Optimal Corporate Taxation Under Financial Frictions*

Eduardo Dávila† Benjamin Hébert‡

September 2019

Abstract

We study optimal corporate taxation when firms are financially constrained. We describe a corporate taxation principle: taxes should be levied on unconstrained firms, which value resources inside the firm less than constrained firms. Under complete information, this principle completely characterizes optimal corporate tax policy. With incomplete information, the government can use payout policy to elicit whether a firm is constrained, and tax accordingly. In our static model, optimal corporate taxation can be implemented by a corporate payout tax, and in our dynamic model, the optimal sequence of mechanisms can also be implemented by a corporate payout tax.

JEL numbers: G38, H21, H25

Keywords: corporate taxation, dividend taxation, financial frictions

*We would like to thank our discussant Vish Viswanathan, Anat Admati, Vladimir Asriyan, Andy Atkeson, V. V. Chari, John Cochrane, Peter DeMarzo, Sebastian Di Tella, Peter Diamond, Darrell Duffie, Emmanuel Farhi, Michael Faulkender, Piero Gottardi, Oleg Itskhoki, Anton Korinek, Hanno Lustig, Tim McQuade, Holger Mueller, Adriano Rampini, Martin Schneider, Andy Skrzypacz, Jeremy Stein, Felipe Varas, as well as seminar and conference participants for helpful comments. All remaining errors are our own.

†Yale University and NBER. Email: eduardo.davila@yale.edu

‡Stanford University and NBER. Email: bhebert@stanford.edu

1 Introduction

Virtually every developed country collects taxes from corporations. In this paper, we take as given the desire to tax firms, and ask how firms should be taxed or, equivalently, which firms should be taxed. We consider economies with financial frictions, in which, for some firms, the marginal value of funds inside the firm is higher than outside the firm. We refer to these firms as constrained firms.

By studying the problem of a government that seeks to maximize the total value of the corporate sector in the economy subject to a tax revenue target, we identify an elementary but general principle that shapes any optimal corporate taxation exercise. We refer to it as the *corporate taxation principle*: corporate taxes should be designed to minimize the tax burden faced by constrained firms. In other words, whenever possible, corporate taxes should be levied on unconstrained firms. If the government had full information about firms' financing and investment opportunities, this principle would be enough to implement an optimal allocation in which only unconstrained firms are taxed. However, it is not easy for the government to determine whether a firm is constrained or not. When firms have private information about their future investment opportunities, the government must design an optimal incentive compatible mechanism to induce firms to reveal whether or not they are constrained. The optimal mechanism must account for the fact that firms understand that, if they reveal themselves to be unconstrained, they will be taxed. The set of possible mechanisms includes very complex policies. However, we show that the optimal mechanism features a simple implementation: a corporate payout tax (a tax on dividends, share repurchases, and other payouts to the firms' shareholders).

The existence of a rich literature that seeks to identify financially constrained firms supports our premise that it is difficult for the government to determine whether a firm is financially constrained. Our results build on this literature by making use of the notion that the payout policy of a firm can help reveal whether it is financially constrained. This argument, which arises endogenously in our model, goes back at least to Fazzari, Hubbard and Petersen (1988), who a priori argue that firms that are consistently paying large dividends are not likely to be financially constrained. Kaplan and Zingales (1997) provide direct evidence relating dividend payments to financial constraints. They show that, even within a sample of low-dividend-paying firms, paying relatively more dividends is associated with a reduced likelihood of reporting being financially constrained. Consistent with these ideas, the optimal mechanism in our model uses the payout policy of a firm to determine whether a firm should be taxed or not.

Throughout the paper, and motivated by the richness of our framework, we define corporate taxes as "taxes collected from corporations." Our framework allows for conventional corporate income taxes, which in practice are based on a firm's profit, adjusted for numerous credits and deductions. The optimal policy in our model endogenously determines whether taxes on corporations should be based on profits or alternative variables. We show that *corporate payout taxes*, defined as taxes that are paid by corporations in proportion to their payouts to shareholders, characterize the optimal mechanism. The corporate payout tax that we identify is distinct from dividend taxation in the personal income tax system, both because it treats share repurchases and dividends symmetrically and because some shareholders do not pay personal income taxes (e.g., endowments). As a result of this second difference, dividend taxes in the personal income tax code can generate clientele effects that are absent from the dividend taxes in the corporate tax system that we describe.[1]

In our model, firms make investment and financing decisions at the beginning of each date, before production materializes. As in Rampini and Viswanathan (2010), we consider an environment with limited enforcement and no exclusion after default. This constrains the ability of firms to raise financing, as well as the ability of the government to set taxes. After production occurs, firms make a dividend decision, face taxes, and contemplate the possibility of defaulting. We sequentially study i) a static model with perfect information, which illustrates the corporate taxation principle; ii) a static model in which firms have private information about their future investment opportunities, and iii) a dynamic model in which firms have private information about future investment opportunities.

In the first static model that we study, we characterize optimal tax policy in a single-date model under perfect information about the firms future investment opportunities. If a government who seeks to raise revenue can tax, but not subsidize, firms, the optimal policy is to tax only unconstrained firms, provided that this policy generates sufficient revenue. This is the corporate taxation principle. Intuitively, the government, valuing the welfare of all firms equally, wishes to collect taxes from the firms for whom paying taxes is least costly. Even without a revenue raising goal, if the government could both tax and subsidize firms – a scenario that we do not study – the optimal policy involves undoing all financial frictions by taxing unconstrained firms and subsiding constrained firms.

Next, we study a two-date model, in which the second date corresponds to the solution to the single-date full-information optimal tax problem. In the first date, firms privately learn their productivity in the second date. The government can use the firm's payout policy to elicit this information. Our main result shows that the optimal mechanism can be implemented with a corporate payout tax, provided that this policy raises sufficient revenue. The key intuition is that the desire to pay dividends separates firms that will be unconstrained in the second date from

[1] See Allen and Gale (2000) for a theory of dividend determination under clientele effects.

firms that will be constrained.[2] Firms that will be unconstrained in the second date anticipate that they will be taxed, and have low marginal products, and therefore prefer paying dividends in the first date to retaining earnings. Firms that will be constrained in the second date are in the opposite situation. They will have high marginal products and will not be taxed, and therefore prefer to not pay dividends. This difference between constrained and unconstrained firms allows the government to raise taxes in an incentive-compatible way by taxing dividends. Other choices by the firm (in our model, capital/investment during the first date) are not distorted in the optimal mechanism, because they are determined by firms' current productivity, not its future productivity. As a result, they cannot be used to separate firms that have good or bad investment opportunities in the future.

Our most general set of results arise in the context an infinite-horizon dynamic model in which the productivity of a given firm is time varying, and at each date the firm but not the government knows its productivity in the next date. This model allows for entry and exit, has to deal with the evolution of the distribution of firms, and must account for the ability of the government to borrow and save. Although, in general, it is well-known that solving dynamic models with asymmetric information is challenging (e.g., Pavan, Segal and Toikka (2009)), our results from the static model allow us to guess and verify the optimal policies in the dynamic model. We show that the optimal sequence of mechanisms can also be implemented by a corporate payout tax.

The corporate taxation principle that underlies our results has a strong analogy to the personal taxation principle in Mirrlees (1971).[3] However, there are also two key differences. First, we adopt the view that the government has no particular desire to equalize the value of various firms. In the classic non-linear optimal income taxation literature, the goal of the government is to transfer resources from individuals with low marginal utility to individuals with high marginal utility. In contrast, the issue in our model is not "incentives vs. equality," but rather "plucking the goose as to obtain the largest amount of feathers with the least possible amount of hissing."[4] Second, the financial frictions in our model arise from the firms' ability to default or restructure. The tax authority does not have any special power to circumvent these constraints. As a result, the possibility of defaulting acts as an interim participation constraint in our model that limits the ability of the government to extract as much as it would like from

[2]To simplify the exposition, we will often describe firm payouts as dividends (as opposed to share repurchases). These two payout policies are equivalent in our model.

[3]Salanie (2011) traces back the personal taxation principle to Edgeworth (1897), who refers to it as the principle of equimarginal sacrifice: "The richer should be taxed for the benefit of the poorer up to the point where complete equality of fortunes is attained."

[4]Although we use the term optimal policy to describe our analysis, the term efficient policy would apply equally well, since our results abstract by design from redistributive motives across firms. In particular, when there are no revenue needs, optimal taxes are zero.

unconstrained firms.

Our approach integrates several well-developed literatures. There is an extensive literature on corporate taxation, surveyed by Auerbach and Hines Jr (2002) and Graham (2013). This literature includes both theoretical models and empirical work, but largely takes as given the existing structure of corporate taxes. One strand of this literature that is particularly close to our work is the literature on dividend taxation in the personal income tax system. The "old view" (e.g., Poterba and Summers (1985)) is that dividend taxes raise the cost of equity financing, distorting firms' investment decisions. The "new view" (expressed, for example, in Korinek (2009)) is that firms, except at the beginning of their life-cycle, do not actively issue equity, and as a result dividend taxes are not distortionary for existing firms. Our model embeds this perspective – the optimality of dividend taxes, as opposed to some other kinds of corporate taxation, is closely related to this fact.

In our model, firms maximize the expected value of dividends, in contrast to the "agency view" advocated by Chetty and Saez (2010), in which corporate taxes can be used to correct managerial distortions. Given the additivity property of corrective and revenue raising policy objectives (Sandmo, 1975; Kopczuk, 2003), it seems reasonable to study optimal revenue raising policies in an environment in which there is no role for corrective policies. We further discuss how our results relate to some of the existing literature in Section 5.

Our approach shares the emphasis on optimal allocations, rather than particular taxes, with the existing work that studies optimal non-linear taxation, following Mirrlees (1971). Our dynamic model shares some features with the work in dynamic public finance, recently surveyed in Golosov, Tsyvinski and Werquin (2016). The key difference between our paper and these large literatures is our focus on firms and financial frictions. Dynamic Mirrleesian models of optimal taxation usually focus on the behavior of households and treat firms as a veil (see, e.g., Farhi and Werning (2012)). In contrast, we adopt a partial equilibrium perspective that emphasizes how financial frictions create a meaningful distinction between corporate and personal taxes. The financial frictions we employ build, in particular, on the work of Kehoe and Levine (1993), Alvarez and Jermann (2000), and, most closely, Rampini and Viswanathan (2010). Like Li, Whited and Wu (2016), we add taxes to a financial frictions model of the firm, but study optimal mechanisms rather than particular tax instruments.

Our results also relate to the literature on dynamic contracting under financial frictions. For instance, in Albuquerque and Hopenhayn (2004), the optimal contract for an entrepreneur who faces financial frictions only features payments to the entrepreneur (dividends) once the project is operating at the optimal level of capital. The equilibrium under optimal policy in our model shares has a similar flavor. One major difference between our model and most of this literature is our focus on government policy and a population of firms, as opposed to a single firm and its outside investors.[5] A second key difference concerns commitment: our results require that the government lack commitment, whereas Albuquerque and Hopenhayn (2004) require that the outside investors be able to commit to a long-term contract. Put another way, the firm dynamics in our model without commitment are similar to those of Albuquerque and Hopenhayn (2004) with commitment; we speculate that the dynamics in our model with commitment or their model without commitment are likely quite different. Our model emphasizes asymmetric information, a feature it shares with Clementi and Hopenhayn (2006). However, the asymmetric information in our model is about investment opportunities, as opposed to profit realizations. This distinction is important when studying taxation– the government in our model can measure past profits perfectly but has difficulty predicting future profitability. Policy in our model focuses on the allocation of the tax burden, not minimizing tax evasion, although both issues are surely important real-world concerns for governments.

Because we assume that the government must raise revenues by taxing firms, our results cannot speak to the question of whether taxing firms is ever optimal. Much (but not all, Straub and Werning (2014)) of the work on capital taxation under full information (Judd (1985); Chamley (1986); Chari and Kehoe (1999)) argues that, at least in the long-run, capital taxes should be zero. With asymmetric information, this exact result is overturned, but the welfare gains of capital taxation might be quite small (Farhi and Werning, 2012). To our knowledge, there are no results about the optimality or sub-optimality of corporate taxation with asymmetric information and financial frictions in a general equilibrium model. There is also a large literature on the incidence of corporate taxation, going back to Harberger (1962), and on the related issue of the choice of organizational form. Our partial equilibrium approach can be thought of as a building block towards addressing the more general questions of whether corporate taxation is desirable at all, its incidence, and its interactions with the taxation of households.

Section 2 describes the common environment that applies to a single date in all of our models. Section 3 studies static models with perfect information and asymmetric information, introducing our main results, while Section 4 extends our results to a richer dynamic environment. Section 5 interprets our results and discusses policy implications, and Section 6 concludes. The Appendix contains all proofs and derivations.

[5] One strand of this literature, starting with Cooley, Marimon and Quadrini (2004), extends the framework of Albuquerque and Hopenhayn (2004) to a general equilibrium context with a population of firms. Our paper is similar in that it studies a population of firms, but differs by emphasizing partial equilibrium and focusing on optimal taxation as opposed to firm and business cycle dynamics in general equilibrium.

2 Environment

We initially describe the common structure of a single date that applies to both the static and dynamic environments that we consider. There are three groups of agents in the economy: firms, outside investors, and a government. There is a single consumption good (dollars), which serves as numeraire. Both firms and outside investors are risk-neutral, and discount cash-flows within the date, between the beginning and end, at a gross real interest rate of R.[6]

Figure 1 illustrates the timeline of events within a date. At the beginning of the date, before production occurs, firms make financing and investment decisions. At the end of the date, after production and depreciation materialize, firms choose dividend payments, pay taxes, make repayments to outside investors, and consider the possibility of defaulting.

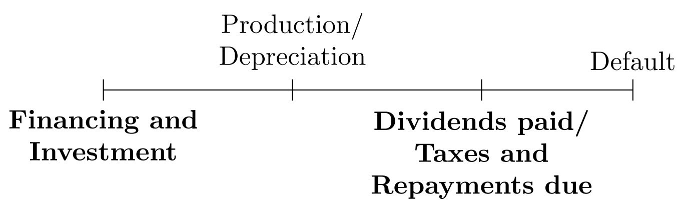

Figure 1: Single Date Timeline

Financing/Investment stage Firms are initially endowed with resources w_t, and can raise additional funds from outside investors, $m_t \geq 0$. We assume that the government cannot subsidize firms. Firms invest these resources in capital k_t, broadly defined, satisfying the following budget constraint:

$$k_t \leq w_t + m_t. \tag{1}$$

An investment of k_t dollars at the beginning of date t yields $f(k_t, \theta_t)$ dollars when production occurs. Firms' date t productivity depends on their date t type, θ_t, which is observed by all agents as of the beginning of date t. We assume that $f(k_t, \theta_t)$ is increasing and differentiable in both arguments, and concave in capital. We also assume that the marginal product of capital is increasing in the firm's type. Formally, the derivative $f_{k\theta}$ exists almost everywhere and is weakly positive wherever it exists. Firms that employ no capital receive no output; that is, $f(0, \theta_t) = 0$ for all types. Capital depreciates neo-classically at a rate $\delta \in [0, 1]$. Initial wealth, financing and investment choices, and production outcomes are observable to the government and outside

[6]Note that the interest rate is constant and invariant to policy – this is what we mean when we say that our analysis is partial equilibrium. One might expect the structure of corporate taxation to affect the real interest rate and the stochastic discount factor. Indeed, these sort of effects are central to the literature on capital taxation. As mentioned above, we view our analysis as building towards a consideration of these general equilibrium effects.

investors. We assume that there exists a first-best level of capital, $k^*(\theta_t)$, which is the smallest level of capital such that[7]

$$f_k(k^*(\theta_t), \theta_t) + 1 - \delta = R.$$

We further assume that once capital exceeds the first-best level, the marginal product of capital remains constant:

$$f_k(k, \theta_t) + 1 - \delta = R, \quad \forall k > k^*(\theta_t).$$

This assumption mimics the ability of the firm, after exhausting its ability to invest in physical capital, to invest at the risk-free rate. We discuss this assumption in more detail below.

The source of asymmetric information in this model arises from the privately known *future* investment opportunities of the firm. At the beginning of date t, firms learn their type for the next date, θ_{t+1}, which determines their future productivity. This generates asymmetric information, since outside investors and the government do not learn the firm's future productivity until the beginning of date $t+1$. Because the firm learns its next period type at the beginning of the period, it may condition its dividend, capital, and default decisions on this information. Repayments to outside investors and taxes can be assessed based on the firm's decisions, and therefore could be conditioned on the firm's type next period. However, the next period type itself is private to the firm. Therefore, repayments to outside investors and tax payments must satisfy incentive-compatibility conditions.

We assume that the space of types is the unit interval, $\theta_t \in [0,1]$. Focusing on a one-dimensional type space simplifies the model but is not without loss of generality. In particular, it prevents us from formally distinguishing between average and marginal products. In what follows, we interpret firms' types as controlling marginal, not average, productivity.

Dividend/Taxes stage After production and depreciation take place, firms declare a weakly positive dividend, $d_t \geq 0$. After this dividend is declared, firms must pay back $b_t \geq 0$ to outside investors and pay taxes $\tau_t \geq 0$ to the government. As we formally describe below when introducing the possibility of defaulting, the government/outside investors can allow or block the proposed dividend. Blocking the dividend prevents the money from leaving the firm, although it cannot prevent default. The firm decides whether or not to default to maximize the payoff to its owners.

The continuation wealth of a firm that repays its obligations is given by w_{t+1}, which can be expressed as

$$w_{t+1} = f(k_t, \theta_t) + (1 - \delta) k_t - d_t - b_t - \tau_t, \tag{2}$$

[7]That is, the first-best level of capital for a firm of date t type θ_t corresponds to the smallest solution of the problem: $\max_k f(k, \theta) + (1 - \delta) k - Rk$.

provided that this quantity is weakly positive. If this quantity is not weakly positive, repayment is not feasible, and the firm is forced to default, a situation that we describe next.

As in Rampini and Viswanathan (2010), we consider an environment with limited enforcement but no exclusion. If a firm defaults, it receives its continuation wealth w^D, which corresponds to the flow production and a fraction $1-\varphi$ of the depreciated capital stock. The continuation wealth $w^D(k_t,\theta_t)$ can be expressed as a function of the firm's capital choice, k_t, and its date t productivity, θ_t, as follows:

$$w^D(k_t,\theta_t) = f(k_t,\theta_t) + (1-\varphi)(1-\delta)k_t, \tag{3}$$

where $\varphi \in (0,1]$. The value of φ captures a form of limited enforcement that restricts the amount of funds that other parties (outside investors and the government) can receive from a firm. We require that a firm declare a dividend no larger than its continuation wealth in the event of default, so $d_t \leq w^D(k_t,\theta_t)$, which prevents the firm from continuing with negative wealth.

In case of default, firm owners cannot be excluded from starting a new firm with the same type as the defaulting firm. To ensure the government cannot circumvent this friction, it is natural to suppose that the government cannot condition on a firms' history, only its current wealth level and type. Consistent with this assumption, we will assume that the government lacks commitment. Otherwise, the government might commit to treating new firms, which result from a default, differently, and thereby discourage default. Formally, a firm will not default if

$$d_t + V(w_{t+1},\theta_{t+1}) \geq \max\left\{d_t + V\left(w^D(k_t,\theta_t) - d_t, \theta_{t+1}\right), V\left(w^D(k_t,\theta_t),\theta_{t+1}\right)\right\}, \tag{4}$$

where $V(w,\theta)$ denotes the continuation value for a firm that start a new date with resources w and type θ. The *max* operator in the right-hand side of Equation (4) reflects the government and creditors' ability to block the firm's proposed dividend. The firm can either propose a dividend acceptable to the government and creditors, in which case the first term in the maximization is the relevant constraint, or propose an unacceptable dividend, which will be blocked, making the second term in the maximization the relevant constraint. To avoid default, both of these deviations to default by the firm must be unprofitable. In what follows, we study mechanisms that avoid default. In our setting, it is without loss of generality to avoid default, and strictly optimal in the presence of inefficiencies associated with default.

We allow the repayments b_t to be contingent on the firm's capital, k_t, and the firm's current type, θ_t, but not the firm's dividend payment. The simplest interpretation of this restriction is that we allow our firms to issue only one-period debt. As a result, we set $m_t = R^{-1}b_t$. An alternative interpretation is that the outside investors also know the firm's next period type, and hence must break-even for each type. However, this interpretation requires the outside investors to have an informational advantage over the government, and the government may be able to extract this information almost costlessly Cremer and McLean (1988).

Remarks on the environment The environment considered here is meant to be the simplest one that allows for meaningful financing, investment, dividend, and default decisions, all of them necessary to study corporate taxation. We now briefly discuss several simplifications and conjecture how our results might apply to richer environments.

First, our environment has a single production input, capital, and decreasing returns to scale. Including labor or intermediate inputs should not affect our results. Our results also hold when constant returns to scale technology firms face decreasing demand for each firms' differentiated product.

Second, our environment features no uncertainty, with the exception of the process governing the firm's type, θ_t. Our results remain unaltered by the addition of observable and contractible shocks, as in Rampini and Viswanathan (2010), under the assumption that both creditors and the government can condition their payments/taxes on this shock. Adding such a shock would allow us to discuss issues like security design in more detail, at the expense of additional notation.

Third, and related to the previous point about security design, there are multiple ways of interpreting the environment. In the model, the agent receiving dividends is the agent controlling the firm's decisions. If firms maximize value for their shareholders, then dividends in the model map indeed to payments to shareholders in reality. Alternatively, one might interpret the model under the assumption of managerial control of the firm. In this case, "dividends" are really managerial compensation, which may include in-kind benefits or wages, and the outside investors might hold debt or equity claims. In other words, our results implicitly assume that it is possible to identify payments to the agents who control the firm. Our model also assumes that dividends cannot be negative, meaning that the shareholders or manager are unable to inject additional funds into the firm. We conjecture that our results would hold if negative dividends were feasible but costly, as in many models of financial frictions (e.g., Bolton, Chen and Wang (2011)). We leave the development of a richer model, in which there are conflicts of interest between manager and shareholders and both influence the firm's decisions, to future work.

Fourth, note that taxes and payments enter symmetrically in the model, operating entirely through the continuation wealth w_{t+1}. That is, substituting repayments b_t for taxes τ_t does not change either the default constraints or the initial budget constraint. This formulation ensures that the government cannot circumvent the financial frictions by assessing taxes that are not subject to default. Note also that any tax increase will tighten firms' financing constraint – this follows immediately from Equations 2 through 4. The key intuition behind our result that a payout tax is optimal is that firms for whom the no-default constraint is binding will not pay dividends, and hence will not be taxed if taxes are proportional to dividends.

Lastly, our information structure is different from the one usually assumed in dynamic public finance (e.g., Golosov, Tsyvinski and Werquin (2016)), which emphasizes the "Inverse Euler Equation." Capital taxation in those models arises through a Jensen's inequality effect, which requires uncertainty about the household's future type or other relevant variables. In contrast, in our model, the firms are perfectly informed about their type (the only exogenous state variable) in the next period. A second difference is that the curvature of the value functions in our model arises from financial frictions, instead of preferences, and a third difference is that in our model the possibility of default constrains the set of mechanisms that can be employed.

3 Static Optimal Taxation

We initially study a single-date model with full information and linear continuation values. Subsequently, we study a single-date model with asymmetric information about firms future continuation values.

3.1 Full Information

We refer to the date in this single-date model as $t = 1$. We assume that the firm simply consumes whatever wealth remains inside the firm at the end of date one. Formally, the continuation value for a firm at the end of date 1 corresponds to

$$V_2\left(w_2, \theta_2\right) = w_2.$$

The firm's value function at the beginning of date 1, given its initial wealth w_1 and type θ_1, is

$$V_1\left(w_1, \theta_1\right) = R^{-1}\left\{d_1\left(w_1, \theta_1\right) + w_2\left(w_1, \theta_1\right)\right\},$$

where $d_1\left(w_1, \theta_1\right)$ and $w_2\left(w_1, \theta_1\right)$ correspond to optimal allocations. The no-default constraint, defined in Equation (4), can be expressed as follows:

$$d_1\left(w_1, \theta_1\right) + w_2\left(w_1, \theta_1\right) \geq w^D\left(k_1\left(w_1, \theta_1\right), \theta_1\right),$$

where $k_1\left(w_1, \theta_1\right)$ denotes the optimal level of capital.

We adopt the language of mechanism design to describe the government's optimal policy, even though, with full information, there are no incentive compatibility constraints. In the rest of the paper, we will introduce asymmetric information and exploit direct revelation mechanisms to characterize the optimal policy. We proceed as if the government directly determines choice variables $\{k_1, b_1, m_1, d_1\}$ and taxes τ_1 for each firm, given its levels of wealth w_1 and type θ_1, which are both observable variables. Each of these choice variables is required to be weakly positive, and the upper bound on dividends, $d_1 \leq w^D(k_1, \theta_1)$, must also be satisfied. The government must respect the financing/investment budget constraint (1) and the definition of continuation wealth and wealth after default, introduced in Equations (2) and (3).

The government must also respect the no-default constraint, which can be interpreted as an interim participation constraint. The no-default constraint arises from the possibility of the firm complying with the government's mechanism, but then defaulting instead of paying its obligations. There is a second interim participation constraint that arises from the possibility of the firm disregarding the government's mechanism entirely. If a firm does this, the government can respond by assigning the firm infinite taxes, inducing default and preventing outside borrowing. As a result, the firm would be limited to investing its initial wealth in capital and then defaulting. The constraint to ensure this deviation is unprofitable is

$$d_1\left(w_1, \theta_1\right) + w_2\left(w_1, \theta_1\right) \geq w^D\left(k_1\left(w_1, \theta_1\right), \theta_1\right).$$

This constraint will always be satisfied, so long as the firm's level of capital, $k_1\left(w_1, \theta_1\right)$, is weakly greater than its wealth, w_1, and the no-default constraint is satisfied. This will always be the case, and consequently this constraint is redundant.

Finally, the government must, across the population of firms, raise sufficient resources through taxation. Let $\mu_1\left(w_1, \theta_1\right)$ denote the measure of firms with wealth $w_1 \in [0, \infty]$ and type $\theta_1 \in [0, 1]$. Any tax policy must satisfy

$$R^{-1} \int_0^\infty \int_0^1 \tau_1\left(w_1, \theta_1\right) \mu_1\left(w_1, \theta_1\right) d\theta_1 dw_1 \geq G + B_1 > 0,$$

where G denotes the (strictly positive) required expenditure and B_1 is government debt that must be repaid. In this one-period model, G and B_1 play exactly the same role; we introduce government debt to connect this model to our two-period and dynamic models.

Subject to all of these constraints, the government maximizes the welfare of firms,[8]

$$\int_0^\infty \int_0^1 V_1 (w_1, \theta_1) \mu_1 (w_1, \theta_1) d\theta_1 dw_1.$$

Although the government's problem features numerous constraints, it can be simplified as follows. First, note that in the one-period problem, the firm's choice of dividends is irrelevant – the wealth of the firm at the end of the period will be consumed, one way or another. It is also straightforward to observe that the financing/investment budget constraint always binds. As a result, there is really only a single choice variable for the firm (in Lemma 1 below, we choose capital, but this is arbitrary). The problem can also be simplified by introducing a multiplier, χ_1, on the government's revenue-raising constraint. Aside from this constraint, there is no interaction between the firms, and the Lagrangian version of the problem can be studied firm by firm. The multiplier χ_1 has a simple interpretation: it is the marginal cost to the firms of raising an additional unit of revenue through taxation. The following lemma summarizes these claims.

Lemma 1. *The government's mechanism design problem can be written as*

$$J_1 (\mu_1, B_1) = \min_{\chi_1 \geq 0} \left\{ \int_0^\infty \int_0^1 U_1 (w_1, \theta_1; \chi_1) \mu_1 (w_1, \theta_1) d\theta_1 dw_1 - \chi_1 (G + B_1) \right\},$$

where

$$U_1 (w_1, \theta_1; \chi_1) = \max_{k_1 \geq 0, \tau_1 \geq 0} \left\{ R^{-1} \{ f (k_1, \theta_1) + (1 - \delta) k_1 - R k_1 \} + w_1 + R^{-1} (\chi_1 - 1) \tau_1 \right\},$$

subject to the constraint that

$$w_1 \leq k_1 \leq \frac{w_1 - R^{-1} \tau_1}{1 - R^{-1} \varphi (1 - \delta)}.$$

Proof. See the Appendix, Section A.1. □

First, note that, if constraints on the government's choice of capital do not bind, the government can choose to set the capital equal to its first-best level, $k^* (\theta_1)$. If the firm has so much wealth that it can achieve more than the first-best capital level without any borrowing, $w > k^* (\theta)$, the government can choose to set capital equal to wealth. Because the marginal product of capital is equal to the risk-free rate for all $k > k^* (\theta)$, this is equally good from the perspective of government. In contrast, if wealth is insufficient to reach the first-best level of capital, in the absence of taxes, then the government cannot set $k_1 \geq k^* (\theta_1)$. As a result, the

[8] For simplicity, we assume that the government has a utilitarian objective that maximizes the sum of all firm values. We believe the optimality of payout taxes remains valid under arbitrary social welfare weights, although the tax rate will depend on the welfare weights.

marginal product of capital, in equilibrium, will be greater than the risk-free rate, and we will call the firm constrained.

The marginal benefit to the government of taxation, $R^{-1}(\chi_1 - 1)$, is the same for all firms. When $\chi_1 = 1$, for the unconstrained firms, there is no particular reason to tax one firm instead of another; so the optimal policy is not determined. However, it will never be optimal to tax a constrained firm if there exists an unconstrained firm that could be taxed instead. As a result, if the government can raise a sufficient quantity of revenue from the unconstrained firms, it will not tax the constrained firms. The following proposition summarizes these results.

Proposition 1. *(Single-date optimal tax) If the level of required expenditure G is sufficiently small and there exists a positive mass of firms with $k^*(\theta_1) < \frac{w_1}{1-R^{-1}\varphi(1-\delta)}$, there exists an optimal policy in which $\chi_1 = 1$ and the government sets*

$$\tau_1(w_1, \theta_1) = \tau R \max\{w_1 - \bar{w}(\theta_1), 0\},$$

for some $\tau \in [0, \varphi(1-\delta)]$, where $\bar{w}(\theta)$ is defined as the level of wealth required to achieve the first-best level of capital in the absence of taxes,

$$\bar{w}(\theta) = \left(1 - R^{-1}\varphi(1-\delta)\right) k^*(\theta). \tag{5}$$

Proof. See the Appendix, Section A.2. □

Note that $\bar{w}(\theta_1)$ corresponds to the level of wealth required to achieve the first-best level of capital in the absence of taxes. We have chosen to focus on a particular policy – a linear tax on "excess wealth" – because of its simplicity and because it generates certain properties in the government and firm's date one value functions that resemble the results of our dynamic model without commitment. Specifically, the following corollary describes the properties of the functions V_1 and U_1 under the particular policy we have chosen.

Corollary 1. *The functions $V_1(w, \theta)$ and $U_1(w, \theta)$ are increasing and concave in wealth, and Lipschitz continuous in firm type. For all $w > \bar{w}(\theta)$, $V_w(w, \theta) = 1 - \tau$ and $U_w(w, \theta) = 1$, and for all $w < \bar{w}(\theta)$, $V_w(w, \theta) > 1$ and $U_w(w, \theta) > 1$. For $w < \bar{w}(\theta)$, U and V are strictly concave in wealth and have strictly positive cross-partials $V_{w\theta}$ and $U_{w\theta}$.*

Proof. See the Appendix, Section A.3. □

These properties summarize the intuitive idea that firms with wealth levels below $\bar{w}(\theta)$ are constrained and untaxed, whereas firms with wealth levels above $\bar{w}(\theta)$ pay a tax rate τ on excess wealth. In the next subsection, we will use this full-information equilibrium as the second date in a two-date model with asymmetric information. The date one value functions described in this section will be the continuation value functions at date zero (the first date in the model). The properties of the marginal value of wealth, for firms and for the government, just described will lead to a particular optimal mechanism at date zero, a dividend tax.

Our choice to study this particular full-information policy and to use it as the continuation value in our two-period model anticipates the results we will show in the infinite-horizon model. In particular, the results of Corollary 1 are almost identical to the properties that we will conjecture in our dynamic model.[9]

Figure 2 graphically illustrates the firm's continuation value function.

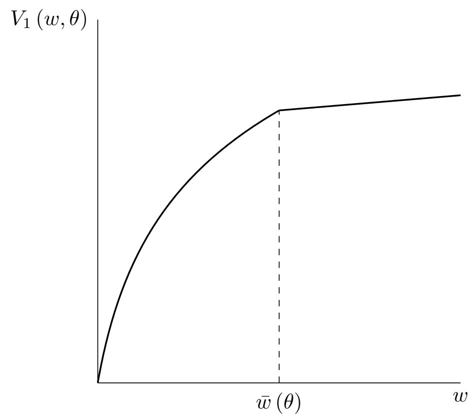

Figure 2: Continuation value

Remark. (Corporate Taxation Principle) Proposition 1 illustrates the principle that optimal corporate taxation under financial frictions implies taxing unconstrained firms, which have a low marginal value of funds inside the firm, and leaving untaxed constrained firms, which have a high marginal value of funds.

We next study how this principle manifests in a richer model in which the government has asymmetric information about firms' investment opportunities.

3.2 Asymmetric Information

We now study a two-date model in which firms have private information about their *future* productivity/investment opportunities. We refer to the first date in this two-date model as $t = 0$. We assume for simplicity that all firms start with the same values of initial wealth w_0 and current productivity θ_0, although it is straightforward to introduce observable heterogeneity along both

[9]The difference is that, in the dynamic model, $V_w(w,\theta)$ can be in the interval $(1-\tau, 1)$ for unconstrained firms. This difference is a consequence of the fact that in the one and two-date models, firms can pay out wealth to their shareholders at the end of date one without ever paying a "dividend."

dimensions, and we consider such heterogeneity in our dynamic model. We assume that the second date of the model is the full-information model described in the previous sub-section.

We will impose some assumptions on the initial wealth w_0, the initial productivity θ_0, and the distribution of types at date one, which we denote $\mu\left(\theta_1|\theta_0\right)$. These assumptions simplify our analysis of the two-period model, but play no role in the dynamic model.

Assumption 1. *The most productive type θ_1 in the support of $\mu\left(\theta_1|\theta_0\right)$ will be constrained at date one, and the least productive type in the support of $\mu\left(\theta_1|\theta_0\right)$ cannot be constrained at date one, but has some productive investment opportunities. Formally,*

$$f\left(k^*\left(\theta_0\right),\theta_0\right)+\left(\left(1-\delta\right)-R\right)k^*\left(\theta_0\right)+Rw_0<\max_{\theta_1:\mu\left(\theta_1|\theta_0\right)>0}\bar{w}\left(\theta_1\right) \tag{6}$$

$$f\left(w_0,\theta_0\right)+\left(1-\varphi\right)\left(1-\delta\right)w_0>\min_{\theta_1:\mu\left(\theta_1|\theta_0\right)>0}\bar{w}\left(\theta_1\right)\geq\varphi\left(1-\delta\right)w_0 \tag{7}$$

We also assume that reaching the date zero first-best level of capital is possible but not required, $w_0<k^*\left(\theta_0\right)<\frac{1}{R-\varphi(1-\delta)}w_0$.

Note that $\bar{w}\left(\theta\right)$ (defined in Equation (5)) is increasing in θ, meaning that firms with higher types become constrained at higher levels of wealth. Under assumption 1, we show that it is feasible to raise funds without distortions, and that the optimal government policy can be implemented using a dividend tax.

We are now ready to describe the optimal mechanism design problem of the government. We allow for the possibility of a "double" deviation in the mechanism, in which a firm reports some type θ_1' initially, when choosing investment/financing, and then reports another type θ_1'' when declaring a dividend/making a default decision. Formally, we use the notation $d_0\left(\theta_1',\theta_1''\right)$ to refer to the dividend allocated to a firm that reports θ_1' at the investment/financing, and then reports θ_1'' at the dividend/tax stage. We use the same two-argument notation for other variables. For variables that exclusively depend on the first report, like k_0 or b_0, we use a single argument notation, of the form $k_0\left(\theta_1'\right)$ and $b_0\left(\theta_1'\right)$. We consider incentive-compatible direct revelation mechanisms with IC constraints at the financing/investment stage and the dividend/taxes stage.

The government faces a revenue-raising constraint across the population of firms. However, we allow the government to borrow and save at the risk-free rate. Government borrowing and saving plays a crucial role in the model: it allows the date zero government to control the tax rate set by the date one government, even though the government lacks commitment.

As described in Proposition 1, assuming truthful reporting by firms in date zero, the date one government will set

$$B_1 + G = \tau \int_0^1 \max\left\{w_1\left(\theta_1, \theta_1\right) - \bar{w}\left(\theta_1\right), 0\right\} \mu\left(\theta_1 | \theta_0\right) d\theta_1$$

for some $\tau \in [0, \varphi(1-\delta)]$, if such a policy is feasible. By controlling firms' continuation wealth $w_1(\theta_1, \theta_1)$ and the level of government debt B_1, the date zero government can effectively determine the next period's tax rate.

The date one level of government debt B_1 is determined (again assuming truthful reporting by firms at date zero) by the flow budget constraint

$$B_0 + G = R^{-1}\left(B_1 + \int_0^1 \tau_0\left(\theta_1, \theta_1\right) \mu\left(\theta_1 | \theta_0\right) d\theta_1\right).$$

We divide the government's problems into two steps, consisting of a choice of tax rate τ to implement at date one and an optimal mechanism given that tax rate. The government maximizes the current value of the firms,

$$J_0\left(w_0, \theta_0\right) = \max_{\tau \in [0, \varphi(1-\delta)]} \max_{\cdots} R^{-1} \int_0^1 \left\{d_0\left(\theta_1, \theta_1\right) + V_1\left(w_1\left(\theta_1, \theta_1\right), \theta_1; \tau\right)\right\} \mu\left(\theta_1 | \theta_0\right) d\theta_1,$$

subject to a single intertemporal budget constraint

$$B_0 + G\left(1 + R^{-1}\right) = R^{-1} \int_0^1 \tau_0\left(\theta_1, \theta_1\right) \mu\left(\theta_1 | \theta_0\right) d\theta_1 + R^{-1}\tau \int_0^1 \max\left\{w_1\left(\theta_1, \theta_1\right) - \bar{w}\left(\theta_1\right), 0\right\} \mu\left(\theta_1 | \theta_0\right) d\theta_1,$$

maximizing over the capital, dividends, debt, continuation wealth, and date zero taxes for firms. We use the notation $V_1(w_1, \theta_1; \tau)$ to emphasize that the firms' continuation value function depends on the tax rate being implemented in the next period.

The government is subject to multiple constraints. In particular, it must respect firms' budget constraints, which imply that investment must be funded with internal resources and outside investors funds,

$$k_0\left(\theta_1'\right) \le w_0 + R^{-1} b_0\left(\theta_1'\right), \ \forall \theta_1'. \qquad \text{(Budget Constraint)} \tag{8}$$

The government also faces the no-default constraint as well as the upper limit on dividend payments, which can be expressed as

$$w_0^D\left(\theta_1'\right) \leq d_0\left(\theta_1', \theta_1''\right) + w_0^R\left(\theta_1', \theta_1''\right), \ \forall \theta_1', \theta_1'', \qquad \text{(No Default)} \qquad (9)$$

$$d_0\left(\theta_1', \theta_1''\right) \leq w_0^D\left(\theta_1'\right), \ \forall \theta_1', \theta_1''. \qquad \text{(Upper Limit on Dividends)} \qquad (10)$$

Note that the no-default constraint has a simpler form than Equation (4), because we will separately consider deviations to a blocked dividend. In fact, it follows from the definitions of continuation wealth that the no-default constraint can be reformulated as

$$b_0\left(\theta_1'\right) + \tau_0\left(\theta_1', \theta_1''\right) \leq \varphi\left(1-\delta\right) k_0\left(\theta_1'\right), \ \forall \theta_1', \theta_1'',$$

which effectively limits the amount of capital. This constraint is the constraint of Rampini and Viswanathan (2010) with taxes treated as additional debt payments.

The government needs to account for three sets of incentive constraints. The first set of constraints apply at the financing/investment stage. These constraints guarantee that firms find it optimal not to deviate at the beginning of the date, when investment is determined and financing from outside investors is received. The second set of constraints prevents firms from deviating at the time in which dividends are paid and taxes are collected. Finally, the last set of constraints guarantees that firms declare a dividend that will not be blocked by the government or outside investors. Formally, these

$$d_0\left(\theta_1', \theta_1\right) + V_1\left(w_1\left(\theta_1', \theta_1\right), \theta_1; \tau\right) \leq d_0\left(\theta_1, \theta_1\right) + V_1\left(w_1\left(\theta_1, \theta_1\right), \theta_1; \tau\right), \ \forall \theta_1, \theta_1', \qquad \text{(Financing/Investment IC)} \qquad (11)$$

$$d_0\left(\theta_1', \theta_1''\right) + V_1\left(w_1\left(\theta_1', \theta_1''\right), \theta_1; \tau\right) \leq d_0\left(\theta_1', \theta_1\right) + V_1\left(w_1\left(\theta_1', \theta_1\right), \theta_1; \tau\right), \ \forall \theta_1, \theta_1', \theta_1'', \qquad \text{(Dividend/Taxes IC)} \qquad (12)$$

$$V_1\left(w^D\left(k_0(\theta_1'), \theta_0\right), \theta_1; \tau\right) \leq d_0\left(\theta_1', \theta_1\right) + V_1\left(w_1\left(\theta_1', \theta_1\right), \theta_1; \tau\right), \ \forall \theta_1, \theta_1'. \qquad \text{(Blocked Dividend IC)} \qquad (13)$$

Finally, the government is subject to non-negativity constraints regarding dividend payments $d_0\left(\theta_1', \theta_1''\right) \geq 0$, tax levels, $\tau_0\left(\theta_1', \theta_1''\right) \geq 0$, outside funding, $b_0\left(\theta_1'\right)$, and continuation wealth, and must have $k_0\left(\theta_1'\right) \geq w_0$.

Summing up, the problem of the government is to maximize its objective, subject to constraints (8), (9), (10), (11), (12) and (13). Our description of the problem assumes that it is optimal for the government at date zero to induce the government at date one to implement the full-information allocation we described in the previous section; we prove this as part of the proposition below. We state a more general version of the full mechanism design problem in Appendix Section C.3.

Proposition 2 introduces our main result.

Proposition 2. *Under Assumption 1, if the amount of funds to be raised $B_0 + G\left(1 + R^{-1}\right)$ is not too large, the optimal mechanism can be implemented by a payout tax.*

Proof. See the Appendix, Section A.4. □

Proposition 2 shows that the principle of taxing only unconstrained firms remains valid even when the government has asymmetric information about firms' continuation values. The dividend tax allows the government to cleanly separate constrained and unconstrained firms while raising revenue.

4 Dynamic Optimal Taxation

In this section, we extend the static analysis of the previous section to an infinite-horizon context. To do so, we introduce firm entry and exit, firm heterogeneity, and government borrowing and saving. We discuss and justify each of these elements before formally describing the government's problem. After introducing the government's problem, we conjecture and subsequently verify a functional form for the government's value function that assumes a structure for continuation values that corresponds to the one derived for the static model in the previous section. We eventually show that most of the conclusions of the previous section carry over to the dynamic model. Our main result shows again that a payout tax is optimal, provided that the amount of money required by the government is not too large.

4.1 Firm Entry and Exit

In our dynamic model, we consider firm entry and firm exit. We initially describe firm exit. Exit is not default; it is possible for a firm to shut down by liquidating and paying off its taxes and creditors. In our model, because default will not occur in equilibrium, exit is the only way that firms will leave the economy.

Each period, some firms are exogenously forced to become "exiting," meaning that they will be forced to exit next period. Firms that are not forced to become exiting can nevertheless choose to become exiting. This will not occur in equilibrium, but the threat of premature firm exit will limit the government's ability to tax the firms.

Formally, we assume that the lowest type, $\theta = 0$, is exiting. For this type, $k^*(0) = 0$, meaning that the firm has no productive opportunities and can only hold cash. Once the firm is exiting, it will remain exiting forever. Formally, in the notation we use below, $\mu(\theta|0)$ is the Dirac delta function. As a result, the value function of an exiting firm with wealth w is

$$V_t(w,0) = R^{-1}(d_t(w,0) + V_{t+1}(w_{t+1}(w,0),0)).$$

A firm of type θ will be forced to become exiting next period with probability $\mu(0|\theta)$. As in the static model, firms privately know their current and next period type, including whether or not they are exiting, at the beginning of the current period. The government only observes the firms' current type. As a result, the government can identify which firms are exiting, but will not directly observe which firms will become exiting next period. Firms that are not forced to become exiting can choose to become exiting instead of or in addition to defaulting. If the government does not want a firm to voluntarily exit (which it will not), the government must ensure that the firm has an incentive to continue if it does not default,

$$V_{t+1}(w_{t+1},\theta_{t+1}) \geq V_{t+1}(w_{t+1},0),$$

and that it has no incentive to deviate by both defaulting and exiting,

$$d_t + V_{t+1}(w_{t+1},\theta_{t+1}) \geq \max\left\{d_t + V_{t+1}\left(w_t^D - d_t, 0\right), V_{t+1}\left(w_t^D, 0\right)\right\}.$$

In the solution to the government's problem that we will guess and verify, these constraints will be satisfied in any allocation satisfying the no-default constraint, and the government does not want any firm to exit prematurely.[10]

We now describe firm entry. At the end of each period, a measure of potential entrants, $e(\bar{w},\theta)$, are born with outside wealth $\bar{w}$ and type $\theta > 0$ (no firm is born exiting). Each potential entrant faces the same fixed cost of entry, F, and can choose how much wealth to put into the firm. We define $w_E(\bar{w},\theta)$ as an entry wealth that maximizes utility,

$$w_E(\bar{w},\theta) \in \arg\max_{w\in[0,\bar{w}]} V_{t+1}(w,\theta) - w,$$

and firms will choose to enter if

$$V_{t+1}(w_E(\bar{w},\theta),\theta) - w_E(\bar{w},\theta) \geq F.$$

[10]The form of exit we contemplate involves shutting down the firm. We do not model the possibility that firms may shift their activities to a different tax jurisdiction. This possibility may limit the taxes the government could collect but otherwise leaves the problem unchanged.

The measure of firms entering the economy in period t with wealth w and type θ is therefore

$$e_t(w,\theta) = \int_0^\infty \int_0^\infty \delta_{dirac}\left(w_E(\bar{w},\theta) - w\right) \mathbf{1}\left\{V_{t+1}\left(w_E(\bar{w},\theta),\theta\right) - w_E(\bar{w},\theta) \geq F\right\} e(\bar{w},\theta)\, d\bar{w}dw, \tag{14}$$

where $\delta_{dirac}(\cdot)$ denotes the Dirac delta function and $\mathbf{1}\{\cdot\}$ denotes the indicator function.

We have assumed that, if there are multiple levels of entry wealth that maximize utility, $w_E(\bar{w},\theta)$ will implement an arbitrary rule to determine which level of wealth entrants choose. In the equilibrium we construct, there will be only one optimal level of entry wealth. Our assumption that the mass of potential entrants, $e(\bar{w},\theta)$, is the same each period simplifies the exposition but is not necessary. For simplicity, we have also assumed that entry is a one-shot option, rather than allowing potential entrants to wait before entering.

Because entry happens at the end of the period, the government in the next period will take whatever entry decisions are made as given. That is, the government lacks commitment regarding entry. This assumption simplifies the problem, and brings it closer to the static problem analyzed in the previous section. Having described the entry and exit of firms, we describe next the dynamics of the population of firms.

4.2 The Population of Firms

To simplify the exposition, we assumed in the static model that there was a single common type and initial wealth level. In the dynamic model, there will be a population of firms with different wealth levels, current types, and types next period. We will use the measure $\mu_t(w,\theta)$ to denote the mass of firms at time t with wealth w and current (observable) type θ.

We assume that the type structure is Markov, and denote by $\mu(\theta'|\theta)$ the likelihood that a firm of current type θ will have type θ' in the next period. As in our static model, the next period's type θ' is known to the firm – but not to the government – in the current period. This is the key form of private information in the model. We assume that $\mu(\theta'|\theta)$ is Lipschitz continuous with respect to θ for $\theta \in [0,1]$, and that $\mu(\theta'|\theta_1)$ strictly first-order stochastically dominates $\mu(\theta'|\theta_2)$ for all $\theta_1 > \theta_2$.[11]

At each period t, the government inherits a measure μ_t, and through its policies will influence the measure μ_{t+1} that carries over to the next period. Let $w_{t+1}(w,\theta,\theta')$ denote the continuation wealth for a firm with current wealth w, current type θ, and future type θ', assuming truthful reporting of the firms' future type θ'. Assuming that the government avoids premature exit by

[11] Our first-order stochastic dominance assumption implies, at least to some degree, persistence of types. It is a sufficient but not necessary assumption to prove our result; in particular, the case of IID types could be handled through a slightly different argument.

firms, the next period's distribution of firm types is

$$\mu_{t+1}\left(w',\theta'\right) = e_t\left(w',\theta'\right) + \int_0^\infty \int_{0^+}^1 \delta\left(w_{t+1}\left(w,\theta,\theta'\right) - w'\right)\mu\left(\theta'|\theta\right)\mu_t\left(w,\theta\right)d\theta dw \qquad (15)$$
$$+\mathbf{1}\left\{\theta'=0\right\}\int_0^\infty \delta_{dirac}\left(w_{t+1}\left(w,0\right)-w'\right)\mu_t\left(w,0\right)dw.$$

Having described the evolution of the population of the firms in the model, we are now able to describe the government's problem.

4.3 The Government's Problem

In our dynamic model, the government is able to borrow and save freely, but lacks commitment. Each period, the government inherits a level of debt B_t and a population of firms μ_t. The government must spend G each period, so next period's debt is

$$B_{t+1} = R\left(B_t+G\right) - \int_0^\infty \int_{0^+}^1 \int_0^1 \tau_t\left(w,\theta,\theta'\right)\mu\left(\theta'|\theta\right)\mu_t\left(w,\theta\right)d\theta' d\theta dw - \int_0^\infty \tau_t\left(w,0\right)\mu_t\left(w,0\right)dw, \qquad (16)$$

where $\tau_t\left(w,\theta,\theta'\right)$ is the tax revenue raised under truthful reporting from a firm with wealth w, current type $\theta > 0$, and next period's type θ', and $\tau_t\left(w,0\right)$ is the revenue raised from exiting firms.

We denote the government's value function by $J_t\left(B,\mu\right)$ and the firm's value function (equivalently, the net present value of dividends under the current and future optimal policies of the government) by $V_t\left(w,\theta\right)$. At each time t, the government takes the policies of future governments, and hence $J_{t+1}\left(B',\mu'\right)$ and $V_{t+1}\left(w',\theta';B',\mu'\right)$, as given. Note that the firm continuation values V_{t+1} depend on both the firm specific variables $\left(w',\theta'\right)$ and government debt B' and the measure μ', because the latter two could determine future government policies and hence future firm continuation values.

Each period, the government designs an incentive-compatible direct-revelation mechanism for the current period, of the sort described in our static model, which allocates debt, capital, taxes, dividends, and continuation wealth to each of the firms. Let $\mathcal{M}$ denote the set of such mechanisms, and let $m \in \mathcal{M}$ be a particular mechanism, including the functions τ_t and w_{t+1} that are necessary to define the evolution of the population of firms and of government debt, and the function d_t that describes the dividends paid in the current period. For a detailed description of this mechanism design problem, which is a slight generalization of the mechanism design problem for the static model, see Appendix Section C.3.

We study a Markov sub-game perfect equilibrium of the game between the current government and its future selves. The government's problem, under the sustained assumption that no firms exit prematurely, in recursive form, is

$$J_t\left(B_t,\mu_t\right)=\max_{m_t\in\mathcal{M}}R^{-1}\int_0^{\infty}\int_{0^+}^{1}\int_0^1 d_t\left(w,\theta,\theta'\right)\mu\left(\theta'|\theta\right)\mu_t\left(w,\theta\right)d\theta' d\theta dw$$
$$+R^{-1}\int_0^{\infty}d_t\left(w,0\right)\mu_t\left(w,0\right)dw+R^{-1}J_{t+1}\left(B_{t+1},\mu_{t+1}\right), \tag{17}$$

subject to the equations describing the evolution of the population firms (15), the level of debt (16), a transversality condition,

$$\lim_{s\to\infty}R^{-s}J_{t+s}\left(B_{t+s},\mu_{t+s}\right)=0,$$

and a No-Ponzi condition,

$$\lim_{s\to\infty}R^{-s}B_{t+s}\leq 0,$$

taking all future policies as given Markov functions of B_{t+1} and μ_{t+1}.

The solution to this problem induces a value function for the non-exiting firms,

$$V_t\left(w,\theta;B_t,\mu_t\right)=R^{-1}\int_0^1\left\{d_t^*\left(w,\theta,\theta'\right)+V_{t+1}\left(w_{t+1}^*\left(w,\theta,\theta'\right),\theta';B_{t+1},\mu_{t+1}\right)\right\}\mu\left(\theta'|\theta\right)d\theta'$$

and for exiting firms,

$$V_t\left(w,0;B_t,\mu_t\right)=R^{-1}\left\{d_t^*\left(w,0\right)+V_{t+1}\left(w_{t+1}^*\left(w,0\right),0;B_{t+1},\mu_{t+1}\right)\right\},$$

where d_t^* and w_{t+1}^* denote policies under the government's optimal mechanism.

In this equilibrium, sub-game perfection requires that each government optimizes, meaning that

$$J_t\left(B,\mu\right)=J_{t+1}\left(B,\mu\right),$$

and we impose the additional stationarity restriction that $V_t\left(w,\theta;B,\mu\right)=V_{t+1}\left(w,\theta;B,\mu\right)$.[12]

The equilibrium we study is equivalent to a problem in which an infinitely-lived government maximizes the net present value of the dividends for all current and future firms, subject to a single inter-temporal budget constraint (arising from the No-Ponzi restriction) and a restriction to Markov policies in the set of history-independent mechanisms. This equivalence follows from that fact that our equilibrium satisfies the transversality condition mentioned above. The restriction to the set of history-independent mechanisms is important, and a consequence of

[12]This restriction is mild, but rules out equilibria in which the government oscillates between different policies that achieve the same value for the government but have different continuation values for specific firms.

the government's lack of commitment. A government that could punish defaulting firms that re-enter the economy (which is ruled out by history-independence) could relax the financial frictions in the economy.[13]

The government's ability to borrow and save is critical for our main result in the presence of any kind of transition dynamics (that is, if the population of firms does not start in a steady state). If the government were required to raise a particular amount of tax revenue each period, with no ability to smooth distortions across periods, the government would need to impose different dividend tax rates in each period (assuming a dividend tax is the optimal policy). However, because firms can defer dividend payments at no cost by accumulating cash, if dividend taxes will be lower in the future, the government will not be able to collect any dividend taxes today. Consequently, a dividend tax would not be optimal in this case. In our model, because the government can borrow and save, today's government can leave tomorrow's government with a debt level that ensures tomorrow's government will implement the same dividend tax rate that today's government implements. Consequently, it is possible for the government to maintain a constant dividend tax rate. This discussion should also make it clear that our assumption of constant government spending G is not important– only the net present value of the stream of government spending matters.

Having introduced the new elements of the model, we next describe our assumption that ensures the feasibility of dividend taxation and our conjecture about the form of the government's value function, and then discuss our result.

4.4 Assumptions, Conjecture, and Result

To discuss our assumptions that ensure the finiteness of value functions, we introduce the value function $\bar{V}(w, \theta; \tau_d)$, which is the value function for a firm in a market equilibrium with wealth w and type θ given a dividend tax rate τ_d. By "market equilibrium" we mean that the firm chooses its own borrowing, capital, dividends, and continuation wealth, subject to all of the limits and financial frictions in our model, in the presence of a dividend tax. We provide a more precise description of this function in Appendix C.1. Our assumptions will also be defined using the mass of firms that enter each period in the presence of a dividend tax, $\bar{e}(w, \theta; \tau_d)$, which is defined by (14) with the value function $\bar{V}(w, \theta; \tau_d)$ in the place of the continuation value.

Our first assumption ensures that the firm and government value functions we study are finite. We use the notation $\mu_0(w, \theta)$ to denote the initial population of firms.

[13] We leave open the question of whether there are non-Markov "sustainable" equilibria (Chari and Kehoe, 1990) in which a government without commitment nevertheless punishes defaulting firms and circumvents the financial frictions. We thank Felipe Varas for pointing out this possibility.

Assumption 2. *There exists a constant, V_{max}, such that, for all $\theta \in [0,1]$ and $w \in [0,\infty)$ with either $\mu_0(w,\theta) > 0$ or $\bar{e}(w,\theta;0) > 0$, $\bar{V}(w,\theta;0) < V_{max}$, and the total value of all present and future dividends in the economy is strictly positive and finite for all $\tau \in [0,\bar{\tau})$, for some $\bar{\tau} > 0$:*

$$0 < \int_0^\infty \int_0^1 \bar{V}(w,\theta;\tau)\,\mu_0(w,\theta)\,d\theta dw + \frac{1}{R-1}\int_0^\infty \int_0^1 \bar{V}(w,\theta;\tau)\,\bar{e}(w,\theta;\tau)\,d\theta dw < \infty.$$

This assumption is necessary, from a technical perspective, to ensure that the optimization problems we study are well-behaved. It implicitly embeds restrictions on the transition probabilities of types, $\mu(\theta'|\theta)$, along with the production functions for each type, to ensure that firm's net present value of dividends is finite. It also requires that the initial mass of firms and entering firms have finite wealth, and that the total mass of firms be finite. The requirement that the net present value of all dividends in the economy is strictly positive and finite ensures that the government's value function is finite and that the government can raise at least some revenue from the firms through a dividend tax.

Next, we conjecture a form for the government's value function. Our proof strategy is to derive the optimal mechanism, using the continuation values from this conjecture, and then show that the resulting value functions satisfy the conditions of our conjecture.

Conjecture. *The optimal mechanism $m \in \mathcal{M}$ and firm value function $V(w,\theta;B,\mu) = V(w,\theta)$ do not depend on B or μ or time. There exists a $\tau \in [0,1)$, a continuous increasing function $\bar{w}(\theta)$, and a function $U(w,\theta)$ such that, for all $w > \bar{w}(\theta)$, $V_w(w,\theta) = 1-\tau$ and $U_w(w,\theta) = 1$, and for all $w < \bar{w}(\theta)$, $V_w(w,\theta) > 1-\tau$ and $U_w(w,\theta) > 1$, with V and U concave in w and with $V(0,\theta) = U(0,\theta) = 0$, such that*

$$J(B_t,\mu_t) = J_E - \frac{G}{1-R^{-1}} - B_t + \int_0^\infty \int_0^1 U(w,\theta)\,\mu_t(w,\theta)\,d\theta dw,$$

where J_E is a positive constant. Furthermore, U and V are Lipschitz continuous in θ, and, for $w < \bar{w}(\theta)$, U and V are strictly concave in wealth and have strictly positive cross-partials $V_{w\theta}$ and $U_{w\theta}$. The function $U(w,\theta)$, for $\theta > 0$, is the net present value of future taxes and dividends,

$$U(w,\theta) = R^{-1}\int_0^1 \left\{d_t(w,\theta,\theta') + \tau_t(w,\theta,\theta') + U\left(w_{t+1}(w,\theta,\theta'),\theta'\right)\right\}\mu(\theta'|\theta)\,d\theta',\ \forall\theta > 0,$$
$$U(w,0) = R^{-1}\left\{d_t(w,0) + \tau_t(w,0) + U(w_{t+1}(w,0),0)\right\}.$$

The function $U(w,\theta)$ can be thought of as the "social value" of the firm. It is always greater than the private value of the firm, $V(w,\theta)$. The constant J_E represents the value of the future entering firms.

Verifying the conjecture involves several steps. First, we solve for the optimal mechanism that characterizes d_t, τ_t, w_{t+1}, and other variables, assuming that the continuation value functions U and V have the properties described in the conjecture. This step is essentially identical to our analysis of the static model, and also allows us to show that the optimal mechanism can be implemented by a dividend tax. Second, having solved for the optimal mechanism, we show that the value functions U and V indeed have the conjectured properties. Third, we show that the conjectured value function J satisfies the recursive equation characterizing equilibrium (17), and state a condition under which the No-Ponzi condition will be satisfied.

The following proposition states our main result. The proposition is expressed in terms of the "funding need" at date zero, $B_0 + \frac{G}{1-R^{-1}}$, which is the net present value of the funds the government needs to raise. Note that there is a maximum amount of revenue that can be raised by the government, since higher taxes are associated with lower firm entry, which reduces the total tax base and generates a revenue Laffer curve.

Proposition 3. *There exists a non-empty interval of the funding need, $B_0 + \frac{G}{1-R^{-1}}$, such that Conjecture 4.4 holds, the equilibrium satisfies the No-Ponzi condition and the transversality condition, and the optimal sequence of mechanisms can be implemented by a constant dividend tax.*

Proof. See the Appendix, Section A.5. □

5 Interpretation and Policy Implications

In this section, we discuss a variety of practical and theoretical issues related to our results. We initially discuss the relationship between our corporate payout tax and corporate taxes as they are implemented in the United States and most other developed countries. We then discuss forces that are absent from our model but have been used to explain firms' dividend policies, and conjecture how our results might change in the presence of these forces. We also emphasize that the optimal policy may be sensitive to the nature of financial frictions assumed in the model, and discuss general equilibrium considerations that are absent from our analysis. Finally, we will mention a proposed tax reform in France along the lines of our optimal policy.

First, at various time periods in the United States, dividends and share repurchases were treated differently for individual income tax purposes. In our model, "dividends" include all payments to the agents controlling the firm, and hence should be interpreted to include share repurchases. That is, our optimal policy would tax dividends and share repurchases at the same rate.[14] Related to this, in the United States and most other countries, corporate taxes are assessed

[14]The global rise in corporate saving (e.g., Chen, Karabarbounis and Neiman (2017)) makes our results more relevant. For instance, Apple Inc., the world's largest company by market capitalization as of 2015, has experienced a large increase in profits, but not on dividends, over the last few decades. At the same time, it has accumulated

on earnings, net of various deductions. Our results show that this is in fact a reasonable way to structure taxes, provided that deductions are given for all retained earnings. After all, earnings less retained earnings is by definition equal to dividends plus share repurchases.[15]

Intriguingly, at a high level, the current tax system, modified to have full expensing of investment, may not be too far from this policy. This conclusion is reminiscent of the "new view" of dividend taxation (described in, for example, Auerbach and Hines Jr (2002)). Perhaps more strikingly, the model justifies taxing dividends but not interest payments on debt (interpreting the claims of outside investors in the model as debt). The key distinction, from the model's perspective, is control. A time-invariant tax on payments to the agents controlling the firm (which we interpret as dividends) does not distort the intertemporal decisions of firms, only entry (a point made by Korinek (2009)). Any other tax would both distort intertemporal decisions and entry. Our model currently does not allow the firm to issue equity except on entry, so we cannot say if dividend taxes distort debt/equity choices after entry. However, in our dynamic model, dividend taxes do distort equity issuance downwards at entry, and hence lead to firms having more debt and less equity at various points in their lifecycle. Despite this, dividend taxes are optimal, because all taxes that raised the same amount of value from the firm would distort the equity issuance decision by at least as much. This emphasizes the point that interest deductibility might both distort debt/equity decisions and be optimal.

This discussion leads naturally to the second point we wish to emphasize, which is related to forces that are absent from the model but perhaps related to firms' dividend policies. Our assumption that shareholders control the firm, and hence that "payments to the agent controlling the firm" are indeed dividends, may not apply to all firms. Managerial entrenchment (Zwiebel, 1996) or other agency conflicts between shareholders and managers (Chetty and Saez, 2010) might result in managerial, as opposed to shareholder, control of the firm. In this case, we speculate that our model could be reinterpreted to justify a tax on managerial compensation. However, if managers and shareholders have the same information and can engage in an optimal contract, then it may not matter which one is taxed. The simplest case in which this is true is when the manager is paid in proportion to the dividends shareholders receive. Regardless of who controls the firm, it might also be possible for the agent controlling the firm to extract value from the firm through in-kind benefits, such as markups for goods or services provided, or disguised as payments to unrelated parties. If these payments destroy value relative to dividend

large holdings of cash and it has recently repurchased its equity. The optimal tax that we characterize implies that these equity repurchases should be taxed.

[15] Implementing the corporate payout tax found to be optimal in this paper within the current personal tax system may impact individuals' investment decisions. For instance, households may have incentives to use firms as tax-deferred investment vehicles in order to avoid capital taxes. We thank Michael Faulkender for pointing out this issue.

payments (along the lines of, e.g., DeMarzo and Urošević (2006)), then the possibility of such payments will limit the dividend tax rate the government can implement but otherwise leave our results unchanged.

Our model also omits signaling and catering explanations for firms' dividend choices, both of which have been discussed in the literature and may be relevant for optimal policy. With regards to catering, we have little to say; welfare analysis in the presence of behavioral biases on the part of investors is beyond the scope of our model. With regards to signaling, recall that our model does feature asymmetric information. The key missing dimension is a reason for the firm to care about creditors' (as opposed to the government's) perceptions, because our creditors only purchase debt in the firm. However, the firm is concerned about signaling its marginal value of wealth to the government – this is the essence of the model. Augmenting the model with choices about which securities to issue to outside investors would allow us to speak more directly to the signaling value of dividends.

One change to our model that would drastically change results is altering the nature of the financial friction. In a previous version of the paper, we considered a model with full commitment by the government and creditors, including the ability to prevent defaulting firms from re-entering. The resulting financial friction was equivalent to the one introduced by Kehoe and Levine (1993). In this model, only the firm's continuation value (the present value of its future dividends) matters for the financial friction, and the firm's current wealth is irrelevant. Under these assumptions, if the government could assess a tax on firms entering the economy and then commit to never tax them again, this would be the optimal policy. The principle of avoiding taxing constrained firms still applies in this model, but because of the forward-looking nature of the constraint, dividend taxes in the future exacerbate the constraints of firms today. As a result, dividend taxes do not avoid taxing constrained firms, and hence are not optimal. The key distinction between the two models concerns firms' ability to borrow against their entire continuation value (which can be interpreted as future earnings) or only their tangible assets. Rampini and Viswanathan (2013) provide evidence that tangible assets are an important determinant of firm borrowing capacity, but which of these constraints apply to which firms is, to our knowledge, an open question.[16]

Our model takes a partial equilibrium approach, neglecting general equilibrium considerations. In practice, corporate taxes are likely to influence equilibrium interest rates (and stochastic discount factors more generally), and these concerns are in fact central to the extensive literature on capital taxation. Our model features both asymmetric information and financial

[16]Considering a scenario with limited enforcement and no exclusion after default, as in Rampini and Viswanathan (2010), but with a government with full commitment, is problematic, since the government can use its commitment ability to overcome financial frictions.

frictions, and hence the classic capital taxation results cannot be directly applied to the model. Other taxes, such as labor income or consumption taxes, will likely affect both constrained and unconstrained firms, and (speculatively) our corporate taxation principle suggests that a corporate payout tax may be preferable to these taxes. We intend to pursue this question in future research.

Finally, we should highlight that France in 2012 implemented a corporate dividend tax somewhat along the lines suggested by our analysis (in addition to the more traditional corporate taxes). France assessed a 3% tax on all dividends paid. Court rulings eventually caused the tax to be declared unconstitutional, apparently because of double taxation, and in 2017 the tax was repealed. A recent analysis of other dividend tax changes in France (Boissel and Matray (2019)) finds, consistent with our model, that dividend tax changes induced some changes in payout policy but only small changes in real outcomes, consistent with patterns documented for the personal dividend tax in the United States (Yagan (2015)).

6 Conclusion

We have provided a normative analysis of optimal corporate taxation under financial frictions. Without financial frictions, all firms operate at the first-best level. With financial frictions, if the government has full information, it is optimal to tax only the unconstrained firms, consistent with the corporate taxation principle that we introduce in this paper. When firms' future investment opportunities are private information, the government must use a mechanism to elicit which firms are unconstrained in an incentive compatible way. Although the structure of the problem appears complicated, we show that a tax on dividends/share repurchases can both separate constrained and unconstrained firms and raise revenue, and hence is optimal. Practically, this dividend tax could be implemented in our current tax system by allowing for full deductibility of retained earnings.

To our knowledge, this paper is the first to address the question of how corporate taxes should be structured in an environment in which any feasible non-negative tax instrument can be employed. Future work taking into account manager-shareholder conflict, security design, interactions between personal and corporate taxation, and general equilibrium considerations should result in a deeper understanding of optimal corporate taxation.

References

Albuquerque, Rui, and Hugo A Hopenhayn. 2004. "Optimal lending contracts and firm dynamics." *The Review of Economic Studies*, 71(2): 285–315.

Allen, F., and D. Gale. 2000. "Financial contagion." *Journal of political economy*, 108(1): 1–33.

Alvarez, Fernando, and Urban J. Jermann. 2000. "Efficiency, equilibrium, and asset pricing with risk of default." *Econometrica*, 68(4): 775–797.

Auerbach, Alan J, and James R Hines Jr. 2002. "Taxation and economic efficiency." *Handbook of public economics*, 3: 1347–1421.

Boissel, Charles, and Adrien Matray. 2019. "Higher Dividend Taxes, No Problem! Evidence from Taxing Entrepreneurs in France." *Evidence from Taxing Entrepreneurs in France (July 10, 2019)*.

Bolton, Patrick, Hui Chen, and Neng Wang. 2011. "A unified theory of Tobin's q, corporate investment, financing, and risk management." *The journal of Finance*, 66(5): 1545–1578.

Chamley, Christophe. 1986. "Optimal taxation of capital income in general equilibrium with infinite lives." *Econometrica: Journal of the Econometric Society*, 607–622.

Chari, VV, and Patrick J. Kehoe. 1999. "Optimal Fiscal and Monetary Policy." *Handbook of Macroeconomics, edited by John B. Taylor and Michael Woodford, Vol. 1C. Elsevier*.

Chari, V.V., and P.J. Kehoe. 1990. "Sustainable plans." *Journal of Political Economy*, 783–802.

Chen, Peter, Loukas Karabarbounis, and Brent Neiman. 2017. "The global rise of corporate saving." *Journal of Monetary Economics*, 89: 1–19.

Chetty, Raj, and Emmanuel Saez. 2010. "Optimal Taxation and Social Insurance with Endogenous Private Insurance." *American Economic Journal: Economic Policy*, 2(2): 85–116.

Clementi, G.L., and H.A. Hopenhayn. 2006. "A Theory of Financing Constraints and Firm Dynamics*." *Quarterly Journal of Economics*, 121(1): 229–265.

Cooley, Thomas, Ramon Marimon, and Vincenzo Quadrini. 2004. "Aggregate consequences of limited contract enforceability." *Journal of Political Economy*, 112(4): 817–847.

Cremer, Jacques, and Richard P McLean. 1988. "Full extraction of the surplus in Bayesian and dominant strategy auctions." *Econometrica: Journal of the Econometric Society*, 1247–1257.

DeMarzo, P.M., and B. Urošević. 2006. "Ownership dynamics and asset pricing with a large shareholder." *Journal of Political Economy*, 114(4): 774–815.

Edgeworth, Francis Ysidro. 1897. "The pure theory of taxation." *The Economic Journal*, 7(28): 550–571.

Farhi, Emmanuel, and Iván Werning. 2012. "Capital taxation: Quantitative explorations of the inverse Euler equation." *Journal of Political Economy*, 120(3): 398–445.

Fazzari, Steven M, R Glenn Hubbard, and Bruce C Petersen. 1988. "Financing Constraints and Corporate Investment." *Brookings Papers on Economic Activity*, 141–206.

Fudenberg, Drew, and Jean Tirole. 1991. *Game Theory.* MIT Press.

Golosov, Mikhail, Aleh Tsyvinski, and Nicolas Werquin. 2016. "Recursive contracts and endogenously incomplete markets." *Handbook of Macroeconomics*, 2: 725–841.

Graham, John R. 2013. "Do taxes affect corporate decisions? A review." *Handbook of the Economics of Finance SET: Volumes 2A & 2B: Corporate Finance and Asset Pricing*, 123.

Harberger, Arnold C. 1962. "The incidence of the corporation income tax." *Journal of Political economy*, 70(3): 215–240.

Judd, K.L. 1985. "On the performance of patents." *Econometrica: Journal of the Econometric Society*, 567–585.

Kaplan, Steven N, and Luigi Zingales. 1997. "Do investment-cash flow sensitivities provide useful measures of financing constraints?" *The Quarterly Journal of Economics*, 112(1): 169–215.

Kehoe, Timothy J., and David K. Levine. 1993. "Debt-constrained asset markets." *The Review of Economic Studies*, 60(4): 865.

Kopczuk, Wojciech. 2003. "A note on optimal taxation in the presence of externalities." *Economics Letters*, 80(1): 81–86.

Korinek, Anton. 2009. "Systemic risk-taking: amplification effects, externalities and regulatory responses." *Working Paper, Maryland*.

Li, Shaojin, Toni M Whited, and Yufeng Wu. 2016. "Collateral, taxes, and leverage." *Review of Financial Studies*, hhw008.

Mirrlees, J.A. 1971. "An exploration in the theory of optimum income taxation." *The review of economic studies*, 38(2): 175–208.

Pavan, A., I. Segal, and J. Toikka. 2009. "Dynamic mechanism design: Incentive compatibility, profit maximization and information disclosure."

Poterba, James M, and Lawrence H Summers. 1985. "The economic effects of dividend taxation." National Bureau of Economic Research.

Rampini, A., and S. Viswanathan. 2010. "Collateral, risk management, and the distribution of debt capacity." *The Journal of Finance*, 65(6): 2293–2322.

Rampini, Adriano A, and S Viswanathan. 2013. "Collateral and capital structure." *Journal of Financial Economics*, 109(2): 466–492.

Salanie, Bernard. 2011. *The Economics of Taxation.* MIT Press.

Sandmo, Agnar. 1975. "Optimal taxation in the presence of externalities." *The Swedish Journal of Economics*, 86–98.

Straub, Ludwig, and Iván Werning. 2014. "Positive long run capital taxation: Chamley-Judd revisited." National Bureau of Economic Research.

Yagan, Danny. 2015. "Capital tax reform and the real economy: The effects of the 2003 dividend tax cut." *American Economic Review*, 105(12): 3531–63.

Zwiebel, Jeffrey. 1996. "Dynamic capital structure under managerial entrenchment." *The American Economic Review*, 1197–1215.

APPENDIX

Section A of the Appendix contains all proofs referenced in the paper. Section B includes several auxiliary results. Section C provides explicit definitions of value functions and mechanism design problems used in our proofs and referenced in the main text.

A Proofs

A.1 Proof of Lemma 1

The government solves the problem

$$\max \int_0^\infty \int_0^1 R^{-1} \left\{ d_1 \left(w_1, \theta_1 \right) + w_2 \left(w_1, \theta_1 \right) \right\} \mu_1 \left(w_1, \theta_1 \right) dw_1 d\theta_1,$$

subject to the following set of constraints that apply to each level of initial wealth w_1 and type θ_1

$$\begin{aligned} k_1 \left(w_1, \theta_1 \right) &\leq R^{-1} b \left(w_1, \theta_1 \right) + w_1, \ \forall w_1, \forall \theta_1 && \text{(Financing/Investment)} \\ w_2 \left(w_1, \theta_1 \right) &\leq f \left(k_1 \left(w_1, \theta_1 \right), \theta_1 \right) + \left(1 - \delta \right) k_1 \left(w_1, \theta_1 \right) && \text{(Dividend/Taxes)} \\ &\qquad - d_1 \left(w_1, \theta_1 \right) - b_1 \left(w_1, \theta_1 \right) - \tau_1 \left(w_1, \theta_1 \right), \ \forall w_1, \forall \theta_1 \\ w_1^D \left(k_1 \left(w_1, \theta_1 \right), \theta_1 \right) &\leq d_1 \left(w_1, \theta_1 \right) + w_2 \left(w_1, \theta_1 \right), \ \forall w_1, \forall \theta_1 && \text{(No Default)} \\ d_1 \left(w_1, \theta_1 \right) &\leq w_1^D \left(k_1 \left(w_1, \theta_1 \right), \theta_1 \right), \ \forall w_1, \forall \theta_1 && \text{(Upper Limit on Dividends)} \end{aligned}$$

as well as the revenue-raising constraint,

$$B_1 + G \leq \int_0^\infty \int_0^1 R^{-1} \tau_1 \left(w_1, \theta_1 \right) \mu_1 \left(w_1, \theta_1 \right) dw_1 d\theta_1, \qquad \text{(Revenue Raising)}$$

and the non-negativity constraints for k_1, d_1, b_1, and τ_1.

Observe that firms' financing/investment budget constraint must bind. Therefore, we can express $d_1 \left(w_1, \theta_1 \right) + w_2 \left(w_1, \theta_1 \right)$ as follows:

$$d_1 \left(w_1, \theta_1 \right) + w_2 \left(w_1, \theta_1 \right) = f \left(k_1 \left(w_1, \theta_1 \right), \theta_1 \right) + \left(1 - \delta \right) k_1 \left(w_1, \theta_1 \right) - b_1 \left(w_1, \theta_1 \right) - \tau_1 \left(w_1, \theta_1 \right).$$

When combined with the definition of continuation wealth after default in Equation (3), the no-default constraint simplifies to

$$b_1 \left(w_1, \theta_1 \right) + \tau_1 \left(w_1, \theta_1 \right) \leq \varphi \left(1 - \delta \right) k_1 \left(w_1, \theta_1 \right).$$

As a result, $d_1 \left(w_1, \theta_1 \right)$ enters only in upper limit on dividends, and therefore it is without loss of generality to assume $d_1(w_1,\theta_1) = 0$ and ignore the limit on dividends. Moreover, if the financing/investment budget constraint does not bind, the government can increase $k_1(w_1,\theta_1)$, increasing the objective and relaxing the no-default constraint, which implies that financing/investment constraint must bind at the optimum.

Therefore, the government's problem can be reformulated in simplified form as follows:

$$\max \int_0^\infty \int_0^1 R^{-1}\left\{f\left(k_1(w_1,\theta_1),\theta_1\right) + (1-\delta)k_1(w_1,\theta_1) - b_1(w_1,\theta_1) - \tau_1(w_1,\theta_1)\right\}\mu_1(w_1,\theta_1)\,dw_1 d\theta_1$$

subject to

$$k_1(w_1,\theta_1) = R^{-1}b_1(w_1,\theta_1) + w_1,\ \forall w_1, \forall\theta_1 \qquad \text{(Financing/Investment)}$$

$$b_1(w_1,\theta_1) + \tau_1(w_1,\theta_1) \le \varphi(1-\delta)k_1(w_1,\theta_1),\ \forall w_1,\forall\theta_1 \qquad \text{(No Default)}$$

$$B_1 + G \le \int_0^\infty \int_0^1 R^{-1}\tau_1(w_1,\theta_1)\mu_1(w_1,\theta_1)\,dw_1 d\theta_1, \qquad \text{(Revenue Raising)}$$

in addition to the the non-negativity constraints. This problem has affine constraints and a concave objective function, and therefore the infinite dimensional analog of the Karush-Kuhn-Tucker conditions are necessary and sufficient to find an optimum. Solving for $b_1(w_1,\theta_1)$ in the financing/investment constraint and substituting in, allows us to rewrite the objective function as follows:

$$\int_0^\infty \int_0^1 \left\{R^{-1}\left(f\left(k_1(w_1,\theta_1),\theta_1\right) + (1-\delta-R)k_1(w_1,\theta_1) - \tau_1(w_1,\theta_1)\right) + w_1\right\}\mu_1(w_1,\theta_1)\,dw_1 d\theta_1.$$

Solving for $b_1(w_1,\theta_1)$ in the financing/investment constraint and substituting in, we can rewrite the no default constraint as follows,

$$k_1(w_1,\theta_1) \le \frac{w_1 - R^{-1}\tau_1}{1 - R^{-1}\varphi(1-\delta)}.$$

The non-negativity constraint on b_1 implies that $w_1 \le k_1$. By defining the Lagrange multiplier on the revenue raising constraint by χ_1, we can express the new objective function as

$$\int_0^\infty \int_0^1 \left\{R^{-1}\left(f\left(k_1(w_1,\theta_1),\theta_1\right) + (1-\delta-R)k_1(w_1,\theta_1) + (\chi_1 - 1)\tau_1(w_1,\theta_1)\right) + w_1\right\}\mu_1(w_1,\theta_1)\,dw_1 d\theta_1 - \chi_1(G + B_1).$$

We can simplify the formulation of the government's problem by defining $U_1(w_1,\theta_1;\chi_1)$ as follows:

$$U_1(w_1,\theta_1;\chi_1) = \max_{k_1\ge 0,\tau_1\ge 0}\left\{R^{-1}\left\{f(k_1,\theta_1) + (1-\delta-R)k_1\right\} + w_1 + R^{-1}(\chi_1 - 1)\tau_1\right\},$$

subject to the constraint that

$$w_1 \leq k_1 \leq \frac{w_1 - R^{-1}\tau_1}{1 - R^{-1}\varphi\left(1-\delta\right)}.$$

So the government's problem can be expressed as

$$J_1\left(\mu_1, B_1\right) = \min_{\chi_1 \geq 0} \left\{ \int_0^\infty \int_0^1 U_1\left(w_1, \theta_1; \chi_1\right) \mu_1\left(w_1, \theta_1\right) d\theta_1 dw_1 - \chi_1\left(G + B_1\right) \right\},$$

which concludes the proof. Note that we use the measure of types μ_1 and the level of debt B_1 as arguments of the government's value function to be consistent with our formulation of the dynamic model. Since G and B_1 enter symmetrically, it is without loss of generality to only include B_1 as an argument of the government's value function.

A.2 Proof of Proposition 1

First, note that if $\chi_1 < 1$, it will be optimal to set $\tau_1 = 0$ always, and therefore raise no revenue. Therefore, it must be that $\chi_1 \geq 1$, and that, if $\chi_1 = 1$ is feasible, it will be optimal.

If $\chi_1 = 1$, the government's problem corresponds to

$$U_1\left(w_1, \theta_1; 1\right) = \max_{k_1 \geq 0, \tau_1 \geq 0} R^{-1}\left\{f\left(k_1, \theta_1\right) + \left(1-\delta\right)k_1 - Rk_1\right\} + w_1,$$

subject to the following constraint on capital allocations

$$w_1 \leq k_1 \leq \frac{w_1 - R^{-1}\tau_1}{1 - R^{-1}\varphi\left(1-\delta\right)},$$

and non-negativity constraint for k_1 and τ_1. Let μ and ϕ be the multipliers on the upper and lower bounds for capital, and let ν be the multiplier on the constraint that $\tau_1 \geq 0$. We can therefore define the following Lagrangian:

$$\mathcal{L} = R^{-1}\left\{f\left(k_1, \theta_1\right) + \left(1-\delta\right)k_1 - Rk_1\right\} + w_1 - \mu\left(k_1 - \frac{w_1 - R^{-1}\tau_1}{1 - R^{-1}\varphi\left(1-\delta\right)}\right) + \phi\left(k_1 - w_1\right) + \nu\tau_1.$$

Note that, under the assumption that $w_1 \geq 0$ for all firms, the capital non-negativity constraint becomes redundant. The optimality conditions for k_1 and τ_1 respectively are

$$R^{-1}\left\{f_k\left(k_1, \theta_1\right) + \left(1-\delta\right) - R\right\} - \mu + \phi = 0 \Rightarrow R^{-1}\left\{f_k\left(k_1, \theta_1\right) + \left(1-\delta\right) - R\right\} = \mu - \phi \quad \text{(A1)}$$

$$-\frac{1}{R - \varphi\left(1-\delta\right)}\mu + \nu = 0 \Rightarrow \nu = \frac{1}{R - \varphi\left(1-\delta\right)}\mu \quad \text{(A2)}$$

Equation (A2) implies that if $\mu > 0$, then $\nu > 0$ and $\tau_1 = 0$. In general, $\mu > 0$ and $\phi > 0$ are mutually exclusive. Therefore, if $\tau_1 = 0$, that is, $\nu > 0$, it must be that $\mu > 0$ and $\phi = 0$.

Because $f_k(k_1, \theta_1) + (1-\delta) - R \geq 0, \forall k_1$, with equality if and only if $k_1 \geq k^*(\theta_1)$, it follows that $k_1 < k^*(\theta_1)$ implies that $\mu > 0$ and $\phi = 0$, and also $\nu > 0$ and $\tau_1 = 0$. This case requires that

$$k^*(\theta_1) > \frac{w_1}{1 - R^{-1}\varphi(1-\delta)}.$$

If $k_1 \geq k^*(\theta_1)$, then we must have $\mu = \phi = 0$ and $\nu = 0$. This case requires that

$$k^*(\theta_1) \leq \frac{w_1}{1 - R^{-1}\varphi(1-\delta)}.$$

In this case, the tax is indeterminate, but must necessarily satisfy

$$0 \leq \tau_1 \leq \varphi(1-\delta) w_1,$$

so that the feasible set for capital is non-empty.[17] In the statement of Proposition 1, we propose the following tax function

$$\tau_1(w_1, \theta_1) = \tau R \max\left\{w_1 - \left(1 - R^{-1}\varphi(1-\delta)\right) k^*(\theta_1), 0\right\},$$

where $\tau \in [0, \varphi(1-\delta)]$. The proposed functional form satisfies the restrictions on τ_1 in both cases, and raises positive revenue. Therefore, as long as the revenue raised by the proposed policy satisfies the revenue-raising constraint, our claim holds.

A.3 Proof of Corollary 1

Properties of V_1

The function $V_1(w_1, \theta_1) = R^{-1}\{d_1(w_1, \theta_1) + w_2(w_1, \theta_1)\}$ corresponds to an indirect utility function for a firm with wealth w_1 and type θ_1. Note that

$$\begin{aligned} d_1(w_1, \theta_1) + w_2(w_1, \theta_1) &= f(k_1(w_1, \theta_1), \theta_1) + (1-\delta) k_1(w_1, \theta_1) - \tau_1(w_1, \theta_1) - R k_1(w_1, \theta_1) + R w_1, \\ &= f(k_1(w_1, \theta_1), \theta_1) + (1-\delta) k_1(w_1, \theta_1) - \tau R \max\{w_1 - \bar{w}(\theta_1), 0\} - R k_1(w_1, \theta_1) + R w_1, \end{aligned}$$

where we use the fact that the optimal policy satisfies $\tau_1(w_1, \theta_1) = \tau \max\{w_1 - \bar{w}(\theta_1), 0\}$, and where the optimal choice of capital is given by

$$k_1(w_1, \theta_1) = \min\left\{k^*(\theta_1), \frac{w_1}{1 - R^{-1}\varphi(1-\delta)}\right\}.$$

[17] Note that feasibility requires that τ_1 satisfies $w_1 \leq \frac{w_1 - R^{-1}\tau_1}{1 - R^{-1}\varphi(1-\delta)}$.

Note first that $V_1(w,\theta)$ is continuous. Anywhere $w_1 > \bar{w}(\theta)$,

$$V_{1,w}(w_1,\theta_1) = 1-\tau.$$

Anywhere $w_1 < \bar{w}(\theta)$,

$$V_{1,w}(w_1,\theta_1) = R^{-1}\frac{f_k\left(\frac{w_1}{1-R^{-1}\varphi(1-\delta)},\theta_1\right)+(1-\delta)-R}{1-R^{-1}\varphi(1-\delta)}+1>1$$

by the assumption that if $w_1 < \bar{w}(\theta)$, $\frac{w_1}{1-R^{-1}\varphi(1-\delta)} < k^*(\theta_1)$. The properties of the cross-partial and concavity follow from the cross-partial and concavity of the production function. Lipschitz continuity follows from the differentiability of the production function (and hence $k^*(\theta_1)$) with respect to θ_1.

Properties of U_1

The function $U_1(w_1,\theta_1;1)$, which corresponds to an indirect utility function for firms from the perspective of the government, is given by

$$U_1(w_1,\theta_1;1) = f(k_1(w_1,\theta_1),\theta_1)+(1-\delta)k_1(w_1,\theta_1)-Rk_1(w_1,\theta_1)+Rw_1,$$

where $k_1(w_1,\theta_1)$ corresponds to

$$k_1(w_1,\theta_1) = \min\left\{k^*(\theta_1),\frac{w_1}{1-R^{-1}\varphi(1-\delta)}\right\}.$$

The argument above for V_1 applies from this point, essentially unmodified.

A.4 Proof of Proposition 2

The government's problem in the static (two-date) model, assuming they wish to implement the tax rate τ described by Proposition 1, is

$$J_0(w_0,\theta_0) = \max_{\tau\in[0,\varphi(1-\delta)]}\max_{m_t\in\mathcal{M}(\tau)} R^{-1}\int_0^1\{d_0(\theta_1,\theta_1)+V_1(w_1(\theta_1,\theta_1),\theta_1;\tau)\}\,\mu(\theta_1|\theta_0)\,d\theta_1$$

subject to the constraints

$$B_1+G = R^{-1}\tau R\int_0^1\max\{w_1(\theta_1,\theta_1)-\bar{w}(\theta_1),0\}\,\mu(\theta_1|\theta_0)\,d\theta_1,$$

$$R^{-1}B_1 = B_0+G-R^{-1}\int_0^1\tau_0(\theta_1,\theta_1)\,\mu(\theta_1|\theta_0)\,d\theta_1$$

which can be combined into a single government budget constraint

$$B_0 + G\left(1 + R^{-1}\right) = R^{-1} \int_0^1 \tau_0(\theta_1, \theta_1) \mu(\theta_1|\theta_0) d\theta_1 + R^{-1}\tau \int_0^1 \max\{w_1(\theta_1, \theta_1) - \bar{w}(\theta_1), 0\} \mu(\theta_1|\theta_0) d\theta_1.$$

Note that the set of feasible mechanisms is influenced by the tax rate, through the continuation value function of the firms.

The Lagrangian version is

$$\begin{aligned} J_0(w_0, \theta_0) = & \max_{\tau \in [0, \varphi(1-\delta)]} \max_{m_t \in \mathcal{M}(\tau)} \min_{\chi_0 \geq 0} R^{-1} \int_0^1 \{d_0(\theta_1, \theta_1) + \chi_0 \tau_0(\theta_1, \theta_1)\} \mu(\theta_1|\theta_0) d\theta_1 \\ & + R^{-1} \int_0^1 \{V_1(w_1(\theta_1, \theta_1), \theta_1; \tau) + \chi_0 \tau \max\{w_1(\theta_1, \theta_1) - \bar{w}(\theta_1), 0\}\} \mu(\theta_1|\theta_0) d\theta_1. \end{aligned}$$

We begin by proving that if $\chi_0 < 1$, the optimal mechanism collects no taxes. This is shown in Lemma 5.

Consequently, if a solution with $\chi_0 = 1$ is optimal, it is feasible. Note also that if an infeasibly high tax rate τ is required, the government's continuation value will be lower than if such a tax rate were feasible, and hence if the tax rate associated with a $\chi_0 = 1$ solution is feasible, that solution is optimal even without the requirement that τ be implementable.

Hence, we conjecture that $\chi_0 = 1$, and write the Lagrangian as

$$J_0(w_0, \theta_0) = \max_{\tau \in [0, \varphi(1-\delta)]} \max_{m_t \in \mathcal{M}(\tau)} R^{-1} \int_0^1 \{d_0(\theta_1, \theta_1) + \tau_0(\theta_1, \theta_1) + U_1(w_1(\theta_1, \theta_1), \theta_1; \tau)\} \mu(\theta_1|\theta_0) d\theta_1,$$

noting that in the solution to the full information problem,

$$U_1(w_1, \theta_1; \tau) = V_1(w_1, \theta_1; \tau) + \tau \max\{w_1 - \bar{w}(\theta_1), 0\}.$$

By Proposition 5, the optimal mechanism uses either maximal or first-best capital, and sets taxes proportional to dividends. By Proposition 6, it can be implemented as a dividend tax.

Under our assumptions for the static model, capital is first-best and the top type is constrained. Consequently,

$$\tau_0(\theta_1, \theta_1) = \tau \max\{\pi(k^*(\theta_0), \theta_0) + Rw_0 - w_1(\theta_1, \theta_1), 0\}$$

The funds raised are

$$\begin{aligned} R^{-1} \int_0^1 \tau_0(\theta_1, \theta_1) \mu(\theta_1|\theta_0) d\theta_1 + R^{-1}\tau \int_0^1 \max\{w_1(\theta_1, \theta_1) - \bar{w}(\theta_1), 0\} \mu(\theta_1|\theta_0) d\theta_1 = \\ R^{-1}\tau \int_0^1 \max\{\pi(k^*(\theta_0), \theta_0) + Rw_0 - \bar{w}(\theta_1), 0\} \mu(\theta_1|\theta_0) d\theta_1. \end{aligned}$$

By Assumption 1 (in particular, Equation (7)), this is a positive quantity, and hence there is a non-empty interval of $B_0 + G\left(1 + R^{-1}\right)$ with a positive, feasible dividend tax.

A.5 Proof of Proposition 3

In this subsection, we prove Proposition 3. First, under our conjecture, using Lemma 2, Proposition 4, and Proposition 5, we have shown that the optimal mechanism assigns zero dividends and taxes to the top type, and taxes proportional to the dividends paid to all other types, given a particular w_t, θ_t, with the same constant of proportionality,

$$\tau_t\left(w, \theta_t, \theta_{t+1}, \theta_{t+1}\right) = \frac{\tau}{1-\tau} d_t\left(w, \theta_t, \theta_{t+1}, \theta_{t+1}\right).$$

In addition, the capital level chosen is either maximal or first-best. The implementation claim follows from Proposition 6.

To complete the proof, we must show that the functions $U_t\left(w, \theta\right)$ and $V_t\left(w, \theta\right)$ have the claimed properties. Combining the result of the aforementioned propositions,

$$\begin{aligned} U_t\left(w, \theta\right) &= R^{-1} \int_0^1 \left\{d_t(w, \theta, \theta', \theta') + \tau_t\left(w, \theta, \theta', \theta'\right) + U_{t+1}\left(w_{t+1}\left(w, \theta, \theta', \theta'\right), \theta'\right)\right\} \mu\left(\theta'|\theta\right) d\theta' \\ &= R^{-1} \int_0^1 U_{t+1}\left(\pi\left(\min\left\{w \frac{R}{R - \varphi\left(1-\delta\right)}, k^*\left(\theta\right)\right\}, \theta\right) + Rw, \theta'\right) \mu\left(\theta'|\theta\right) d\theta'. \end{aligned}$$

Observe that $U_t\left(0, \theta\right) = 0$ and that $U_t\left(w, \theta\right)$ inherits the concavity in wealth from U_{t+1}. Lipschitz continuity in θ follows from the differentiability of the production function (and hence of $k^*\left(\theta\right)$) and the Lipschitz continuity of $\mu\left(\theta'|\theta\right)$. If $w\frac{R}{R-\varphi(1-\delta)} < k^*\left(\theta\right)$, then $U_{t,w^+}\left(w, \theta\right) > 1$. If $w\frac{R}{R-\varphi(1-\delta)} \geq k^*\left(\theta\right)$ but $\pi\left(k^*\left(\theta\right), \theta\right) + Rw < \sup_{\theta' : \mu(\theta'|\theta) > 0} \bar{w}\left(\theta'\right)$, then $U_{t,w^+}\left(w, \theta\right) > 1$. Otherwise, $U_{t,w^+}\left(w, \theta\right) = 1$ as required. Hence, we have

$$\bar{w}\left(\theta\right) = \max\left\{\frac{R - \varphi\left(1-\delta\right)}{R} k^*\left(\theta\right), \sup_{\theta' : \mu(\theta'|\theta) > 0} R^{-1}\left(\bar{w}\left(\theta'\right) - \pi\left(k^*\left(\theta\right), \theta\right)\right)\right\}.$$

Note that if $w > \bar{w}\left(\theta\right)$, we have $U_{t,w\theta}\left(w, \theta\right) = 0$. If $w < \bar{w}\left(\theta\right)$, then $U_{t,w\theta}\left(w, \theta\right) > 0$ by the fact that $f_{k\theta}\left(k, \theta\right) > 0$ for $k < k^*\left(\theta\right)$ (if $w\frac{R}{R-\varphi(1-\delta)} < k^*\left(\theta\right)$) and the fact that $\mu\left(\theta'|\theta\right)$ strictly first-order stochastically dominates $\mu\left(\theta'|\theta''\right)$ for $\theta > \theta'$.

The same argument applies essentially unmodified to $V_t\left(w, \theta\right)$, and hence both functions have the claimed properties. Lemma 6 proves the corresponding results for exiting firms.

Observe that, due the conjecture, $V_{t+1}\left(w, \theta\right) \geq V_{t+1}(w, 0)$ and $U_{t+1}\left(w, \theta\right) \geq U_{t+1}(w, 0)$ and hence the government never wants a firm to voluntarily exit, and firms never choose to voluntarily exit, assuming that the no-default constraints are satisfied.

To conclude, we must verify that our conjectured value function satisfies the Bellman equation. Plugging in the conjectured function,

$$
\begin{aligned}
J_t(B_t, \mu_t) = {} & R^{-1} \int_0^\infty \int_{0^+}^1 \int_0^1 d_t^* \left(w, \theta, \theta'\right) \mu\left(\theta'|\theta\right) \mu_t\left(w, \theta\right) d\theta' d\theta dw \\
& + R^{-1} \int_0^\infty d_t^*(w,0)\mu_t(w,0)dw \\
& + R^{-1} \left\{ J_E - \frac{G}{1-R^{-1}} - B_{t+1} + \int_0^\infty \int_0^1 U\left(w, \theta\right) \mu_{t+1}\left(w, \theta\right) d\theta dw \right\}.
\end{aligned}
$$

Plugging in the debt evolution, given by Equation (16),

$$
\begin{aligned}
J_t(B_t, \mu_t) = {} & R^{-1} \int_0^\infty \int_{0^+}^1 \int_0^1 d_t^* \left(w, \theta, \theta'\right) \mu\left(\theta'|\theta\right) \mu_t\left(w, \theta\right) d\theta' d\theta dw \\
& + R^{-1} \int_0^\infty d_t^*(w,0)\mu_t(w,0)dw \\
& + R^{-1} \left\{ J_E - \frac{G}{1-R^{-1}} + \int_0^\infty \int_0^1 U\left(w, \theta\right) \mu_{t+1}\left(w, \theta\right) d\theta dw \right\} \\
& - B_t - G + R^{-1} \int_0^\infty \int_{0^+}^1 \int_0^1 \tau_t^* \left(w, \theta, \theta'\right) \mu\left(\theta'|\theta\right) \mu_t\left(w, \theta\right) d\theta' d\theta dw \\
& + R^{-1} \int_0^\infty \tau_t^*(w,0)\mu_t(w,0)dw.
\end{aligned}
$$

Plugging in the population evolution, given by Equation (15), and doing some simplification

$$
\begin{aligned}
J_t(B_t, \mu_t) = {} & R^{-1} \int_0^\infty \int_{0^+}^1 \int_0^1 (d_t^* \left(w, \theta, \theta'\right) + \tau_t^* \left(w, \theta, \theta'\right)) \mu\left(\theta'|\theta\right) \mu_t\left(w, \theta\right) d\theta' d\theta dw \\
& + R^{-1} \int_0^\infty \{d_t^*(w,0) + \tau_t^*(w,0)\}\mu_t(w,0)dw \\
& + R^{-1} J_E + R^{-1} \int_0^\infty \int_0^1 U(w', \theta') e_t(w', \theta') d\theta' dw' \\
& + R^{-1} \int_0^\infty \int_0^1 \int_0^\infty \int_{0^+}^1 U(w', \theta') \delta(w_{t+1}^* \left(w, \theta, \theta'\right) - w') \mu\left(\theta'|\theta\right) \mu_t\left(w, \theta\right) d\theta dw d\theta' dw' \\
& + R^{-1} \int_0^\infty \int_0^\infty U(w', 0) \delta(w_{t+1}^*(w,0) - w') \mu_t(w,0) dw dw' \\
& - B_t - G \left(1 + \frac{R^{-1}}{1-R^{-1}}\right)
\end{aligned}
$$

which further simplifies to

$$
\begin{aligned}
J_t(B_t,\mu_t) = R^{-1}&\int_0^\infty\int_{0^+}^1\int_0^1 (d_t^*\left(w,\theta,\theta'\right)+\tau_t^*\left(w,\theta,\theta'\right)+U(w_{t+1}^*\left(w,\theta,\theta'\right),\theta'))\mu\left(\theta'|\theta\right)\mu_t\left(w,\theta\right)d\theta' d\theta dw \\
&+R^{-1}\int_0^\infty \{d_t^*(w,0)+\tau_t^*(w,0)+U(w_{t+1}^*(w,0),0)\}\mu_t(w,0)dw \\
&+R^{-1}J_E+R^{-1}\int_0^\infty\int_0^1 U(w',\theta')e_t(w',\theta')d\theta' dw' \\
&-B_t-\frac{G}{1-R^{-1}}.
\end{aligned}
$$

Defining

$$
J_E=\frac{R^{-1}}{1-R^{-1}}\int_0^\infty\int_0^1 U(w',\theta')e_t(w',\theta')d\theta' dw'
$$

and noting that $e_t(w',\theta')$ is time-invariant, and using the equations

$$
U\left(w,\theta\right)=\int_0^1 (d_t^*\left(w,\theta,\theta'\right)+\tau_t^*\left(w,\theta,\theta'\right)+U(w_{t+1}\left(w,\theta,\theta'\right),\theta'))\mu\left(\theta'|\theta\right)d\theta',
$$

$$
U(w,0)=R^{-1}\left\{d_t^*(w,0)+\tau_t^*(w,0)+U(w_{t+1}^*(w,0),0)\right\},
$$

the conjecture for the form of the function $J(B,\mu)$ is verified.

Lastly, we must show that the transversality and No-Ponzi conditions hold. Observe, however, that under the dividend tax implementation, we have

$$
\begin{aligned}
V\left(w,\theta\right) &= \bar{V}\left(w,\theta;\tau\right)=\left(1-\tau\right)\bar{V}(w,\theta;0) \\
U\left(w,\theta\right) &= \frac{1}{1-\tau}V\left(w,\theta\right)=\bar{V}(w,\theta;0).
\end{aligned}
$$

Thus, it must the case that the net present value of taxes collected is

$$
\frac{\tau}{1-\tau}\int_0^\infty\int_0^1\bar{V}\left(w,\theta;\tau\right)\mu_0\left(w,\theta\right)d\theta dw+\frac{\tau}{1-\tau}\frac{1}{R-1}\int_0^\infty\int_0^1\bar{V}\left(w,\theta;\tau\right)\bar{e}\left(w,\theta;\tau\right)d\theta dw,
$$

which is strictly positive by Assumption 2. It follows that there is a non-empty interval of $B+G\frac{1}{1-R^{-1}}$ such that

$$
\frac{\tau}{1-\tau}\int_0^\infty\int_0^1\bar{V}\left(w,\theta;\tau\right)\mu_0\left(w,\theta\right)d\theta dw+\frac{\tau}{1-\tau}\frac{1}{R-1}\int_0^\infty\int_0^1\bar{V}\left(w,\theta;\tau\right)\bar{e}\left(w,\theta;\tau\right)d\theta dw=B_0+G\frac{1}{1-R^{-1}}
$$

for $\tau\in[0,\bar{\tau})$. It follows immediately that if this equation holds, the No-Ponzi condition holds with equality.

Observe that, under the optimal mechanism, we always have $w_{t+1}\left(w_t,\theta_t,\theta_{t+1}\right)\leq\bar{w}\left(1\right)$, and hence $\mu_t(w, \theta)$ will have bounded support for all $t > 1$. This implies that

$$\lim_{t \to \infty} R^{-t} \int_0^\infty \int_0^1 U(w, \theta)\, \mu_t(w, \theta)\, d\theta dw = 0,$$

which proves the transversality condition, concluding the proof.

B Auxiliary Results

Lemma 2. *[Same capital/same allocation] In the mechanism design problem (C.3), the financing/investment and dividend/taxes constraints bind. Moreover, for any type types $\tilde{\theta}_{t+1}$ and $\hat{\theta}_{t+1}$ with $k_t(w_t, \theta_t, \tilde{\theta}_{t+1}) = k_t(w_t, \theta_t, \hat{\theta}_{t+1})$ in an optimal allocation, it is without loss of generality to suppose that $d_t(w_t, \theta_t, \hat{\theta}_{t+1}, \theta'_{t+1}) = d_t(w_t, \theta_t, \tilde{\theta}_{t+1}, \theta'_{t+1})$ for all $w_t, \theta_t, \theta_{t+1}, \theta'_{t+1}$, and likewise for $\tau_t(\cdot)$ and $w_{t+1}(\cdot)$.*

Proof. Suppose otherwise: that the financing/investment constraint was slack for some types. By substituting debt for taxes one-to-one, it is possible to weakly increase the objective function and leave all other constraints unchanged. Similarly, if the dividend/taxes constraint is slack, it would be possible to raise taxes and weakly improve welfare.

Next, we rewrite the problem in terms of dividends, continuation wealth, and capital. Note that the dividend limit and no-default constraint together imply weakly positive continuation wealth. The reformulated problem corresponds to

$$\max_{d_t(\cdot), w_{t+1}(\cdot), k_t(\cdot)} R^{-1} \int_0^1 \{(1-\chi_t) d_t(w_t, \theta_t, \theta_{t+1}, \theta_{t+1}) + U(w_{t+1}(w_t, \theta_t, \theta_{t+1}, \theta_{t+1}), \theta_{t+1})\} \mu(\theta_{t+1} | \theta_t)\, d\theta_{t+1}$$
$$+ R^{-1} \chi_t \int_0^1 \{\pi(k_t(w_t, \theta_t, \theta_{t+1}), \theta_t) + R w_t - w_{t+1}(w_t, \theta_t, \theta_{t+1}, \theta_{t+1})\} \mu(\theta_{t+1} | \theta_t)\, d\theta_{t+1},$$

subject to the constraints (no default, upper limit on dividends, positive taxes)

$$w^D(k_t(w_t, \theta_t, \theta'_{t+1}), \theta_t) \leq d_t(w_t, \theta_t, \theta'_{t+1}, \theta''_{t+1}) + w_{t+1}(w_t, \theta_t, \theta'_{t+1}, \theta''_{t+1}),\ \forall w_t, \theta_t, \theta'_{t+1}, \theta''_{t+1},$$
$$d_t(w_t, \theta_t, \theta'_{t+1}, \theta''_{t+1}) \leq w^D(k_t(w_t, \theta_t, \theta'_{t+1}), \theta_t),\ \forall w_t, \theta_t, \theta'_{t+1}, \theta''_{t+1},$$
$$w_{t+1}(w_t, \theta_t, \theta'_{t+1}, \theta''_{t+1}) + d_t(w_t, \theta_t, \theta'_{t+1}, \theta''_{t+1}) \leq \pi(k_t(w_t, \theta_t, \theta'_{t+1}), \theta_t) + R w_t,\ \forall w_t, \theta_t, \theta'_{t+1}, \theta''_{t+1},$$

as well as (dividend/taxes IC, financing/investment IC, blocked dividend)

$$d_t(w_t, \theta_t, \theta'_{t+1}, \theta''_{t+1}) + V(w_{t+1}(w_t, \theta_t, \theta'_{t+1}, \theta''_{t+1}), \theta_{t+1}) \leq d_t(w_t, \theta_t, \theta'_{t+1}, \theta_{t+1}) + V(w_{t+1}(w_t, \theta_t, \theta'_{t+1}, \theta_{t+1}), \theta_{t+1}),\ \forall w_t, \theta_t, \theta_{t+1}, \theta'_{t+1}, \theta''_{t+1},$$
$$d_t(w_t, \theta_t, \theta'_{t+1}, \theta_{t+1}) + V(w_{t+1}(w_t, \theta_t, \theta'_{t+1}, \theta_{t+1}), \theta_{t+1}) \leq d_t(w_t, \theta_t, \theta_{t+1}, \theta_{t+1}) + V(w_{t+1}(w_t, \theta_t, \theta_{t+1}, \theta_{t+1}), \theta_{t+1}),\ \forall w_t, \theta_t, \theta_{t+1}, \theta'_{t+1},$$
$$V\left(w^D(k_t(w_t, \theta_t, \theta'_{t+1}), \theta_t), \theta_{t+1}\right) \leq d_t(w_t, \theta_t, \theta_{t+1}, \theta_{t+1}) + V(w_{t+1}(w_t, \theta_t, \theta_{t+1}, \theta_{t+1}), \theta_{t+1}),\ \forall w_t, \theta_t, \theta_{t+1}, \theta'_{t+1}$$

as well as non-negativity constraints and the lower bound on capital,

$$d_t(w_t, \theta_t, \theta'_{t+1}, \theta''_{t+1}) \geq 0,\ k_t(w_t, \theta_t, \theta'_{t+1}) \geq w_0.$$

Now suppose that in a optimal policy, there were two types, $\tilde{\theta}$ and $\hat{\theta}$, such that

$k_t\left(w_t,\theta_t,\tilde{\theta}\right) = k_t\left(w_t,\theta_t,\hat{\theta}\right)$ but, for some θ, $d_t\left(w_t,\theta_t,\tilde{\theta},\theta\right) \neq d_t\left(w_t,\theta_t,\hat{\theta},\theta\right)$ (and thus, by the ICs, $w_{t+1}\left(w_t,\theta_t,\tilde{\theta},\theta\right) \neq w_{t+1}\left(w_t,\theta_t,\hat{\theta},\theta\right)$). Observe that it would be feasible to switch all of the allocations of type $\tilde{\theta}$ to those of $\hat{\theta}$, $d_t\left(w_t,\theta_t,\tilde{\theta},\theta\right) \to d_t\left(w_t,\theta_t,\hat{\theta},\theta\right)$ and $w_{t+1}\left(w_t,\theta_t,\tilde{\theta},\theta\right) \to w_{t+1}\left(w_t,\theta_t,\hat{\theta},\theta\right)$, or vice versa. Doing so would relax (weakly) the financing/investment ICs and would still satisfy the dividend/taxes ICs and the blocked-dividend no default constraint. Because the capital level is the same, the no-default, upper limit on dividends, and taxes positive constraints would still be satisfied.

Consequently, for the objective function, the values for the reports $(\hat{\theta},\hat{\theta})$ and $(\hat{\theta},\tilde{\theta})$ must be the same, and likewise for $(\tilde{\theta},\tilde{\theta})$ and $(\tilde{\theta},\hat{\theta})$. Therefore, it is without loss of generality to switch one to the other. $\square$

Lemma 3. *[Monotone allocations] It is without loss of generality, in the capital sub-problem (C.5), to restrict attention to "monotone" allocations that have dividends weakly decreasing in type and continuation wealth weakly increasing in type, treating "about to exit" firms as the lowest type.*

Proof. This result is standard (e.g., Fudenberg and Tirole (1991)), except that the cross-partial of our firm continuation value, $V_{w\theta}$, is only weakly positive. Formally,

$$J\left(\hat{\mu},w_t,\theta_t,k,y_1,V^*\right) = \max_{d_t(\theta),y_t(\theta)} \frac{R^{-1}}{\int_0^1 \hat{\mu}\left(\theta;k\right)d\theta} \int_0^1 \left\{ \begin{array}{c} \chi_t(\pi\left(k,\theta_t\right)+Rw_t) - \chi_t y_t\left(\theta\right) \\ +U\left(y_t\left(\theta\right)-d_t\left(\theta\right),\theta\right)+d_t\left(\theta\right) \end{array} \right\} \hat{\mu}\left(\theta;k\right)d\theta,$$

subject to the constraints

$$\begin{aligned} w^D\left(k,\theta_t\right) &\leq y_t\left(\theta\right), \ \forall\theta && \text{(No Default)} \\ d_t\left(\theta\right) &\leq w^D\left(k,\theta_t\right), \ \forall\theta && \text{(Upper Limit on Dividends)} \\ 0 &\leq d_t\left(\theta\right), \ \forall\theta && \text{(Non-Negative Dividends)} \\ y_t\left(\theta\right) &\leq \pi\left(k,\theta_t\right)+Rw_t, \ \forall\theta, && \text{(Non-Negative Taxes)} \end{aligned}$$

and the incentive constraints

$$\begin{aligned} d_t\left(\theta'\right)+V\left(y_t\left(\theta'\right)-d_t\left(\theta'\right),\theta\right) &\leq d_t\left(\theta\right)+V\left(y_t\left(\theta\right)-d_t\left(\theta\right),\theta\right), \ \forall\theta,\theta' && \text{(Dividend/Taxes IC)} \\ V\left(w^D,\theta\right) &\leq d_t\left(\theta\right)+V\left(y_t\left(\theta\right)-d_t\left(\theta\right),\theta\right), \ \forall\theta, && \text{(Blocked Dividend No Default)} \\ d_t\left(\theta\right)+V\left(y_t\left(\theta\right)-d_t\left(\theta\right),\theta\right) &\leq V^*\left(\theta\right), \ \forall\theta \text{ s.t. } \hat{\mu}\left(\theta;k\right)=0, && \text{(Financing/Investment IC)} \end{aligned}$$

and the constraints on the top type allocation, $y_t\left(1\right)=y_1$.

The dividend/taxes IC constraint for types $\theta > \theta'$ are

$$\begin{aligned} d_t\left(\theta\right)-d_t\left(\theta'\right) &\geq V\left(w_{t+1}\left(\theta'\right),\theta\right)-V\left(w_{t+1},\theta\right) \\ d_t\left(\theta\right)-d_t\left(\theta'\right) &\leq V\left(w_{t+1}\left(\theta'\right),\theta'\right)-V\left(w_{t+1}\left(\theta\right),\theta'\right) \end{aligned}$$

and therefore

$$V\left(w_{t+1}\left(\theta'\right),\theta'\right)-V\left(w_{t+1}\left(\theta\right),\theta'\right) \geq V\left(w_{t+1}\left(\theta'\right),\theta\right)-V\left(w_{t+1}\left(\theta\right),\theta\right).$$

Suppose that $w_{t+1}(\theta') > w_{t+1}(\theta)$. Then we would have

$$\int_{w_{t+1}(\theta)}^{w_{t+1}(\theta')} V_w\left(w,\theta'\right) dw \geq \int_{w_{t+1}(\theta)}^{w_{t+1}(\theta')} V_w\left(w,\theta\right) dw,$$

which can be expressed as

$$0 \geq \int_{w_{t+1}(\theta)}^{w_{t+1}(\theta')} \int_{\theta'}^{\theta} V_{w\theta}\left(w,\theta''\right) d\theta'' dw.$$

This would require, by $V_{w\theta}(w,\theta'') > 0$ if $w < \bar{w}(\theta)$, that $w \geq \bar{w}(\theta'')$ for all the domain of integration, a constraint that is tightest at $w_{t+1}(\theta) \geq \bar{w}(\theta)$. In this case, $V_w(w,\theta') = V_w(w,\theta) = 1-\tau$, and we would have

$$d_t(\theta) - d_t\left(\theta'\right) = (1-\tau)\left(w_{t+1}\left(\theta'\right) - w_{t+1}(\theta)\right),$$

and indifference for the firms between these two types.Note that in this case we must have $d_t(\theta) > d_t(\theta')$ and

$$d_t(\theta) + w_{t+1}(\theta) < w_{t+1}\left(\theta'\right) + d_t\left(\theta'\right),$$

which is $y_t(\theta') > y_t(\theta)$. Imagine first that $\hat{\mu}(\theta;k) = 0$. In this case, setting the allocation of type θ to the allocation of type θ' has no effect on the objective or budget/positive taxes constraints, and removes a global IC, and hence is without loss of generality. Likewise, if $\hat{\mu}(\theta';k) = 0$, setting the allocation of type θ' to the allocation of type θ is w.l.o.g.

Now suppose that neither of these conditions hold, and define weighted averages:

$$w_A^R = \frac{\hat{\mu}(\theta;k)\, w_{t+1}(\theta) + \hat{\mu}(\theta';k) w_{t+1}(\theta')}{\hat{\mu}(\theta;k) + \hat{\mu}(\theta';k)},$$

and

$$d_A = \frac{\hat{\mu}(\theta;k)\, d_t(\theta) + \hat{\mu}(\theta';k) d_t(\theta')}{\hat{\mu}(\theta;k) + \hat{\mu}(\theta';k)},$$

and define $y_A = w_A^R + d_A$.

Consider an alternative allocation which sets $y_t(\theta) = y_t(\theta') = y_A$ and $d_t(\theta) = d_t(\theta') = d_A$, leaving all other types unchanged. Note that, because all of these types are unconstrained at all of the relevant wealth levels, for all $\theta'' \in [\theta', \theta]$,

$$-\chi_t y_t\left(\theta''\right) + U\left(y_t\left(\theta''\right) - d_t\left(\theta''\right), \theta''\right) + d_t\left(\theta''\right) = U\left(\bar{w}(\theta), \theta''\right) - \chi_t y_t\left(\theta''\right) + d_t\left(\theta''\right) + \left(y_t\left(\theta_t''\right) - d_t\left(\theta''\right) - \bar{w}\left(\theta''\right)\right).$$

Hence the objective function is linear in y_t and d_t and it follows that this alternative allocation achieves the same utility. Likewise, by construction, the alternative allocation leaves the budget/positive taxes constraint unchanged.

Note also that $w_{t+1}(\theta') > w_A$ and

$$d_t(\theta') - d_A = -(1-\tau)\left(w_{t+1}(\theta') - w_A\right).$$

For any type θ'', we must have

$$\begin{gathered} d_t(\theta') + V\left(w_{t+1}(\theta'), \theta''\right) - d_A - V_1\left(w_A, \theta_1''\right) = \\ V\left(w_{t+1}(\theta'), \theta''\right) - V\left(w_A, \theta''\right) - (1-\tau)\left(w_{t+1}(\theta') - w_A\right) \geq 0. \end{gathered}$$

Therefore, if the global IC was satisfied in the original allocation,

$$d_t(\theta'') + V\left(w_{t+1}(\theta''), \theta''\right) \geq d_t(\theta') + V\left(w_{t+1}(\theta'), \theta''\right),$$

it will be satisfied in the new allocation,

$$d_t(\theta'') + V\left(w_{t+1}(\theta''), \theta''\right) \geq d_A + V\left(w_A, \theta''\right).$$

Lastly, the deviation to default constraint can be written as

$$d_t(\theta) + (1-\tau)\left(w_{t+1}(\theta) - \bar{w}(\theta)\right) + V\left(\bar{w}(\theta), \theta\right) \geq V\left(w^D(k, \theta_t), \theta\right)$$

and therefore

$$d_A + (1-\tau)\left(w_A - \bar{w}(\theta)\right) + V\left(\bar{w}(\theta), \theta\right) \geq V\left(w^D(k, \theta_t), \theta\right),$$

and the constraint is still satisfied. The same argument applies for type θ_1', and for the financing/investment IC constraint. It follows that the new allocation is at least as good, and has weakly monotone increasing levels of continuation wealth after repayment. By the IC constraint, this implies that dividends must be weakly decreasing in type. $\square$

Lemma 4. *[Local IC's] Any allocation that is monotone (in the sense of Lemma 3) and satisfies the global IC conditions satisfies the local IC condition for all $\theta \in (0,1)$:*

$$\frac{d}{d\theta}\left[d_t(\theta) + V(w_{t+1}(\theta), \theta)\right] = V_\theta(w_{t+1}(\theta), \theta).$$

Proof. The result is standard (e.g., Fudenberg and Tirole (1991)), except that the cross-partial of our firm continuation value, $V_{w\theta}$, is only weakly positive, and the function has a kink in wealth at $\bar{w}(\theta)$. Formally, consider first the possibility of jumps in the level of continuation wealth/dividends. Note that both of these are bounded and monotone, and hence converge to some limit and are differentiable almost everywhere. Suppose for some $\theta_1 > \theta_1'$,

$$\lim_{\theta_1' \to \theta_1^-} w_0^R(\theta_1') = w_0^R(\theta_1^-) < w_0^R(\theta_1).$$

By the global IC, we must have

$$d_0\left(\theta_1'\right) - d_0\left(\theta_1\right) \le V_1\left(w_0^R\left(\theta_1\right), \theta_1\right) - V_1\left(w_0^R\left(\theta_1'\right), \theta_1\right)$$
$$d_0\left(\theta_1'\right) - d_0\left(\theta_1\right) \ge V_1\left(w_0^R\left(\theta_1\right), \theta_1'\right) - V_1\left(w_0^R\left(\theta_1'\right), \theta_1'\right)$$

and hence

$$\lim_{\theta_1' \to \theta_1^-} d_0\left(\theta_1'\right) - d_0\left(\theta_1\right) = V_1\left(w_0^R\left(\theta_1\right), \theta_1\right) - V_1\left(w_0^R\left(\theta_1^-\right), \theta_1\right) > 0.$$

Hence, jumps in $w_0^R\left(\theta_1\right)$ imply jumps in $d_0\left(\theta_1\right)$, and the argument can be used in reverse to show that jumps in dividends imply jumps in continuation wealth, as well. □

We can rewrite the IC, using the Lipschitz continuity of V with respect to θ as

$$\begin{aligned} d_t\left(\theta_1'\right) + V\left(w_{t+1}\left(\theta_1'\right), \theta_1\right) &= d_t\left(\theta_1'\right) + V\left(w_{t+1}\left(\theta_1'\right), \theta_1'\right) + \int_{\theta_1'}^{\theta_1} V_{1,\theta}\left(w_{t+1}\left(\theta_1'\right), \theta\right) d\theta \\ &\le d_t\left(\theta_1\right) + V\left(w_{t+1}\left(\theta_1\right), \theta_1\right) \end{aligned}$$

and

$$\begin{aligned} d_t\left(\theta_1\right) + V\left(w_{t+1}\left(\theta_1\right), \theta_1'\right) &= d_t\left(\theta_1\right) + V\left(w_{t+1}\left(\theta_1\right), \theta_1\right) - \int_{\theta_1'}^{\theta_1} V_{\theta}\left(w_{t+1}\left(\theta_1\right), \theta\right) d\theta \\ &\le d_t\left(\theta_1'\right) + V\left(w_{t+1}\left(\theta_1'\right), \theta_1'\right). \end{aligned}$$

Hence we must have

$$\begin{aligned} \lim_{\theta_1' \to \theta_1} \frac{\int_{\theta_1}^{\theta_1'} V_{\theta}\left(w_{t+1}\left(\theta_1'\right), \theta\right) d\theta}{\theta_1' - \theta_1} &\ge \lim_{\theta_1' \to \theta_1^+} \frac{d_t\left(\theta_1'\right) + V\left(w_{t+1}\left(\theta_1'\right), \theta_1'\right) - d_t\left(\theta_1\right) - V\left(w_{t+1}\left(\theta_1\right), \theta_1\right)}{\theta_1' - \theta_1} \\ &\ge \lim_{\theta_1' \to \theta_1} \frac{\int_{\theta_1}^{\theta_1'} V_{\theta}\left(w_{t+1}\left(\theta_1\right), \theta\right) d\theta}{\theta_1' - \theta_1} \end{aligned}$$

and

$$\begin{aligned} \lim_{\theta_1' \to \theta_1^-} \frac{\int_{\theta_1'}^{\theta_1} V_{1,\theta}\left(w_{t+1}\left(\theta_1\right), \theta\right) d\theta}{\theta_1 - \theta_1'} &\ge \lim_{\theta_1' \to \theta_1^-} \frac{d_t\left(\theta_1\right) + V\left(w_{t+1}\left(\theta_1\right), \theta_1\right) - d_t\left(\theta_1'\right) - V\left(w_{t+1}\left(\theta_1'\right), \theta_1'\right)}{\theta_1 - \theta_1'} \\ &\ge \lim_{\theta_1' \to \theta_1^-} \frac{\int_{\theta_1'}^{\theta_1} V_{\theta}\left(w_{t+1}\left(\theta_1'\right), \theta\right) d\theta}{\theta_1 - \theta_1'} \end{aligned}$$

Anywhere $w_{t+1}\left(\theta_1\right)$ is left continuous,

$$\lim_{\theta_1' \to \theta_1^-} \frac{\int_{\theta_1'}^{\theta_1} V_{1,\theta}\left(w_{t+1}\left(\theta_1\right), \theta\right) d\theta}{\theta_1 - \theta_1'} = V_{\theta}\left(w_{t+1}\left(\theta_1\right), \theta_1\right),$$

and anywhere it is right-continuous,

$$\lim_{\theta_1' \to \theta_1} \frac{\int_{\theta_1}^{\theta_1'} V_{1,\theta}\left(w_{t+1}\left(\theta_1\right),\theta\right) d\theta}{\theta_1' - \theta_1} = V_\theta\left(w_{t+1}\left(\theta_1\right),\theta_1\right),$$

and therefore $d_t\left(\theta_1\right) + V\left(w_{t+1}\left(\theta_1\right),\theta_1\right)$ is continuous, and differentiable anywhere $w_{t+1}\left(\theta_1\right)$ is continuous, which is almost everywhere. Moreover, observe by the properties of V that

$$V_\theta\left(\bar{w}\left(\theta\right),\theta\right) \geq V_\theta\left(w_{t+1}\left(\theta_1'\right),\theta\right) \geq V_\theta\left(0,\theta\right),$$

and hence we must have

$$\lim_{\theta_1' \to \theta_1} \frac{\int_{\theta_1}^{\theta_1'} V_\theta\left(w_{t+1}\left(\theta_1'\right),\theta\right) d\theta}{\theta_1' - \theta_1} \leq V_\theta\left(\bar{w}\left(\theta_1\right),\theta_1\right)$$

and

$$\lim_{\theta_1' \to \theta_1^-} \frac{\int_{\theta_1'}^{\theta_1} V_{1,\theta}\left(w_{t+1}\left(\theta_1'\right),\theta\right) d\theta}{\theta_1 - \theta_1'} \geq V_{1,\theta}\left(0,\theta_1\right).$$

Therefore, the function $d_t\left(\theta_1\right) + V\left(w_{t+1}\left(\theta_1\right),\theta_1\right)$ is Lipschitz-continuous, implying absolute continuity, and hence

$$d_t\left(\theta_1\right) + V\left(w_{t+1}\left(\theta_1\right),\theta_1\right) = d_t\left(\theta_1'\right) + V\left(w_{t+1}\left(\theta_1'\right),\theta_1'\right) + \int_{\theta_1'}^{\theta_1} V_\theta\left(w_{t+1}\left(\theta\right),\theta\right) d\theta.$$

Lemma 5. *[Multiplier less than one] In the static model, if the multiplier χ_0 is less than one, no taxes will be collected.*

Proof. Consider the capital sub-problem for the static model. The problem is, ignoring the financing/investment ICs,

$$\max_{d_0(\theta_1)\geq 0, y_0(\theta_1)\geq 0} R^{-1}\int_0^1 \{d_0(w_0,\theta_0,\theta_1) + \chi_0\tau_0(w_0,\theta_0,\theta_1)\}\hat{\mu}(\theta_1;k)d\theta_1$$
$$+ R^{-1}\int_0^1 \left\{V_1(w_1(w_0,\theta_0,\theta_1),\theta_1;\tau) + \chi_0 R^{-1}\tau \max\left\{w_1\left(w_0,\theta_0,\theta_1\right) - \bar{w}\left(\theta_1\right),0\right\}\right\}\hat{\mu}(\theta_1;k)d\theta_1.$$

subject to the constraints

$$\begin{aligned}
w^D &\leq y_0\left(\theta_1\right),\ \forall\theta_1 && \text{(No Default)}\\
d_0\left(\theta_1\right) &\leq w^D\left(k,\theta_0\right),\ \forall\theta_1 && \text{(Upper Limit on Dividends)}\\
0 &\leq d_0\left(\theta_1\right),\ \forall\theta_1 && \text{(Non-Negativity of Dividends)}\\
y_0\left(\theta_1\right) &\leq \pi\left(k,\theta_0\right) + Rw_0,\ \forall\theta_1 && \text{(Debt/Taxes Positive)}
\end{aligned}$$

and the incentive constraints

$$d_0(\theta_1') + V_1\left(y_0\left(\theta_1'\right) - d_0(\theta_1'), \theta_1; \tau\right) \leq d_0\left(\theta_1\right) + V_1\left(y_0\left(\theta_1\right) - d_0\left(\theta_1\right), \theta_1; \tau\right), \ \forall \theta_1, \theta_1'.$$

$$d_0\left(\theta_1\right) + V_1\left(y_0\left(\theta_1\right) - d_0\left(\theta_1\right), \theta_1; \tau\right) \geq V_1(w^D, \theta_1; \tau).$$

Observe that when $\chi_0 < 1$, if the firm is constrained ($w_1 < \bar{w}\left(\theta\right)$), the objective is increasing in $y_0\left(\theta_1\right)$ and decreasing in $d_0\left(\theta_1\right)$. If the firm is unconstrained, the derivative with respect to $y_0\left(\theta_1\right)$ is

$$-\chi_0 + \left(1 - \tau + \tau\chi_0\right) = \left(1 - \tau\right)\left(1 - \chi_0\right) > 0$$

and with respect to $d_0\left(\theta_1\right)$ is

$$1 - \left(1 - \tau + \tau\chi_0\right) = \tau(1 - \chi_0) > 0.$$

Hence, it follows that the best possible allocation for a type (ignoring incentive compatibility) is

$$y_0\left(\theta_1\right) = \pi\left(k, \theta_0\right) + Rw_0$$

and

$$d_0\left(\theta_1\right) = \min(\max(\pi\left(k, \theta_0\right) + Rw_0 - \bar{w}\left(\theta_1\right), 0), w^D\left(k, \theta_0\right)).$$

Observe that these dividends are increasing in type, and that $w_1\left(\theta_1\right) = \bar{w}\left(\theta_1\right)$ for all dividend-payers with $d_0\left(\theta_1\right) < w^D\left(k, \theta_0\right)$. Hence, over this interval, by the assumption that $y_0\left(\theta_1\right)$ is constant and all types are at the kink, no type wants to deviate to a higher type or a lower type, and therefore the dividend/taxes IC constraint is satisfied.

In this case, there is no revenue raised initially, and

$$\begin{aligned} \tau(w_1\left(\theta_1\right) - \bar{w}\left(\theta_1\right)) &= \tau \max(\pi\left(k, \theta_0\right) + Rw_0 - \bar{w}\left(\theta_1\right) - w^D\left(k, \theta_0\right), 0) \\ &= \tau \max(Rw_0 - (R - \varphi\left(1 - \delta\right))k - \bar{w}\left(\theta_1\right), 0) \end{aligned}$$

By assumption, all types in the support of $\mu\left(\theta_1 | \theta_0\right)$, and hence $\hat{\mu}(\theta_1; k)$, satisfy

$$\varphi\left(1 - \delta\right) w_0 \leq \bar{w}\left(\theta_1\right),$$

and by $k \geq w_0$ it follows that no revenue is raised. It always follows that first-best capital, or maximum capital if first-best is infeasible, is optimal in the capital choice problem. Consequently, the financing/investment ICs do not bind and this solution is optimal.

However, this solution raises no revenue, and hence we must have $\chi_0 \geq 1$. $\square$

Lemma 6. *[Optimal policy for exiting types] The optimal policy for exiting types can be implemented by a constant dividend tax with* $\tau_t(w, 0) = \frac{\tau}{1-\tau} d_t(w, 0)$.

Proof. Consider the mechanism design problem for exiting types. Observing immediately that both financing/investment and dividend/taxes constraints bind, we have

$$d_t(w_t, 0) + w_{t+1}(w_t, 0) = Rw_t - \tau_t(w_t, 0).$$

Thus, the government is free to choose any feasible allocation (by $U(w, 0) = w$). To implement a tax rate

$$\tau_t(w_t, 0) = \frac{\tau}{1-\tau} d_t(w_t, 0),$$

we need for the no-default constraint

$$Rw_t - \frac{\tau}{1-\tau} d_t(w_t, 0) \geq (R - \varphi(1-\delta)) k_t(w_t, 0)$$

and for the blocked dividend constraint

$$Rw_t \geq (R - \varphi(1-\delta)) k_t(w_t, 0)$$

Choose $k_t(w_t, 0) = w_t$ relaxes these constraints as much as possible, and hence we must have

$$\frac{\tau}{1-\tau} d_t(w_t, 0) \leq \varphi(1-\delta) w_t$$

and the upper limit on dividends,

$$d_t(w_t, 0) \leq (R - \varphi(1-\delta)) w_t.$$

The firm value is

$$\begin{aligned} V(w, 0) &= R^{-1}\{d_t(w_t, 0) + (1-\tau)(Rw_t - \frac{1}{1-\tau} d_t(w_t, 0)\} \\ &= (1-\tau) w_t. \end{aligned}$$

Observe also that implementation via a dividend tax is immediate, by the firm's indifference with regards to paying dividends. Hence any dividend policy that satisfies the transversality condition is optimal– say,

$$d_t(w_t, 0) = \min\{\frac{1-\tau}{\tau} \varphi(1-\delta), R - \varphi(1-\delta)\} w_t.$$

□

Proposition 4. *[Optimal Policy in the relaxed capital sub-problem] Suppose that the arguments to the Capital Sub-Problem satisfy the following necessary conditions for feasibility:*

$$w^D(k, \theta_t) \leq y_1 \leq \pi(k, \theta_t) + Rw_t,$$

that the financing/investment IC has been relaxed ($V^*(\theta) = \infty$), and that $\chi_t = 1$. Then in the Relaxed

Capital Sub-Problem an optimal mechanism is characterized by, for all θ,

$$\begin{aligned}
d_t(\theta) - d_1 &= (1-\tau)\max\{y_1 - d_1 - w_{t+1}(\theta), 0\}, \\
w_{t+1}(\theta) &= \min\{y_1 - d_1, \max\{\bar{w}(\theta), w_{t+1}(0)\}\}, \\
w_{t+1}(0) &= \max\left\{y_1 - d_1 - \frac{w^D - d_1}{1-\tau}, \frac{w^D - y_1}{\tau} + y_1 - d_1\right\}, \\
d_t(\theta) + V(w_{t+1}(\theta), \theta) &= d_1 + V(y_1 - d_1, \theta), \\
d_1 &= \begin{cases} 0 & y_1 \leq \bar{w}(1) \\ \min\left\{\frac{1-\tau}{\tau}(\pi(k,\theta_t) + Rw_t - y_1), y_1 - \bar{w}(1), w^D(k,\theta_t)\right\} & y_1 > \bar{w}(1) \end{cases}.
\end{aligned}$$

Moreover, this allocation is also a solution to the Capital Sub-Problem with the financing/investment IC relaxed.

Proof. Observe first that it is without loss of generality to suppose that $w_{t+1}(\theta) < y_1 - d_1$ for all $\theta < \theta^H$, otherwise θ^H can be decreased without altering the objective function. Note also that for any types with the same allocation as the top type, the proposition holds.

We first consider the case in which $\theta^H > 0$, so some non-exiting types have an allocation different from the top type. In the $\chi_t = 1$ case, the objective function is

$$\begin{aligned}
J^R(\hat{\mu}, w_t, \theta_t, k, y_1, V^*) = \max_{\theta^H \in [0,1], w_{t+1}(\theta) \geq 0, y_t(\theta) \geq 0} & R^{-1} \int_{\theta^{H+}}^{1} \left\{ \begin{array}{c} \pi(k,\theta_t) + Rw_t \\ +U(y_1 - d_1, \theta) - y_1 + d_1 \end{array} \right\} \hat{\mu}(\theta; k)\, d\theta \\
& + R^{-1} \int_0^{\theta^H} \{\pi(k,\theta_t) + Rw_t + U(w_{t+1}(\theta), \theta) - w_{t+1}(\theta)\} \hat{\mu}(\theta; k)\, d\theta.
\end{aligned}$$

Because there are no constraints on $y_t(\theta)$ for $\theta \in (0, \theta^H)$, the local IC constraint can be thought of as definition. The FOC for increasing $w_{t+1}(\theta)$ for $\theta \in (0, \theta^H)$ is

$$(U_{w^+}(w_{t+1}(\theta), \theta) - 1)\hat{\mu}(\theta) = \nu_H(\theta),$$

where ν_H is the multiplier on the relaxed monotonicity constraint and U_{w^+} denotes the directional derivative in the increasing wealth direction. This multiplier does not bind, so it immediately follows that $w_{t+1}(\theta) \geq \bar{w}(\theta)$, and it is without loss of generality to suppose that $w_{t+1}(\theta) > \bar{w}(\theta)$. Applying this argument to the type θ^H, it follows by the continuity of $\bar{w}(\theta)$ that $w_{t+1}(\theta_H) \geq \bar{w}(\theta_H)$ as well.

Consequently, the local IC requires that (using $V_{\theta w} = 0$ along the integral), for all $\theta \in [0, \theta^H)$,

$$y_t(\theta) + V(w_{t+1}(\theta), \theta) - w_{t+1}(\theta) + \int_\theta^{\theta^H} V_\theta(y_1 - d_1, \theta')\, d\theta' = d_1 + V(y_1 - d_1, \theta)$$

and hence

$$y_t(\theta) + V(w_{t+1}(\theta), \theta) - w_{t+1}(\theta) = d_1 + V(y_1 - d_1, \theta),$$

which is

$$d_t(\theta) - d_1 = (1-\tau)(y_1 - d_1 - w_{t+1}(\theta)),$$

or

$$y_1 - y_t(\theta) = \tau(y_1 - d_1 - w_{t+1}(\theta)).$$

The local IC also directly implies that

$$y_t(\theta_H) + V(w_{t+1}(\theta_H), \theta_H) - w_{t+1}(\theta_H) = d_1 + V(y_1 - d_1, \theta_H),$$

and hence these properties hold for θ_H as well. It follows that the positive dividend and positive taxes constraints for these all of these types (from the Capital Sub-Problem) are satisfied, given a feasible (y_1, d_1).

Observing that

$$d_t(\theta) + V(w_{t+1}(\theta), \theta) - V(w^D(k, \theta_t), \theta) = d_1 + V(y_1 - d_1, \theta) - V(w^D(k, \theta_t), \theta),$$

if $y_1 - d_1 \geq w^D$, the deviation to default constraint is satisfied, and if $y_1 - d_1 < w^D$, then

$$\begin{aligned} d_1 + V(y_1 - d_1, \theta) - V\left(w^D(k, \theta_t), \theta\right) &= d_t(1) - \int_{y_1-d_1}^{w^D(k,\theta_t)} V_{w^+}(w, \theta)\, dw \\ &\geq d_t(1) - \int_{y_1-d_1}^{w^D(k,\theta_t)} V_{w^+}(w, 1) dw, \end{aligned}$$

and thus $d_t(\theta_1) + V(w_{t+1}(\theta_1), \theta_1) - V(w^D, \theta_1) \geq d_1 + V(y_1 - d_1, 1) - V(w^D, 1) \geq 0$. Hence, the deviation to default constraint (from the Capital Sub-Problem) is satisfied, assuming a feasible allocation for the top type.

Now consider the bottom type. We have, from earlier,

$$y_t(0) + V(w_{t+1}(0), 0) - w_{t+1}(0) = d_1 + V(y_1 - d_1, 0),$$

noting that this must also be true in the case where $\theta^H = 0$.

This simplifies (by $\bar{w}(0) = 0$) to

$$y_t(1) - \tau w_{t+1}(1) = y_t(0) - \tau w_{t+1}(0),$$

and hence for all $\theta \in (0, 1]$,

$$\begin{aligned} y_t(\theta) - \tau w_{t+1}(\theta) &= y_t(0) - \tau w_{t+1}(0) \\ d_t(\theta) + (1-\tau) w_{t+1}(\theta) &= d_t(0) + (1-\tau) w_{t+1}(0) \\ y_t(\theta) &\geq y_t(0) \\ d_t(\theta) &\leq d_t(0) \end{aligned}$$

implying that the no-default constraint and the dividend upper bound do not bind (from the

Capital Sub-Problem, for types $\theta \in (0, \theta^H]$). By these results, we have

$$\begin{aligned}\tau_t(\theta) &= \pi(k, \theta_t) + Rw_t - y_t(\theta) \\ &= \pi(k, \theta_t) + Rw_t - y_1 + \tau \max\{y_1 - d_1 - w_{t+1}(\theta), 0\},\end{aligned}$$

and also

$$\begin{aligned}U(w_{t+1}(\theta), \theta) &= U(y_1 - d_1, \theta) \\ d_t(\theta) &= d_1 + (1-\tau)\max\{y_1 - d_1 - w_{t+1}(\theta), 0\},\end{aligned}$$

and therefore

$$\begin{aligned}\pi(k, \theta_t) + Rw_t + U(w_{t+1}(\theta), \theta) - w_{t+1}(\theta) &= \\ \tau_t(\theta) + d_t(\theta) + U(w_{t+1}(\theta), \theta) &= \\ \pi(k, \theta_t) + Rw_t - y_1 + d_1 + U(y_1 - d_1, \theta)&.\end{aligned}$$

Any allocation satisfying $w_{t+1}(\theta) \geq \bar{w}(\theta)$ for all types $\theta \leq \theta^*$ is optimal (given allocations for the top and bottom types), so it is without loss of generality to choose a monotone one. We choose

$$w_{t+1}(\theta) = \max\{\bar{w}(\theta), w_{t+1}(0)\}.$$

Now consider the allocation of the bottom type. We must have $y_1 - d_1 \geq w_{t+1}(0) \geq \bar{w}(0) = 0$,

$$w^D(k, \theta_t) \geq d_t(0) = d_1 + (1-\tau)(y_1 - d_1 - w_{t+1}(0)),$$

$$y_1 - \tau(y_1 - d_1 - w_{t+1}(0)) = y_t(0) \geq w^D(k, \theta_t).$$

Observe that $y_1 - d_1 = w_{t+1}(0)$ is always feasible (given a feasible top type allocation), but the lowest possible value is determined by

$$w_{t+1}(0) \geq \max\{0, y_1 - d_1 - \frac{w^D(k, \theta_t) - d_1}{1-\tau}, \frac{w^D(k, \theta_t) - y_1}{\tau} + y_1 - d_1\}.$$

Note, however, that

$$\frac{w^D(k, \theta_t) - y_1}{\tau} + y_1 - d_1 < 0$$

implies

$$w^D(k, \theta_t) - d_1 - (1-\tau)(y_1 - d_1) < 0,$$

and hence

$$y_1 - d_1 - \frac{w^D(k, \theta_t) - d_1}{1-\tau} > 0.$$

Therefore, the lower bound is always

$$w_{t+1}(0) \geq \max\{y_1 - d_1 - \frac{w^D(k, \theta_t) - d_1}{1-\tau}, \frac{w^D(k, \theta_t) - y_1}{\tau} + y_1 - d_1\} \geq 0.$$

For the value d_1, if $y_1 \leq \bar{w}(1)$, then $d_1 = 0$ is optimal, because the top types are constraint and reducing d_1 increases the objective. Note in this case the reducing dividends can in change $w_{t+1}(0)$, but this has no effect on the government's objective function. The deviation-to-default constraint is satisfied in this case by $y_1 \geq w^D(k, \theta_t)$. If $y_1 > \bar{w}_1$, which can only happen in the dynamic case, the government is indifferent between all feasible allocations. Setting

$$d_1 = \min \left\{ \frac{1-\tau}{\tau} (\pi(k, \theta_t) + R w_t - y_1), y_1 - \bar{w}(1), w^D(k, \theta_t) \right\}$$

ensures that

$$\begin{aligned} 0 \leq d_1 &\leq w^D(k, \theta_t) \\ V(w^D(k, \theta_1), 1) &\leq d_1 + V(y_1 - d_1, 1). \end{aligned}$$

As a result, if the capital sub-problem is feasible, there always exists a monotone, locally IC (and hence globally IC) allocation satisfying all of the constraints in the capital sub-problem, with a θ^* such that $\bar{w}(\theta) \leq w_{t+1}(1)$ for all $\theta \leq \theta^*$, satisfying the stated properties. □

Proposition 5. *[Uniqueness of the capital level in the capital choice problem] There is an optimal allocation in the capital choice problem in which, for all θ, θ',*

$$\begin{aligned} k_t(w_t, \theta_t, \theta) &= \min \left\{ w_t \frac{R}{R - \varphi(1-\delta)}, k^*(\theta_t) \right\} \\ d_t(w_t, \theta_t, \theta, \theta') &= \frac{\tau}{1-\tau} \tau_t(w_t, \theta_t, \theta, \theta'). \end{aligned}$$

Proof. Note that

$$J^R(\hat{\mu}, w_t, \theta_t, k, y_1, V^*) = R^{-1} \{ \pi(k, \theta_t) + R w_t - y_1 + d_1 + \int_0^1 U(y_1 - d_1, \theta) \hat{\mu}(\theta; k) d\theta \}$$

where

$$d_1 = \max \left\{ \min \left\{ \frac{1-\tau}{\tau} (\pi(k, \theta_t) + R w_t - y_1), y_1 - \bar{w}(1), w^D(k, \theta_t) \right\}, 0 \right\}.$$

In the relaxed version of the capital choice problem, in which we relax ICs, it is immediately apparent that the objective function is increasing in y_1 and increasing in k, at least weakly, and is strictly increasing in k if $k < k^*(\theta_t)$.

Hence, if $\frac{R}{R-\varphi(1-\delta)} w_t < k^*(\theta_t)$ (meaning that achieving first-best is not possible), we will set $k_t(w_t, \theta_t, \theta_{t+1}) = \frac{R}{R-\varphi(1-\delta)} w_t$ and

$$y_t(w_t, \theta_t, \theta_{t+1}, 1) = \pi \left(\frac{R}{R - \varphi(1-\delta)} w_t, \theta_t \right) + R w_t.$$

In this case, $d_t(w_t, \theta_t, \theta_{t+1}, 1) = 0$.

Otherwise, set $k = k^*(\theta_t)$. In this case, if $\pi(k^*(\theta_t), \theta_t) + R w_t \leq \bar{w}(1)$, we will set

$y_t(w_t, \theta_t, \theta_{t+1}, 1) = \pi(k^*(\theta_t), \theta_t) + Rw_t$ and $d_t(w_t, \theta_t, \theta_{t+1}, 1) = 0$.

If $\pi(k^*(\theta_t), \theta_t) + Rw_t > \max\{\bar{w}(1), w^D(k^*(\theta_t), \theta_t)\}$, then any value of $\pi(k^*(\theta_t), \theta_t) + Rw_t \geq y_t(w_t, \theta_t, \theta_{t+1}, 1) \geq \max\{\bar{w}(1), w^D(k^*(\theta_t), \theta_t)\}$ is optimal. There exists a value of y_1 with positive dividends such that

$$\frac{1-\tau}{\tau}(\pi(k, \theta_t) + Rw_t - y_1) \leq \min\left\{y_1 - \bar{w}(1), w^D(k, \theta_t)\right\}$$

and using this value, dividends $d_t(w_t, \theta_t, \theta_{t+1}, 1)$ will proportional to taxes,

$$d_t(w_t, \theta_t, \theta_{t+1}, 1) = \frac{1-\tau}{\tau}\tau_t(w_t, \theta_t, \theta_{t+1}, 1) = \frac{1-\tau}{\tau}(\pi(k, \theta_t) + Rw_t - y_1).$$

It follows also that

$$\begin{aligned}
\tau_t(w_t, \theta_t, \theta_{t+1}, \theta) &= \pi(k_t(w_t, \theta_t, \theta_{t+1}), \theta_t) + Rw_t - y_t(w_t, \theta_t, \theta_{t+1}, \theta) \\
&= \pi(k_t(w_t, \theta_t, \theta_{t+1}), \theta_t) + R - d_t(w_t, \theta_t, \theta_{t+1}, \theta) \\
&- w_{t+1}(w_t, \theta_t, \theta_{t+1}, 1) + \\
&+ \frac{1}{1-\tau}(d_t(w_t, \theta_t, \theta_{t+1}, \theta) - d_t(w_t, \theta_t, \theta_{t+1}, 1))
\end{aligned}$$

which simplifies to

$$\tau_t(w_t, \theta_t, \theta_{t+1}, \theta) = \frac{\tau}{1-\tau}d_t(w_t, \theta_t, \theta_{t+1}, \theta).$$

Because the optimal allocation in the relaxed problem is the same for all types, it is incentive compatible. Therefore, this allocation is optimal in the capital choice problem, and hence in the original mechanism design problem. □

Proposition 6. *[Implementation with linear dividend tax] The allocation described in Proposition 5 can be implemented using only a dividend tax at a rate*

$$\tau_t(w_t, \theta_t, \theta_{t+1}, \theta_{t+1}) = \frac{\tau}{1-\tau}d_t(w_t, \theta_t, \theta_{t+1}, \theta_{t+1}).$$

Proof. Consider the problem defining $\bar{V}(w, \theta; \tau)$ (C.1), and conjecture that $\bar{V}(w, \theta; \tau) = V(w, \theta)$ in the optimal mechanism. Observe that if $d = 0$, we will have

$$V(w', \theta') \geq V\left(\pi(k) + k(R - \varphi(1-\delta)), \theta'\right)$$

by the borrowing constraint, and hence this constraint cannot bind. The budget constraint binds, and hence we have

$$w' = \pi(k, \theta) + Rw - \left(1 + \frac{\tau}{1-\tau}\right)d.$$

$$\frac{\tau}{1-\tau}d \leq Rw_0 - (R - \varphi(1-\delta)k)$$

The simplified optimization problem is

$$\bar{V}(w,\theta,\theta';\tau) = \max_{k,d\geq 0} R^{-1}\left\{d + V(w,\theta')\right\}$$

subject to

$$\begin{aligned} w' &= \pi\left(k,\theta\right) + Rw - \left(1+\frac{\tau}{1-\tau}\right)d, \\ \frac{\tau}{1-\tau}d &\leq Rw_0 - \left(R - \varphi\left(1-\delta\right)k\right) \\ d &\leq \pi\left(k,\theta\right) + k(R - \varphi\left(1-\delta\right)). \end{aligned}$$

If the firm is constrained but has a positive dividend, the FOC for reducing d is

$$-1 + V\left(w',\theta'\right)\left(1+\frac{\tau}{1-\tau}\right) \leq 0,$$

a contradiction. Therefore, constrained firms will pay no dividends. Such firms will also choose at least first-best capital if possible, and maximum capital otherwise, and have

$$w' = \bar{\pi} + Rw.$$

Such types must have $\bar{w}\left(\theta'\right) > \bar{\pi} + Rw$. For unconstrained firms, the FOC for reducing dividends is satisfied exactly. It follows that the upper bound constraints don't bind, and therefore that these types will also employ first best capital if possible, and maximal capital if not, and are free to choose the same dividends required by the optimal mechanism. As a consequence, it follows immediately that

$$\bar{V}\left(w,\theta;\tau\right) = V\left(w,\theta\right)$$

is verified. This argument applies essentially unmodified to the static problem, with slightly different notation for the value functions. □

C Problem Definitions

In this section, we formally define several value functions and mechanism design problems. These definitions are used in our proofs and referenced in the main text. Section C.1 defined the function $\bar{V}$ that described the net present value of a firm's dividends in the dynamic model under a constant dividend tax. Section C.2 defines $\bar{V}_0$ and $\bar{V}_1$, which are the net present value of a firm's dividends under a constant dividend tax in the static model, at dates zero and one respectively.

Section C.3 describes the full mechanism design problem for a non-exiting firm in the dynamic model, and Section C.4 describes the mechanism design problem for an exiting firm. Sections C.5 and C.6 describe the "capital sub-problem" and "capital choice problem," which break down the full mechanism design problem of Section C.3 into two parts. The capital choice problem corresponds to the first report of the firms, which determines their debt and capital levels, and the capital sub-problem corresponds to the second report, which occurs after capital is observed and determines the firm's dividend payment and taxes.

C.1 Definition of $\bar{V}(w, \theta; \tau)$

Define

$$\pi(k, \theta) = f(k, \theta) + (1-\delta) k - Rk.$$

Consider the firm's problem in the presence of a dividend tax, assuming that firm knows its next period type θ':

$$\bar{V}\left(w, \theta, \theta'; \tau\right) = \max_{b \geq 0, k \geq 0, w' \geq 0, d \geq 0} R^{-1}\left\{d + \bar{V}\left(w', \theta'; \tau\right)\right\}$$

subject to

$$\begin{aligned} w' &= \pi(k, \theta) + Rk - \left(1 + \frac{\tau}{1-\tau}\right) d - b \\ w &\leq k \leq w + R^{-1} b \\ \frac{\tau}{1-\tau} d + b &\leq \varphi (1-\delta) k \\ d &\leq \pi(k, \theta) + k (R - \varphi(1-\delta)) \\ d + \bar{V}\left(w', \theta'; \tau\right) &\geq \bar{V}\left(\pi(k, \theta) + k(R - \varphi(1-\delta)), \theta'; \tau\right) \end{aligned}$$

and note that voluntary exit is never optimal ($\bar{V}(w', \theta'; \tau) \geq \bar{V}(w', 0; \tau)$ for all w', θ).

The firm's continuation value function prior to learning its type is

$$\bar{V}(w, \theta; \tau) = \int_0^1 \bar{V}\left(w, \theta, \theta'; \tau\right) \mu\left(\theta' | \theta\right) d\theta'.$$

This defines a Bellman equation that, along with the transversality condition

$$\lim_{t \to \infty} R^{-t} V\left(w_t, \theta_t; \tau\right) = 0$$

defines a unique value function $\bar{V}(w, \theta; \tau)$.

C.2 Definition of $\bar{V}_0(w,\theta;\tau)$ and $\bar{V}_1(w,\theta;\tau)$

In the static (two-date) problem, the firm's value function in the presence of a dividend tax rate τ in period 1 is given by

$$\bar{V}_1(w,\theta;\tau) = \max_{b\geq 0, k\geq 0, w'\geq 0, d\geq 0} R^{-1}\left\{d + w'\right\}$$

subject to

$$\begin{aligned} w' &= \pi(k,\theta) + Rk - \left(1 + \frac{\tau}{1-\tau}\right) d - b \\ w &\leq k \leq w + R^{-1} b \\ \frac{\tau}{1-\tau} d + b &\leq \varphi(1-\delta) k \\ d &\leq \pi(k,\theta) + k(R - \varphi(1-\delta)) \end{aligned}$$

The firm's value function in period 0 is given by

$$\bar{V}_0\left(w,\theta,\theta';\tau\right) = \max_{b\geq 0, k, w'\geq 0, d\geq 0} R^{-1}\left\{d + \bar{V}_1\left(w',\theta';\tau\right)\right\}$$

subject to

$$\begin{aligned} w' &= \pi(k,\theta) + Rk - \left(1 + \frac{\tau}{1-\tau}\right) d - b \\ w &\leq k \leq w + R^{-1} b \\ \frac{\tau}{1-\tau} d + b &\leq \varphi(1-\delta) k \\ d &\leq \pi(k,\theta) + k(R - \varphi(1-\delta)) \\ d + \bar{V}_1\left(w',\theta';\tau\right) &\geq \bar{V}_1\left(\pi(k,\theta) + k(R - \varphi(1-\delta)), \theta';\tau\right), \end{aligned}$$

with

$$\bar{V}_0(w,\theta;\tau) = \int_0^1 \bar{V}_0\left(w,\theta,\theta';\tau\right) \mu\left(\theta'|\theta\right) d\theta'.$$

C.3 Statement of the Mechanism Design Problem for Non-Exiting Types

In this subsection, we describe the mechanism design problem given observable wealth w_t and current type $\theta_t > 0$. The multiplier χ_t controls the relative value of taxes and dividends, from the government's perspective. In the dynamic problem, our conjecture implicitly assumes that $\chi_t = 1$. In the static problem, we will show that $\chi_t \geq 1$, and the proposition will apply in the case in which $\chi_t = 1$ by focusing on situations in which required spending G is sufficiently small.

Given our timing assumptions, illustrated in Figure 1, firms sequentially choose capital and financing, and then make dividend and default decisions. We allow for the possibility of a "double" deviation in the mechanism, in which a firm reports some type θ'_{t+1} initially, when choosing capital/financing, and then reports another type θ''_{t+1}, when declaring a dividend/making a default decision. Formally, we use the notation $d_t\left(\theta'_{t+1}, \theta''_{t+1}\right)$ to refer to the dividend allocated to a firm that reports θ'_{t+1} at the investment/financing, and then reports θ''_{t+1} at the dividend/tax accrual stage. We use the same two-argument notation for other variables. We consider incentive-compatible direct revelation mechanisms, with both initial and interim IC constraints.

Formally, the government solves the following problem:

$$\max_{d_t(\cdot),b_t(\cdot),w_{t+1}(\cdot),k_t(\cdot),\tau_t(\cdot)} R^{-1}\int_0^1 \left\{ \begin{array}{c} d_t\left(w_t,\theta_t,\theta_{t+1},\theta_{t+1}\right) + U\left(w_{t+1}\left(w_t,\theta_t,\theta_{t+1},\theta_{t+1}\right),\theta_{t+1}\right) \\ +\chi_t\tau_t\left(w_t,\theta_t,\theta_{t+1},\theta_{t+1}\right) \end{array} \right\} \mu\left(\theta_{t+1}\middle|\,\theta_t\right) d\theta_{t+1}$$

subject to the following constrains:

Financing/Investment Budget Constraint:

$$k_t\left(w_t,\theta_t,\theta'_{t+1}\right) \leq w_t + R^{-1}b_t\left(w_t,\theta_t,\theta'_{t+1}\right),\ \forall w_t,\theta_t,\theta'_{t+1},$$

Dividend/Taxes Budget Constraint:

$$\begin{aligned} w_{t+1}\left(w_t,\theta_t,\theta',\theta''_{t+1}\right) \leq{}& \pi\left(k_t\left(w_t,\theta_t,\theta'_{t+1}\right),\theta_t\right) + Rk_t\left(w_t,\theta_t,\theta'_{t+1}\right) - d_t\left(w_t,\theta_t,\theta'_{t+1},\theta''_{t+1}\right) \\ & - b_t\left(w_t,\theta_t,\theta'_{t+1}\right) - \tau_t\left(w_t,\theta_t,\theta'_{t+1},\theta''_{t+1}\right),\ \forall w_t,\theta_t,\theta_{t+1},\theta'_{t+1},\theta''_{t+1}, \end{aligned}$$

No Default:

$$w^D\left(k_t\left(w_t,\theta_t,\theta'_{t+1}\right),\theta_t\right) \leq d_t\left(w_t,\theta_t,\theta'_{t+1},\theta''_{t+1}\right) + w_{t+1}\left(w_t,\theta_t,\theta'_{t+1},\theta''_{t+1}\right),\ \forall w_t,\theta_t,\theta'_{t+1},\theta''_{t+1},$$

Upper Limit on Dividends:

$$d_t\left(w_t,\theta_t,\theta'_{t+1},\theta''_{t+1}\right) \leq w^D\left(k_t\left(w_t,\theta_t,\theta'_{t+1}\right),\theta_t\right),\ \forall w_t,\theta_t,\theta'_{t+1},\theta''_{t+1}$$

Dividend/Taxes IC:

$$\begin{aligned} d_t\left(w_t,\theta_t,\theta'_{t+1},\theta''_{t+1}\right) + V\left(w_{t+1}\left(w_t,\theta_t,\theta'_{t+1},\theta''_{t+1}\right),\theta_{t+1}\right) \leq{}& d_t\left(w_t,\theta_t,\theta'_{t+1},\theta_{t+1}\right) \\ & + V\left(w_{t+1}\left(w_t,\theta_t,\theta'_{t+1},\theta_{t+1}\right),\theta_{t+1}\right),\ \forall w_t,\theta_t,\theta_{t+1},\theta'_{t+1},\theta''_{t+1}, \end{aligned}$$

Financing/Investment IC:

$$\begin{aligned} d_t\left(w_t,\theta_t,\theta'_{t+1},\theta_{t+1}\right) + V\left(w_{t+1}\left(w_t,\theta_t,\theta'_{t+1},\theta_{t+1}\right),\theta_{t+1}\right) \leq{}& d_t\left(w_t,\theta_t,\theta_{t+1},\theta_{t+1}\right) \\ & + V\left(w_{t+1}\left(w_t,\theta_t,\theta_{t+1},\theta_{t+1}\right),\theta_{t+1}\right),\ \forall w_t,\theta_t,\theta_{t+1},\theta'_{t+1}, \end{aligned}$$

Blocked Dividend No Default:

$$V\left(w^D\left(k_t\left(w_t,\theta_t,\theta'_{t+1}\right),\theta_t\right),\theta_{t+1}\right) \leq d_t\left(w_t,\theta_t,\theta_{t+1},\theta_{t+1}\right)$$
$$+ V\left(w_{t+1}\left(w_t,\theta_t,\theta_{t+1},\theta_{t+1}\right),\theta_{t+1}\right), \ \forall w_t,\theta_t,\theta_{t+1},\theta'_{t+1}$$

as well as non-negativity constraints and the lower bound on capital,

$$d_t\left(w_t,\theta_t,\theta'_{t+1},\theta''_{t+1}\right) \geq 0, \ \tau_t\left(w_t,\theta_t,\theta'_{t+1},\theta''_{t+1}\right) \geq 0, \ w_{t+1}\left(w_t,\theta_t,\theta'_{t+1},\theta''_{t+1}\right) \geq 0,$$

$$b_t\left(w_t,\theta_t,\theta'_{t+1}\right) \geq 0, \ k_t\left(w_t,\theta_t,\theta'_{t+1}\right) \geq w_0$$

and the definitions

$$\pi\left(k,\theta\right) = f\left(k,\theta\right) + \left(1-\delta\right)k - R$$
$$w^D\left(k,\theta\right) = f\left(k,\theta\right) + \left(1-\varphi\right)\left(1-\delta\right)k.$$

Note that this formulation omits deviations to exit, because we do not permit such deviations in the static model, and under our conjecture in the dynamic model will never be occur or be a binding constraint.

C.4 Statement of the Mechanism Design Problem for Exiting Types

For exiting types, there is no asymmetric information, and no production. That is,

$$f\left(k,0\right) = Rk - \left(1-\delta\right)k,$$

and therefore $\pi\left(k,0\right) = 0$ and $w^D\left(k,0\right) = (R-\varphi\left(1-\delta\right))k$.

The government solves

$$U(w,0) = \max_{d_t(\cdot),b_t(\cdot),w_{t+1}(\cdot),k_t(\cdot),\tau_t(\cdot)} R^{-1}\left\{d_t\left(w_t,0\right) + U(w_{t+1}\left(w_t,0\right),0) + \tau_t\left(w_t,0\right)\right\},$$

subject to the constraints

$$\begin{aligned}
k_t(w_t,0) &\leq w_t + R^{-1} b_t(w_t,0), \ \forall w_t, && \text{(Financing/Investment Constraint)} \\
w_{t+1}(w_t,0) &\leq R k_t(w_t,0) - d_t(w_t,0) && \\
&\quad - b_t(w_t,0) - \tau_t(w_t,0), \ \forall w_t, && \text{(Dividend/Taxes Constraint)} \\
w^D(k_t(w_t,0),0) &\leq d_t(w_t,0) + w_{t+1}(w_t,0), \ \forall w_t, && \text{(No Default)} \\
d_t(w_t,0) &\leq w^D(k_t(w_t,0),0), \ \forall w_t, && \text{(Upper Limit on Dividends)} \\
V\left(w^D(k_t(w_t,0),0),0\right) &\leq d_t(w_t,0) + V(w_{t+1}(w_t,0),0), \ \forall w_t, && \text{(Blocked Dividend No Default)}
\end{aligned}$$

observing that the two no-default constraints are identical, as well as non-negativity constraints and the lower bound on capital,

$$d_t(w_t,0) \geq 0, \ \tau_t(w_t,0) \geq 0, \ w_{t+1}(w_t,0) \geq 0, \ b_t(w_t,0) \geq 0, \ k_t(w_t,0) \geq w_0.$$

C.5 Statement of the Capital Sub-Problem

In this subsection, we consider the mechanism design problem described in the Appendix Section C.3, taking as given the initial wealth w_t, initial type $\theta_t > 0$, and initial capital level k chosen. The initial capital choice decisions generate an endogenous measure of types,

$$\hat{\mu}(\theta_{t+1};k) = \mu(\theta_{t+1}|\theta_t)\, \mathbf{1}\{k_t(w_t,\theta_t,\theta_{t+1}) = k\}.$$

To keep notation compact, in what follows we will use the shorthand $w_{t+1}(\theta)$ to mean $w_{t+1}(w_t,\theta_t,\theta'',\theta)$ for some θ_1'' such that $k_t(w_t,\theta_t,\theta_{t+1}) = k$, and likewise for $d_t(\theta)$ and $\tau_t(\theta)$. Note by Lemma 2 that it is without loss of generality to assume all such θ_1'' have the same allocations. We define the variable $y_t(\theta)$ as

$$y_t(\theta_t) = w_{t+1}(\theta) + d_t(\theta).$$

We also constrain the problem to have a particular value of $y_t(1)$ for the top type, $y_t(1) = y_1$. That is, we will consider the value of $y_t(1)$ for the top type as part of the capital choice problem (see below).

The sub-problem can be expressed as a function of two controls, $d_t(\theta)$ and $y_t(\theta)$:

$$J(\hat{\mu}, w_t, \theta_t, k, y_1, V^*) = \max_{d_t(\theta), y_t(\theta)} \frac{R^{-1}}{\int_0^1 \hat{\mu}(\theta;k)\, d\theta} \int_0^1 \left\{ \begin{array}{c} \chi_t(\pi(k,\theta_t) + R w_t) - \chi_t y_t(\theta) \\ +U(y_t(\theta) - d_t(\theta), \theta) + d_t(\theta) \end{array} \right\} \hat{\mu}(\theta;k)\, d\theta,$$

subject to the constraints

$$w^D(k,\theta_t) \le y_t(\theta), \ \forall\theta \qquad \text{(No Default)}$$
$$d_t(\theta) \le w^D(k,\theta_t), \ \forall\theta \qquad \text{(Upper Limit on Dividends)}$$
$$0 \le d_t(\theta), \ \forall\theta \qquad \text{(Non-Negative Dividends)}$$
$$y_t(\theta) \le \pi(k,\theta_t) + Rw_t, \ \forall\theta, \qquad \text{(Non-Negative Taxes)}$$

and the incentive constraints

$$d_t(\theta') + V(y_t(\theta') - d_t(\theta'),\theta) \le d_t(\theta) + V(y_t(\theta) - d_t(\theta),\theta), \ \forall\theta,\theta' \qquad \text{(Dividend/Taxes IC)}$$
$$V\left(w^D,\theta\right) \le d_t(\theta) + V(y_t(\theta) - d_t(\theta),\theta), \ \forall\theta, \qquad \text{(Blocked Dividend No Default)}$$
$$d_t(\theta) + V(y_t(\theta) - d_t(\theta),\theta) \le V^*(\theta), \ \forall\theta \ s.t.\ \hat{\mu}(\theta;k) = 0, \qquad \text{(Financing/Investment IC)}$$

and the constraints on the top type allocation, $y_t(1) = y_1$.

The function $V^*(\theta)$ from the financing/investment IC is shorthand: $V^*(\theta) = d_t(w_t,\theta_t,\theta_{t+1},\theta_{t+1}) + V(w_{t+1}(w_t,\theta_t,\theta_{t+1},\theta_{t+1}),\theta_{t+1})$ in the mechanism of Appendix Section C.3. In other words, for types θ that do not receive this capital level in equilibrium, the financing/investment IC is relaxed by reducing that type's payoff from this capital level.

The problem is well-defined given a feasible value of $y_1 \in \left[w^D(k,\theta_t), \pi(k,\theta_t) + Rw_t\right]$ and a function $V^*(\theta)$ such that is it possible to satisfy all of the constraints. We adopt the convention that $J(\hat{\mu}, w_t, \theta_t, k, d_1, y_1, V^*) = -\infty$ for infeasible problems.

C.6 Statement of the Capital Choice Sub-Problem

In this subsection, we rewrite the mechanism design problem of C.3 by separating it into the capital choice problem and the capital sub-problem described in (C.7). We simplify the problem using the result of Lemma 2 that both budget constraints bind, and hence eliminate the debt and taxes choice variables.

The problem is

$$\max_{d_t(\cdot),y_t(\cdot),k_t(\cdot),V^*(\cdot)} \int_0^1 J(\hat{\mu}(\cdot;k_t(w_t,\theta_t,\theta)), w_t, \theta_t, k_t(w_t,\theta_t,\theta), y_t(w_t,\theta_t,\theta,1), V^*(\theta))\, \mu(\theta|\theta_t)\, d\theta$$

subject to the constraints

$$w^D(k_t(w_t,\theta_t,\theta_{t+1}),\theta_t) \le y_t(w_t,\theta_t,\theta_{t+1},1) \ \ \forall w_t,\theta_t,\theta_{t+1}, \qquad \text{(No Default at Top)}$$
$$y_t(w_t,\theta_t,\theta_{t+1},1) \le \pi(k_t(w_t,\theta_t,\theta_{t+1}),\theta_t) + Rw_t \ \forall w_t,\theta_t,\theta_{t+1}, \qquad \text{(Non-Negative Debt/Taxes at Top)}$$

and lower bounds on dividends at the top and capital,

$$k_t\left(w_t,\theta_t,\theta'_{t+1}\right)\geq w_0,$$

and the requirement that $d_t(w_t,\theta_t,\theta'_{t+1},\cdot)$, $y_t(w_t,\theta_t,\theta'_{t+1},\cdot)$ are in the set of optimal policies of the capital sub-problem for all θ'_{t+1}, the form of the financing/investment IC constraint,

$$V^*(\theta_{t+1})\leq d_t\left(w_t,\theta_t,\theta_{t+1},\theta_{t+1}\right)+V\left(y_{t+1}\left(w_t,\theta_t,\theta_{t+1},\theta_{t+1}\right)-d_t\left(w_t,\theta_t,\theta_{t+1},\theta_{t+1}\right),\theta_{t+1}\right),$$

and using the definitions

$$\hat{\mu}\left(\theta';k_t\left(w_t,\theta_t,\theta\right)\right)=\mu\left(\theta'|\theta_t\right)\mathbf{1}\left\{k_t\left(w_t,\theta_t,\theta\right)=k_t\left(w_t,\theta_t,\theta'\right)\right\}.$$

Note that we can write $V^*(\theta)$ as a choice variable and use the inequality constraint because the function $J(\cdot,V^*)$ is weakly greater than $J(\cdot,V^{*\prime})$ if $V^*(\theta)\geq V^{*\prime}(\theta)$ for all θ.

C.7 Statement of the Relaxed Capital Sub-Problem

In this subsection, we describe a relaxed version of the capital sub-problem described in in C.5, taking the value d_1 as given. In our proofs , we will solve this problem, and then show that its solution is a solution to the relaxed problem.

Motivated by our monotonicity results (see Lemma 3 above), we restrict attention to monotone allocations in the capital sub-problem. We define $\theta^H\in[0,1]$ as the value for which $d_t(\theta)>d_t(1)$ for all $\theta<\theta^H$ and $d_t(\theta)=d_t(1)$ for all $\theta>\theta^H$.

We also impose "relaxed monotonicity conditions,"

$$w_{t+1}(0)\leq w_{t+1}(\theta)\leq y_1-d_1,$$

and use only the local IC condition (Lemma 4). Furthermore, for all types $0<\theta\leq\theta^*$, we ignore all of the limits on dividends, the positive taxes constraint, the deviation to default constraint, and the no-default constraint. We also ignore the financing/investment IC constraint.

The relaxed problem

$$J^R\left(\hat{\mu},w_t,\theta_t,k,y_1\right)=\max_{\theta^H\in[0,1],w_{t+1}(\theta)\geq0,y_t(\theta)\geq0}R^{-1}\int_{\theta^{H+}}^{1}\left\{\begin{array}{c}\chi_t(\pi\left(k,\theta_t\right)+Rw_t)-\chi_t y_1\\+U\left(y_1-d_1,\theta\right)+d_1\end{array}\right\}\hat{\mu}\left(\theta;k\right)d\theta$$
$$+R^{-1}\int_0^{\theta^H}\left\{\begin{array}{c}\chi_t(\pi\left(k,\theta_t\right)+Rw_t)+(1-\chi_t)y_t\left(\theta\right)\\+U\left(w_{t+1}\left(\theta\right),\theta\right)-w_{t+1}\left(\theta\right)\end{array}\right\}\hat{\mu}\left(\theta;k\right)d\theta,$$

subject to the constraints

$$w_{t+1}(0) \le w_{t+1}(\theta) \le y_1 - d_1,\ \forall \theta \in [0, \theta^H]$$

$$y_t(\theta) + V(w_{t+1}(\theta), \theta) - w_{t+1}(\theta) + \int_{\theta}^{\theta^H} V_\theta(w(\theta'), \theta') d\theta' = d_1 + V(y_1 - d_1, \theta),\ \forall \theta \in [0, \theta^H],$$

and the constraints for the exiting (bottom) types,

$$w^D(k, \theta_t) \le y_t(0), \qquad \text{(No Default)}$$

$$d_t(0) \le w^D(k, \theta_t), \qquad \text{(Upper Limit on Dividends)}$$

$$0 \le d_t(0), \qquad \text{(Non-Negative Dividends)}$$

$$y_t(0) \le \pi(k, \theta_t) + R w_t. \qquad \text{(Non-Negative Debt/Taxes)}$$

Observe that this problem is a strictly relaxed version of the capital sub-problem, and hence if a solution to this problem is feasible in the capital sub-problem, it is optimal.