

Optimal Corporate Taxation Under Financial Frictions*

Eduardo Dávila[†]

Benjamin Hébert[‡]

September 8 2017

Abstract

We study the optimal design of corporate taxation when firms are subject to financial constraints. We find that corporate taxes should be levied on unconstrained firms, since those firms value resources inside the firm less than financially constrained firms. When the government has complete information about which firms are and are not constrained, this principle is sufficient to characterize optimal corporate tax policy. When the government (and other outsiders) do not know which firms are and are not constrained, the government can use the payout policies of firms to elicit whether or not the firm is constrained, and assess taxes accordingly. Using this insight, we discuss conditions under which a tax on dividends paid is the optimal corporate tax. We then extend this result to a dynamic setting, showing that, if the government lacks commitment, the optimal sequence of tax mechanisms can be implemented with a dividend tax. With commitment, we reach a very different conclusion— a lump sum tax on firm entry is optimal. We argue that these two models demonstrate an underlying principle, that optimal corporate taxes should avoid exacerbating financial frictions, and demonstrate that the structure of the financial frictions can drastically change the optimal policy.

JEL numbers: G18, K35, G33

Keywords: corporate taxation, dividend taxation, financial frictions, interest tax deductibility

*Work in progress. We would like to thank Anat Admati, Peter DeMarzo, Peter Diamond, Emmanuel Farhi, Piero Gottardi, Anton Korinek, Adriano Rampini, Jeremy Stein, and seminar participants at Duke Fuqua for helpful comments and discussions.

[†]New York University and NBER. Email: edavila@nyu.edu

[‡]Stanford University. Email: bhebert@stanford.edu

1 Introduction

Virtually every developed country collects taxes from corporations. In this paper, we take as given that such taxation is necessary, and ask how firms should be taxed, or, equivalently, which firms should be taxed. In an economy with financial frictions, some firms value marginal internal resources more than other firms. We refer to these firms as constrained. A government that seeks to maximize the total value of the firms in the economy should tax constrained firms less than other firms (or even subsidize them, if possible). However, it is not easy for the government, or any other outside investor, to determine if a firm is constrained. The government must employ a mechanism to induce the firms to reveal whether or not they are constrained. Of course, firms are aware that, if they reveal they are not constrained, they will be taxed. The mechanism the government employs must therefore be incentive compatible.

At a high level, our description of the government's problem resembles the classic non-linear optimal (household) taxation model of [Mirrlees \(1971\)](#). We will show this resemblance formally. 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. As a result, 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." 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 participation constraint in our model that limits the ability of the government to extract as much as it would like from unconstrained firms.

We argue that the payout policy of firms can help reveal whether firms are financially constrained. This argument has a long history, going back to at least [Fazzari, Hubbard and Petersen \(1988\)](#). Firms consistently paying large dividends are, to those authors, *a priori* unlikely to be financially constrained. Even among "low-dividend-paying" firms (the sample of [Kaplan and Zingales \(1997\)](#)), paying relatively more dividends is associated with a reduced likelihood of reporting being financially constrained. The existence of a literature on identifying financially constrained firms also supports our assumption that it is difficult for the government and other outsiders to determine whether a firm is financially constrained.

To summarize, the optimal mechanism uses the payout policy to help determine whether or not a firm should be taxed. Put another way, something akin to a dividend tax¹ is part of the implementation of the optimal mechanism. This, however, is contrary to what is typically meant by "corporate taxes," which in practice are based on a firm's profits, adjusted for numerous credits and deductions. Moreover, it is not what is usually meant by "dividend taxes," which in the United States are collected as part of personal income taxes. For our purposes, throughout

¹We will not distinguish between share buy-backs and dividend payments, even though the current tax code treats them differently.

this paper, we will define corporate taxes as “taxes collected from corporations.” Our question, in some sense, is whether these taxes should be based on profits, or something else. Our answer focuses in part on “dividend taxes,” which we will define to mean taxes that are paid by corporations in proportion to the amount of dividends they pay. These taxes are not necessarily the same as the dividend taxes in the personal income tax system, for two reasons. First, with financial constraints, whether the firm or its shareholders pay the tax can change the incidence of the tax. Second, because there are shareholders who do not pay personal income taxes (e.g. endowments), dividend taxes in the personal income tax code can generate clientele effects (Allen, Bernardo and Welch (2000)) that are absent from the dividend taxes in the corporate tax system that we describe.

We analyze both static and dynamic models. In the static model, we first study the optimal tax policy in a single-date model with perfect information. If the government can both subsidize and tax firms, the optimal policy is to undo the financial frictions. If the government can only tax, but not subsidize, firms, the optimal policy is to tax only unconstrained firms, provided that this policy generates sufficient revenue. Intuitively, the government, valuing the welfare of all firms equally, wishes to collect taxes from the firms for whom paying taxes is least costly. We next consider a two-date model, without commitment for the government. In the second date, we assume that the equilibrium is the solution to the single-date full-information optimal tax problem. In the first date, firms learn about their marginal product in the second date before the government does. The government can use the firm’s payout policy to elicit this information. In fact, under certain conditions, the optimal mechanism can be implemented with a dividend tax. The key intuition is that the desire to pay dividends separates firms that will be unconstrained in the second date from firms that will be constrained. 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 today 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/output at date zero) 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.

One complication that arises in the two period static model is the propagation of the effects of taxation on financial constraints both “forwards” and “backwards.” Taxes collected today reduce firms’ wealth, which propagates “forward” and tightens financial constraints in the future. Taxes collected today also reduce the value of the firm today, which propagates “backwards,” reducing the value of the firm in past, making default more likely, and tightening financial constraints in

the past. Depending on the nature of the financial frictions, one or both of these effects may be operative. To study these effects separately, we develop infinite horizon dynamic models, in the spirit of the dynamic public finance models surveyed by [Golosov, Tsyvinski and Werquin \(2016\)](#). These models feature the entry and exit of firms.

In the first model, we assume that the government lacks commitment, and creditors cannot exclude defaulting firms from re-entering. The setup of this model builds on the work of [Rampini and Viswanathan \(2010\)](#), and the resulting financial frictions depend only on current-date variables. Consequently, in this model taxes tighten financial constraints at the current date and in the future (the effects propagate forward, through wealth) but there is no backwards propagation. We show, in results similar to those for the static model, that a dividend tax is optimal. Lump-sum taxes on entering firms, while feasible, are strictly sub-optimal.

In the second model, we assume the government can commit and that creditors can exclude firms from re-entering. The resulting financial friction builds on the work of [Kehoe and Levine \(1993\)](#), and depends on the firms' continuation value. In fact, the problem can be rewritten entirely in terms of continuation values, and as a result, the "wealth" of a firm is irrelevant. Consequently, taxes propagate backwards (by reducing continuation values) but not forwards. As a result, it is optimal to "back-load" the payoffs to the firm, meaning that a lump-sum tax on entry is optimal, and a dividend tax is strictly sub-optimal. The contrast between the optimal policies in these two models illustrates the importance of designing taxes that avoid exacerbating financial constraints, and the dependence of the structure of optimal taxes on the structure of the relevant financial constraints.

Our approach integrates several well-developed literatures. There is an extensive literature on corporate taxation, surveyed by [Auerbach \(2002\)](#) 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 and Stiglitz \(2009\)](#)) 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, in effect, embeds this perspective— the optimality of dividend taxes, as opposed to some other kinds of corporate taxation, is closely related to this fact. Our model assumes that firms maximize the expected value of dividends, and therefore is not obviously compatible with the "agency view" advocated by [Chetty and Saez \(2010\)](#). We discuss in the text how our model could be extended in this direction.

Our approach emphasizes optimal allocations, rather than particular taxes, because we adopt the optimal non-linear taxation perspective pioneered by [Mirrlees \(1971\)](#). Our dynamic models resemble, in many respects, the dynamic public finance models surveyed in [Golosov, Tsyvinski](#)

and Werquin (2016). The key difference between our paper and these large literatures is our focus on firms, and on financial frictions. Dynamic Mirrleesian models of optimal taxation focus on the behavior of households and treat firms as a “veil” (e.g. Farhi and Werning (2012)). We adopt, for simplicity, a partial equilibrium perspective that emphasizes how financial frictions create a meaningful distinction between corporate and household taxes. The financial frictions we employ build, in particular, on the work of Kehoe and Levine (1993) and 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. Because we simply assume that the government must raise revenues by taxing firms, we have nothing to say on the topic of whether taxing firms is ever optimal. Much of the work on capital taxation under full information (Judd (1985); Chamley (1986); Chari and Kehoe (1999)) emphasizes 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.

We begin, in Section 2, by describing a single date in our models. Our static models will feature one or two dates with this structure, and our dynamic models will have infinitely many dates with the same structure. We describe our results for static models in Section 3, and our results for dynamic models in Section 4. Section 5 concludes.

2 Environment

In this section, we introduce the basic structure of a single “date” in our models. We will modify this structure slightly in each of the specific models we study, for example, by limiting the government’s ability to collect taxes or issue subsidies.

There are three groups of agents in the economy: firms, outside investors, and the government. There is a single consumption good (dollar), which serves as numeraire. Outside investors and firms are both risk-neutral, and discount cash-flows within the date, between the beginning and end, at a gross real interest rate of R . This interest rate is constant and invariant to policy—this is what we mean when we say that our analysis is partial equilibrium. One might, in general, 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.

Figure 1 illustrates the timeline of events within a given date. Firms face a financing/investment decision, a dividend payment decision, and a default decision.

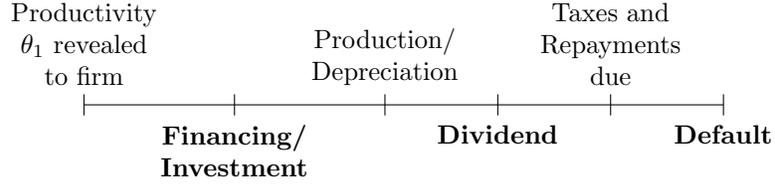


Figure 1: Within-date timeline

Firms are initially endowed with resources w_t . Firms can raise additional resources from outside investors, $r_t \geq 0$, or the government, $s_t \geq 0$ (if the government is allowed to subsidize firms). Firms invest these resources in capital k_t , broadly defined, satisfying the following budget constraint:

$$k_t \leq w_t + r_t + s_t. \quad (1)$$

An investment of k_t dollars at the beginning of date t yields $f(k_t, \theta_t)$ dollars by the end of date t . Firms productivity depends on their 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, concave, and differentiable almost everywhere in capital. Firms which employ no capital receive no output, that is, $f(0, \theta_t) = 0$. Capital investment depreciates neoclassically at a rate δ . Investment choices and production outcomes are observable to the government and outside creditors.

We will assume that there exists a “first-best” level of capital, $k^*(\theta)$, which is the smallest level of capital such that

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

We will say, for a given date, that a firm is constrained if it has capital strictly less than this level, and otherwise call it unconstrained. We will also assume that, for all $k > k^*(\theta)$,

$$f_k(k, \theta) + 1 - \delta = R.$$

That is, the marginal product of capital reaches its first-best level of capital and then remains constant. This assumption, in certain respects, mimics the ability of the firm, after exhausting its ability to invest in physical capital, to invest at the risk-free rate. We will discuss this assumption in more detail below.

After production occurs, the firm declares a (weakly positive) dividend, $d_t \geq 0$. After this dividend is declared, the firms’ obligations to outside investors, $b_t \geq 0$, and to the government, $\tau_t \geq 0$, are determined. The government/creditors can also allow or block the proposed dividend. Blocking the dividend prevents the money from leaving the firm. The firm can default on its

obligations, and will decide whether or not to default to maximize the payoff to the shareholders. We consider alternative assumptions about the consequences of default, which we detail below.

If a firm repays its obligations, it receives continuation wealth

$$w_t^R = f(k_t, \theta_t) + (1 - \delta)k_t - d_t - b_t - \tau_t, \quad (2)$$

as long as that this quantity is weakly positive. If this quantity is not weakly positive, repayment is not feasible, and the firm defaults. If the firm defaults, its shareholders can receive continuation wealth

$$w_t^D(k_t, \theta_t) = f(k_t, \theta_t) + (1 - \omega)(1 - \delta)k_t, \quad (3)$$

less any dividends paid. The particular functional form of the continuation wealth under default will vary across our models, for tractability reasons. We require that the firm declare a dividend no larger than its continuation wealth in the event of default ($d_t \leq w_t^D$), which prevents the firm from continuing with negative wealth.

At the beginning of date t , firms privately learn their type for the next date, θ_{t+1} , which determines their future productivity. This generates asymmetric information – the outside investors and the government do not learn the firm’s type until the beginning of date $t + 1$. The outside creditors and government must therefore assess debt and tax payments that satisfy incentive-compatibility conditions for the firm. These incentive compatibility conditions will differ across the models we consider, and we will discuss them in detail below.

Default assumptions In the sections that follow, we will study mechanisms that avoid default. We believe that, in our settings, it is without loss of generality to avoid default, and strictly optimal in the presence of inefficiencies associated with default. We consider two alternative assumptions about default. First, we consider the case in which default requires liquidation. We refer to this default constraint as the *liquidation default constraint*. Let $V_{t+1}(w_t^R, \theta_{t+1}, h_{t+1})$ denote the continuation value of a firm with wealth conditional on repayment of w_t^R , with type θ_{t+1} and “history” h_{t+1} , at date $t + 1$. In this context, the history might include past profitability, investment levels, and other relevant publicly available information. If defaulting requires liquidation, then the firm will not default if

$$d_t + V_{t+1}(w_t^R, \theta_{t+1}, h_{t+1}) \geq w_t^D(k_t, \theta_t). \quad (4)$$

This constraint implies that the default decision could be history-dependent.

Second, as in [Rampini and Viswanathan \(2010\)](#), we also consider the case in which the shareholders, after defaulting, cannot be excluded from starting a new firm with the same type as the defaulting firm. We will refer to this default constraint as the *no exclusion default constraint*.

In this case, it is natural to suppose that the government cannot condition on a firm's history, only its current wealth level and type. When we use this constraint, we will also assume that the government lacks commitment, which is consistent with this assumption. Otherwise, the government might commit to treating new firms, which result from a default, differently, and thereby discourage default. The firm will not default if

$$d_t + V(w_t^R, \theta_{t+1}, t+1) \geq \max\{d_t + V(w_t^D(k_t, \theta_t) - d_t, \theta_{t+1}), V(w_t^D(k_t, \theta_t), \theta_{t+1})\}. \quad (5)$$

The maximization in this problem reflects the government and creditor's 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.

Discussion of the Environment This environment is designed to be as simple as possible, while allowing us to demonstrate the main results of the paper. We will briefly discuss four simplifications, in particular, and conjecture how our results might change in richer environments.

First, our environment has only a single asset, capital, that firms can invest in. As discussed above, our assumption that the marginal product of capital reaches and then remains equal to the risk-free rate is meant to capture the ability of the firm to purchase securities that earn the risk-free rate. It may appear to interact awkwardly with the default conditions— in particular, it implies that securities and physical capital can serve equally well as collateral, but a firm's earnings cannot serve as collateral. However, in the models we analyze, firms will either face a binding default constraint, or have a marginal product of capital equal to the risk-free rate, but never both at the same time. As a result, this assumption leads to the same conclusions as a richer model that explicitly modeled cash as distinct from capital, without adding an additional choice variable.

Second, our environment has only a single input into production, capital. Our preliminary analysis of richer models, with labor and intermediate goods inputs, appears to reach similar conclusions. The key assumption, in those models, is decreasing returns to scale. If the firm's production function were constant returns to scale in its various inputs, firms would aggregate and the distribution of wealth across firms would be irrelevant. Alternatively, we speculate that in a model in which the goods produced by firms are not perfect substitutes (e.g. the New Keynesian setup of [Woodford \(2003\)](#)), firms could have constant returns to scale in production and our main results would still hold.

Third, our environment has no uncertainty apart from the process governing the firm's type, θ_t . Our results are unaltered by the addition of an observable and contractible shock (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.

Fourth, and related to previous point about security design, there are multiple ways of interpreting the model. The agent receiving "dividends" is the agent controlling the firm's decisions. If firms maximize value for their shareholders, then dividends is indeed the correct term. This does not imply, however, that the outside investors have debt claims, in the sense of claims that are constant in the absence of default. In fact, in one particular case that we will discuss, the outside investors receive payments that are proportional to the "dividends," as if they were shareholders without (*de facto* or *de jure*) control rights. Alternatively, one might interpret the model under the assumption of managerial control of the firm. In this case, "dividends" are really managerial compensation, and the outside investors might hold debt or equity claims. Our model assumes that dividends cannot be negative, meaning that the shareholders or manager are unable to inject additional funds into the firm. We believe 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 a manager and shareholders both influence the firm's decisions, to future work.

Lastly, note that in both of the default constraints we discuss, taxes and debt repayment have identical effects, operating entirely through the continuation wealth w_t^R . The model is setup to ensure that the government cannot circumvent the financial frictions by assessing taxes that are not subject to default. That is, substituting debt issuance r_t and repayment b_t for subsidies s_t and taxes τ_t does not change either of the default constraints or the initial budget constraint.

Having described the basic environment of a single date, we will turn next to our discussion of one and two-date models.

3 Static formulation

In this section, we describe our static models. We will begin by describing a single-date model with full information, and use this model as the second date in a two-date model with asymmetric information. In the single-date full information model, we will assume that government cannot subsidize firms. If we did not make this assumption, the government would simply undo all of the financial frictions.

For our static models, we will assume that the space of types is the unit interval, $[0, 1]$, and that marginal products are increasing in the firms type. Formally, the derivative $f_{k\theta}(k, \theta)$ exists almost everywhere and, anywhere it does exist, is weakly positive. Our choice to focus on a

one-dimensional type space is not without loss of generality; in particular, there may be useful distinctions to be drawn between average and marginal products. However, this assumption allows us to convey the intuition of the model in a straightforward way.

3.1 A Single Date, Full-Information Model

We will refer to the date in this single-date model as date $t = 1$. The firm will simply “consume” whatever wealth remains inside the firm at the end of date one,

$$V_2(w_1^R, \theta_2, h_2) = w_1^R.$$

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

$$V_1(w_1, \theta_1) = R^{-1}\{d_1(w_1, \theta_1) + w_1^R(w_1, \theta_1)\}, \quad (6)$$

where $d_1(w_1, \theta_1)$ and $w_1^R(w_1, \theta_1)$ are the equilibrium allocations. Both of the no-default constraints described in the previous section ((4) and (5)) simplify to the same constraint,

$$d_1(w_1, \theta_1) + w_1^R(w_1, \theta_1) \geq w_1^D(k_1(w_1, \theta_1), \theta_1),$$

where $k_1(w_1, \theta_1)$ denotes the equilibrium 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. We use this language because, in the models we will discuss in the following sections, there is asymmetric information, and we discuss direct revelation mechanisms as a means of implementing optimal policy. We imagine that the government is choosing all of firm’s choice variables, $\{k_1, b_1, r_1, d_1\}$, and taxes τ_1 , for each level of wealth w_1 and type θ_1 . Each of these choice variables is required to be weakly positive, and the upper bound on dividends must also be satisfied. The government must respect the initial budget constraint ((1)) and production functions ((2) and (3)), as well as the outside creditor’s full-information participation constraint,

$$R^{-1}b(w_1, \theta_1) \geq r(w_1, \theta_1).$$

The government must also respect the no-default constraint, which forms a sort of 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

$$d_1(w_1, \theta_1) + w_1^R(w_1, \theta_1) \geq w_1^D(w_1, \theta_1).$$

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

Finally, the government is constrained, across the population of firms, to raise sufficient resources through taxation. Let $dF(w_1, \theta_1)$ denote the density of firms with wealth w_1 and type θ_1 . The government's tax policies must satisfy

$$R^{-1} \int \int \tau_1(w_1, \theta_1) dF(w_1, \theta_1) \geq R^{-1} G_1 > 0,$$

where G_1 denotes the (strictly positive) required expenditure.

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

$$\int \int V_1(w_1, \theta_1) dF(w_1, \theta_1).$$

Although the problem appears to have numerous constraints, it simplifies to a straightforward problem. First, note that 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 creditor's participation constraint and initial budget constraint will always bind. As a result, there is really only a single choice variable for the firm (in the lemma below, we choose capital, but this is arbitrary). The problem can also be simplified by introducing a multiplier, χ , 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 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*

$$\min_{\chi \geq 0} \chi R^{-1} G_1 + \int \int U_1(w_1, \theta_1; \chi) dF(w_1, \theta_1),$$

where

$$U_1(w_1, \theta_1; \chi) = \max_{k_1 \geq 0, \tau_1 \geq 0} R^{-1} \{f(k_1, \theta_1) + (1 - \delta)k_1 - Rk_1\} + w_1 + R^{-1}(\chi - 1)\tau_1, \quad (7)$$

subject to the constraint that

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

Proof. See the appendix, 5.1. □

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 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, $R^{-1}(\chi - 1)$, is the same for all firms. For the unconstrained firms, there is no particular reason to tax one firm instead of another; the optimal policy is not determined. However, it will never be optimal to tax a constrained firm, if there exists an unconstrained firm who 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. *If G_1 is sufficiently small and there exists a positive mass of firms with $(1 - R^{-1}\omega(1 - \delta))k^*(\theta_1) < w_1$, $\chi = 1$, and there exists an optimal policy in which the government sets*

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

for some $\tau \in [0, \omega(1 - \delta)]$.

Proof. See the appendix, 5.2. □

We have chosen to focus on a particular policy— a linear tax on “excess wealth”— because it is straightforward, 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. When wealth is insufficient to reach the first-best level of capital, the government's marginal value of wealth, $U_{1,w}(w_1, \theta_1; 1)$, is strictly greater than one, and when there is sufficient wealth, is equal to one. Meanwhile, the firm's marginal value of wealth, $V_{1,w}(w_1, \theta)$, is equal to $1 - \tau$ if w_1 is sufficient to reach the first-best capital level, and strictly greater than $1 - \tau$ otherwise.

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.

3.2 A Two-Date, Asymmetric Information Model

We now introduce the first date of the model in which firms have private information about their types. We assume that the government solves the mechanism design problem at date 0 taking as given the solution to the date 1 mechanism design problem just described. In that sense, our date 0 mechanism features no-commitment. However, the mechanism design problem solved by the government at date 0 takes into account that the outcomes of that mechanism affect the government's ability to raise revenue at date 1, so the date 0 government is not myopic.

Our formal statements will presume that taxing unconstrained firms is sufficient to satisfy the government's budget constraint, that is, we focus on scenarios in which $\chi_0 = \chi_1 = 1$. We assume that the government cannot subsidize firms, and that outside investors cannot commit to financing a firm after learning their type. As a result, there are interim participation constraints in the government's mechanism for the outside investors.

The government's objective function corresponds to

$$R^{-1} \int \left\{ d_0(\theta_1) + U_1(w_0^R, \theta_1) \right\} dF(\theta_1 | \theta_0),$$

where $U_1(w_0^R, \theta_1)$, which is defined in Equation 7, incorporates the private continuation of a firm of type θ_1 and the revenue raised from a firm of type θ_1 , and where $w_0^R(\theta_1)$ satisfies

$$w_0^R(\theta_1) = f(k_0(\theta_1), \theta_0) + (1 - \delta)k_0(\theta_1) - d_0(\theta_1) - b_0(\theta_1) - \tau_0(\theta_1), \forall \theta_1.$$

The government must respect firms' initial budget constraints, which imply that investment must be funded with internal funds and the government subsidy. Formally,

$$k_0(\theta_1) \leq w_0 + r_0(\theta_1), \forall \theta_1.$$

The government must also satisfy the creditor's participation constraints,

$$r_0(\theta_1) \leq R^{-1}b_0(\theta_1), \forall \theta_1$$

As in date 1, the government faces a revenue raising constraint across the population of firms. It must raise, net of subsidies, a total of G_0 dollars

$$\int R^{-1}\tau_0(\theta_1) dF(\theta_1 | \theta_0) \geq G_0 > 0.$$

The government must respect the no-default constraint, which assume it corresponds to that in Equation (5). Formally, this imposes an upper bound on the tax level that depends exclusively on the level of capital:

$$\tau_0(\theta_1) + b_0(\theta_1) \leq \omega(1 - \delta)k_0(\theta_1), \forall \theta_1.$$

Finally, the government must satisfy firms' interim incentive compatibility constraints, which can be expressed as

$$d_0(\theta_1) + V_1(w_0^R(\theta_1), \theta_1) \geq d(\theta'_1) + V_1(w_0^R(\theta'_1), \theta_1), \forall \theta_1, \theta'_1,$$

where $V_1(w_0^R(\theta_1), \theta_1)$ is defined in Equation (6) above. The functions $U_1(\cdot)$ and $V_1(\cdot)$ are taken as given from the perspective of the government that faces the date 0 mechanism design problem.

Define

$$\bar{w}(\theta) = \left(1 - R^{-1}\omega(1 - \delta)\right)k^*(\theta)$$

as the level of wealth required to achieve the first-best level of capital in the absence of taxes.

We will assume that the firms initial wealth is sufficient to reach the first-best level of capital in date zero, but not so large that no borrowing is required:

$$\bar{w}(\theta_0) < w_0 < k^*(\theta_0).$$

We will also assume that the initial level of wealth, when combined with the first-best level of production at date zero, are strictly sufficient to achieve the first-best level in date one for the lowest type, but strictly insufficient for the highest type:

$$\bar{w}(1) > f(k^*(\theta_0), \theta_0) + (1 - \delta - R)k^*(\theta_0) + R w_0 > \bar{w}(0).$$

Under these assumptions, we show that it is feasible to raise funds without distortions, using a dividend tax.

Proposition 2. *If G_1 is sufficiently small, $\chi_0 = \chi = 1$, the optimal mechanism uses the first-best level of capital for all firms at date zero, and taxes firms in proportion to their dividends. The optimal mechanism can be implemented by a dividend tax, with tax rate*

$$\tau_0(\theta_1) = \frac{\tau_e}{1 - \tau_e} d_0(\theta_1).$$

Proof. See the appendix, 5.2. □

4 Dynamic Models

In this section, we will describe infinite-horizon versions of our basic model. Each date follows the structure described in Section 2. We will study optimal policy for a government in two models that are, in some sense, polar opposites. In the first model we study, the government does not have commitment across dates, firms in default cannot be excluded from re-entering but can be excluded from liquidating, and creditors can renege on promises to lend. These features give rise to a default constraint similar to the one studied by Rampini and Viswanathan (2010). In the second model we study, the government and creditors can commit to long-term contracts, and firms in default can be excluded from re-entry but not from liquidation. These features give rise to a default constraint similar to that of Kehoe and Levine (1993) and related papers.

In the background of our model, we have in mind an economy in which there exists, at all times, a population of firms, with firms exiting and entering over time. We will study the steady states of such a model, in which the government must raise at least $G > 0$ from the population of firms. When we study models without commitment, we will study stationary equilibria of the model. When we study models with commitment, we take an approach analogous to “optimal policy from a timeless perspective” (Woodford (2003)). That is, we will consider the problem of a government which commits to tax policies for a new firm entering the economy. When this new firm enters (date zero), there are other firms in the economy, which entered on previous dates. These firms, however, have already received commitments with respect to their tax policies.

In both models, an entrepreneur enters with initial wealth w_0 . The government can charge a lump-sum tax on entry, T_0 , collected on entry, and therefore the firm will enter with at most wealth $w_0 - T_0$, depending on whether the entrepreneur puts all of her wealth in the firm or not. The firm begins operations immediately, learning whether it is normal or about-to-exit. At that point, the firm is free to contract with private creditors to raise additional funds. Firms enter at a rate ψ , which will also be the rate at which they exit.

We will discuss two different dynamic models, which correspond to the two default assumptions discussed previously. Both of these models will have the same, simple structure for types. Firms are either “normal,” “about to exit,” or “exiting.” A firm that is exiting has a constant returns to scale technology, with a marginal product of capital equal to the gross risk-free rate, and no depreciation. If a firm is exiting, its type is observable to the government and does not change. A firm that is “about to exit” will become an “exiting” firm at the next date. Normal firms are identical to about to exit firms, except that they will not (except by choice) be exiting in the next date. If a normal firm chooses not to become exiting (which we will assume is optimal), it will be either normal (with probability $1 - \psi$) or about to exit (with probability ψ). Whether the firm is “normal” or “about to exit” is private information for the firm, which the firm learns at the beginning of the date, as in the structure described in Section 2.

We will use $dF(w, t)$ and $dF^e(w, t)$ to denote the densities of firms that are normal/about-to-exit and of firms that are exiting in the current period, respectively. The density $dF(w, t)$ includes newly entering firms. The government must set tax policy to collect at least $G > 0$, meaning that

$$R^{-1} \int [\psi \tau_{t,a}(w) + (1 - \psi) \tau_{t,n}(w)] dF(w, t) + R^{-1} \int \tau_{t,e}(w) dF^e(w, t) + \psi T_0 \geq R^{-1} G.$$

Here, $\tau_{t,a}$, $\tau_{t,n}$, and $\tau_{t,e}$ denote the taxes collected from about-to-exit, normal, and exiting firms, respectively. Subject to this constraint, the government maximizes the welfare of the firms. As in the static model, this can be written in Lagrangian form, with a multiplier on the taxation constraint. We will use χ_t to denote this multiplier. It has the same interpretation as in the static models: it corresponds to the marginal cost to firms of paying a marginal unit of taxes. For this reason, in our models, it will always be weakly greater than one (otherwise, no taxes would be collected).

In the first subsection, we will discuss models without commitment by the government, and in the second subsection, models with commitment by the government.

4.1 The Model with No Commitment, Exclusion, or Subsidies

In this subsection, we will discuss the model without commitment by the government or creditors, and with re-entry following default. We will also impose the assumption of no subsidies, for reasons that we will explain below. Because wealth is observable, the government is free to treat firms differently based on their level of wealth.

As in the static model, the government's mechanism in the current date will influence the distribution of firm wealth in the next date, and as a result influence the multiplier on the fundraising constraint, χ_{t+1} . We will construct a steady state equilibrium in which $\chi_t = 1$, which will allow us to neglect this effect. Put another way, we will guess and verify that such an equilibrium exists. In the discuss that follows, we will simply assume that $\chi = 1$, and ignore the effects of the current date's mechanism on the wealth distribution.

We will begin our description of the environment by describing the problem for exiting firms. For simplicity, and without loss of generality, we will write the problem assuming exiting firms cannot borrow. For these firms, the government's problem can be written recursively,

$$U_t^e(w_t) = \max_{d_{t,e} \geq 0, \tau_{t,e} \geq 0} R^{-1} (d_{t,e} + U_{t+1}^e(w_t^R, \chi_{t+1}) + \tau_{t,e}),$$

where $U_{t+1}^e(w)$ is the government's continuation value function, which the government takes as given, due to the lack of commitment. The government is constrained by the exiting firm's production function,

$$w_{t,e}^R = R w_t - d_{t,e} - \tau_{t,e},$$

and the no-default constraint,

$$d_{t,e} + V_{t+1}^e(w_{t,e}^R) \geq \max\{d_{t,e} + V_{t+1}^e(R(1-\omega)w_t - d_{t,e}), V_{t+1}^e(R(1-\omega)w_t)\},$$

where $V_{t+1}^e(w)$ is the firm's continuation value function. The mechanism is also constrained by the limit on dividends, $d_{t,e} \leq R(1-\omega)w_t$, and the requirement that dividends and taxes be weakly positive.

We will consider stationary equilibria of this model, in which, for all levels of wealth,

$$U_t^e(w_t) = U_{t+1}^e(w_t)$$

and

$$V_t^e(w_t) = R^{-1}(d_{t,e}^* + V_{t+1}^e(w_{t,e}^{R*})) = V_{t+1}^e(w_t),$$

where $d_{t,e}^*$ and $w_{t,e}^{R*}$ denote the dividend and continuation wealth under the optimal mechanism.

Intuitively, because this firm cannot create resources, the only question for this sequence of mechanisms is how to "divide the pie." The government (at least weakly) prefers to collect as much taxes as possible. However, the possibility of default limits how much the government can take. The following lemma summarizes these results.

Lemma 2. *If there exist steady-state equilibria of the model in which $\chi_t = 1$, then there exist equilibrium in which, for exiting firms,*

$$U^e(w_t) = w_t$$

and

$$V^e(w_t) = (1 - \tau_e)w_t,$$

for any $\tau_e \in [0, 1)$. These equilibria can be implemented with a tax on dividends at rate $\frac{\tau_e}{1-\tau_e}$.

Proof. See the appendix, 5.4. □

Using the results of this lemma, we now turn to the mechanism design problem for normal and about to exit firms. Both of these firms have the same standard production function described in Section 2, $f(k)$, with a strictly positive first-best capital level $k^* > 0$.

We will take the rate τ_e described in the previous lemma as given. In this model, the government designs at each date a direct revelation mechanism, for each level of wealth, that maximizes a weighted sum of the value function of the firm's shareholders and revenue raised. The objective of the mechanism design problem is to maximize

$$\begin{aligned} U_t(w_t) = \max & R^{-1}(1 - \psi)(d_{t,n} + U_{t+1}(w_{t,n}^R)) + R^{-1}\psi(d_{t,a} + (1 - \tau_e)w_{t,a}^R) \\ & + R^{-1}((1 - \psi)\tau_{t,n} + \psi\tau_{t,a} + \psi\tau_e w_{t,a}^R), \end{aligned}$$

where $d_{t,n}$ is the dividend of a normal firm, $d_{t,a}$ is the dividend of an “about to exit” firm, and the continuation wealth, taxes, and subsidies use the same convention. Here, $U_{t+1}(w)$ represents the government’s continuation value for a firm which was, at time t , normal, but at time $t + 1$ may be either normal or about to exit. We will use $V_{t+1}(w_t)$ to denote the equilibrium continuation value for such a firm. The no-default constraints are

$$d_{t,n} + V_{t+1}(w_{t,n}^R) \geq \max\{d_{t,n} + V_{t+1}(w_{t,n}^D - d_{t,n}), V_{t+1}(w_{t,n}^D)\},$$

$$d_{t,a} + (1 - \tau_e)w_{t,a}^R \geq \max\{d_{t,a} + (1 - \tau_e)(w_{t,a}^D - d_{t,a}), (1 - \tau_e)w_{t,a}^D\}.$$

There are also incentive-compatibility constraints. A normal firm can pretend to be about to exit, and then either default, prematurely exit, or repay its debt. The incentive compatibility constraints for the normal firm are

$$d_{t,n} + V_{t+1}(w_{t,n}^R) \geq d_{t,a} + V_{t+1}(w_{t,a}^R),$$

$$d_{t,n} + V_{t+1}(w_{t,n}^R) \geq d_{t,a} + (1 - \tau_e)w_{t,a}^R,$$

$$d_{t,n} + V_{t+1}(w_{t,n}^R) \geq d_{t,a} + V_{t+1}(w_{t,a}^D - d_{t,a}).$$

The normal firm also has the option to become exiting (although this will be redundant with other constraints), and therefore we must have

$$d_{t,n} + V_{t+1}(w_{t,n}^R) \geq d_{t,n} + (1 - \tau_e)w_{t,n}^R.$$

The constraint for deviating to default is redundant– any allocation which does not induce an about to exit firm to default will cause a normal firm to prefer non-defaulting deviations to defaulting deviations.

Similarly, for the firm that is about to exit, there are incentive compatibility constraints that involve deviating to the normal firm’s allocation. These constraints are

$$d_{t,a} + (1 - \tau_e)w_{t,a}^R \geq d_{t,n} + (1 - \tau_e)w_{t,n}^R$$

and

$$d_{t,a} + (1 - \tau_e)w_{t,a}^R \geq d_{t,n} + (1 - \tau_e)(w_{t,n}^D - d_{t,n}).$$

Again, the constraint that involves deviating to default is redundant.

There is also a participation (or individual rationality) constraint. The firms can disregard entirely the government’s mechanism, and default. If firms attempt this, the government can punish them by assigning them infinite taxes and prevent them from paying dividends. Because taxes are senior to debt payments, the firms will not be able to raise outside funding, and therefore

will have capital equal to wealth (the initial budget constraint binds). Define

$$\hat{w}_t^D = f(w_t) + (1 - \omega)(1 - \delta)w_t.$$

To induce the participation of normal firms, the government must satisfy

$$d_{t,n} + V_{t+1}(w_{t,n}^R) \geq \max\{V_{t+1}(\hat{w}_t^D), (1 - \tau_e)\hat{w}_t^D\},$$

and to induce the participation of about-to-exit firms, the government must satisfy

$$d_{t,a} + (1 - \tau_e)w_{t,a}^R \geq (1 - \tau_e)\hat{w}_t^D.$$

It is straightforward to observe that the initial budget constraint binds (the marginal product of capital is always positive), and therefore capital is always weakly greater than wealth. In any equilibrium in which capital is greater than wealth, if the no-default constraints are satisfied, and normal firms do not choose to exit, the participation constraints will be satisfied, and are therefore redundant.

The government's mechanism must also satisfy the initial budget constraints,

$$w_t + r_{t,n} \geq k_{t,n},$$

$$w_t + r_{t,a} \geq k_{t,a}.$$

In this model, we assume that creditors cannot commit to lending. As a result, there are interim participation constraints,

$$r_{t,n} \leq R^{-1}b_{t,n},$$

$$r_{t,a} \leq R^{-1}b_{t,a}.$$

There is also an ex-ante participation constraint,

$$(1 - \psi)r_{t,n} + \psi r_{t,a} \leq (1 - \psi)R^{-1}b_{t,n} + \psi R^{-1}b_{t,a},$$

but this constraint will always be satisfied if the creditor's interim participation constraints are satisfied.

Finally, recall that $(k_{t,n}, k_{t,a}, r_{t,n}, r_{t,a}, b_{t,n}, b_{t,a}, d_{t,n}, d_{t,a}, \tau_{t,n}, \tau_{t,a})$ are all weakly positive, and that dividends are bounded above.

The initial budget and outside creditor participation constraints illustrate an important difference, in this model, between the government and private creditors. Suppose, for the sake of argument, that the private creditors enter a long-term contract with the firms. If the firm

defaults, the private creditors can recover whatever assets are left behind, but cannot make any claims against the new firm. This assumption is at the center of proposition one of Rampini and Viswanathan (2010), which demonstrates that any incentive-compatible long term contract with outside creditors can be implemented as a sequence of short-term contracts. In the mechanism design problem described above, we have imposed the assumption that contracts with outside creditors are short-term, but this is without loss of generality.

In contrast, if the firm defaults and creates a new firm, the government will receive taxes either way. The new firm may be more or less wealthy, and pay more or less taxes, but it cannot dodge taxes entirely. Implicitly, we are assuming that there is no “alternative government” the firm can contract with. As a result, the government can extract more resources from the firm than private creditors. Relatedly, creditors must at least break even on each firm they lend to, conditional on observables (wealth). The government, in contrast, might be willing to give money to some firms, and tax others. As a result, if the government were able to subsidize the firms, it could circumvent the financial frictions. To prevent the government from circumventing the financial constraints, we will assume that the government cannot put money into firms.

The mechanism design problem takes the continuation value functions $U_{t+1}(w)$ and $V_{t+1}(w)$ as given, due to the government’s lack of commitment. In equilibrium, however, the value functions are stationary, $U_t(w) = U_{t+1}(w)$ and $V_t(w) = V_{t+1}(w)$, where

$$V_t(w_t) = R^{-1}\psi(d_{t,a}^* + (1 - \tau_e)w_{t,a}^{R*}) + R^{-1}(1 - \psi)(d_{t,n}^* + V_{t+1}(w_{t,n}^{R*})).$$

Here, “starred” variables denote the variable’s value under the optimal mechanism design, given initial wealth level w_t . We will consider equilibria in which the firm’s value function is weakly increasing in wealth.

Finally, we discuss entry. When the firm enters, an entrepreneur with initial wealth w_0 can choose how much of this wealth to put into firm, after paying the lump sum tax, or to not enter at all. That is, the entrepreneur solves

$$\max_{w \in [0, w_0 - T_0]} \{w_0 - T_0 - w + V(w), w_0\}.$$

The government maximizes a weighted combination of the entering entrepreneur’s utility and tax revenues (both the lump sum revenues and future revenues),

$$\max_{T_0 \geq 0} \{w_0 - w(T_0) + U(w), w_0\}.$$

Although this problem looks formidable, the optimal tax structure is quite simple. Suppose,

in the mechanism design problem above, taxes are proportional to dividends,

$$\tau_{t,i} = \frac{\tau_e}{1 - \tau_e} d_{t,i},$$

for $i \in \{a, n\}$. In this case, we will have

$$d_{t,i} + \tau_{t,i} = \frac{1}{1 - \tau_e} d_{t,i},$$

and it follows immediately that, for all levels of wealth,

$$U_t(w) = \frac{1}{1 - \tau_e} V_t(w).$$

Consider the mechanism design problem facing the firms if they are taxed on dividends paid. The firm's objective, in this mechanism design problem, is to maximize

$$\tilde{V}_t(w_t) = \max_{\dots} R^{-1}(1 - \psi)(d_{t,n} + \tilde{V}_{t+1}(w_{t,n}^R)) + R^{-1}\psi(d_{t,a} + (1 - \tau_e)w_{t,a}^R),$$

subject to the exact same constraints described above (with \tilde{V}_t replacing V_t). If we define

$$\tilde{U}(w_t) = \frac{1}{1 - \tau_e} \tilde{V}(w_t),$$

it is straightforward to observe that the same mechanism is optimal from the firm and government's perspective.

This argument shows that a dividend tax can implement the optimal mechanism, conditional on desiring an allocation in which taxes are proportional to dividends. We finish the argument by showing that, conditional on using a dividend tax in the future, such an allocation is optimal today.

Proposition 3. *In the model without commitment, if the revenue required is sufficiently small, there is a steady-state equilibrium in which $\chi_t = 1$ and the optimal sequence of mechanisms can be implemented by a dividend tax.*

Proof. See the appendix, 5.5. □

Intuitively, the firm's choice to issue dividends reveals that it is not financially constrained. As a result, this is also the ideal time to tax the firm. Moreover, because the firm must pay dividends eventually, to return money to its owners, there is no way for it to evade the dividend tax. As a result, a dividend tax is the ideal way for the government to extract resources from the "right" firms.

Returning to the problem of an entering firm, note that if the entrepreneur has sufficiently little wealth, lump sum taxes will crowd out wealth in the firm one-to-one. Moreover, for sufficiently low levels of wealth, the government can gain more taxing firms later, when they are unconstrained, than initially, when they are constrained. The following lemma summarizes this result.

Corollary 1. *For sufficiently low levels of initial wealth w_0 , the optimal tax policy without commitment, exclusion, or subsidies can be implemented as a dividend tax, with no lump sum taxes. In this case, any lump sum taxes are strictly sub-optimal.*

Proof. See the appendix, 5.6. □

These results are closely related to the “new,” or “trapped equity,” view of dividend taxation (Auerbach (1981, 2002); Korinek and Stiglitz (2009)), in which, because of the lack of equity issuance, dividend taxes do not distort investment decisions by the firm. Consistent with this view, the only distortion that arises in the model is a reduction in the amount of wealth initially invested in the firm. This distortion causes the firm, in the beginning of its life, to use less than the first-best level of capital, and ultimately reduces steady-state output. However, any other taxation scheme would also have these effects, by reducing $V_w(w)$ for entering firms.

4.2 The Model with Commitment, Exclusion, and Subsidies

In the model with commitment, firms can be induced to avoid defaulting by a promise that they will be given high continuation values, and punished by receiving low continuation values if they do default. We will suppose that both the government and outside creditors are capable of this sort of commitment. In the government’s case, it really is commitment– the government can condition its taxes and subsidies on past behavior. In the creditor’s case, this also involves exclusion– a defaulting firm cannot reenter. It seems natural, in this case, to assume that if the government can tax and subsidize based on a firm’s history, including whether or not it defaulted in the past, it can also enforce creditor rights via exclusion.

We will begin by studying the problem of outside creditors. We will assume, as is standard in the literature (Golosov, Tsyvinski and Werquin (2016)), that the creditors promise a continuation value, V_t , to the firms, and run a direct revelation mechanism to elicit firm types and maximize their own payments. When a new firm enters the economy, it maximizes this promise, subject to the creditor’s participation constraint. Because we have setup our model without any kind of adjustment costs in capital, and creditors have full commitment, it is without loss of generality to suppose that the firm’s resources are entirely removed at the end of each period, and put back in at the beginning of the next period. Because the creditors are assumed to have an unlimited

supply of resources and full commitment, it follows that there is no state variable, aside from the promised continuation value V_t , in the creditor's problem.

We will begin by considering exiting firms. Creditors have promised V_t^e . It follows that, upon exit, creditors must simply pay the firm's shareholders V_t^e as a dividend. For other firms, creditors run a direct revelation mechanism, whose objective is

$$C(V_t) = \max_{\dots} (1 - \psi)(R^{-1}b_{t,n} - r_{t,n}) + \psi(R^{-1}b_{t,a} - r_{t,a}) - R^{-1}\psi V_t^e + R^{-1}(1 - \psi)C(V_{t+1}).$$

The direct revelation mechanism of the creditors must satisfy the no-default constraints (with exclusion),

$$d_{t,n} + V_{t+1} \geq w_{t,n}^D$$

$$d_{t,a} + V_t^e \geq w_{t,a}^D,$$

the incentive compatibility constraint for a normal firm pretending to be an about to exit firm and exiting,

$$d_{t,n} + V_{t+1} \geq d_{t,a} + V_t^e,$$

the incentive compatibility constraint for a normal firm pretending to be an about to exit firm and defaulting,

$$d_{t,n} + V_{t+1} \geq w_{t,a}^D$$

the incentive compatibility constraint for an about to exit firm pretending to be a normal firm and defaulting,

$$d_{t,a} + V_t^e \geq w_{t,n}^D,$$

the promise-keeping constraint,

$$V_t \leq R^{-1}(1 - \psi)(d_{t,n} + V_{t+1}) + R^{-1}\psi(d_{t,a} + V_t^e),$$

and the initial budget constraints (for a firm with zero wealth),

$$r_{t,n} \geq k_{t,n},$$

$$r_{t,a} \geq k_{t,a},$$

and the resource constraints arising from the production functions,

$$d_{t,n} + b_{t,n} \leq f(k_{t,n}) + (1 - \delta)k_{t,n},$$

$$d_{t,a} + b_{t,a} \leq f(k_{t,a}) + (1 - \delta)k_{t,a}.$$

At the beginning of the period, the firm could exit the mechanism. But, having no wealth, it would receive a value of zero. As a result, we must have $V_{t+1} \geq 0$ and $V_t^e \geq 0$. The other choice variables $(r_{t,n}, r_{t,a}, k_{t,n}, k_{t,a}, d_{t,n}, d_{t,a}, b_{t,n}, b_{t,a})$ must also be greater than zero.

When entering, the firm with initial wealth w_0 , facing lump sum taxes T_0 , solve

$$\max_{w \leq w_0 - T_0, V_0 \geq 0} \{V_0 + w_0 - w - T_0, w_0\}$$

subject to the outside creditors participation constraint,

$$w + C(V_0) \geq 0.$$

Now consider the government's problem. The government can replicate, through taxes and subsidies, anything that the creditors can accomplish, and the creditors can implement anything the government can accomplish. As a result, a government that needs to raise at least $\psi^{-1}R^{-1}G > 0$ in expectation from entering firms would like, if it could control entering firm's entry and initial investment decisions, which it cannot, to solve

$$\max_{w \leq w_0 - T_0, V_0 \geq 0} V_0 + w_0 - w - T_0$$

subject to

$$w + C(V_0) + T_0 \geq \psi^{-1}R^{-1}G.$$

Suppose the firm will choose to enter if the government assigns a lump-sum tax $T_0 = \psi^{-1}R^{-1}G$. It would follow immediately that this lump sum tax, and no other taxes, will induce the creditors to implement the optimal mechanism design. If G is sufficiently small, the firm will indeed choose to enter, and this taxation scheme will be optimal. Of course, if firms are not financially constrained, this is not a particularly interesting result. The proposition below shows that, in this model, firms can be financially constrained, and yet the optimal tax policy is a lump sum tax on entry.

Proposition 4. *In the model with commitment and exclusion, under the optimal mechanism, there exists a non-empty interval of "net wealth" $w_0 - \psi^{-1}R^{-1}G$ in which firms enter but are financially constrained, meaning that they do not achieve the first-best level of investment upon entry, and the optimal tax policy in this case is a lump sum tax. In this case, a dividend tax is strictly sub-optimal.*

Proof. See the appendix, 5.7. □

The intuition for the result is standard in the literature. In one-sided commitment problems, it is generally optimal to defer "consumption" for the agent with the commitment problem as much as possible. In this model, there is no force in the other direction— the firm's shareholders

are not less patient than the government, for example. As a result, raising taxing at any time after the firm enters is sub-optimal. Doing so would reduce continuation values, which would cause constraints to tighten. Another way to view this result is through a complete markets lens. Because outside creditors and shareholders can enter into an optimal mechanism, they efficiently share risk, and there is nothing the government can do to improve this risk-sharing.

The contrast between 4 and 1 is striking. If the government can commit and exclude, it should tax firms right away, whereas if the government cannot commit, exclude, or subsidize, it should implement a dividend tax. Moreover, the optimal policy in one case is strictly sub-optimal in the other. This is not an artefact of commitment and exclusion eliminating financial frictions– in both models, assuming the entrepreneur’s initial wealth is sufficiently low, the firm does not achieve the first best level of investment initially. This contrast is also troubling because empirical work (Li, Whited and Wu (2016)) that has attempted to determine which of these types of models of financial frictions is most applicable to actual firms has reached inconclusive results.

5 Conclusion

We have provided a normative analysis of optimal corporate taxes. We emphasize the interaction between taxation policy and financial frictions, demonstrating that, in certain models in which the government lacks commitment, dividend taxes implement the optimal tax policy. Under commitment, we reach very different policy conclusions, finding that lump sum taxes at firm entry are optimal. We argue that these two results are consistent with the principle that taxes should avoid exacerbating financial frictions, and demonstrate the importance to public policy of research determining the nature of these constraints.

Appendix

References

- Allen, Franklin, Antonio E Bernardo, and Ivo Welch.** 2000. "A theory of dividends based on tax clienteles." *The Journal of Finance*, 55(6): 2499–2536.
- Auerbach, Alan J.** 1981. "TAX INTEGRATION AND THE "NEW VIEW" OF THE CORPORATE TAX: A 1980'S PERSPECTIVE." 21–27, JSTOR.
- Auerbach, Alan J.** 2002. "Taxation and corporate financial policy." *Handbook of public economics*, 3: 1251–1292.
- 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.
- Chetty, Raj, and Emmanuel Saez.** 2010. "Dividend and corporate taxation in an agency model of the firm." *American Economic Journal: Economic Policy*, 2(3): 1–31.
- 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.
- 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, Kenneth L.** 1985. "Redistributive taxation in a simple perfect foresight model." *Journal of public Economics*, 28(1): 59–83.
- 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.
- Korinek, Anton, and Joseph E Stiglitz.** 2009. "Dividend taxation and intertemporal tax arbitrage." *Journal of Public Economics*, 93(1): 142–159.
- 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.

- 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.
- Woodford, Michael.** 2003. *Interest and Prices: Foundations of a Theory of Monetary Policy*. Princeton University Press.

Proofs

5.1 Proof of Lemma 1

The government solves

$$\int \int R^{-1} \{d_1(w_1, \theta_1) + w_1^R(w_1, \theta_1)\} dF(w_1, \theta_1)$$

subject to the constraint, for each level of wealth w_1 and type θ_1 , that

$$d_1(w_1, \theta_1) + w_1^R(w_1, \theta_1) \geq w_1^D(k_1(w_1, \theta_1), \theta_1),$$

the initial budget constraints,

$$r(w_1, \theta_1) + w_1 \geq k_1(w_1, \theta_1),$$

the creditor participation constraints,

$$r(w_1, \theta_1) \leq R^{-1}b(w_1, \theta_1),$$

the limit on dividends,

$$d_1(w_1, \theta_1) \leq w_1^D(k_1(w_1, \theta_1), \theta_1),$$

the fund-raising constraint,

$$\int \int R^{-1} \tau_1(w_1, \theta_1) dF(w_1, \theta_1) \geq R^{-1}G,$$

and the postivity constraints for $k_1, d_1, b_1, \tau_1, r_1$.

Using the definitons of the continuation and default wealth,

$$d_1(w_1, \theta_1) + w_1^R(w_1, \theta_1) = f(k_1(w_1, \theta_1), \theta_1) + (1 - \delta) k_1(w_1, \theta_1) - b_1(w_1, \theta_1) - \tau_1(w_1, \theta_1).$$

Moreover, the no-default constraint simplifies to

$$\omega (1 - \delta) k_1(w_1, \theta_1) \geq b_1(w_1, \theta_1) + \tau_1(w_1, \theta_1).$$

As a result, $d_1(w_1, \theta_1)$ enters only in the 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.

If the initial budget constraint does not mind, the government can increase $k_1(w_1, \theta_1)$, increasing the objective and relaxing the no-default constraint. Therefore, it must bind. By essentially the same argument, the creditor participation constraint must bind.

The problem can therefore be simplified to

$$\max \int \int R^{-1} \{f(k_1(w_1, \theta_1), \theta_1) + (1 - \delta) k_1(w_1, \theta_1) - b_1(w_1, \theta_1) - \tau_1(w_1, \theta_1)\} dF(w_1, \theta_1)$$

subject to

$$R^{-1}b_1(w_1, \theta_1) + w_1 = k_1(w_1, \theta_1),$$

$$\omega(1-\delta)k_1(w_1, \theta_1) \geq b_1(w_1, \theta_1) + \tau_1(w_1, \theta_1),$$

$$\int \int R^{-1}\tau_1(w_1, \theta_1)dF(w_1, \theta_1) \geq R^{-1}G,$$

and the positivity constraints. This problem has entirely affine constraints and a concave objective function, and therefore the (infinite dimensional analog of the) KKT conditions are necessary and sufficient. Substituting the equality constraint and the positivity constraint for b_1 yields the result.

5.2 Proof of Proposition 1

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

The problem, if $\chi = 1$, is

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

subject to the constraint that

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

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 assume that $w_1 \geq 0$ for all firms, and therefore the capital positivity constraint is redundant. The FOCs of the Lagrangian are

$$-\frac{1}{R - \omega(1-\delta)}\mu + \nu = 0,$$

$$R^{-1}\{f'(k_1, \theta_1) + (1-\delta) - R\} - \mu + \phi = 0.$$

Note that $\mu > 0$ implies $\nu > 0$ and therefore $\tau_1 = 0$. If $\tau_1 = 0$, $\mu > 0$ and $\phi > 0$ are mutually exclusive, and therefore $\mu > 0$ implies $\phi = 0$.

Because

$$f'(k_1, \theta_1) + (1-\delta) - R \geq 0,$$

with equality if and only if $k_1 \geq k^*(\theta_1)$, it follows that $k_1 < k^*(\theta_1)$ implies $\mu > 0$, $\nu > 0$, $\tau_1 = 0$, and $\phi = 0$. This case requires that

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

If $k_1 \geq k^*(\theta_1)$, then we must have $\mu = \phi = 0$, and

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

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

$$0 \leq \tau_1 \leq \omega(1-\delta)w_1.$$

The proposed functional form satisfies the restrictions on τ_1 in both cases, and raises positive revenue, proving the result.

5.3 Proof of Proposition 2

The government's problem is to solve

$$\max_{\dots} R^{-1} \int \left\{ d_0(\theta_1) + U_1(w_0^R, \theta_1) \right\} dF(\theta_1 | \theta_0),$$

subject to the constraints

$$k_0(\theta_1) \leq w_0 + r_0(\theta_1),$$

$$r_0(\theta_1) \leq R^{-1}b_0(\theta_1),$$

$$\int R^{-1}\tau_0(\theta_1) dF(\theta_1 | \theta_0) \geq R^{-1}G_0 > 0,$$

$$b_0(\theta_1) + \tau_0(\theta_1) \leq \omega(1 - \delta)k_0(\theta_1),$$

$$d_0(\theta_1) + V_1(w_0^R(\theta_1), \theta_1) \geq d_0(\theta'_1) + V_1(w_0^R(\theta'_1), \theta_1),$$

and the positivity constraints for the choice variables s_0, k_0, d_0, τ_0 , the limit on dividends,

$$d_0(\theta_1) \leq f(k_0(\theta_1), \theta_0) + (1 - \omega)(1 - \delta)k_0(\theta_1)$$

and the definition

$$w_0^R(\theta_1) = f(k_0(\theta_1), \theta_0) + (1 - \delta)k_0(\theta_1) - d_0(\theta_1) - \tau_0(\theta_1).$$

Note that it is without loss of generality to assume that $r_0(\theta_1) = R^{-1}b_0(\theta_1)$.

First, we discuss the feasibility, ignoring incentive compatibility, of achieving the first-best level of capital in the second date for all types. This is easiest to accomplish without paying dividends or taxes. Suppose that

$$f(k_0(\theta_1), \theta_0) + (1 - \delta)k_0(\theta_1) - b_0(\theta_1) \geq \bar{w}(\theta_1),$$

where $\bar{w}(\theta_1) = (1 - R^{-1}\omega(1 - \delta))k^*(\theta_1)$. Therefore, we would need

$$f(k_0(\theta_1), \theta_0) + (1 - \delta)k_0(\theta_1) - Rk_0(\theta_1) + Rw_0 \geq \bar{w}(\theta_1).$$

The initial borrowing constraint requires that

$$k_0(\theta_1) \leq \frac{w_0}{1 - R^{-1}\omega(1 - \delta)}.$$

By the concavity of the production function,

$$f(k_0(\theta_1), \theta_0) + (1 - \delta - R)k_0(\theta_1) \leq f(k^*(\theta_0), \theta_0) + (1 - \delta - R)k^*(\theta_0).$$

Therefore, to achieve first-best, we must have

$$f(k^*(\theta_0), \theta_0) + (1 - \delta - R)k^*(\theta_0) + R w_0 \geq \bar{w}(\theta_1)$$

for all types. Assume this is not the case, so that at least one type must be constrained in the second date. Assume, in addition, that it is strictly feasible for the lowest marginal product type,

$$\bar{w}(1) > f(k^*(\theta_0), \theta_0) + (1 - \delta - R)k^*(\theta_0) + R w_0 > \bar{w}(0)$$

We begin by considering a related problem with the multiplier χ_0 on the fund-raising constraint. We also define the variables

$$y_0(\theta_1) = d_0(\theta_1) + \tau_0(\theta_1) + b_0(\theta_1),$$

$$w_0(\theta_1) = f(k_0(\theta_1), \theta_0) + (1 - \delta)k_0(\theta_1).$$

By the fact that the marginal product of capital is always positive, the relationship between $k_0(\theta_1)$ and $w_0(\theta_1)$ is one-to-one, and $w_0(\theta_1) = 0$ is equivalent to $k_0(\theta_1) = 0$.

The related problem is

$$\max_{\dots} R^{-1} \int \{y_0(\theta_1) - b_0(\theta_1) + U_1(w_0(\theta_1) - y_0(\theta_1), \theta_1) + (\chi - 1)\tau_0(\theta_1)\} dF(\theta_1 | \theta_0),$$

subject to the constraints

$$k_0(\theta_1) \leq w_0 + R^{-1}b_0(\theta_1),$$

$$\int (R^{-1}\tau_0(\theta_1)) dF(\theta_1 | \theta_0) \geq G_0 > 0,$$

$$\tau_0(\theta_1) + b_0(\theta_1) \leq \omega(1 - \delta)k_0(\theta_1),$$

$$\tau_0(\theta_1) + b_0(\theta_1) \leq y_0(\theta_1) \leq f(k_0(\theta_1), \theta_0) + (1 - \omega)(1 - \delta)k_0(\theta_1) + \tau_0(\theta_1) + b_0(\theta_1)$$

$$y_0(\theta_1) - \tau_0(\theta_1) + V_1(w_0(\theta_1) - y_0(\theta_1), \theta_1) \geq y_0(\theta'_1) - \tau_0(\theta'_1) + V_1(w_0(\theta'_1) - y_0(\theta'_1), \theta_1),$$

$$w_0(\theta_1) = f(k_0(\theta_1), \theta_0) + (1 - \delta)k_0(\theta_1).$$

The function V_1 is strictly increasing in wealth. It follows that if $d_0(\theta) > d_0(\theta')$, the IC constraints require that $w_0(\theta) - y_0(\theta) < w_0(\theta') - y_0(\theta')$. This argument also applies in reverse. As a result, there exists a function $D(w, y)$ such that

$$d_0(\theta) = D(w_0(\theta), y_0(\theta)),$$

and another function, $T(w, y) = y - D(w, y)$, such that

$$b_0(\theta) + \tau_0(\theta) = T(w_0(\theta), y_0(\theta)).$$

Anywhere the derivatives $w'_0(\theta)$, $y'_0(\theta)$, $T_w(w_0(\theta), y_0(\theta))$, and $T_y(w_0(\theta), y_0(\theta))$ all exist, there

is a local IC constraint,

$$\begin{aligned} & [1 - V_{1,w}(w_0(\theta) - y_0(\theta), \theta) - T_y(w_0(\theta), y_0(\theta))]y'_0(\theta) + \\ & [V_{1,w}(w_0(\theta) - y_0(\theta), \theta) - T_w(w_0(\theta), y_0(\theta))]w'_0(\theta) = 0. \end{aligned}$$

As a result, at any point satisfying differentiability and the local IC,

$$\frac{d}{d\theta} [y_0(\theta) - T(w_0(\theta), y_0(\theta)) + V_1(w_0(\theta) - y_0(\theta), \theta)] = V_{1,\theta}(w_0(\theta) - y_0(\theta), \theta).$$

Suppose that there exist firms θ, θ' , with $\theta' > \theta$ and $d_0(\theta) < d_0(\theta')$. It follows by the argument above that $w(\theta) - y(\theta) > w(\theta') - y(\theta')$. The IC constraints require that

$$V_1(w(\theta) - y(\theta), \theta) - V_1(w(\theta') - y(\theta'), \theta) \geq d_0(\theta') - d_0(\theta) \geq V_1(w(\theta) - y(\theta), \theta') - V_1(w(\theta') - y(\theta'), \theta').$$

This can be rewritten in integral form as

$$\int_{w(\theta')-y(\theta')}^{w(\theta)-y(\theta)} V_{1,w}(x, \theta) dx \geq d_0(\theta') - d_0(\theta) \geq \int_{w(\theta')-y(\theta')}^{w(\theta)-y(\theta)} V_{1,w}(x, \theta') dx.$$

Combining these inequality,

$$\int_{w(\theta')-y(\theta')}^{w(\theta)-y(\theta)} (V_{1,w}(x, \theta) - V_{1,w}(x, \theta')) dx \geq 0.$$

The function V_1 has a cross-partial $V_{1,w,\theta}$ that is weakly positive and exists almost everywhere, and is strictly positive if $V_{1,w} > 1 - \tau_e$. Expressed as an integral (which is permissible by the existence a.e. of the derivative),

$$- \int_{w(\theta')-y(\theta')}^{w(\theta)-y(\theta)} \int_{\theta}^{\theta'} V_{1,w\theta}(x, \hat{\theta}) dx \geq 0.$$

It follows that this must hold with equality, and that $V_{1,w}(x, \hat{\theta}) = 1 - \tau_e$ over the domain of integration. In this case, it follows that both IC constraints hold with equality.

In this case, consider the alternative policy of switching the allocations of an equal mass of firms of type θ and θ' . By construction, the welfare of the two types of firms is unchanged, the IC constraints for θ deviating to θ' and vice versa are still satisfied, and the IC constraints for all other types are unconstrained. Moreover, the government collects the same amount of taxes. All other constraints are also satisfied. The government's objective function is under unchanged, because both types of firms are unconstrained at the lower level of continuation wealth. Therefore, it is without loss of generality to perform this switch. Because multiple allocations for a single type cannot relax any constraints, it follows that it is without loss of generality to assume that $d_0(\theta)$ is decreasing in θ .

If any firm pays zero dividends, define

$$\theta^* = \inf\{\theta \in [0, 1] : d_0(\theta) = 0\}.$$

If no such firm exists, define $\theta^* = 1$. By construction, $d_0(\theta) = 0$ if $\theta < \theta^*$ and $d_0(\theta) > 0$ if $\theta < \theta^*$.

Note that, for all firms with $d_0(\theta) = 0$, the quantity $w_0(\theta) - y_0(\theta)$ must be identical, by the IC constraint. Moreover, by the IC constraint, the quantity

$$y_0(\theta) + V_1(w_0(\theta) - y_0(\theta), \theta) - T(w_0(\theta), y_0(\theta))$$

must be continuous in θ . Therefore, it will always be the case that

$$\begin{aligned} y(\theta^*) + V_1(w_0(\theta^*) - y_0(\theta^*), \theta^*) - T(w_0(\theta^*), y_0(\theta^*)) = \\ y(1) + V_1(w_0(1) - y_0(1), \theta^*) - T(w_0(1), y_0(1)), \end{aligned}$$

and, if any firm pays zero dividends, then

$$y(1) + V_1(w_0(1) - y_0(1), \theta^*) - T(w_0(1), y_0(1)) = V_1(w_0(1) - y_0(1), \theta^*).$$

It follows, if the derivatives of y_0, w_0, T exist almost everywhere, that

$$\begin{aligned} [y_0(\theta) + V_1(w_0(\theta) - y_0(\theta), \theta) - T(w_0(\theta), y_0(\theta))] = \\ [y(1) + V_1(w_0(1) - y_0(1), \theta^*) - T(w_0(1), y_0(1))] - \int_{\theta}^{\theta^*} V_{1,\theta}(w_0(\hat{\theta}) - y_0(\hat{\theta}), \hat{\theta}) d\hat{\theta}. \end{aligned}$$

Lastly, note that we must have

$$T(w_0(\theta_1), y_0(\theta_1)) \leq y_0(\theta_1) \leq w_0(\theta_1).$$

Define a relaxed problem using only the positivity constraints for w_0 , the limit above, the local IC constraints, and the constraint that $T(w_0(1), y_0(1)) = T^* \geq 0$, and the no-default constraint. Let $w^* = w_0(1) - y_0(1)$. Additionally, partition the objective function into the types above and below θ^* . The objective function can be written as

$$\begin{aligned} \max_{w^* \geq 0, \theta^* \in [0, 1], y_0(\theta), w_0(\theta), k_0(\theta), s_0(\theta)} R^{-1} \int_{\theta^*}^1 \{ \chi(w_0(\theta_1) - w^*) - \chi b_0(\theta_1) + U_1(w^*, \theta_1) \} dF(\theta_1 | \theta_0) \\ + R^{-1} \int_0^{\theta^*} \{ y_0(\theta_1) + (\chi - 1)T_0(\theta_1) + V_1(w_0(\theta_1) - y_0(\theta_1), \theta_1) \} dF(\theta_1 | \theta_0) \\ + R^{-1} \int_0^{\theta^*} (U_1(w_0(\theta_1) - y_0(\theta_1), \theta_1) - V_1(w_0(\theta_1) - y_0(\theta_1), \theta_1) - \chi b_0(\theta_1)) dF(\theta_1 | \theta_0). \end{aligned}$$

Subtracting and adding,

$$\begin{aligned}
& \max_{\dots} R^{-1} \int_{\theta^*}^1 \{\chi(w_0(\theta_1) - w^*) - \chi b_0(\theta_1) + U_1(w^*, \theta_1)\} dF(\theta_1 | \theta_0) \\
& R^{-1}(1 - \chi) \int_0^{\theta^*} \{y_0(\theta_1) - T_0(\theta_1) + V_1(w_0(\theta_1) - y_0(\theta_1), \theta_1)\} dF(\theta_1 | \theta_0) \\
& R^{-1}\chi \int_0^{\theta^*} \{y_0(\theta_1) + V_1(w_0(\theta_1) - y_0(\theta_1), \theta_1)\} dF(\theta_1 | \theta_0) \\
& + R^{-1} \int_0^{\theta^*} (U_1(w_0(\theta_1) - y_0(\theta_1), \theta_1) - V_1(w_0(\theta_1) - y_0(\theta_1), \theta_1) - \chi b_0(\theta_1)) dF(\theta_1 | \theta_0).
\end{aligned}$$

which simplifies to

$$\begin{aligned}
& \max_{\dots} R^{-1} \int_{\theta^*}^1 \{\chi(w_0(\theta_1) - w^*) - \chi b_0(\theta_1) + U_1(w^*, \theta_1)\} dF(\theta_1 | \theta_0) \\
& + R^{-1}(1 - \chi)[y(1) + V_1(w_0(1) - y_0(1), \theta^*) - T^*]F(\theta^* | \theta_0) \\
& + R^{-1}(\chi - 1) \int_0^{\theta^*} dF(\theta_1 | \theta_0) \int_{\theta_1}^{\theta^*} V_{1,\theta}(w_0(\theta) - y_0(\theta), \theta) d\theta \\
& + R^{-1}\chi \int_0^{\theta^*} \{y_0(\theta_1) + V_1(w_0(\theta_1) - y_0(\theta_1), \theta_1)\} dF(\theta_1 | \theta_0) \\
& + R^{-1} \int_0^1 \{\tau_e \max(w_0(\theta_1) - y_0(\theta_1) - \bar{w}(\theta_1), 0) - \chi b_0(\theta_1)\} dF(\theta_1 | \theta_0).
\end{aligned}$$

Note also that

$$R^{-1} \int_0^{\theta^*} dF(\theta_1 | \theta_0) \int_{\theta_1}^{\theta^*} V_{1,\theta}(w_0(\theta) - y_0(\theta), \theta) d\theta = R^{-1} \int_0^{\theta^*} F(\theta_1 | \theta_0) V_{1,\theta}(w_0(\theta_1) - y_0(\theta_1), \theta_1) d\theta_1.$$

The remaining constraints are

$$0 \leq b_0(\theta_1) \leq T(w_0(\theta_1), y_0(\theta_1)) \leq \omega(1 - \delta)k_0(\theta_1)$$

$$0 \leq k_0(\theta_1) \leq w_0 + R^{-1}b_0(\theta_1),$$

$$w_0(\theta_1) = f(k_0(\theta_1), \theta_0) + (1 - \delta)k_0(\theta_1),$$

$$T(w_0(\theta_1), y_0(\theta_1)) \leq y_0(\theta_1) \leq w_0(\theta_1),$$

The objective and other constraints aside from the initial budget constraint are increasing in capital, and therefore the initial budget constraint binds. We can therefore write

$$0 \leq R(k_0(\theta_1) - w_0) \leq T(w_0(\theta_1), y_0(\theta_1)) \leq \omega(1 - \delta)k_0(\theta_1).$$

The problem is therefore

$$\begin{aligned}
& \max_{\dots} R^{-1} \int_{\theta^*}^1 \{ \chi(w_0(\theta_1) - w^*) - \chi R(k_0(\theta_1) - w_0) + U_1(w^*, \theta_1) \} dF(\theta_1 | \theta_0) \\
& \quad + R^{-1}(1 - \chi)[y(1) + V_1(w_0(1) - y_0(1), \theta^*) - T^*] F(\theta^* | \theta_0) \\
& \quad + R^{-1}(\chi - 1) \int_0^{\theta^*} F(\theta_1 | \theta_0) V_{1,\theta}(w_0(\theta_1) - y_0(\theta_1), \theta_1) d\theta_1 \\
& \quad + R^{-1} \chi \int_0^{\theta^*} \{ y_0(\theta_1) + V_1(w_0(\theta_1) - y_0(\theta_1), \theta_1) \} dF(\theta_1 | \theta_0) \\
& + R^{-1} \int_0^1 \{ \tau_e \max(w_0(\theta_1) - y_0(\theta_1) - \bar{w}(\theta_1), 0) - \chi R(k_0(\theta_1) - w_0) \} dF(\theta_1 | \theta_0),
\end{aligned}$$

subject to

$$0 \leq R(k_0(\theta_1) - w_0) \leq T(w_0(\theta_1), y_0(\theta_1)) \leq \omega(1 - \delta)k_0(\theta_1),$$

$$T(w_0(\theta_1), y_0(\theta_1)) \leq y_0(\theta_1) \leq w_0(\theta_1).$$

$$w_0(\theta_1) = f(k_0(\theta_1), \theta_0) + (1 - \delta)k_0(\theta_1).$$

Note that these constraints apply for $\theta_1 = 1$, and therefore apply to w^* , $T^* = y(\theta_1)$, and $k(\theta_1)$.

We can partition θ into the set for which dividends are zero and the set for which dividends are positive ($\theta > \theta^*$ and $\theta < \theta^*$). Taking first-order conditions in the set for which dividends are positive, the FOC for y_0 is (if $w_0(\theta_1) - y_0(\theta_1) \neq \bar{w}(\theta_1)$, avoiding the kink),

$$\begin{aligned}
& -R^{-1}(\chi - 1)F(\theta_1 | \theta_0)V_{1,w\theta}(w_0(\theta_1) - y_0(\theta_1), \theta_1) + \\
& R^{-1}\chi(1 - V_{1,w}(w_0(\theta_1) - y_0(\theta_1), \theta_1))dF(\theta_1 | \theta_0) - \\
& R^{-1}(U_{1,w}(w_0(\theta_1) - y_0(\theta_1), \theta_1) - V_{1,w}(w_0(\theta_1) - y_0(\theta_1), \theta_1))dF(\theta_1 | \theta_0) \geq 0,
\end{aligned}$$

with equality if $y_0(\theta_1) < w_0(\theta_1)$. If $y_0(\theta_1) = w_0(\theta_1)$, then we would have

$$\begin{aligned}
& -R^{-1}(\chi - 1)F(\theta_1 | \theta_0)V_{1,w\theta}(0, \theta_1) + \\
& R^{-1}\chi(1 - V_{1,w}(0, \theta_1)) \geq 0,
\end{aligned}$$

a contradiction for any $\chi \geq 1$. Therefore, the FOC holds with equality.

At the kink, the derivative $V_{1,w\theta}(\bar{w}(\theta_1), \theta_1)$ does not exist. In this case, we must have

$$R^{-1}(\chi - 1)F(\theta_1 | \theta_0)V_{1,w\theta}^-(\bar{w}(\theta_1), \theta_1) - R^{-1}\chi\tau_e dF(\theta_1 | \theta_0) \leq 0,$$

where $V_{1,w\theta}^-$ denotes the derivative in the decreasing wealth direction, and

$$R^{-1}(\chi - 1)\tau_e dF(\theta_1 | \theta_0) \leq 0.$$

If $\chi = 1$, the FOC is satisfied for any $w_0(\theta_1) - y_0(\theta_1) \geq \bar{w}(\theta_1)$, but not for any firm constrained at date one ($V_{1,w} > 1 - \tau_e$). Therefore, to pay positive dividends, the firm must be unconstrained at date one. Therefore, because there is a constrained firm, there is a firm paying zero dividends.

It follows that

$$y(1) + V_1(w_0(1) - y_0(1), \theta^*) - T(w_0(1), y_0(1)) = V_1(w^*, \theta^*).$$

In this case, we can simplify the objective function to

$$\begin{aligned} & \max_{\dots} R^{-1} \int_{\theta^*}^1 \{(w_0(\theta_1) - w^*) - R(k_0(\theta_1) - w_0) + U_1(w^*, \theta_1)\} dF(\theta_1 | \theta_0) \\ & + R^{-1} \int_0^{\theta^*} \{y_0(\theta_1) + U_1(w_0(\theta_1) - y_0(\theta_1), \theta_1) - R(k_0(\theta_1) - w_0)\} dF(\theta_1 | \theta_0). \end{aligned}$$

Recall that

$$y_0(\theta) + V_1(w_0(\theta) - y_0(\theta), \theta) - V_1(w^*, \theta^*) + \int_{\theta}^{\theta^*} V_{1,\theta}(w_0(\hat{\theta}) - y_0(\hat{\theta}), \hat{\theta}) d\hat{\theta} = T(w_0(\theta), y_0(\theta)).$$

For non-dividend payers, this simplifies to

$$V_1(w^*, \theta) - V_1(w^*, \theta^*) + \int_{\theta}^{\theta^*} V_{1,\theta}(w_0(\hat{\theta}) - y_0(\hat{\theta}), \hat{\theta}) d\hat{\theta} = 0,$$

which is always satisfied. We can treat $T(\theta) = T(w_0(\theta), y_0(\theta))$ as a choice variable, and write the constraints (for dividend payers)

$$0 \leq R(k_0(\theta_1) - w_0) \leq T(w_0(\theta_1), y_0(\theta_1)) \leq \omega(1 - \delta)k_0(\theta_1),$$

$$T(w_0(\theta_1), y_0(\theta_1)) \leq y_0(\theta_1) \leq w_0(\theta_1),$$

$$w_0(\theta_1) = f(k_0(\theta_1), \theta_0) + (1 - \delta)k_0(\theta_1),$$

and for non-dividend payers,

$$0 \leq R(k_0(\theta_1) - w_0) \leq w_0(\theta_1) - w^* \leq \omega(1 - \delta)k_0(\theta_1),$$

$$w_0(\theta_1) = f(k_0(\theta_1), \theta_0) + (1 - \delta)k_0(\theta_1),$$

and the constraints specialized to $\theta = 1$,

$$0 \leq R(k_0(1) - w_0) \leq T^* \leq \omega(1 - \delta)k_0(1),$$

$$T^* \leq w^*,$$

$$w^* = f(k_0(1), \theta_0) + (1 - \delta)k_0(1) - T^*.$$

In the absence of binding constraints, we would have (by optimality for w^*)

$$R^{-1} \int_{\theta^*}^1 (U_{1,w}(w^*, \theta_1) - 1) dF(\theta_1 | \theta_0) = 0,$$

implying all non-dividend-payers are unconstrained, contradicting the assumption that some

firms are constrained. The only constraint tightened by increasing w^* (and therefore k^*) is

$$R(k^* - w_0) \leq T^*,$$

and therefore this constraint must bind, implying that this firm is not taxed. Note that if any non-dividend payer is constrained in the next date (and one must be), then type $\theta = 1$ is constrained.

For the non-dividend payers, if the marginal product of capital at date zero is less than R , increasing the level of capital relaxes constraints and increase the objective function. It follows that all non-dividend payers achieve the first-best level of capital, $k_0(\theta_1) \geq k^*(\theta_0)$. Note that using more than the first-best level does not change the firm's utility. Because type $\theta = 1$ achieves the first-best level of capital and pays no taxes, it must be the case that this holds for all non-dividend payers.

For the dividend payers, who must be unconstrained, using the first-best level of capital will maximize their welfare if feasible. Conjecture that this is the case, implying that $w_0(\theta)$ is constant for all types. Then anywhere the $y_0(\theta)$ is decreasing (it cannot increase),

$$\tau_e = T_y(w_0(\theta), y_0(\theta)).$$

Because taxes are zero for $\theta \geq \theta^*$ (by the fact that $w_0(\theta)$ is continuous under this conjecture), we must have

$$T(w_0(\theta), y_0(\theta)) = \tau_e(y_0(\theta) - y_0(\theta^*)) + y_0(\theta^*),$$

implying taxes proportion to dividends in this region. To satisfy the constraints under this conjecture, we need

$$R(k^*(\theta_0) - w_0) \leq y_0(\theta^*) + \tau_e(y_0(\theta) - y_0(\theta^*)) \leq \omega(1 - \delta)k^*(\theta_0).$$

Note that

$$y_0(\theta^*) = b_0(\theta^*) = R(k^*(\theta_0) - w_0).$$

Therefore, the lower bound is always satisfied. The upper bound will be satisfied if

$$\tau_e(y_0(\theta) - y_0(\theta^*)) \leq R w_0 - (1 - R^{-1}\omega(1 - \delta))k^*(\theta_0).$$

This is feasible for some $y_0(\theta) > y_0(\theta^*)$, and therefore the conjecture is verified.

We have shown that this tax scheme is feasible. For the lowest type, supposing that this type is a dividend payer, we have

$$w_0(0) - y_0(0) = f(k^*(\theta_0), \theta_0) + (1 - \delta)k^*(\theta_0) - y_0(0).$$

It follows by the assumption of strict feasibility that $y_0(0) > y_0(\theta^*)$ is feasible, and therefore optimal. Note that this implies the government can raise strictly positive taxes.

Finally, we return to the original problem. Note that the limit on dividends is satisfied. It remains to consider the global IC constraints. Any non-dividend-payer would not switch to the type just below θ^* . Because it follows that they would not switch to a lower type than pays more dividends. Similarly, all unconstrained types below θ^* are indifferent, and therefore indifferent to

switching to type θ^* and paying no dividends. The global ICs follow.

The implementation with a dividend tax follows from the strict preference, under such a tax, of firms that will be constrained to not paying dividends, and the indifference of dividend-paying firms.

5.4 Proof of Lemma 2

Conjecture the equilibrium,

$$\begin{aligned} U_{t+1}^e(w_t) &= w_t \\ V_{t+1}^e(w_t) &= (1 - \tau_e)w_t \end{aligned}$$

The objective simplifies to

$$U_t^e(w_t) = \max_{d_{t,e} \geq 0, \tau_{t,e} \geq 0} w_t,$$

verifying that part of the conjecture. The constraints simplify to

$$\tau_{t,e} \leq \omega R w_t$$

and the limit on dividends,

$$d_{t,e} \leq R(1 - \omega)w_t.$$

The value function is

$$V_t^e(w_t) = R^{-1}(\tau_e d_{t,e} - (1 - \tau_e)\tau_{t,e}) + (1 - \tau_e)w_t.$$

We can set

$$\tau_{t,e} = \frac{\tau_e}{1 - \tau_e} d_{t,e}$$

to satisfy the value function, so long as

$$\omega R w_t \geq \frac{\tau_e}{1 - \tau_e} d_{t,e}.$$

It immediately follows that, facing a dividend tax of this proportion, the firm could choose whatever policy the government finds optimal.

The revenue raised per period is

$$\min\left(\frac{\tau_e}{1 - \tau_e} R(1 - \omega)w_t, \tau_e R w_t\right),$$

depending on whether τ_e is greater or smaller than ω .

5.5 Proof of 3

Conjecture an equilibrium in which V_{t+1} is weakly increasing in wealth.

We begin by listing the (potentially) relevant constraints, assuming that the initial budget

constraint binds. The no-default constraint for the about to exit firms simplifies to

$$(1 - \tau_e)w_{t,a}^R \geq (1 - \tau_e)(w_{t,a}^D - d_{t,a}),$$

which further simplifies to

$$\omega(1 - \delta)k_{t,a} \geq b_{t,a} + \tau_{t,a}.$$

For the normal firms, the no-default constraint can be expressed as two constraints,

$$d_{t,n} + V_{t+1}(w_{t,n}^R) \geq d_{t,n} + V_{t+1}(w_{t,n}^D - d_{t,n}),$$

which simplifies (under the conjecture that V_{t+1} is increasing in wealth) to

$$\omega(1 - \delta)k_{t,n} \geq b_{t,n} + \tau_{t,n},$$

and

$$d_{t,n} + V_{t+1}(w_{t,n}^R) \geq V_{t+1}(w_{t,n}^D).$$

We also have the IC constraints,

$$d_{t,n} + V_{t+1}(w_{t,n}^R) \geq d_{t,a} + V_{t+1}(w_{t,a}^R)$$

and

$$d_{t,a} + (1 - \tau_e)w_{t,a}^R \geq d_{t,n} + (1 - \tau_e)w_{t,n}^R,$$

and the initial budget constraints, merged with the creditor participation constraints (which always bind due to the positive marginal product of capital),

$$w_t + R^{-1}b_{t,n} \geq k_{t,n},$$

$$w_t + R^{-1}b_{t,a} \geq k_{t,a}.$$

The objective is

$$U_t(w_t) = \max_{\dots} R^{-1}(1 - \psi)(d_{t,n} + U_{t+1}(w_{t,n}^R)) + R^{-1}\psi(d_{t,a} + (1 - \tau_e)w_{t,a}^R), \\ + R^{-1}((1 - \psi)\tau_{t,n} + \psi\tau_{t,a} + \psi\tau_e w_{t,a}^R)$$

Rescaling by $(1 - \tau_e)$ and simplifying,

$$(1 - \tau_e)U_t(w) = \max_{\dots} R^{-1}(1 - \psi)((1 - \tau_e)(d_{t,n} + \tau_{t,n}) + V_{t+1}(w_{t,n}^R)) \\ + \psi R^{-1}(1 - \tau_e)(f(k_{t,a}) + (1 - \delta)k_{t,a} - b_{t,a}).$$

We will conjecture that the solution to a relaxed version of this problem is the solution to our problem, and then verify. The result will be that taxes can be (optimally) proportional to dividends, and that this does not cause any distortions. As a result, the government can raise a strictly positive amount through a dividend tax. We will then finish the proof by arguing that there exists a steady state in which $\chi_t = 1$.

5.5.1 Conjecture

Conjecture the following properties for $V_{t+1}(w)$: it is concave, differentiable everywhere, and there exists a point

$$\bar{w} = k^*(1 - R^{-1}\omega(1 - \delta))$$

at which $V'_{t+1}(w) = 1 - \tau_e$ for all $w \geq \bar{w}$, and $V'_{t+1}(w) > 1 - \tau_e$ and is strictly increasing for all $w < \bar{w}$. Note that, by concavity,

$$f(k^*) + (1 - \delta)k^* = Rk^* > \bar{w}.$$

Conjecture that only the following constraints are relevant:

$$\omega(1 - \delta)k_{t,n} \geq b_{t,n} + \tau_{t,n},$$

$$\omega(1 - \delta)k_{t,a} \geq b_{t,a},$$

$$w_t + R^{-1}b_{t,n} \geq k_{t,n},$$

$$w_t + R^{-1}b_{t,a} \geq k_{t,a}.$$

The objective is

$$U_t(w) = \max_{\dots} R^{-1}(1 - \psi)((1 - \tau_e)(d_{t,n} + \tau_{t,n}) + V_{t+1}(w_{t,n}^R)) \\ + \psi R^{-1}(1 - \tau_e)(f(k_t) + (1 - \delta)k_{t,a} - b_{t,a})$$

The FOC for $d_{t,n}$ is

$$-V'_{t+1}(w_{t,n}^R) + (1 - \tau_e) + v_{d,n,t} = 0$$

The FOC for $\tau_{t,n}$ is

$$-V'_{t+1}(w_{t,n}^R) + (1 - \tau_e) - \mu_{n,t} + v_{\tau,n,t} = 0$$

The FOC for $b_{t,n}$ is

$$\lambda_{t,n} - V'_{t+1}(w_{t,n}^R) - \mu_{n,t} + v_{b,n,t} = 0$$

The FOC for $b_{t,a}$ is

$$\lambda_{t,a} - (1 - \tau_e) - \mu_{a,t} + v_{b,a,t} = 0$$

The FOC for $k_{t,n}$ is

$$-\lambda_{t,n} + R^{-1}[f'(k_{t,n}) + (1 - \delta)]V'_{t+1}(w_{t,n}^R) + \mu_{n,t}R^{-1}\omega(1 - \delta) = 0$$

The FOC for $k_{t,a}$ is

$$-\lambda_{t,a} + R^{-1}[f'(k_{t,a}) + (1 - \delta)](1 - \tau_e) + \mu_{a,t}R^{-1}\omega(1 - \delta) = 0$$

The envelope theorem implies that

$$U'(w_t) = (1 - \psi)\lambda_{t,n} + \psi\lambda_{t,a}.$$

The KKT conditions are necessary and sufficient (in the relaxed problem) by concavity.

Note that

$$v_{\tau,n,t} = \mu_{n,t} + v_{d,n,t}$$

and that

$$\lambda_{t,n} + v_{b,n,t} = (1 - \tau_e) + v_{\tau,n,t}.$$

Suppose that $\lambda_{t,a} < 1 - \tau_e$. We must have $v_{b,a,t} > 0$ and $\mu_{a,t} = 0$, since it is mutually exclusive with $v_{b,a,t}$. The capital FOCs in this case is

$$R^{-1}[f'(k_{t,a}) + (1 - \delta)](\lambda_{t,a} + v_{b,a,t}) = \lambda_{t,a},$$

contradicting the assumption that

$$R^{-1}[f'(k) + (1 - \delta)] \geq 1$$

for all k . Therefore, $\lambda_{t,a} \geq 1 - \tau_e$. Similarly, suppose that $\lambda_{t,n} < 1 - \tau_e$. We have

$$v_{b,n,t} = \mu_{n,t} + (V'_{t+1}(w_{t,n}^R) - (1 - \tau_e)) + (1 - \tau_e - \lambda_{t,n}),$$

and therefore $v_{b,n,t} > 0$. It would follow that $v_{\tau,n,t} > 0$, and hence that $\mu_{n,t} = 0$ and $v_{d,t} > 0$. The capital FOC would be

$$R^{-1}[f'(k_{t,n}) + (1 - \delta)](\lambda_{t,n} + v_{b,n,t}) = \lambda_{t,n},$$

which is a contradiction. Therefore, $\lambda_{t,n} \geq 1 - \tau_e$.

Suppose that $\lambda_{t,a} = \lambda_{t,n} = 1 - \tau_e$. We must have $v_{b,a,t} = \mu_{a,t}$, and because they are mutually exclusive, they are both equal to zero. Therefore, by the FOC for $k_{t,a}$, $k_{t,a} \geq k^*$. We also have $v_{b,n,t} = v_{\tau,n,t} = \mu_{n,t} + v_{d,n,t}$. Because $\mu_t = 0$ if $v_{\tau,n,t} = v_{b,n,t} > 0$, we must have $\mu_{n,t} = 0$. Therefore,

$$v_{b,n,t} = v_{\tau,n,t} = v_{d,n,t} = V'_{t+1}(w_{t,n}^R) - (1 - \tau_e).$$

The FOC for $k_{t,n}$ is

$$R^{-1}[f'(k_{t,n}) + (1 - \delta)]\left(\frac{v_{b,n,t}}{\lambda_{t,n}} + 1\right) = 1,$$

implying that $k_{t,n} \geq k^*$ and $v_{b,n,t} = 0$. Therefore, all constraints are slack, and $V'_{t+1}(w_{t,n}^R) = (1 - \tau_e)$, and $k_{t,n} \geq k^*$. For this to be feasible, we must have

$$k_{t,n} \leq w_t + R^{-1}b_{t,n} \leq w_t + R^{-1}\omega(1 - \delta)k_{t,n},$$

and therefore

$$w_t \geq (1 - R^{-1}\omega(1 - \delta))k^*.$$

In this region, suppose that we consider alternative policies

$$\tilde{d}_{t,n} = \min\left(d_{t,n}, \frac{1 - \tau_e}{\tau_e} \tau_{t,n}\right)$$

and

$$\tilde{\tau}_{t,n} = \frac{\tau_e}{1 - \tau_e} \tilde{d}_{t,n}.$$

It follows that continuation wealth is weakly greater, $\tilde{w}_{t,n}^R \geq w_{t,n}^R \geq \bar{w}$, and therefore achieves the same utility. Moreover, taxes are lower, and therefore the solution remains feasible in the relaxed problem. It follows that it is without loss of generality, in the relaxed problem, to assume that taxes are proportional to dividends for normal firms. It is also without loss of generality, in this region, to assume that the two capital levels are equal, and that the two debt levels are equal.

Suppose that $\lambda_{t,n} > (1 - \tau_e)$. We must have

$$\lambda_{t,n} - (1 - \tau_e) + v_{b,n,t} = \mu_{n,t} + (V'_{t+1}(w_{t,n}^R) - (1 - \tau_e)) = v_{\tau,n,t} > 0.$$

Suppose that $v_{b,n,t} > 0$. Then we would have $\mu_{n,t} = 0$, and therefore

$$-\lambda_{t,n} + R^{-1}[f'(k_{t,n}) + (1 - \delta)](\lambda_{t,n} + v_{b,n,t}) = 0,$$

a contradiction. It follows that $v_{b,n,t} = 0$. Now suppose that $\mu_{n,t} = 0$, which implies $v_{\tau,n,t} = v_{d,n,t} > 0$. We would have $k_{t,n} \geq k^*$, and

$$\begin{aligned} w_{t,n}^R &= f(k_{t,n}) + (1 - \delta)k_{t,n} - b_{t,n} \\ &\geq f(k_{t,n}) + (1 - \omega)(1 - \delta)k_{t,n} \\ &\geq f(k^*) + (1 - \omega)(1 - \delta)k^* \end{aligned}$$

and

$$V'_{t+1}(w_{t,n}^R) = \lambda_t > (1 - \tau_e),$$

implying

$$w_{t,n}^R < \bar{w}.$$

By definition,

$$\bar{w} = k^*(1 - R^{-1}\omega(1 - \delta)),$$

and therefore we would have

$$k^*(1 - R^{-1}\omega(1 - \delta)) > f(k^*) + (1 - \omega)(1 - \delta)k^* \geq R^{-1}(f(k^*) + (1 - \omega)(1 - \delta)k^*),$$

which is

$$R^{-1}(f(k^*) + (1 - \delta)k^* - Rk^*) < 0.$$

By the concavity of the production function, and the fact that $f(0) = 0$,

$$f(k^*) + (1 - \delta)k^* \geq Rk^*,$$

a contradiction.

Therefore, $\mu_{n,t} > 0$. It follows in this case that $\tau_{t,n} = 0$, but it is possible that $d_{t,n} > 0$. However, because (in this case) $w_{t,n}^R \geq \bar{w}$, it is without loss of generality to set $d_{t,n} = 0$. The capital FOC in

this case is

$$-\lambda_{t,n} + R^{-1}[f'(k_{t,n}) + (1 - \delta)](\lambda_{t,n} - \mu_{n,t}) + \mu_{n,t}R^{-1}\omega(1 - \delta) = 0,$$

which implies that $k_{t,n} < k^*$, that $w_t < k^*(1 - R^{-1}\omega(1 - \delta))$. It is therefore not feasible for $k_{t,a} \geq k^*$, implying that $\mu_{t,a} > 0$, $\lambda_{t,a} > 1 - \tau_e$, and therefore that $k_{t,a} = k_{t,n}$.

We have shown that an optimal policy in the relaxed problem has taxes proportional to dividends for normal firms. Next, we discuss the dividend and tax policies for the about to exit firms that implement these allocations feasibly in the original problem. In the low wealth region, debt and capital are maximal (and the same for both firms), and dividends and taxes are zero for the normal firms. The government can set dividends and taxes for the about to exit firms to zero and satisfy the IC constraints.

Now consider the higher wealth region, in which $\lambda_{t,n} = \lambda_{t,a} = 1 - \tau_e$. If $\tau_{t,n} + b_{t,n} \geq b_{t,a}$, the government can set

$$\tau_{t,a} = \tau_{t,n} + b_{t,n} - b_{t,a}$$

and $d_{t,a} = d_{t,n}$, which will generate the same dividends and continuation wealth for both types and therefore satisfy the IC constraint. If $\tau_{t,n} + b_{t,n} < b_{t,a}$, consider the alternative policy of setting $\tilde{d}_{t,n} = 0$ and $\tilde{\tau}_{t,n} = 0$ and $\tilde{b}_{t,n} = \tilde{b}_{t,a} = \psi b_{t,a} + (1 - \psi)b_{t,n}$. Under such a policy, debt levels are equal, and feasible by construction. Such a policy also raises the same initial wealth, and therefore $k_{t,n} \geq k^*$. Moreover, we have

$$\begin{aligned} \tilde{w}_{t,n}^R &= f(k^*) + (1 - \delta)k^* - \psi b_{t,a} - (1 - \psi)b_{t,n} \\ &\geq f(k^*) + (1 - \omega)(1 - \delta)k^*. \end{aligned}$$

By the concavity of the production function,

$$f(k^*) + (1 - \omega)(1 - \delta)k^* \geq Rk^* - \omega(1 - \delta)k^* \geq Rk^*(1 - \omega R^{-1}(1 - \delta)) \geq \bar{w}.$$

It follows that such a policy does not change welfare, and is therefore without loss of generality. Hence, the IC can be satisfied by setting dividends and taxes identically for the two types. Moreover, if it is possible to increase dividends and taxes proportionally, this does not change welfare, and is therefore this also feasible and optimal, up to the borrowing constraint.

We also need to check the default constraint for normal firms which was dropped in the relaxed problem. That constraint is

$$d_{t,n} + V_{t+1}(w_{t,n}^R) \geq V_{t+1}(w_{t,n}^D).$$

When $d_{t,n} = 0$ ($\lambda_t > 1 - \tau_e$), then this constraint is implied by the no-default constraint in the relaxed problem. Consider the case when $d_{t,n} > 0$, which is the higher wealth region. It must be the case that $w_{t,n}^R \geq \bar{w}$ (because $V'_{t+1}(w_{t,n}^R) = 1 - \tau_e$ in this region). Therefore,

$$d_{t,n} + V_{t+1}(w_{t,n}^R) \geq V_{t+1}(w_{t,n}^R + d_{t,n}),$$

and the condition is satisfied by the no-default constraint in the relaxed problem.

Finally, we verify the conjectured properties of the value function. Note that $w_t \geq \bar{w}$ implies

$\lambda_t = 1 - \tau_e$, as required, and $w_t < \bar{w}$ implies $\lambda_t > 1 - \tau_e$. When $\lambda_t > 1 - \tau_e$, we have

$$\begin{aligned} k_{t,n}(1 - R^{-1}\omega(1 - \delta)) &= w_t, \\ w_{t,n}^R &= f(k_{t,n}) + (1 - \omega)(1 - \delta)k_{t,n}, \\ -\lambda_{t,n} + R^{-1}[f'(k_{t,n}) + (1 - \delta)]V'_{t+1}(w_{t,n}^R) + \\ &(\lambda_t - V'_{t+1}(w_{t,n}^R))R^{-1}\omega(1 - \delta) = 0 \end{aligned}$$

and therefore

$$\lambda_t(1 - R^{-1}\omega(1 - \delta)) = V'_{t+1}(w_{t,n}^R)R^{-1}[f'(k_t) + (1 - \omega)(1 - \delta)].$$

By the concavity of V_{t+1} and of f , and the fact that $w_{t,n}^R$ is increasing in w_t , it follows that λ_t is decreasing in w_t , and therefore V_{t+1} is strictly concave in this region, as required.

It follows that our conjectures are verified, completing this portion of the proof.

5.5.2 Steady States and $\chi = 1$

The proof thus far shows that, if $\chi = 1$ in the future, it is possible to raise strictly positive revenue using a mechanism in which taxes are proportional to dividends paid. The argument given in the text demonstrates that such a mechanism can be implemented with a dividend tax.

Assume firms choose to enter, which we will prove is possible below. It follows that firms will be entering and exiting at the same rate, and therefore that the mass of firms in the economy is stable. Moreover, the distribution of firm ages will converge to an exponential distribution. Under the optimal mechanism, there is a one-to-one map between firm age and wealth, and therefore there exists a steady state distribution of firm wealth.

By the fact that $V(0) = 0$, the concavity of $V(w)$, and the fact that $V_w(w) \geq 1 - \tau_e$, it must be the case that

$$V(w) \geq (1 - \tau_e)w.$$

By similar arguments,

$$U(w) \geq w.$$

It follows that in any steady state in which firms enter with initial wealth $\bar{w} > 0$, it is possible to raise at least $\tau_e\psi\bar{w}$ each period. Therefore, for G sufficiently small, there so long as entry occurs, there exists a steady state in which $\chi = 1$, regardless of the dynamics of the wealth distribution. The condition to ensure entry, in the absence of lump sum taxes, is

$$V(w) \geq w$$

for some $w \in (0, w_0]$. It follows by concavity, that

$$\lim_{w \rightarrow 0^+} V'(w) > 1$$

is necessary and sufficient. By the proportionality of $U(w)$ and $V(w)$,

$$(1 - \tau_e)U'(w) = V'(w) = \psi\lambda_{t,a} + (1 - \psi)\lambda_{t,n}.$$

It follows from the envelope theorem, capital FOCs, the borrowing limit, the concavity of the production function, and the fact that $V'(w) \geq 1 - \tau_e$, that

$$V'(w) \geq R^{-1}[f(w(1 - R^{-1}\omega(1 - \delta)) + (1 - \delta))(1 - \tau_e).$$

Therefore,

$$\lim_{w \rightarrow 0^+} V'(w) \geq R^{-1}[f(0) + (1 - \delta)](1 - \tau_e).$$

By the definition of k^* and the assumption that $k^* > 0$,

$$R^{-1}[f(0) + (1 - \delta)] > 1.$$

Therefore there exists a $\tau_e > 0$ such that entry will occur.

5.6 Proof of 1

Assume that $\chi = 1$. As shown in the proof of 3, for $w_t < k^*(1 - R^{-1}\omega(1 - \delta))$, we have $\lambda_t > (1 - \tau_e)$, implying that $U'(w_t) > 1$. If $w_0 < k^*(1 - R^{-1}\omega(1 - \delta))$, then the initial wealth will be less than this amount. It follows immediately that $T_0 = 0$ is strictly optimal for the government if the entrepreneur enters. Because $U(0) = 0$, the result follows by the concavity of U :

$$U(w) - U(0) \geq U'(w)w > w,$$

and therefore

$$w_0 + U(w) - w > w_0$$

for all $w > 0$. The government therefore strictly prefers entry.

5.7 Proof of 4

We begin by discussing sufficient conditions for entry. Under the proposed lump sum tax, the entering entrepreneur solves

$$\max_{w \leq w_0 - \psi^{-1}R^{-1}G, V_0 \geq 0} \{V_0 + w_0 - w - \psi^{-1}R^{-1}G, w_0\}$$

subject to the outside creditors participation constraint,

$$w + C(V_0) \geq 0.$$

If the constraint did not bind, the firm would achieve infinite continuation utility and choose to enter. Assume, therefore, that the constraint binds. For entry, we must have

$$V_0 + C(V_0) \geq \psi^{-1}R^{-1}G,$$

with

$$C(V_0) \geq \psi^{-1}R^{-1}G - w_0$$

and

$$V_0 \geq 0.$$

As w_0 and G become small, with $\psi^{-1}R^{-1}G - w_0 \geq 0$,

$$C(0) > 0$$

is a sufficient condition to guarantee entry.

Below, I will prove the following properties for $C(V)$: it is weakly decreasing in V , differentiable, concave, always satisfies $C'(V) \in [-1, 0)$ for $V > 0$, strictly if $V < \bar{V}$, for some $\bar{V} > 0$, and $C(0) > 0$. Additionally, if $V < \bar{V}$, both types of firm use less than the first-best level of capital in the first period. Using these properties, and assuming the constraint is binding and entry occurs, the firm will maximize $V_0 + C(V_0)$. There are two possibilities:

$$C'(V_0^*) = 1,$$

implying that $V_0^* \geq \bar{V}$, or $C(V_0^*) = \psi^{-1}R^{-1}G - w_0$ and $C'(V_0^*) < 1$, implying that $V_0^* < \bar{V}$.

By the fact that $C'(V) < 0$, $C(0) > C(\bar{V})$. It follows that if

$$C(\bar{V}) < \psi^{-1}R^{-1}G - w_0 < C(0),$$

we must have $C(V_0^*) > C(\bar{V})$, and therefore $V_0^* < \bar{V}$, implying that the firms are constrained but still enter.

Lastly, we show that a dividend tax is strictly sub-optimal. If a dividend tax were optimal, it would achieve the same revenue as the lump-sum tax,

$$\tau V_0^* = \psi^{-1}R^{-1}G,$$

and while requiring the firm to put in the same amount of wealth, and inducing entry. That is, the entrepreneur solves

$$\max_{V_0 \geq 0} \{(1 - \tau)V_0 + C(V_0) + w_0, w_0\},$$

subject to

$$C(V_0) \geq -w_0.$$

It follows immediately that the dividend tax is sub-optimal, in the sense that it will induce a choice by the entering firm, regardless of whether the constraint binds or not.

5.7.1 Analysis of the Creditor's Problem

Note that the initial budget constraints and production constraints bind. Note also that only the some of $d_{t,a} + V_{t+1}^e$ matters, so assume without loss of generality that $d_{t,a} = 0$.

The problem is

$$C(V_t) = \max_{\dots} (1 - \psi)(R^{-1}[f(k_{t,n}) + (1 - \delta)k_{t,n}] - k_{t,n}) + \psi(R^{-1}[f(k_{t,a}) + (1 - \delta)k_{t,a}] - k_{t,a}) \\ - (1 - \psi)R^{-1}d_{t,n} - \psi R^{-1}V_t^e + R^{-1}(1 - \psi)C(V_{t+1}).$$

subject to the non-redundant incentive compatibility constraints,

$$V_t^e \geq f(k_{t,a}) + (1 - \omega)(1 - \delta)k_{t,a},$$

$$d_{t,n} + V_{t+1} \geq V_t^e,$$

$$V_t^e \geq f(k_{t,n}) + (1 - \omega)(1 - \delta)k_{t,n},$$

the promise-keeping constraint,

$$V_t \leq R^{-1}(1 - \psi)(d_{t,n} + V_{t+1}) + R^{-1}\psi V_t^e,$$

the upper bounds on dividends,

$$d_{t,n} \leq f(k_{t,n}) + (1 - \omega)(1 - \delta)k_{t,n},$$

and the positivity constraints for $(d_{t,n}, k_{t,n}, k_{t,a}, V_t^e, V_{t+1})$. There are also positivity constraints for funds raised and repaid, but they are redundant.

The constraints are entirely affine or concave, and it immediately follows that the KKT conditions are necessary and sufficient. Differentiability, concavity, and the envelope theorem follow by standard arguments:

$$C'(V_t) = -\tilde{\zeta}_t,$$

where $\tilde{\zeta}_t$ is the multiplier on the promise-keeping constraint.

Let $\mu_{a,t}$, ϕ_t , and $\mu_{n,t}$ be the multipliers on the three non-redundant IC constraints, let ρ_t be the bound on dividends, and let $\nu_{d,n,t}$, $\nu_{v,a,t}$, $\nu_{v,n,t}$, $\nu_{k,n,t}$, and $\nu_{k,a,t}$ denote the constraints on dividends, continuation values, and capital being positive.

The FOC for an about to exit firm's terminal payoff and capital are

$$-\psi R^{-1} + \mu_{a,t} + \mu_{n,t} - \phi_t + \psi R^{-1}\tilde{\zeta}_t + \nu_{v,a,t} = 0$$

and

$$\psi R^{-1}(f'(k_{t,a}) + (1 - \delta) - R) - (f'(k_{t,a}) + (1 - \omega)(1 - \delta))\mu_{a,t} + \nu_{k,a,t} = 0.$$

The FOCs for normal firms are

$$-R^{-1}(1 - \psi)\tilde{\zeta}_{t+1} + \phi_t + R^{-1}(1 - \psi)\tilde{\zeta}_t + \nu_{v,n,t} = 0$$

$$-(1 - \psi)R^{-1} + \phi_t + R^{-1}(1 - \psi)\tilde{\zeta}_t + \nu_{d,n,t} - \rho_t = 0$$

$$(1 - \psi)R^{-1}(f'(k_{t,n}) + (1 - \delta) - R) - (f'(k_{t,n}) + (1 - \omega)(1 - \delta))(\mu_{n,t} - \rho_t) + \nu_{k,n,t} = 0.$$

Suppose that $\xi_t > 1$. It would follow that $\xi_{t+j} > 1$ for all $j > 0$, and that $\phi_t > 0$. However,

$$\xi_{t+1} = 1 - \frac{v_{d,n,t}}{(1-\psi)R^{-1}} \leq 1,$$

and therefore we must have $\xi_t \leq 1$.

Suppose that $V_t = 0$. In this case, the firm is free to set $V_{t+1} = 0$ and

$$d_{t,n} = V_t^e = f(k^*) + (1-\omega)(1-\delta)k^*.$$

If this policy were optimal, it would generate a value function

$$C(0)(1 - R^{-1}(1 - \psi)) = \omega(1 - \delta)k^*,$$

and consequently $C(0) > 0$. By the fact that $-\xi_t = C'(V_t) \geq -1$, it will always be the case that

$$V_t + C(V_t) \geq C(0) > 0.$$

Suppose that $\xi_t = 1$. In this region, it must be the case that $\phi_t = v_{v,n,t} = 0$ (otherwise $\xi_{t+1} > 1$), and therefore $\mu_{a,t} = \mu_{n,t} = v_{v,a,t} = 0$. By the capital FOCs, it follows that $v_{k,n,t} = v_{k,a,t} = \rho_t = 0$, and both capital levels are first-best. For this to be feasible,

$$V_t \geq f(k^*) + (1-\omega)(1-\delta)k^* = \bar{V} > 0.$$

Therefore, if $V_t < f(k^*) + (1-\omega)(1-\delta)k^*$, we must have $\xi_t < 1$.

Now suppose that $\xi_t < 1$. We must have $\mu_{a,t} + \mu_{n,t} + v_{v,a,t} > 0$. Suppose that $\mu_{a,t} + \mu_{n,t} > 0$. If $k_{t,a} < k_{t,n}$, then $\mu_{a,t} = 0$, and consequently

$$R^{-1}(f'(k_{t,n}) + (1-\delta) - R) + \frac{v_{k,n,t}}{1-\psi} + \rho_t(f'(k_{t,n}) + (1-\omega)(1-\delta)) > 0.$$

Therefore, $k_{t,n} < k^*$, and $k_{a,t} < k^*$. The exact same argument applies if $k_{t,a} > k_{t,n}$, in reverse, and if they are equal. If $v_{v,a,t} > 0$, then $k_{t,n} = k_{t,a} = 0 < k^*$. Therefore, $\xi_t < 1$ implies $k_{t,n} < k^*$ and $k_{t,a} < k^*$.

Suppose that $V_t > 0$. Then, by concavity and the fact that $\xi_{t+1} \geq \xi_t$, $V_{t+1} > 0$ and $v_{v,n,t} = 0$. Therefore,

$$\begin{aligned} \psi R^{-1}\xi_t - \phi_t - \psi R^{-1} &= R^{-1}(\xi_t - 1) + v_{d,n,t} \\ &= \mu_{a,t} + \mu_{n,t} + v_{v,a,t} \end{aligned}$$

We have

$$R^{-1}(\xi_t - 1) + (1-\psi)R^{-1}(1 - \xi_{t+1}) = \mu_{a,t} + \mu_{n,t} + v_{v,a,t}.$$

If $\xi_t < 1$, then $\mu_{a,t} + \mu_{n,t} + v_{v,a,t} > 0$. If $\xi_t = 0$, then we would have

$$(1-\psi)R^{-1}(1 - \xi_{t+1}) > 1,$$

required $\xi_{t+1} < 0$, a contradiction. Therefore, $\xi_t > 0$ for all $V_t > 0$.