Maziero (2013) study non-exclusive contracting, and thus fully endogenize the side contracts that agents can secretly sign. Broadly, the goal of this literature is to study how agents' ability to secretly trade affects efficient production and risk sharing in economies with asymmetric information. Our purpose is quite different. We study an economy that is trivially Pareto efficient at the outset. There are no gains from social interaction between agents: They receive risk-free endowments of a single private good and do not produce. Tax avoidance and side trades matter only because of the presence of a social planner who uses taxation to implement inequality-averse social views. Our focus is on how tax avoidance and risk shifting stand in the way of this social planner, and may lead him to create two types of inefficiency. First, he may induce tax avoidance, which wastes resources. Second, he may also spur risk shifting, thereby adding non rewarded risk to a risk-averse economy. We show that the commitment power of the planner determines whether he creates such costs for the economy or not. A planner who can commit designs schemes that do not induce tax avoidance nor risk shifting in equilibrium. This crucially relies on increasing returns to avoidance, as shown by Casamatta (2013). Absent commitment, we show that there is tax avoidance in equilibrium in the dynamic version of the model, and that risk shifting also occurs in equilibrium.

Acemoglu, Golosov, and Tsyvinski (2010), and Farhi, Sleet, Werning, and Yeltekin (2012) also study environments in which an inequality-averse planner cannot commit. Perhaps closest to our paper, Bisin and Rampini (2006) introduce anonymous trading in such an environment, and show that it might be preferable to publicly observed trading because the limited information of the governement mitigates time-inconsistency problems. Our result that there is secret consumption in equilibrium in the dynamic case without commitment is related to theirs.

Finally, the prediction that regressive taxation at the top endogenously increases pre-tax inequality because of a risk-shifting friction is novel to our knowledge, although Posner informally made this claim.[11] The growth and development literatures have shown that credit constraints may create poverty traps that amplify income inequality (see, e.g., Aghion and Bolton, 1997, Banerjee and Newman, 1993, Galor and Zeira, 1993, or Greenwood and Jovanovic, 1990). We suggest that another friction - risk shifting - may amplify inequality at the top of the income distribution. No systematic empirical test of this prediction has been carried out to our knowledge. Yet, Gentry and Hubbard (2000) and Cullen and Gordon (2007) document the related fact that entrepreneurial risk taking is reduced when taxation becomes more progressive.

# References


[1] Acemoglu, Daron, Michael Golosov and Aleh Tsyvinski, "Dynamic Mirrless Taxation Under Political Economy Contraints," Review of Economic Studies, 77 (3), 2010.

[2] Aghion, Philippe and Patrick Bolton, "A Theory of Trickle-Down Growth and Development," The Review of Economic Studies, 64 (2), 1997.

[3] Ales, Laurence and priscilla Maziero, "Non-Exclusive Dynamic Contracts and the Limits of Insurance," working paper, 2013.

[4] Aumann, Robert and Micha Perles, "A variational problem arising in economics," Journal of Mathematical Analysis and Applications, 11, 1965.

[5] Banerjee, Abhijit and Andrew Newman, "Occupational Choice and the Process of Development", Journal of Political Economy, 101, 1997.

[6] Basak, Suleyman, Anna Pavlova and Alex Shapiro, "Optimal Asset Allocation and Risk Shifting in Money Management," Review of Financial Studies, 20(5), 2007.

[7] Becker, Gary, Kevin Murphy and Iván Werning, "The Equilibrium Distribution of Income and the Market for Status," Journal of Political Economy , 113 (2), 2005.


[11] http://www.becker-posner-blog.com/2006/12/should-we-worry-about-the-rising-inequality-in-income-and-wealth–posner.html

[8] Bisin, Alberto and Adriano Rampini, "Markets as Beneficial Constraints on the Government," Journal of Public Economics, 90, 2006.

[9] Black, Fischer and Myron Scholes, "The Pricing of Options and Corporate Liabilities," Journal of Political Economy, 81 (3),1973.

[10] Carpenter, Jennifer "Does Option Compensation Increase Managerial Risk Appetite?," Journal of Finance, 55(5), 2000.

[11] Casamatta, Georges, "Optimal Income Taxation with Tax Avoidance," working paper, 2010.

[12] Cole, Harlod and Narayana Kocherlakota, "Efficient Allocations with Hidden Income and Hidden Storage," Review of Economic Studies, 68(3), 2001.

[13] Congress of the United States, Congressional Budget Office, "The Distribution of Major Tax Expenditures in the Individual Income Tax System," report, 2013.

[14] Cullen, Julie Berry and Roger Gordon, "Taxes and Entrepreneurial Risk-Taking: Theory and Evidence for the U.S.," Journal of Public Economics, 91(7-8), 2007.

[15] Diaz-Gimenez, Javier, Andy Glover and Jose-Victor Rios-Rull, "Facts on the Distributions of Earnings, Income, and Wealth in the United States: 2007 Update," Federal Reserve Bank of Minneapolis Quarterly review, 34(1), 2011.

[16] Farhi, Emmanuel, Chris Sleet, Ivan Werning and Sevin Yeltekin, "Nonlinear Capital Taxation without Commitment," Review of Economic Studies 79 (4), 2012.

[17] Galor, Oded and Joseph Zeira, "Income Distribution and Macroeconomics," The Review of Economic Studies, 60 (1), 1993.

[18] Gentry, William and R. Glenn Hubbard, "Tax Policy and Entrepreneurial Entry," American Economic Review, 90(2), 2000.

[19] Golosov, Mikhail and Aleh Tsyvinski, "Optimal Taxation with Endogenous Insurance Markets," Quarterly Journal of Economics, 122 (2), 2007.

[20] Greenwood, Jeremy and Boyan Jovanovic, "Financial Development, Growth, and the Distribution of Income," Journal of Political Economy, 98 (5), 1990.

[21] Grochulski, Borys ,"Optimal Nonlinear Income Taxation with Costly Tax Avoidance," Economic Quarterly - Federal reserve Bank of Richmond, 93 (1), 2007.

[22] Jensen, Michael and William Meckling,"Theory of the Firm: Managerial Behavior, Agency Costs and Ownership Structure," Journal of Financial Economics, 3(4), 1976.

[23] Johannesen, Niels and Gabriel Zucman, "The End of Bank Secrecy? An Evaluation of the G20 Tax Haven Crackdown," forthcoming, American Economic Journal: Economic Policy.

[24] Kleven, Henrik, Camille Landais and Emmanuel Saez, "Taxation and International Mobility of Superstars: Evidence from the European Football Market,", forthcoming, American Economic Review.

[25] Kleven, Henrik, Camille Landais, Emmanuel Saez, and Esben Schultz, "Migration and Wage Effects of Taxing Top Earners: Evidence from the Foreigners' Tax Scheme in Denmark," NBER Working Paper No. 18885, 2013.

[26] Laffont, Jean-Jacques and Jean Tirole, "The Dynamics of Incentive Contracts," Econometrica , 56 (5), 1988.

[27] Landais, Camille, Thomas Piketty and Emmanuel Saez, "Pour une Révolution Fiscale," La République des Idées, Seuil, 2011.

[28] Lang, Oliver, Norhass, Karl-Heinz and Konrad Stahl,"On Income Tax Avoidance; The Case of Germany," Journal of Public Economics, 66, 1997.

[29] Makarov, Igor and Guillaume Plantin, "Rewarding Trading Skills Without Inducing Gambling," forthcoming, Journal of Finance.

[30] Piketty, Thomas, and Emmanuel Saez, "Income Inequality in the United States, 1913-1998" with Thomas Piketty, Quarterly Journal of Economics, 118(1), 2003, table updates found at http://elsa.berkeley.edu/~saez/

[31] Piketty, Thomas, Emmanuel Saez and Stefanie Stantcheva, "Optimal Taxation of Top Labor Incomes: A Tale of Three Elasticities," forthcoming, American Economic Journal: Economic Policy.

[32] Rajan, Raghuram, "Bankers' Pay Is Deeply Flawed," Financial Times, 2010.

[33] Ray, Debraj and Arthur Robson, "Status, intertemporal Choice, And Risk-Taking," Econometrica , 80 (4), 2012.

[34] Arthur Robson, "Status, the Distribution of Wealth, Private and Social Attitudes to Risk," Econometrica , 60 (4), 1992.

[35] Ross, Stephen, "Compensation, Incentives, and the Duality of Risk Aversion and Riskiness," Journal of Finance, 59(1), 2004.

[36] Shackelford, Douglas, "The Tax Environment Facing the Wealthy," in Does Atlas Shrug? The Economic Consequences of Taxing the Rich, edited by Joel B. Slemrod, Russell Sage Foundation and Harvard University (New York, NY), 2000.

[37] Slemrod, Joel and Wojciech Kopczuk, "The Optimal Elasticity of Taxable Income," Journal of Public Economics, 84(1), 2002.

[38] Stiglitz, Joseph, "Utilitarianism and Horizontal Equity: The Case for Random Taxation," NBER Working Paper No. 694, 1981.

[39] Stiglitz, Joseph, "The general theory of tax avoidance", National Tax Journal, 38(3), 1985.

[40] Stiglitz, Joseph, "Pareto Efficient and Optimal Taxation and the New New Welfare Economics," NBER Working Paper No. 2189, 1987.

[41] Zucman, Gabriel, "The Missing Wealth of Nations: Are Europe and the U.S. net Debtors or net Creditors?," forthcoming, Quarterly Journal of Economics.

# 6 Appendix

## 6.1 Proof of Lemma 2

**Proof of i).** For all $t \in [0, 1]$ and $w \geq 0$,

$$g(tw) = g(tw + (1-t)0) \leq tg(w) + (1-t)g(0) = tg(w).$$

Thus for all $w$, $w' \geq 0$ s.t. $w + w' > 0$,

$$\begin{aligned} g(w) + g(w') &= g\left(\frac{w}{w+w'}\left(w+w'\right)\right) + g\left(\frac{w'}{w+w'}\left(w+w'\right)\right) \\ &\leq \frac{w}{w+w'} g\left(w+w'\right) + \frac{w'}{w+w'} g\left(w+w'\right) \\ &= g\left(w+w'\right). \end{aligned}$$

**Proof of ii).** We need to show that

$$\forall\ 0 \leq w' \leq w,\ g\left(w-w'\right) + g\left(w'\right) \leq g\left(w\right), \tag{28}$$

for

$$g(x) = \max\left\{f(x) - k; 0\right\}.$$

Inequality (28) clearly holds if

$$\min\left\{f(w-w'); f(w')\right\} \leq k.$$

Otherwise the inequality becomes

$$f\left(w\right) \geq f\left(w'\right) + f(w-w') - k.$$

Concavity of $f$ implies

$$f\left(w'\right) + f(w-w') \leq 2f\left(\frac{w}{2}\right),$$

and thus

$$f\left(w\right) - \left(f\left(w'\right) + f(w-w')\right) \geq f\left(w\right) - 2f\left(\frac{w}{2}\right) \geq -k$$

from (3).■

## 6.2 Proof of Proposition 3

**Step 1.** We first show that we can without loss of generality restrict the analysis to mechanisms such that

$$\forall\ w \geq 0,\ r(w) = w.$$

Consider an arbitrary scheme $(r, v)$ that satisfies constraints (5). Define the scheme $(\rho, \nu)$ as

$$\begin{aligned} \rho(w) &= w, \\ \nu(w) &= v(r(w)) + g(w - r(w)). \end{aligned}$$

First, the scheme $(\rho,\nu)$ is incentive-compatible: For all $w \geq 0$ and $w'$ s.t. $w' < w$, we have

$$\begin{aligned}
\nu(w) &= v(r(w)) + g(w - r(w)) \geq v(r(w')) + g(w - r(w')), \\
&\geq v(r(w')) + g(w' - r(w')) + g(w - w'), \\
&= \nu(w') + g(w - w').
\end{aligned}$$

The first inequality stems from the fact that $(r, v)$ is incentive-compatible. The second one follows from the fact that $g$ is superadditive.

Second, the scheme $(\rho, \nu)$ is feasible:

$$\begin{aligned}
\int_0^{+\infty} \nu(w) dF(w) &= \int_0^{+\infty} v(r(w)) dF(w) + \int_0^{+\infty} g(w - r(w)) dF(w) \\
&\leq \int_0^{+\infty} r(w) dF(w) + \int_0^{+\infty} (w - r(w))\, dF(w) \\
&\leq \int_0^{+\infty} w dF(w) = \int_0^{+\infty} \rho\,(w)\, dF(w).
\end{aligned}$$

Finally, the scheme $(\rho, \nu)$ delivers the same utility as the scheme $(r, v)$ for each income level. Thus, the restriction to avoidance-free schemes is without loss of generality.■

**Step 2.** Consider the following auxiliary program

$$\begin{aligned}
&\max_v \int_0^{+\infty} u\,(v(w))\, dF(w) \\
&s.t. \begin{cases} \int_0^{+\infty} v(w) dF(w) \leq \int_0^{+\infty} w dF(w), \\ \forall w \geq 0,\ v(w) \geq g\,(w) + v(0). \end{cases}
\end{aligned} \tag{29}$$

This amounts to considering only the deviation of a zero-report $w' = 0$ in the incentive-compatibility constraints of $(\wp)$. We will show that

$$V(w) = g\,(w) + \int_0^{+\infty} (t - g(t))\, dF(t)$$

solves this program. It is easy to see that $V$ satisfies constraints (29).

Consider a function $v$ that solves this program. Clearly, $v$ must be (weakly) increasing. Thus, $v$ admits a left limit $v(x^-)$ and a right limit $v(x^+)$ at each point $x \in (0, +\infty)$. Suppose that for some $x_0 \in (0, +\infty)$, $v(x_0^-) < v(x_0^+)$. Then one could slightly increase $v$ in the left neighborhood of $x$, slightly decrease it in the right neighborhood, and thus strictly increase utilitarian welfare while still satisfying constraints (29). Thus $v$ must be continuous over $(0, +\infty)$ almost surely (and with a similar argument also has a right-limit in 0).

Suppose now that for some $x_1 \in (0, +\infty)$,

$$v(x_1) > g(x_1) + v(0). \tag{30}$$

Since $v$ and $g$ are continuous, inequality (30) actually holds over some neighborhood $\Omega$ of $x_1$. Consider a bounded measurable function $h$ with support within $\Omega$ s.t. $\int hdF = 0$. The function

$$w \to v(w) + th(w)$$

satisfies constraints (29) for $t$ sufficiently small. Thus it must be that

$$\Phi(t) = \int_0^{+\infty} u\left(v(w) + th(w)\right) dF(w)$$

has a local maximum in 0, or that

$$\Phi'(0) = \int_0^{+\infty} u'\left(v(w)\right) h(w) dF(w) = 0. \tag{31}$$

Since (31) holds for any function $h$, and $u$ is strictly concave, it must be that $v$ is constant over $\Omega$. Clearly this implies that $v$ must be constant over $[0, x_1)$, which cannot be unless $g$ is equal to 0 over this interval. In any case, this contradicts (30). Thus $v = V$.

Since constraints (29) are necessary conditions for constraints (5) and $V$ happens to satisfy (5) by superadditivity, this concludes the proof.■

## 6.3 Proof of Corollary 4

The proof of Proposition 3 only uses that $u$ is increasing and strictly concave, and so is $\Omega \circ u$.■

## 6.4 Proof of Corollary 5

Apply Corollary 4 with

$$\Omega(x) = -\frac{e^{-ax}}{a},$$

and let $a \to +\infty$.■

### 6.5 Proof of Corollary 6

Clearly the tax scheme $(r^*, \tau^*)$ satisfies the constraints of $(\wp')$ by superadditivity of $g$. Any scheme $(r, \tau)$ that also satisfies these constraints must in particular satisfy:

$$\forall w \geq 0,\ \tau\left(r(w)\right) \leq r(w) + g\left(w - r(w)\right) - g\left(w\right).$$

The right-hand side is maximal for $r(w) = w$ from (1), in which case it is equal to $\tau^*(w)$, which establishes the result.

Taxation cannot be strictly progressive because this would imply that $g$ is strictly concave and thus strictly subadditive.■

### 6.6 Proof of Proposition 7

**Ex post randomization.** Suppose that mechanisms can be random. The planner offers a random net transfer $\widetilde{v}(r(w))$ to an agent who reports $r(w)$. First, it is easy to see that, as in the deterministic case, restricting the analysis to $r(w) = w$ is without loss of generality. As in the deterministic case, it suffices to replace this mechanism with one in which

$$\begin{aligned} \rho(w) &= w, \\ \widetilde{\nu}(w) &= \widetilde{v}(r(w)) + g(w - r(w)). \end{aligned}$$

This new mechanism is feasible and incentive-compatible by superadditivity of $g$, actually regardless of the properties of the utility function $u$.

Second, for any random mechanism $\widetilde{v}(w)$ that is feasible and incentive compatible, define the deterministic mechanism

$$\gamma(w) = u^{-1}\left(E\left[u\left(\widetilde{v}(w)\right)\right]\right).$$

First, this mechanism is incentive-compatible. For all $w \geq 0$ and $w' < w$, incentive-compatibility of $\widetilde{v}$ implies

$$E\left[u\left(\widetilde{v}(w)\right)\right] \geq E\left[u\left(\widetilde{v}(w') + g(w - w')\right)\right],$$

or

$$\begin{aligned} u^{-1}\left(E\left[u\left(\widetilde{v}(w)\right)\right]\right) &\geq u^{-1}\left(E\left[u\left(\widetilde{v}(w') + g(w - w')\right)\right]\right) \\ &\geq u^{-1}\left(E\left[u\left(\widetilde{v}(w')\right)\right]\right) + g(w - w'). \end{aligned}$$

The second inequality stems from the fact that the function

$$y \rightarrow u^{-1}\left(E\left[u\left(\widetilde{v}(w') + y\right)\right]\right) - y$$

is increasing.

Second, this mechanism is feasible because

$$\int_0^{+\infty} u^{-1}\left(Eu\left(\widetilde{v}(w)\right)\right) dF(w) \leq \int_0^{+\infty} E\left[\widetilde{v}(w)\right] dF(w)$$

from Jensen's inequality.

Third, this mechanism obviously delivers the same utility as the random one $\widetilde{v}(w)$ for each income level. Thus restricting the analysis to deterministic mechanisms is without loss of generality in this case.

**Ex ante randomization.** It consists in initially splitting the population in subgroups indexed by $i \in I$ that differ only with respect to their initial masses, where $I$ may be a finite, infinite, or even uncountable set. The subgroups share the same conditional income distribution since agents are split *ex ante*. It is easy to see that such randomization cannot help. For each group $i$, the optimal tax scheme must be of the form

$$\begin{aligned} r_i(w) &= w, \\ v_i(w) &= v_i(0) + g(w). \end{aligned}$$

The only degree of freedom induced by *ex ante* randomization is that the constant terms $v_i(0)$ may differ across groups as groups may cross-subisdize each other. Jensen's inequality implies, however, that it is optimal to have

$$\forall i, i' \in I,\ v_i(0) = v_i'(0),$$

so that ruling out *ex ante* randomization is without loss of generality.

## 6.7 Proof of Lemma 9

For $w_0 > 0$, define

$$\begin{aligned} W^D(w_0) &= \min_{(z_1,z_2)\in\mathbb{R}^2} z_1 + w_0 z_2, \\ s.t.\ \forall w &\geq 0,\ z_1 + w z_2 \geq u(g(w)). \end{aligned} \tag{32}$$

The program defining $W^D(w_0)$ is the dual of that defining $W(w_0)$. It has a simple graphical interpretation. It consists in finding, among all the straight lines above the graph of $u \circ g$, the one that takes the smallest value in $w_0$. Makarov and Plantin (2013) show that

$$W^D(w_0) = W(w_0).$$

It is graphically intuitive that $W^D$ is the concavification of $u \circ g$. We now prove it formally.

Fix $w_0 > 0$. The function $(z_1, z_2) \to z_1 + w_0 z_2$ is continuous. Thus, there exists at least one $(z_1(w_0),\ z_2(w_0))$ satisfying (32) such that $W^D(w_0) = z_1(w_0) + z_2(w_0)w_0$. Clearly, $z_2(w_0) \geq 0$ since $u \circ g$ is strictly increasing. For such a pair $(z_1(w_0),\ z_2(w_0))$, let

$$S(w_0) = \{w \geq 0 : z_1(w_0) + z_2(w_0)w = u \circ g(w)\}.$$

Continuity of $u \circ g$ implies that $S(w_0)$ is nonempty and closed. It is clearly bounded and therefore compact. Let

$$\underline{\sigma}\,(w_0) = \min S(w_0), \quad \overline{\sigma}\,(w_0) = \max S(w_0).$$

We have:

$$\underline{\sigma}\,(w_0) \leq w_0 \leq \overline{\sigma}\,(w_0)\,. \tag{33}$$

*Proof.* We prove that $w_0 \leq \overline{\sigma}\,(w_0)$. The proof that $\underline{\sigma}\,(w_0) \leq w_0$ is symmetric. Suppose the opposite that $w_0 > \overline{\sigma}\,(w_0)$ then for some $\varepsilon \in (0, w_0 - \overline{\sigma}\,(w_0))\,,$ let

$$\eta\,(\varepsilon) = \min_{y \geq \overline{\sigma}(w_0)+\varepsilon} \left\{ \frac{z_1(w_0) - u \circ g(y)}{y} + z_2(w_0) \right\}.$$

Clearly, $\eta\,(\varepsilon) > 0$.

Define $(z_1', z_2')$ as $z_1' = z_1(w_0) + (\overline{\sigma}\,(w_0) + \varepsilon)\,\eta\,(\varepsilon)\,,\ z_2' = z_2(w_0) - \eta\,(\varepsilon)\,.$ The pair $(z_1', z_2')$ satisfies (32). To see this, notice that $z_1' + yz_2' = z_1(w_0) + yz_2(w_0) + \eta\,(\varepsilon)\,(\overline{\sigma}\,(w_0) + \varepsilon - y)\,.$ Thus $z_1' + yz_2' > z_1(w_0) + yz_2(w_0) \geq u \circ g(y)$ for $y < \overline{\sigma}\,(w_0) + \varepsilon$. Further, $z_1' + yz_2' \geq z_1\,(w_0) + yz_2\,(w_0) - \eta\,(\varepsilon)\,y \geq u \circ g(y)$ for $y \geq \overline{\sigma}\,(w_0) + \varepsilon$ by definition of $\eta\,(\varepsilon)\,.$ At the same time,

$$z_1' + w_0 z_2' = z_1\,(w_0) + w_0 z_2\,(w_0) + (\overline{\sigma}\,(w_0) + \varepsilon - w_0)\,\eta\,(\varepsilon) < z_1\,(w_0) + w_0 z_2\,(w_0)\,,$$

which contradicts the definition of $(z_1\,(w_0)\,, z_2\,(w_0))$.■

Inequalities (33) imply that for each $w_0$, we can define

$$\begin{cases} \underline{w}\,(w_0) = \sup\left\{w \leq w_0 \ s.t. W^D\,(w) = u\,(g\,(w))\ \right\} \\ \overline{w}\,(w_0) = \inf\left\{w \geq w_0 \ s.t. W^D\,(w) = u\,(g\,(w))\ \right\} \end{cases}$$

because these sets are not empty: They respectively contain $\underline{\sigma}\,(w_0)$ and $\overline{\sigma}\,(w_0)\,.$ It must be indeed that

$$\begin{cases} z_1(w_0) + z_2(w_0)\underline{\sigma}\,(w_0) = u \circ g(\underline{\sigma}\,(w_0)) = W^D(\underline{\sigma}\,(w_0)) \\ z_1(w_0) + z_2(w_0)\overline{\sigma}\,(w_0) = u \circ g(\overline{\sigma}\,(w_0)) = W^D(\overline{\sigma}\,(w_0)) \end{cases}.$$

Thus, for any $w_0 > 0$, if $W^D(w_0) \neq u(g(w_0))$, then $[\underline{w}(w_0), \overline{w}(w_0)]$ is not a singleton, and $W_D$ is linear over it. We are now able to prove:

$$W^D(w) = \overline{u \circ g}(w). \tag{34}$$

*Proof.* Notice first that by construction, $W^D \geq u \circ g$. Second, suppose that there exists a concave function $\theta$ such that

$$\begin{aligned} \theta &\geq u \circ g, \\ \exists w_0 \ s.t.\ \theta(w_0) &< W^D(w_0). \end{aligned}$$

In this case, it must be that $u \circ g(w_0) < W^D(w_0)$. But then, this means that $\theta$ is above the line $y = z_1(w_0) + x z_2(w_0)$ in $\underline{w}(w_0)$ and $\overline{w}(w_0)$, and strictly below it in $w_0$: it cannot be concave. Third, $W^D$ is concave. Suppose otherwise that there exists $w_1 < w_2 < w_3$ such that the chord between $\left(w_1, W^D(w_1)\right)$ and $\left(w_3, W^D(w_3)\right)$ is strictly above $\left(w_2, W^D(w_2)\right)$ in $w_2$. This contradicts that there exists a straight line that meets the graph of $W^D$ in $w_2$ and that is above the graph of $W^D$, since such a straight line cannot be above both $\left(w_1, W^D(w_1)\right)$ and $\left(w_3, W^D(w_3)\right)$. ■

Equality (34) also defines the risk-taking choices of an individual with initial income $w_0$. If

$$u(g(w_0) = W^D(w_0) = W(w_0),$$

then the agent reaches $W(w_0)$ by investing at the risk free rate. If

$$u(g(w_0) < W^D(w_0),$$

then we have

$$\begin{aligned} z_1(w_0) + \underline{w}(w_0) z_2(w_0) &= u(g(\underline{w}(w_0))), \\ z_1(w_0) + \overline{w}(w_0) z_2(w_0) &= u(g(\overline{w}(w_0))). \end{aligned}$$

so that

$$\begin{aligned} W^D(w_0) &= z_1(w_0) + z_2(w_0) w_0 = \frac{w_0 - \underline{w}(w_0)}{\overline{w}(w_0) - \underline{w}(w_0)} u(g(\overline{w}(w_0))) \\ &\quad + \frac{\overline{w}(w_0) - w_0}{\overline{w}(w_0) - \underline{w}(w_0)} u(g(\underline{w}(w_0))). \end{aligned}$$

Thus the lottery that pays off $\underline{w}(w_0)$ with probability $\frac{\overline{w}(w_0) - w_0}{\overline{w}(w_0) - \underline{w}(w_0)}$ and $\overline{w}(w_0)$ with probability $\frac{w_0 - \underline{w}(w_0)}{\overline{w}(w_0) - \underline{w}(w_0)}$ attains $W^D(w_0) = W(w_0)$. Other lotteries with support in $S(w_0)$ can also attain it. But in this case their support is on the left of $\underline{w}(w_0)$ and on the right of $\overline{w}(w_0)$. Thus they are dominated by this minimal one in the sense of second-order stochastic dominance because their c.d.f. must single cross that of this minimal lottery. ■

## 6.8 Proof of Corollary 11

*Proof of i).* If $u \circ g = \overline{u \circ g}$, there is no risk shifting in equilibrium and utilitarian welfare is the same as absent risk shifting. If $u \circ g < \overline{u \circ g}$, then a non negligible set of agents shift risk. It is actually as if agents could not shift risk, but the planner offered the random mechanism

$$g(\widetilde{\rho}(w)) + v^{**}(0),$$

which is incentive compatible and feasible absent risk shifting since it is when agents can shift risk. This mechanism is different from (6) for a non negligible set of agents and thus yields a strictly lower utilitarian welfare.

*Proof of ii).* Straightforward computations show that

$$E\left[u\left(v^{**}(0)\right) + g(\widetilde{\rho}(w))\right] > u\left(v^{*}(0) + g(w)\right)$$

if and only if

$$E\left[u\left(g\left(\widetilde{\rho}(w)\right) - g(w) - \int_0^{+\infty}\left(E\left[g\left(\widetilde{\rho}(t)\right)\right] - g(t)\right)dF_0(t)\right)\right] > u(0).$$

Notice that $\widetilde{\rho}(.)$ does not depend on $F_0$.

If an arbitrarily large mass of $F_0$ is concentrated on agents that do not shift risk, then the inequality is satisfied for an agent with income $w$ who shifts risk, because $\int_0^{+\infty}\left(E\left[g\left(\widetilde{\rho}(t)\right)\right] - g(t)\right)dF_0(t)$ becomes arbitrarily small.

Let $t_0 \in \arg\max_{t\geq 0}\left\{\left(E\left[g\left(\widetilde{\rho}(t)\right)\right] - g(t)\right)\right\}$. If the mass of $F_0$ is arbitrarily concentrated on $t_0$, then all agents are worse off because

$$\max_{w\geq 0}\left\{E\left[g\left(\widetilde{\rho}(w)\right)\right] - g(w) - \int_0^{+\infty}\left(E\left[g\left(\widetilde{\rho}(t)\right)\right] - g(t)\right)dF_0(t)\right\},$$

which is positive, is arbitrarily close to 0. Thus lotteries $g\left(\widetilde{\rho}(w)\right) - g(w) - \int_0^{+\infty}\left(E\left[g\left(\widetilde{\rho}(t)\right)\right] - g(t)\right)dF_0(t)$ are at best arbitrarily close to fair in the limit and a risk-averse agent is unwilling to take any of them in the limit.■

## 6.9 Proof of Proposition 12

We skip the proof that due to the superadditivity of $g$, one can restrict the analysis to avoidance-free mechanisms where agents report their entire income at date 1: it is identical to that in the baseline model. A direct mechanism is now a pair $(v(w), H^w)$ for each income level $w$, where $v$ is the net transfer following a report $w$ and $H^w \in \Gamma$, the set of c.d.f. over $[0,+\infty)$. The planner now faces the program $(\widetilde{\wp})$ :

$$\max_{(v(w),H^w)} \int_0^{+\infty} \left( \int_0^{+\infty} u(v(t))dH^w(t) \right) dF_0(w)$$

$$s.t. \begin{cases} \int_0^{+\infty} \left( \int_0^{+\infty} v(t)dH^w(t) \right) dF_0(w) \leq \int_0^{+\infty} w dF_0(w), \\ \forall 0 \leq w' \leq w, v(w) \geq v(w') + g(w-w'), \\ \forall w \geq 0, H^w = \arg\max_{G\in\Gamma} \left\{ \begin{array}{c} \int_0^{+\infty} u(v(t))dG(t) \\ s.t. \int_0^{+\infty} t dG(t) = w \end{array} \right\}. \end{cases}$$

Compared with the problem absent risk shifting $(\wp)$, the program $(\widetilde{\wp})$ has an additional incentive-compatibility constraint - the last one - stating that agents pick an optimal lottery given the net transfer $v$.

**Step 1.** We first show that the scheme $v^{***}(w)$ satisfies the constraints of $(\widetilde{\wp})$. By construction, it satisfies the resource constraint, and does not induce risk taking. Also, for every $0 \leq w' \leq w$,

$$\begin{aligned} u^{-1} \circ \overline{u \circ g}(w') + g(w-w') &= u^{-1}\left(E\left[u\left(g\left(\widetilde{\rho}(w')\right) + g(w-w')\right)\right]\right) \\ &\leq u^{-1}\left(E\left[u\left(g\left(\widetilde{\rho}(w') + w - w'\right)\right)\right]\right) \\ &\leq u^{-1} \circ \overline{u \circ g}(w). \end{aligned}$$

The equality stems from CARA, the first inequality stems from the superadditivity of $g$, and the second inequality stems from the fact that the lottery $\widetilde{\rho}(w') + w - w'$ is not strictly preferable to the optimal lottery $\widetilde{\rho}(w)$ .

**Step 2.** Suppose that $(v(w), H^w)$ is a mechanism that satisfies the constraints of $(\widetilde{\wp})$. Then for all $w \geq 0$,

$$u^{-1}\left(\int_0^{+\infty} u(v(t))dH^w(t)\right) - v(0) \geq u^{-1} \circ \overline{u \circ g}(w).$$

*Proof.* Incentive-compatibility of $v$ requires in particular that

$$\forall w \geq 0, v(w) \geq v(0) + g(w),$$

and therefore

$$\int_0^{+\infty} u(v(t))dH^w(t) \geq u \circ v(w) \geq u(v(0) + g(w)).$$

Since $w \to \int_0^{+\infty} u(v(t))dH^w(t)$ is concave by construction, it is also larger than the concavification of $w \longrightarrow u(v(0) + g(w))$. Since $u$ is CARA, this implies

$$u^{-1}\left(\int_0^{+\infty} u(v(t))dH^w(t)\right) - v(0) \geq u^{-1} \circ \overline{u \circ g}(w).$$

■

**Step 3.** Let $h$ a continuously increasing function such that $h(0) = 0$, and

$$J(h) = \max_{v} \int_0^{+\infty} u\left(v(w)\right) dF_0(w) \tag{35}$$

$$s.t. \left\{ \begin{array}{c} \int_0^{+\infty} v(w) dF_0(w) \leq \int_0^{+\infty} w dF_0(w), \\ \forall w \geq 0, \; v(w) \geq h\left(w\right) + v(0). \end{array} \right.$$

We have

$$J(h) = \int_0^{+\infty} u\left(h(w) + \int_0^{+\infty} \left(t - h(t)\right) dF_0(t)\right) dF_0(w), \tag{36}$$

*Proof***.** Step 2 in the proof of Proposition 3.■

**Proof that $v^{***}(w)$ solves $(\widetilde{\wp})$.** For any mechanism $(v(w), H^w)$ that satisfies the constraints of $(\widetilde{\wp})$, define the mechanism $(\psi\left(w\right), \Phi^w)$ as

$$\left\{ \begin{array}{c} \psi\left(w\right) = u^{-1}\left(\int_0^{+\infty} u(v(t)) dH^w\left(t\right)\right), \\ \Phi^w(t) = 1_{\{t \geq w\}}. \end{array} \right.$$

This mechanism satisfies the resource constraint in (35) from Jensen's inequality. From Step 2, it satisfies the other constraint of (35) for $h = u^{-1} \circ \overline{u \circ g}$. Step 3 shows that it yields a lower utilitarian welfare than $v^{***}$, and thus so does $(v(w), H^w)$.■

Finally, if $\overline{u \circ g} \neq u \circ g$, utilitarian welfare is strictly smaller than absent risk shifting because $v^{***}$ is different from $v^*$ on a non negligible set, and strictly larger than absent commitment because $v^{***}$ is different from $v^{**}$ on a non negligible set.■

## 6.10 Proof of Lemma 13

For brevity, we skip the proof that restricting the analysis to schemes such that the date-0 report is $r(w) = w$ is again without loss of generality. Such schemes are incentive-compatible if the date-1 payment $v(w)$ satisfies

$$u(v(w)) \geq \max_{\substack{0 \leq w' \leq w \\ E\widetilde{s} = w - w'}} E\left[u\left(v\left(w'\right) + g(\widetilde{s})\right)\right].$$

In words, the right-hand side considers all possible reports $w' \in [0, w]$ and all possible gambles on concealed income $w - w'$. It is easy to see that this implies in particular that $v$ satisfies

$$v(w) - v(0) \geq u^{-1} \circ \overline{u \circ g}\left(w\right).$$

That this implies that $v = v^{***}$ follows then from Step 3 in the proof of Proposition 12.

## 6.11 Proof of Lemma 14

The program

$$\max_{w_1} \overline{u \circ g}\,(w - w_1) + \overline{u \circ g}\,(w_1) \tag{37}$$

is concave and is trivially solved by $w_1 = \frac{w}{2}$, that generates the value $2\overline{u \circ g}\left(\frac{w}{2}\right)$. Its objective function is larger than that of (21) for all $w_1$. Thus if there exists a $w_1$ such that the value function of (21) reaches $2\overline{u \circ g}\left(\frac{w}{2}\right)$, then it must be a solution to (21). The $w_1$ described in the Lemma clearly achieves this by linearity of $\overline{u \circ g}$ over $\left(\underline{w}\left(\frac{w}{2}\right), \overline{w}\left(\frac{w}{2}\right)\right)$.■

## 6.12 Proof of Proposition 15

The proof follows the same broad steps as that of the proof of Proposition 3.

**Step 1.** Restricting the analysis to mechanisms such that $r(w) = w$ is without loss of generality.

*Proof.* For any mechanism $(r, v_0, v_1, \widetilde{\sigma}, \sigma)$ that satisfies the constraints (23) of $\left(\wp^D\right)$, define the mechanism

$$\begin{aligned}
\rho\,(w) &= w, \\
\nu_0(w) &= v_0(r(w)) + g\,(w - r(w) - \sigma(w))\,, \\
\nu_1(w) &= u^{-1}\,(E\,[u\,(v_1(r(w)) + g\,(\widetilde{\sigma}(w)))])\,.
\end{aligned}$$

This mechanism is clearly feasible, and delivers the same utility to each agent as that delivered by $(r, v_0, v_1)$. Let us show that it is incentive-compatible. We have

$$\begin{aligned}
& u\left(\nu_0(w') + g\left(w - w' - s\right)\right) + E\left[u\left(\nu_1(w') + g\,(\widetilde{s})\right)\right] \\
= \;& u\begin{pmatrix} v_0(r\,(w')) + g\,(w' - r(w') - \sigma(w')) \\ +g\,(w - w' - s) \end{pmatrix} \\
& +E\left[u\left(u^{-1}\left(E\left[u\left(v_1(r(w')) + g\left(\widetilde{\sigma}(w')\right)\right)\right]\right) + g\,(\widetilde{s})\right)\right].
\end{aligned}$$

Superadditivity and CARA imply

$$\begin{aligned}
u\begin{pmatrix} v_0(r\,(w')) + g\,(w' - r(w') - \sigma(w')) \\ +g\,(w - w' - s) \end{pmatrix} &\leq u\left(v_0(r\left(w'\right)) + g\left(w - r(w') - \sigma(w') - s\right)\right), \\
E\left[u\left(u^{-1}\left(E\left[u\left(v_1(r(w')) + g\left(\widetilde{\sigma}(w')\right)\right)\right]\right) + g\,(\widetilde{s})\right)\right] &\leq E\left[u\left(v_1(r(w')) + g\left(\widetilde{\sigma}(w') + \widetilde{s}\right)\right)\right].
\end{aligned}$$

Incentive-compatibility of $(r, v_0, v_1, \widetilde{\sigma}, \sigma)$ implies in turn

$$\begin{aligned}\left[\begin{array}{c} u\left(v_0(r\left(w'\right)) + g\left(w - r(w') - \sigma(w') - s\right)\right) \\ +E\left[u\left(v_1(r(w')) + g\left(\widetilde{\sigma}(w') + \widetilde{s}\right)\right)\right] \end{array}\right] &\leq \begin{array}{c} u\left(v_0(r(w)) + g\left(w - r(w) - \sigma(w)\right)\right) \\ +E\left[u\left(v_1(r(w)) + g\left(\widetilde{\sigma}(w)\right)\right)\right] \end{array} \\ &= u\left(\nu_0(w)\right) + u\left(\nu_1(w)\right).\end{aligned}$$

■

**Step 2.** Consider the less constrained program $\left(\wp_0^D\right)$ solving for an optimal mechanism $(v_0, v_1)$ that is feasible, and such that each agent prefers to report his entire income rather than none of it:

$$\max_{(v_0,v_1)} \int_0^{+\infty} \left(u\left(v_0(w)\right) + u\left(v_1(w)\right)\right) dF(w) \tag{38}$$

$$s.t. \left\{\begin{array}{c} \int_0^{+\infty} \left(v_0(w) + v_1(w)\right) dF(w) \leq \int_0^{+\infty} w dF(w), \\ \forall w \geq 0, u\left(v_0(w)\right) + u\left(v_1(w)\right) \geq \max_{s,\widetilde{s}} \left\{\begin{array}{c} u(v_0(0) + g(w-s)) \\ +E\left[u\left(v_1(0) + g(\widetilde{s})\right)\right] \end{array}\right\}. \\ s.t.\ 0 \leq s \leq w, \\ E\left[\widetilde{s}\right] = s. \end{array}\right. \tag{39}$$

We show in two steps that the solution to this program is $(v_D, v_D)$.

First, applying the same variational argument as in step 2 in the proof of Proposition 3, one gets that an optimal $(v_0, v_1)$ must satisfy

$$\forall w \geq 0, u\left(v_0(w)\right) + u\left(v_1(w)\right) = \max_{s,\widetilde{s}} \left\{\begin{array}{c} u(v_0(0) + g(w-s)) \\ +E\left[u\left(v_1(0) + g(\widetilde{s})\right)\right] \end{array}\right\}. \quad \begin{array}{c} s.t.\ 0 \leq s \leq w, \\ E\left[\widetilde{s}\right] = s. \end{array} \tag{40}$$

This means that a solution to $\left(\wp_0^D\right)$ is fully characterized by its constant terms $(v_0(0), v_1(0))$.

Second, for any pair of positive numbers $(v_0(0), v_1(0))$, we have

$$\max_{s,\widetilde{s}} \left\{\begin{array}{c} u(v_0(0) + g(w-s)) \\ +E\left[u\left(v_1(0) + g(\widetilde{s})\right)\right] \end{array}\right\} \leq 2e^{-\alpha\left(\frac{v_0(0)+v_1(0)}{2}\right)} \overline{u \circ g}\left(\frac{w}{2}\right). \quad \begin{array}{c} s.t.\ 0 \leq s \leq w, \\ E\left[\widetilde{s}\right] = s. \end{array} \tag{41}$$

*Proof of (41).* We have

$$\max_{s,\widetilde{s}} \left\{\begin{array}{c} u(v_0(0) + g(w-s)) \\ +E\left[u\left(v_1(0) + g(\widetilde{s})\right)\right] \end{array}\right\} \leq \max_{s\in\mathbb{R},\ x\in[0,1]} \Lambda\left(x, s\right), \quad \begin{array}{c} s.t.\ 0 \leq s \leq w, \\ E\left[\widetilde{s}\right] = s. \end{array} \tag{42}$$

where

$$\Lambda(x,s)= \begin{array}{l} e^{-\alpha(xv_0(0)+(1-x)v_1(0))}\overline{u\circ g}(w-s) \\ +e^{-\alpha(xv_1(0)+(1-x)v_0(0))}\overline{u\circ g}(s) \end{array}.$$

(42) holds because the program on the right-hand side is less constrained and has a larger objective than that on the left-hand side. The envelope theorem implies

$$\arg\max_x\left\{\max_s\Lambda(x,s)\right\}=\frac{1}{2}.$$

■

This implies that for any feasible scheme defined by the pair $(v_0(0), v_1(0))$, the scheme $v^D(w)=\frac{v_0(0)+v_1(0)}{2}+u^{-1}\circ\overline{u\circ g}(\frac{w}{2})$ is preferable. It is also feasible because it implies less savings distortion and risk shifting than the one defined by the pair $(v_0(0), v_1(0))$, and thus a higher tax capacity.

**Step 3.** It only remains to show that $(r_D, v_D, v_D)$ satisfies the general incentive-compatibility constraints of $(\wp^D)$. We leave it to the reader to check that it stems, again, from $u$ being CARA and $g$ superadditive.■

## 6.13 Proof of Proposition 16

In this proof we consider a given mechanism that is feasible and incentive-compatible, and show by contradiction that it cannot comprise unbounded date-0 income reports.

**Step 1.** Notice first that if the support of the planner's beliefs about an agent's date-1 pre-tax income has a lower bound $\underline{w}_1$, then *ex post* optimality implies that he promises the agent a utility of $-\infty$ if his date-1 report $r_1(w_1)$ is less than $\underline{w}_1$, and a date-1 variable payment $g(r_1(w_1)-\underline{w}_1)$ otherwise. The date-1 scheme is thus such that $r_1(w_1)=w_1$.

**Step 2.** Denote $k_0$ and $k_1$ the respective date-0 and date-1 transfers to an agent who does not report any income at any date. An agent with income $w$ who reports no income at date 0 can reach at least utility:

$$u\left(k_0+g\left(\frac{w}{2}\right)\right)+u\left(k_1+g\left(\frac{w}{2}\right)\right).$$

Thus as $w\to+\infty$, the utility of an agent with income $w$ tends to $2\lim_{+\infty}u$.

**Step 3.** $\lim\limits_{w\to+\infty}\overline{w}(w)-\underline{w}(w)=0$.

*Proof.* If $\overline{w}(w)>\underline{w}(w)$, then the function $u\circ g$ has a lower right derivative in $\underline{w}(w)$ than left derivative in $\overline{w}(w)$ by construction of its concavification. But marginal utility decreases exponentially, while the left and right derivatives of $g$ are bounded by assumption. Thus it can be that the function $u \circ g$ has a lower right derivative in $\underline{w}(w)$ than left derivative in $\overline{w}(w)$ only when $\overline{w}(w)$ and $\underline{w}(w)$ are arbitrarily close as $w$ becomes large.■

**Proof of the Proposition**. Suppose now that the mechanism features a date-0 report function $r_0(w)$ such that the set $\{r_0(w)\}_{w\geq 0}$ is not bounded. There exists a sequence $(\rho_n)_{n\geq 0}$ within this set such that

$$\rho_n \underset{n\to+\infty}{\to} +\infty.$$

For $n \geq 0$, define

$$\begin{aligned} \Omega_n &= \{w \text{ } s.t. \text{ } r_0(w_n) = \rho_n\}, \\ \omega_n &= \inf \ \Omega_n. \end{aligned}$$

For every $n \geq 1$, there exists a $w_n \in \Omega_n$ s.t.

$$w_n \leq \omega_n + \frac{1}{n}.$$

We now show that the date-1 consumption of an agent with such an income $w_n$ cannot grow without bound as $n \to +\infty$. For this to be the case, it would also have to be the case that the equilibrium savings of this agent $\sigma(w_n)$ grow without bound with $n$. From step 3, $\underline{w}(\sigma(w_n)) \underset{n\to+\infty}{\longrightarrow} \sigma(w_n)$. But then, it must be that the lower bound of the support of the planner's beliefs about this agent's date-1 pre-tax income realization converges to $\underline{w}(\sigma(w_n))$ and thus $\sigma(w_n)$ as $n \to +\infty$. This is because all the agents who make the same report as him have an income larger than $w_n - \frac{1}{n}$. Thus it must be that their savings are at least $\sigma(w_n) - \frac{1}{n}$, invested in arbitrarily small lotteries from step 3.

From step 1, this implies that the date-1 consumption of the agent with income $w_n$ is bounded as $n \to +\infty$. From step 2, the agent with income $w_n$ could reach a strictly higher utility by not reporting any income at date 0 for $n$ sufficiently large, a contradiction.■

## 6.14 Proof of Lemma 17

**Remark.** Infinitely risk-averse agents never find it optimal to gamble, and we can thus ignore the risk-shifting friction.

**Step 1.** If the optimal scheme is such that

$$\forall w \geq 0, \ r_0(w) = 0,$$

then it must be such that the net transfers at dates 0 and 1, $v_0$ and $v_1$, are

$$\begin{aligned} v_0(w) &= K, \\ v_1(w_1) &= K + g(w_1), \end{aligned}$$

where $w_1$ is date-1 income and $K = \frac{1}{2} \int \left( \frac{w}{2} - g(\frac{w}{2}) \right) dF(w)$.

*Proof.* *Ex post* optimality implies that the variable part of the scheme at date 1 is $g(w_1)$. We only have to prove that it is optimal to give equal minimum consumption level across dates. It could be preferable to grant a larger minimum consumption level at date 0, so that savings are higher and thus more taxes are collected at date 1. We show that this is not the case. Denoting $k_0$ and $k_1$ the respective date-0 and date-1 minimum consumption levels, suppose that $k_0 \geq k_1$. An increase $dk$ in $k_0$ financed by a decrease $dk$ in $k_1$ raises date-1 savings $s$ by $ds$ for all agents with income $w$ such that $k_0 \leq k_1 + g(w)$, with

$$dk - g'_L(w-s)\, ds = -dk + g'_R(s)\, ds,$$

or

$$ds = \frac{2dk}{g'_L(w-s) + g'_R(s)}.$$

Thus at each income level the additional collected taxes are

$$d[s - g(s)] = \frac{2dk}{g'_L(w-s) + g'_R(s)} \left( 1 - g'_R(s) \right),$$

so that the net change in date-1 minimum consumption - the one that matters to the social planner - is larger than

$$dk \int \left( \frac{2(1 - g'_R(s))}{g'_L(w-s) + g'_R(s)} \right) dF(w) - dk = dk \int \left( \frac{2 - 3g'_R(s) - g'_L(w-s)}{g'_L(w-s) + g'_R(s)} \right) dF(w).$$

The integrand is negative for all $w$ because

$$\frac{1}{2} \leq g'_L,\ g'_R.$$

■

**Step 2.** Consider a mechanism that features date-0 reports $r_0(w)$ s.t.

$$\sup \{r_0(w)\}_{w \geq 0} = r > 0,$$

There exists a sequence $(\rho_n)_{n \geq 1}$ within the set $\{r_0(w)\}_{w \geq 0}$ such that

$$\rho_n \underset{n \to +\infty}{\to} r.$$

For $n \geq 1$, define

$$\begin{aligned} \Omega_n &= \{w \ s.t. \ r_0(w) \geq \rho_n\}, \\ \omega_n &= \inf \ \Omega_n. \end{aligned}$$

For every $n \geq 1$, there exists a $w_n \in \Omega_n$ s.t.

$$w_n \leq \omega_n + \frac{1}{n}.$$

It must be that the lower bound of the support of the planner's beliefs about the date-1 pre-tax income of agent with income $w_n$ converges to its actual value as $n \to +\infty$. The planner knows that given a report $r_0(w_n)$, an agent's income is at least $\omega_n$, from which he infers a lower bound on the savings of the agent that is arbitrarily close to the actual savings of agent with actual income $w_n$. Thus as $n \to +\infty$, the date-1 consumption of an agent with income $w_n$ is bounded above as $n \to +\infty$. Thus it must be that $(w_n)_{n\geq 1}$ is bounded above otherwise an agent with an income $w_n$ would prefer to report zero income so as to be able to smooth consumption across dates. We can therefore extract a subsequence $(W_n)$ of $(w_n)$ that has a finite limit $W$. Notice that it must be that $W \geq r$ by construction of the sequence $w_n$.

Denote $k_0$ and $k_1$ the fixed payments of the scheme. If the agent with income $W_n$ deviates by reporting no income at date 0 and optimally saving $\sigma_n$, then $(\sigma_n)_{n\geq 1}$ has a limit $\sigma$ because $(W_n)_{n\geq 1}$ converges. If $\sigma > 0$, it means that the date-0 consumption of such a deviating agent is weakly larger than that at date 1 for $n$ sufficiently large, otherwise positive savings would not be optimal. With infinite risk-aversion, the agent has minimum consumption $k_1 + g(\sigma)$ if he deviates, while it is at most $k_1$ from above if he reports $r(W_n)$. Thus such positive reports require that agent with income $W_n$ does not save in equilibrium for $n$ sufficiently large, or that

$$k_1 \geq k_0 + g(W) \geq k_0 + g(r). \tag{43}$$

**Proof of the Proposition.** Compare now a scheme with no initial reports as described in step 1 with a scheme in which agents report up to $r > 0$ at date 0 as in step 2. In the scheme without reports, the minimum consumption level $K$ is the same across dates from Step 1. In the scheme with initial reports, the total constant payments $k_1 + k_0$ are larger than $2K$ by at most $r - g(r)$. This is because agents report at most $r$ , and do so only if they consume at least $k_0 + g(r)$ at date 0. Thus, it must be that

$$k_1 + k_0 \leq 2K + r - g(r).$$

Combined with (43), this implies that

$$k_0 \leq K + \frac{r}{2} - g(r).$$

Since $\frac{r}{2} \leq g(r)$, this scheme cannot be preferable to that in which $r_0 = 0$ with infinitely risk-averse agents.■

## 6.15 Proof of Lemma 18

This is the textbook Black-Scholes model in the particular case in which the risk-free rate and drift are both equal to zero. It is well-known that a self-financed strategy can replicate the cash-or-nothing call. Its initial value is equal to the expected payoff under the risk-neutral measure, which is in this case identical to the subjective measure. It is therefore equal to

$$E_0\left[1_{\{P_1 \geq K\}}\right] = E_0\left[1_{\left\{e^{-\frac{\sigma^2}{2}+\sigma W_1} \geq K\right\}}\right] = \Phi\left(-\frac{\ln K}{\sigma} - \frac{\sigma}{2}\right).$$

From Ito's Lemma, the "delta" of the cash-or-nothing call - the number of units of the risky assets that one must hold at date $t$ in order to implement the replicating strategy - is equal to $\frac{\partial E_t\left[1_{\{P_1 \geq K\}}\right]}{\partial P_t}$, where

$$E_t\left[1_{\{P_1 \geq K\}}\right] = \Pr\left\{P_t e^{-\frac{\sigma^2(1-t)}{2}+\sigma(W_1 - W_t)} \geq K\right\} = \Phi\left(-\frac{\ln \frac{K}{P_t}}{\sigma\sqrt{1-t}} - \frac{\sigma}{2}\sqrt{1-t}\right).$$

Straightforward derivation of this expression w.r.t. $P_t$ yields $q(K, P_{t,} t)$■

## 6.16 Proof of Proposition 19

The agent invests in the risk-free rate and in a cash-or-nothing call that pays off with probability $\frac{w - \underline{w}}{\overline{w} - \underline{w}}$. From Lemma 18, the strike $K$ must solve

$$\Phi\left(-\frac{\ln K}{\sigma} - \frac{\sigma}{2}\right) = \frac{w - \underline{w}}{\overline{w} - \underline{w}},$$

which yields (25). The aggregate demand for the risky asset at date $t$ is

$$(\overline{w} - \underline{w}) \int_{\underline{w}}^{\overline{w}} q(K(w), P_{t,} t) dF_0(w).$$

Plugging in the expression (24) and simplifying yields (26).

Finally, for a given realization $W_1$, the fraction of successful gambles is by definition of the strikes $K(w)$ :

$$F_0\left(\overline{w}\right) - F_0\left(\overline{w} - \left(\overline{w} - \underline{w}\right)\Phi\left(W_1\right)\right),$$

which readily yields (27).■

## 6.17 Proof of Proposition 20

The tax scheme $\tau(w) = w - u^{-1} \circ \overline{u \circ g}(w)$ is clearly the only one that leaves each agent with his reservation utility $\overline{u \circ g}(w)$ without generating risk shifting by definition of $\overline{u \circ g}$ as the smallest concave function above $u \circ g$. It is also easy to see that any tax scheme that induces risk shifting cannot generate higher expected taxes than this one by Jensen's inequality since $u^{-1} \circ \overline{u \circ g}$ is convex.■

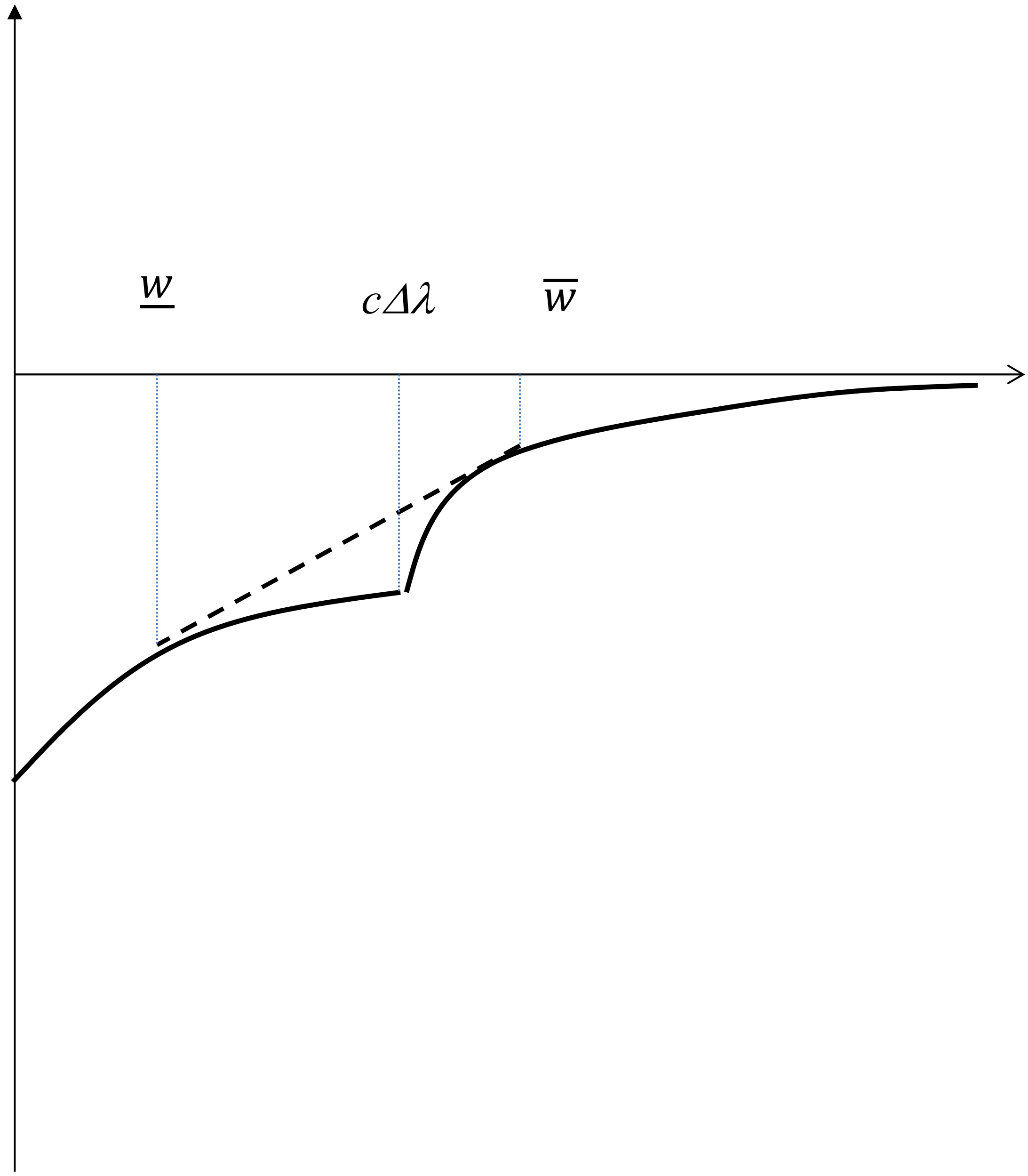


Figure 1. The solid curve represents the graph of $u \circ g$ and the dashed line its concavification over $[\underline{w}, \overline{w}]$. The two functions coincide outside this segment.

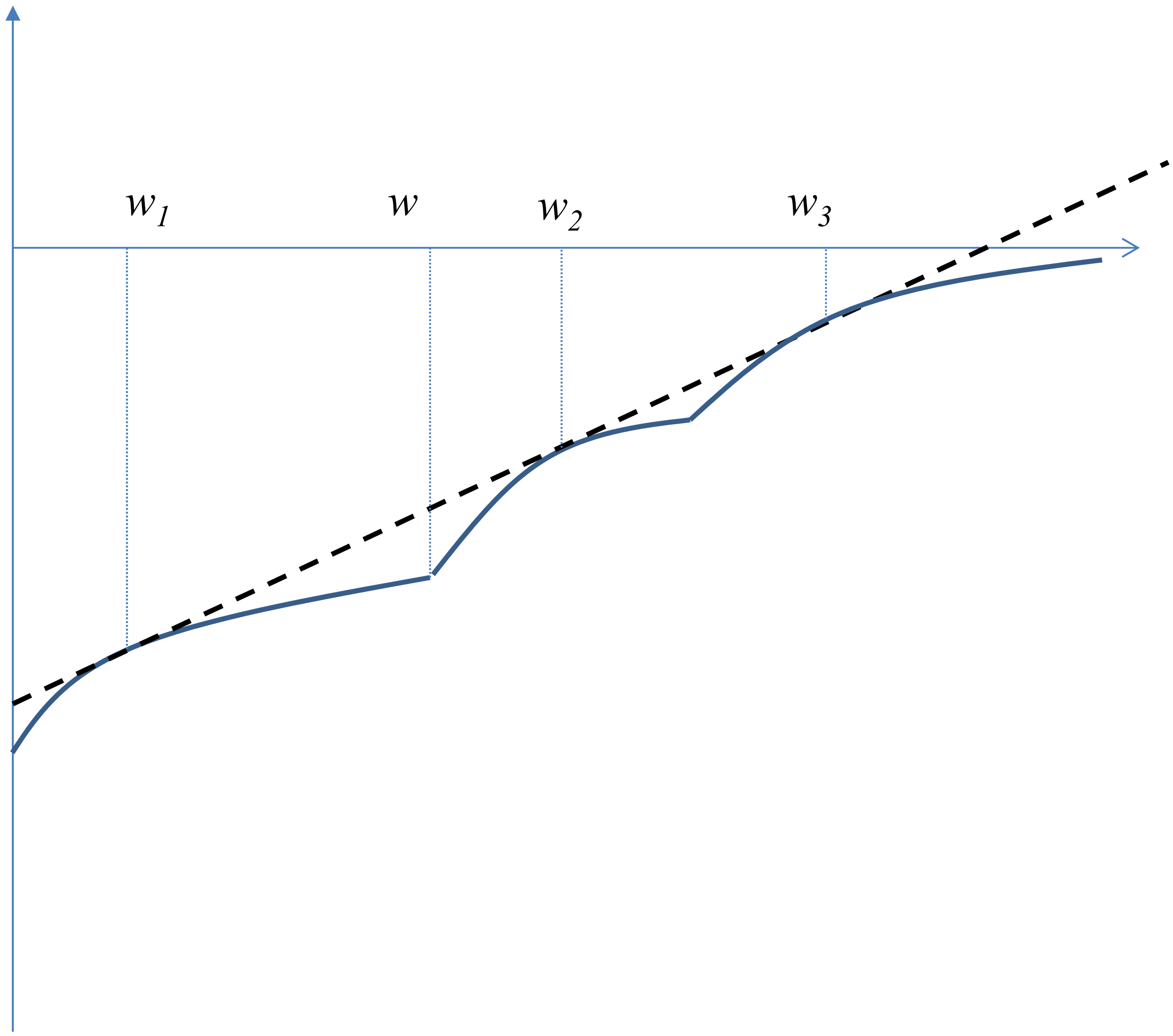


Figure 2. Here the straight line that concavifies $u \circ g$ in $w$ has three points of contact with $u \circ g$, $\{w_1; w_2; w_3\}$. An agent with income $w$ can concavify with a lottery that has support $\{w_1; w_2; w_3\}$, $\{w_1; w_3\}$, or $\{w_1; w_2\}$. The latter is less risky in the sense of second-order stochastic dominance.