Fiscal Rules and Discretion
under Limited Enforcement[*]

Marina Halac[†]   Pierre Yared[‡]

December 31, 2018

Abstract

We study a fiscal policy model in which the government is present-biased towards public spending. Society chooses a fiscal rule to trade off the benefit of committing the government to not overspend against the benefit of granting it flexibility to react to privately observed shocks to the value of spending. Unlike prior work, we examine rules under limited enforcement: the government has full policy discretion and can only be incentivized to comply with a rule via the use of penalties which are joint and bounded. We show that optimal incentives must be bang-bang. Moreover, under a distributional condition, the optimal rule is a maximally enforced deficit limit, triggering the largest feasible penalty whenever violated. Violation optimally occurs under high enough shocks if and only if available penalties are weak and such shocks are rare. If the rule is self-enforced in a dynamic setting, penalties take the form of temporary overspending.

Keywords: Private Information, Fiscal Policy, Deficit Bias, Enforcement Constraints

JEL Classification: C73, D02, D82, E6, H1, P16

1 Introduction

Countries impose rules on their governments to constrain their policy decisions. Increasingly prevalent are fiscal rules, in place in 92 countries in 2015, compared to only seven countries in 1990.[1] Yet, according to the International Monetary Fund, governments comply with their fiscal rules only about 50 percent of the time.[2]

Whenever a rule is violated, a formal or informal enforcement mechanism is triggered. In the European Union, an Excessive Deficit Procedure—a formal procedure specifying a sequence of costly fiscal adjustments and potential sanctions—applies whenever a fiscal limit is breached.[3] In other countries, penalties for rule violation are more informal. In Chile, for example, the government's breach of its fiscal rule in 2009 was punished with lax fiscal policy by the next administration, which continued to ignore the rule despite criticism of fiscal irresponsibility.[4] The use of penalties, formal or informal, is critical to incentivize governments to respect the fiscal constraints set by society. Penalties however are harmful to both the government and society, and they are limited in scope.

In this paper, we study the optimal design of fiscal rules under limited enforcement. What is the optimal structure of fiscal constraints when available penalties for violation are joint and bounded, and how are these penalties used in an optimal rule? Should the government violate fiscal constraints when the economy is in distress? And if rules are self-enforced by the behavior of future governments, what is the form that endogenous penalties optimally take?

Our analysis of fiscal rules builds on the approach used in Amador, Werning, and Angeletos (2006) and Halac and Yared (2014).[5] We consider a government that is present-biased towards public spending and privately informed about shocks affecting the value of this spending. Society chooses a fiscal rule to trade off the benefit of committing the government to not overspend against the benefit of granting it flexibility to react to shocks. Motivated by real-world rules and unlike prior work, we posit that this rule can only be enforced via the use of penalties which are limited and socially costly.

Our environment is a small open economy in which the government makes a borrowing and spending decision. Prior to the choice of policy, a shock to the social value of deficit-financed government spending is realized. The government is present-biased: for any given shock, the government overvalues current spending relative to future welfare compared to

---

[1]See IMF Fiscal Rules Data Set, 2015 and Schaechter et al. (2012).

[2]See Caselli et al. (2018).

[3]See Lledó et al. (2017, p.81) for a description of this procedure.

[4]See Halac and Yared (2017b) for a description of this episode and related references.

[5]See also Athey, Atkeson, and Kehoe (2005), Amador and Bagwell (2013), and Ambrus and Egorov (2013).

society. This preference structure results naturally from the aggregation of heterogeneous, time-consistent citizens' preferences (Jackson and Yariv, 2014, 2015), or as a consequence of turnover in a political economy setting (e.g., Aguiar and Amador, 2011).[6] We assume that the shock to the value of spending is privately observed by the government, capturing the fact that rules cannot explicitly condition on all contingencies in fiscal policy. Furthermore, the government has full discretion when choosing borrowing and spending. Society can only specify a schedule of joint penalties, determining the continuation value to the government and society as a function of the government's policy choice. A fiscal rule in this setting therefore consists of an allocation of debt and continuation value for each shock, where this allocation must satisfy the government's private information and enforcement constraints.

To describe the forces underlying our model, suppose first that fiscal rules could be perfectly enforced. Society then chooses a rule to optimally resolve the tradeoff between commitment and flexibility. Fully committing the government to a contingent debt plan would allow to implement the first best policy in the absence of private information, whereas granting the government full flexibility would yield the first best policy in the absence of a present bias. Given both private information and a present bias, however, a tradeoff arises, and the first best is not implementable. Amador, Werning, and Angeletos (2006) show that, under perfect enforcement and certain distributional assumptions, the solution to this tradeoff is a fiscal rule that takes the form of a deficit limit. The government borrows within the limit if the shock to the value of spending is relatively low and it borrows at the limit if the shock is higher, without triggering any penalties.

Our focus is on understanding the optimal fiscal rule when enforcement is limited. As is also true under perfect enforcement, a rule under limited enforcement must satisfy the government's private information constraints: given a realized shock, the government must prefer its assigned debt level and continuation value to those prescribed for any other shock. In addition, the rule must satisfy the government's enforcement constraints: given a realized shock, the government must prefer its assigned debt level and continuation value to any other level of debt not prescribed for any shock. Any observable (*off-path*) deviation—where the government chooses a debt level corresponding to no shock—is optimally punished with the worst possible continuation value conditional on the choice of debt. In fact, if penalties are severe enough, enforcement constraints are non-binding: the government always prefers to abide to the perfect-enforcement deficit limit to avoid punishment.

Our main result is a characterization of the optimal fiscal rule when enforcement con-

---

[6]See also Persson and Svensson (1989), Alesina and Tabellini (1990), Lizzeri (1999), Battaglini and Coate (2008), and Yared (2018). Our formulation corresponds to the quasi-hyperbolic model; see Laibson (1997).

straints are binding. We show that this rule takes the form of a *maximally enforced deficit limit*, which, if violated, leads to the maximal penalty for the government. The rule is thus similar to that under perfect enforcement, although it differs in two aspects. First, the deficit limit imposed on the government is more relaxed than the perfect-enforcement limit. Second, the possibility of *on-path* penalties emerges, with the government breaching the deficit limit and triggering punishment under sufficiently high shocks.

To obtain this result, we begin by establishing general properties of optimal incentive provision. Society incentivizes the government not to overborrow by using continuation values as reward and punishment. We show that in any optimal rule, continuation values must be *bang-bang*, so the government receives either the maximal reward or the worst punishment (conditional on the level of debt), depending on its policy choice. Using milder penalties would be less socially costly; however, the harshest future punishment maximizes the range of shocks under which society can impose fiscal discipline in the present, and we show that this maximizes social welfare. Our bang-bang result relies only on generic properties of the distribution of shocks, and it applies under both limited and perfect enforcement. This result thus has implications for other models of delegation with money burning, including some of those studied in Amador and Bagwell (2013): we find that optimal delegation requires money burning to take a bang-bang form.

We complete our characterization of optimal incentives by introducing an assumption on the distribution of shocks.[7] We show that under this assumption, optimal bang-bang continuation values must be monotonic, with the government either always receiving the maximal reward or receiving the maximal punishment only under high enough shocks. Moreover, building on monotonicity, we are able to characterize the optimal borrowing allocation. We establish that our distributional assumption is sufficient, as well as necessary, for maximally enforced deficit limits to be the unique optimal fiscal rule.

Specifically, we show that the optimal rule takes one of two forms. On the one hand, society can set a relaxed deficit limit that satisfies the enforcement constraint under all shocks and thus entails no penalties on path. On the other hand, society can prescribe a tighter deficit limit with on-path penalties: the government respects the deficit limit and receives the maximal continuation value under low enough shocks, but it breaches the limit and receives the worst continuation value under higher shocks. We prove that the optimal deficit limit is unique, and we provide a necessary and sufficient condition for this limit to feature on-path penalties. This condition depends on the ease of enforcement and the

---

[7] See Assumption 1 in Subsection 3.3. This assumption holds for a broad range of distributions, including uniform, exponential, log-normal, gamma, and beta for a subset of its parameters. Assumption 1 is similar to, but stronger than, the assumption used in Amador, Werning, and Angeletos (2006).

distribution of shocks. In particular, we find that society chooses a tight deficit limit with punishment on path if and only if available penalties are sufficiently weak and high shocks are sufficiently unlikely. Intuitively, a deficit limit that is respected under all shocks would have to be too lax if penalties for violation are not severe. Moreover, tightening the deficit limit by the use of punishment is optimal when high shocks are rare, as the expected cost of punishing the government following such shocks is then relatively low.

Our baseline model consists of a two-period environment in which the available penalties, and thus the set of feasible continuation values, is exogenously given. We extend our analysis to an infinite horizon economy in which penalties are endogenous. That is, in the absence of an external enforcement authority, we study how a fiscal rule may be self-enforced by the interaction of a sequence of governments. Under an assumption on preferences, we establish that our main result regarding the unique optimality of maximally enforced deficit limits extends to this environment. Furthermore, we show that the worst self-enforcing punishment takes the form of temporary overborrowing, following which the optimal deficit limit is reinstated. Hence, in an optimal rule with on-path punishment, the temporary violation of fiscal constraints, as in the Chilean case previously described, is used as a deterrent for breaking these constraints. The fiscal rule is self-enforced by transitions in and out of the best and worst equilibria, with the economy fluctuating between periods of fiscal responsibility and periods of fiscal irresponsibility.

Related literature. Our paper fits into the aforementioned literature on mechanism design that studies the tradeoff between commitment and flexibility.[8] This literature is mainly concerned with environments with perfect enforcement, whereas we examine the optimal rule under limited enforcement. Closely related is Amador and Bagwell (2016), which considers the problem of regulating a privately informed monopolist in the absence of transfers and given an ex-post participation constraint. The paper shows that optimal regulation takes the form of a threshold which is imposed unless the monopolist chooses to shut down. In contrast to their work, our analysis does not rely on the perfect enforcement of thresholds, and it allows for money burning, which is sometimes optimally used on path.

Also related to our paper is the literature on the political economy of fiscal policy.[9]

---

[8]In addition to the work previously cited, see Sleet (2004). More broadly, our paper relates to the literature on delegation in principal-agent settings, including Holmström (1977, 1984), Alonso and Matouschek (2008), and Ambrus and Egorov (2015).

[9]In addition to Aguiar and Amador (2011) and the work cited in fn. 6, see Krusell and Rios-Rull (1999), Acemoglu, Golosov, and Tsyvinski (2008), Yared (2010), Azzimonti (2011), and Song, Storesletten, and Zilibotti (2012). For quantitative analyses of fiscal rules, see Bassetto and Sargent (2006), Alfaro and Kanczuk (2016), Azzimonti, Battaglini, and Coate (2016), and Hatchondo, Martinez, and Roch (2017), and for a study of coordinated fiscal rules across countries, see Halac and Yared (2018).

Dovis and Kirpalani (2017), in particular, examine deficit limits that are endogenously enforced by reputational concerns. We instead study the enforcement of optimal fiscal rules under privately observed shocks and no restrictions on their structure.

Our analysis of self-enforcement in an infinite horizon setting contributes to the literature on hyperbolic discounting and the benefits of commitment devices.[10] Bernheim, Ray, and Yeltekin (2015) is related in that it studies self-enforcing consumption rules for a consumer with quasi-hyperbolic preferences, albeit without any private information. Finally, the equilibrium dynamics in our infinite horizon economy bear a relationship to the seminal work of Abreu, Pearce, and Stacchetti (1990), who establish the optimality of bang-bang continuation values in a class of repeated games. Their analysis however is constrained to settings with finite actions and a continuous public signal, and thus it does not apply to our environment in which the action is continuous. Also related is Athey, Bagwell, and Sanchirico (2004), which studies a repeated Bertrand game with private information.

2 Model

2.1 Setup

We study a simple model of fiscal policy in which the government makes a borrowing and spending decision. Our setting is similar to that analyzed in Amador, Werning, and Angeletos (2006) and Halac and Yared (2014). Our departure is in examining the design of fiscal rules under only limited, as opposed to perfect, enforcement.

Consider a small open economy. A shock to the economy $\theta > 0$ is drawn from a bounded set $\Theta \equiv \left[\underline{\theta}, \overline{\theta}\right]$, with a continuously differentiable probability density function $f(\theta) > 0$ and associated cumulative density function $F(\theta)$. The realization of this shock is privately observed by the government, so we refer to θ as the government's *type*.

The government chooses debt $b \in [\underline{b}, \overline{b}]$ and public spending g subject to the budget constraint

$$g = \omega + b, \tag{1}$$

where $\omega > 0$ denotes the exogenous resources of the government, representing collected tax revenue net of any debt repayment.[11]

---

[10]See for example Phelps and Pollak (1968), Laibson (1997), Krusell and Smith, Jr. (2003), Krusell, Kruscu, and Smith, Jr. (2010), Lizzeri and Yariv (2014), Bisin, Lizzeri, and Yariv (2015), Cao and Werning (2017), and Moser and de Souza e Silva (2017).

[11]Note that since $g - b$ is independent of θ, cross-subsidization across types is not possible, unlike in other models such as Atkeson and Lucas (1992), Thomas and Worrall (1990), Phelan (1998), Sleet and Yeltekin (2006, 2008), and Farhi and Werning (2007).

Society can influence policy by imposing socially costly penalties on the government as a function of its choice of borrowing and spending. As discussed in the Introduction, these penalties may take the form of formal sanctions, or they may arise from the continuation play of future governments in a dynamic environment (see Section 4). We model penalties by letting society commit ex ante to a continuation value function $V(b) \in [\underline{V}(b), \overline{V}(b)]$, specifying a continuation value for the government and society for each level of debt $b \in [\underline{b}, \overline{b}]$. Here $\overline{V}(b)$ denotes the highest feasible continuation value given debt b, and $\underline{V}(b)$ is the lowest such value.

The timing and payoffs are as follows. First, society chooses a continuation value function $V(b)$ in order to maximize social welfare:

$$\mathbb{E}\left[\theta U(g) + \delta V(b)\right], \tag{2}$$

where $\theta U(g)$ is the social utility from public spending and $\delta \in (0, 1)$ is the discount factor. Second, the government observes the realized shock θ and, given θ and the function $V(b)$, chooses debt b and spending g subject to (1) in order to maximize government welfare:

$$\theta U(g) + \beta \delta V(b), \tag{3}$$

where $\beta \in (0, 1)$.

We take $U(g)$ to be strictly increasing, strictly concave, and continuously differentiable. We also assume that for all $b \in [\underline{b}, \overline{b}]$, $\underline{V}(b)$ and $\overline{V}(b)$ are bounded and continuous and satisfy $\overline{V}(b) > \underline{V}(b)$.

There are three main features of this environment. First, since $\beta < 1$, the government's objective (3) given its realized type does not coincide with the social objective (2). Compared to society, the government overweighs the importance of current spending relative to the future continuation value. The government would thus want to overborrow: for example, if $V(b)$ is decreasing and concave, the government's preferred policy satisfies $\theta U'(g) = -\beta \delta V'(b)$, whereas the social optimum sets $\theta U'(g) = -\delta V'(b)$. As mentioned in the Introduction, this payoff structure arises naturally when the government's preferences aggregate heterogeneous citizens' preferences (see Jackson and Yariv, 2014, 2015). This formulation can also be motivated by political turnover; for instance, preferences such as these emerge in settings with political uncertainty where policymakers place a higher value on public spending when they hold power and can make spending decisions (see Aguiar and Amador, 2011).[12]

---

[12]For a more detailed discussion of our payoff structure, see Yared (2018).

The second feature of our environment is that the realization of θ—which affects the marginal social utility of public spending—is privately observed by the government. One interpretation is that fiscal rules imposed on the government cannot explicitly condition on the value of θ, even if this shock were observable.[13] An alternative interpretation is that the exact cost of public goods is only observable to the policymaker, who may be inclined to overspend on these goods. Another possibility is that citizens have heterogeneous preferences or heterogeneous information regarding the optimal level of public spending, and the government sees an aggregate that the citizens do not see (see Sleet, 2004; Piguillem and Schneider, 2016).

The third and most critical feature of our environment is that the government has full discretion when choosing borrowing and spending. Society can only influence policy by specifying contingent penalties which are limited and socially costly. This is a main distinction from prior work, which assumes that available actions can be restricted arbitrarily and at no cost. The extent of penalties that society can impose on the government is captured in our model by the difference $\overline{V}(b) - \underline{V}(b)$: the larger the difference between the highest and lowest feasible continuation values, the larger the potential for punishment.[14,15] Note that $\underline{V}(b)$ and $\overline{V}(b)$ are exogenous in our benchmark setting and can take on a number of possible forms. For example, these values may satisfy $\underline{V}(b) = \overline{V}(b) - m$ for some $m > 0$, in which case m would represent a maximum additive penalty available for any feasible level of debt b. In the dynamic setting of Section 4, $\underline{V}(b)$ and $\overline{V}(b)$ will be endogenously determined by the set of equilibrium payoffs.

2.2   Fiscal Rules

Using the budget constraint (1), denote the social utility from public spending by $U(\omega + b)$. Given the continuation value function $V(b)$ specified by society and the realization of its type θ, the government chooses a level of debt $b(\theta)$ which maximizes its welfare in (3), with associated continuation value $V(b(\theta))$ (and spending $g(\theta) \equiv \omega + b(\theta)$). We refer to the resulting allocation, $\{b(\theta), V(b(\theta))\}_{\theta \in \Theta}$, as the *fiscal rule* imposed by society. This rule specifies a level of debt and continuation value for each type $\theta \in \Theta$, and it satisfies private information, enforcement, and feasibility constraints, as we describe next.

---

[13]Halac and Yared (2017a) study a delegation problem in which shocks can be verified by a rule-making body at a cost.

[14]A continuation value $V(b) \in (\overline{V}(b), \underline{V}(b))$ corresponds to imposing an intermediate penalty on the government. Intermediate penalties could also be imposed by introducing a public randomization device. However, such a device will not be used in the optimum, which we will show has a bang-bang nature.

[15]As we will show in Section 3, previous analyses in which available actions can be arbitrarily restricted correspond to a special case of our model in which $\underline{V}(b)$ is sufficiently low.

The private information constraint captures the fact that the government can misrepresent its type. A fiscal rule $\{b(\theta), V(b(\theta))\}_{\theta \in \Theta}$ is such that a government of type θ prefers to pursue its assigned policy rather than that of any other type $\theta' \neq \theta$:

$$\theta U(\omega + b(\theta)) + \beta \delta V(b(\theta)) \geq \theta U(\omega + b(\theta')) + \beta \delta V(b(\theta')) \quad \text{for all } \theta, \theta' \in \Theta. \tag{4}$$

The enforcement constraint captures the fact that the government can freely choose any feasible level of debt, including levels not assigned to any other government type. A fiscal rule $\{b(\theta), V(b(\theta))\}_{\theta \in \Theta}$ is such that a government of type θ prefers to pursue its assigned policy rather than any other policy $b' \in [\underline{b}, \overline{b}]$ such that $b' \neq b(\theta')$ for all $\theta' \in \Theta$:

$$\theta U(\omega + b(\theta)) + \beta \delta V(b(\theta)) \geq \theta U(\omega + b') + \beta \delta V(b') \quad \text{for all } \theta \in \Theta \text{ and all } b' \neq b(\theta') \text{ for all } \theta' \in \Theta.$$

Note that since the continuation value satisfies $V(b') \geq \underline{V}(b')$ for all $b' \in [\underline{b}, \overline{b}]$, the above inequality must hold under maximal punishment, i.e. when $V(b') = \underline{V}(b')$. Moreover, since the inequality must then hold for all $b' \in [\underline{b}, \overline{b}]$, it must necessarily hold when b' corresponds to type θ's flexible level of debt conditional on maximal punishment. Specifically, let $b^p(\theta)$ be defined by

$$b^p(\theta) \in \arg\max_{b \in [\underline{b}, \overline{b}]} \{\theta U(\omega + b) + \beta \delta \underline{V}(b)\}.$$

Then a necessary condition for the enforcement constraint to be satisfied is

$$\theta U(\omega + b(\theta)) + \beta \delta V(b(\theta)) \geq \theta U(\omega + b^p(\theta)) + \beta \delta \underline{V}(b^p(\theta)) \quad \text{for all } \theta \in \Theta, \tag{5}$$

where note that the right-hand side is the government's minmax payoff.

Constraints (4) and (5) are clearly necessary for $\{b(\theta), V(b(\theta))\}_{\theta \in \Theta}$ to be incentive compatible. Furthermore, if an allocation satisfies these constraints, then it can be supported by specifying the worst feasible continuation value following any choice $b' \in [\underline{b}, \overline{b}]$ by the government such that $b' \neq b(\theta')$ for all $\theta' \in \Theta$. Since such a choice is off path, it is without loss to assume that it is maximally punished.

Lastly, the feasibility constraints ensure that the continuation value $V(b(\theta))$ is within the lowest and highest feasible values given the level of debt:

$$V(b(\theta)) \geq \underline{V}(b(\theta)) \quad \text{and} \quad V(b(\theta)) \leq \overline{V}(b(\theta)) \quad \text{for all } \theta \in \Theta. \tag{6}$$

A fiscal rule is incentive compatible if it satisfies (4)-(5), and it is incentive compatible and feasible, or incentive feasible for short, if it satisfies (4)-(6). The fiscal rule is optimal

8

if it additionally maximizes social welfare, namely if it solves:

$$\max_{\left\{b(\theta)\in[\underline{b},\overline{b}],V(b(\theta))\right\}_{\theta\in\Theta}} \mathbb{E}\Big[\theta U(\omega+b(\theta))+\delta V(b(\theta))\Big] \tag{7}$$

subject to (4), (5), and (6).

Throughout our analysis, we assume that the solution to (7) admits a piecewise continuously differentiable function $b(\theta)$, which allows us to establish our results by the use of perturbations.[16] This approach follows Athey, Atkeson, and Kehoe (2005), who use perturbation arguments to characterize optimal monetary policy rules under perfect enforcement. An alternative approach would be to use Lagrangian methods, as in the work of Amador, Werning, and Angeletos (2003). Since our problem is not globally concave, however, a Lagrangian approach would not establish uniqueness of the solution, and would make it difficult to identify necessary and sufficient conditions for the optimality of maximally enforced deficit limits. We are able to provide these results using perturbations, as we show in the next section.

3 Optimal Fiscal Rule

We characterize the optimal fiscal rule under limited enforcement by solving program (7). We will establish conditions under which the unique solution to this program is a deficit limit with maximal enforcement.

Let $b^r(\theta)$ denote type θ's flexible level of debt conditional on maximal reward:

$$b^r(\theta) \in \arg\max_{b\in[\underline{b},\overline{b}]}\{\theta U(\omega+b)+\beta\delta\overline{V}(b)\}. \tag{8}$$

We define:

Definition 1. $\{b(\theta),V(b(\theta))\}_{\theta\in\Theta}$ *is a maximally enforced deficit limit if there exist* $\theta^* \in [0,\overline{\theta})$ *and finite* $\theta^{**} > \max\{\theta^*,\underline{\theta}\}$ *such that*

$$\{b(\theta),V(b(\theta))\} = \begin{cases} \{b^r(\theta),\overline{V}(b^r(\theta))\} & if\ \theta<\theta^*, \\ \{b^r(\theta^*),\overline{V}(b^r(\theta^*))\} & if\ \theta\in[\theta^*,\theta^{**}], \\ \{b^p(\theta),\underline{V}(b^p(\theta))\} & if\ \theta>\theta^{**}, \end{cases} \tag{9}$$

---

[16]If the program admits multiple solutions that differ only on a countable set of types, we select the solution that maximizes social welfare for those types.

where

$$\theta^{**}U(\omega + b^r(\theta^*)) + \beta\delta\overline{V}(b^r(\theta^*)) = \theta^{**}U(\omega + b^p(\theta^{**})) + \beta\delta\underline{V}(b^p(\theta^{**})). \tag{10}$$

Figure 1 provides an example.[17] We depict the borrowing allocation (top panel) and the allocation of continuation values (bottom panel) under a maximally enforced deficit limit with $\theta^* > \underline{\theta}$ and $\theta^{**} < \overline{\theta}$. Under this rule, types $\theta \in [\underline{\theta}, \theta^*)$ choose their flexible debt levels conditional on maximal reward, $b^r(\theta)$; types $\theta \in [\theta^*, \theta^{**}]$ choose type θ^*'s flexible debt level conditional on maximal reward, $b^r(\theta^*)$; and types $\theta \in (\theta^{**}, \overline{\theta}]$ choose their flexible debt levels conditional on maximal punishment, $b^p(\theta)$. Types $\theta \leq \theta^{**}$ are maximally rewarded with continuation value $\overline{V}(b(\theta))$ whereas types $\theta > \theta^{**}$ are maximally punished with continuation value $\underline{V}(b(\theta))$. By (10), the enforcement constraint holds with equality for type θ^{**}.

We can verify that the fiscal rule described in Definition 1 is incentive compatible:

Lemma 1. *If $\{b(\theta), V(b(\theta))\}_{\theta \in \Theta}$ is a maximally enforced deficit limit, then it satisfies the private information constraint (4) and the enforcement constraint (5).*

In terms of implementation, this fiscal rule can be implemented using a maximum deficit limit, spending limit, or debt limit, where this limit would be associated with the borrowing level $b^r(\theta^*)$. If the government respects the limit, it receives maximal reward given its level of debt, $\overline{V}(b)$. If the government breaches the limit, it receives maximal punishment given its level of debt, $\underline{V}(b)$. Note that the limit is breached along the equilibrium path if and only if $\theta^{**} < \overline{\theta}$; we will provide conditions under which this inequality holds in an optimal maximally enforced deficit limit.

Our analysis proceeds as follows. First, we present some preliminary results in Subsection 3.1, which yield a convenient formulation of the objective in program (7). Second, we show in Subsection 3.2 that any (interior) solution to this program must feature bang-bang continuation values, so any optimal rule provides high-powered incentives for the government not to overborrow. Third, we show in Subsection 3.3 that under a distributional assumption, optimal bang-bang incentives are monotonic, with either all types receiving the highest continuation value conditional on debt, or only types above an interior point being punished with the lowest continuation value conditional on debt. This result facilitates our characterization of the optimal borrowing allocation in Subsection 3.4, which shows that any solution to (7) is a maximally enforced deficit limit. We further establish that the optimal limit is unique, and provide a necessary and sufficient condition for the

---

[17]The examples in our figures take $\underline{V}(b(\theta))$ and $\overline{V}(b(\theta))$ both strictly decreasing in $b(\theta)$, but this is not required by our model.

(g) ...

ems:

))]

elfare given $\overline{V}(b)$ –

ıd private info θ_t

elfare given $\overline{V}(b)$ –

$)\}_{\theta^t \in \Theta^t}\}_{t=0}^{\infty}$

ınd lower min

$b^r(\theta^*)$

tives

sufficiently binding

$V(b(\theta))$

extreme

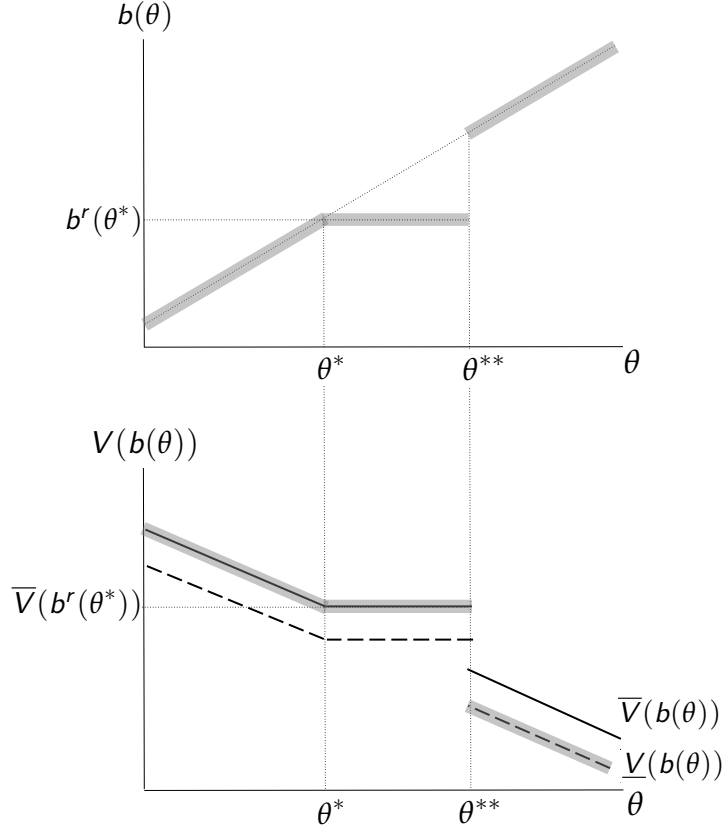ical rectitude and fiscal profligacy sustain each

Figure 1: An example of a maximally enforced deficit limit. The thick grey line depicts the borrowing allocation in the top panel and the allocation of continuation values in the bottom panel. The solid and dashed black lines in the bottom panel depict $\overline{V}(b(\theta))$ and $\underline{V}(b(\theta))$ respectively.

government to violate the limit following high enough shocks. Finally, in Subsection 3.5, we show that the distributional assumption introduced in Subsection 3.3 is not only sufficient but also necessary for any solution to (7) to be a maximally enforced deficit limit.

3.1 Preliminaries

The next lemma follows from standard arguments; see Fudenberg and Tirole (1991):

Lemma 2. $\{b(\theta), V(b(\theta))\}_{\theta \in \Theta}$ *satisfies the private information constraint* (4) *if and only if: (i)* $b(\theta)$ *is nondecreasing, and (ii) the following local private information constraints are satisfied:*

1. *At any point* θ *at which* $b(\cdot)$, *and thus* $V(\cdot)$, *are differentiable,*

$$\frac{db(\theta)}{d\theta}\left(\theta U'(\omega + b(\theta)) + \beta\delta V'(b(\theta))\right) = 0.$$

11

2. At any point θ at which $b(\cdot)$ is not differentiable,

$$\lim_{\theta' \uparrow \theta} \{\theta U(\omega + b(\theta')) + \beta \delta V(b(\theta'))\} = \lim_{\theta' \downarrow \theta} \{\theta U(\omega + b(\theta')) + \beta \delta V(b(\theta'))\}.$$

Observe that since $b(\theta)$ is nondecreasing in θ, satisfaction of (4) requires that $V(b(\theta))$ be nonincreasing in θ.

The local private information constraints imply that the derivative of government welfare with respect to θ is $U(\omega + b(\theta))$. Hence, in an incentive compatible rule, government welfare for type $\theta \in \Theta$ satisfies

$$\theta U(\omega + b(\theta)) + \beta \delta V(b(\theta)) = \underline{\theta} U(\omega + b(\underline{\theta})) + \beta \delta V(b(\underline{\theta})) + \int_{\underline{\theta}}^{\theta} U(\omega + b(\widetilde{\theta})) d\widetilde{\theta}. \qquad (11)$$

Following [Amador, Werning, and Angeletos (2006)](), we can substitute (11) into the social welfare function in (7) to rewrite social welfare as

$$\frac{1}{\beta} \underline{\theta} U(\omega + b(\underline{\theta})) + \delta V(b(\underline{\theta})) + \frac{1}{\beta} \int_{\underline{\theta}}^{\overline{\theta}} U(\omega + b(\theta)) Q(\theta) d\theta, \qquad (12)$$

where

$$Q(\theta) \equiv 1 - F(\theta) - \theta f(\theta)(1 - \beta).$$

This formulation will be useful for our characterization of the optimal fiscal rule, which will appeal to properties of the function $Q(\theta)$. For intuition, note that since the government is biased towards overborrowing relative to society, higher levels of borrowing can be attributed to larger distortions. In this sense, $Q(\theta)$ represents the weight that society places on allowing distortions by a government of type θ: the higher $Q(\theta)$, the lower the social welfare cost of distorting type θ's borrowing. The shape of this function will tell us how society wishes to allocate distortions across different government types.

3.2 Bang-Bang Incentives

Society uses continuation values as rewards and punishments to discipline the government and limit overborrowing. The next proposition shows that in any optimal rule, these rewards and punishments are extreme. That is, continuation values are bang-bang: given the feasible set $[\underline{V}(b), \overline{V}(b)]$, along the equilibrium path $V(b(\theta))$ only travels to the extreme points in this set.

Proposition 1 (necessity of bang-bang)**.** *Assume $Q(\theta)$ satisfies the generic property that $Q'(\theta) \neq 0$ almost everywhere. If $\{b(\theta), V(b(\theta))\}_{\theta \in \Theta}$ is an optimal rule with $b(\theta) \in (\underline{b}, \overline{b})$ for all $\theta \in \Theta$, then $V(b(\theta)) \in \left\{ \underline{V}(b(\theta)), \overline{V}(b(\theta)) \right\}$ for all $\theta \in \Theta$.*

This result shows that the bang-bang property is *necessary* for social welfare maximization. An optimal fiscal rule using only extreme continuation values always exists in our framework; this is true simply because an interior continuation value $V(b(\theta)) \in (\underline{V}(b(\theta)), \overline{V}(b(\theta)))$ can be assigned in expectation by randomizing over $\underline{V}(b(\theta))$ and $\overline{V}(b(\theta))$. Proposition 1 proves a stronger result: any rule that prescribes interior continuation values is strictly dominated by one with high-powered incentives.

Proposition 1 is consistent with other work such as Riley and Zeckhauser (1983) and Fuchs and Skrzypacz (2015), which also find high-powered incentives to be a feature of optimal contracting in adverse selection environments. As in those papers, the intuition for our bang-bang result stems from the linearity of payoffs, together with the richness of the information structure. Linearity emerges here because the penalties captured by $V(b(\theta))$ are joint for the government and society, entering the objective function and constraint set linearly in program (7).[18] Given a continuum of types, the richness of the information structure is guaranteed by the condition that we identify in Proposition 1, which says that the set of types θ for which $Q'(\theta) = 0$ is nowhere dense.[19] Intuitively, since $Q(\theta)$ is then either strictly decreasing or strictly increasing over any sufficiently small interval, society benefits from moving borrowing distortions towards either lower types or higher types in the interval. As such, spreading out continuation values always allows to reduce distortions. These values however must be spread out in an incentive feasible manner, satisfying the constraints in (11) which must hold type by type in our problem together with the monotonicity of the allocation.

Before we describe the proof of Proposition 1, it is worth commenting on the scope of this result. The proposition holds regardless of the tightness of enforcement constraints, and it therefore continues to hold absent constraint (5). Hence, our bang-bang characterization applies more generally to other delegation problems with money burning. One example is the widely studied delegation setting in which the agent's bias and private information take a multiplicative form; see Amador and Bagwell (2013). In our context, that coincides with the special case in which the highest and lowest feasible continuation values satisfy $\underline{V}(b) = \overline{V}(b) - m$ for some penalty $m > 0$. Proposition 1 implies that in any interior

---

[18]In Riley and Zeckhauser (1983) and Fuchs and Skrzypacz (2015), the bang-bang property applies to the probability of trade, which enters symmetrically in the objective functions of the seller and buyers.

[19]Given $f(\theta)$ continuously differentiable, this condition holds generically. Specifically, this condition fails only if $\theta f'(\theta)/f(\theta) = -(2 - \beta)/(1 - \beta)$ for a positive mass of types θ, but then any arbitrarily small perturbation of β would render the condition true.

solution to this problem, each agent type is assigned either no money burning or maximal money burning.[20],[21]

We next provide a summary of the proof of Proposition 1. The proof makes use of perturbation arguments and proceeds in three steps.[22]

Step 1 shows that in any optimal rule, $V(b(\theta))$ is left-continuous at each $\theta \in (\underline{\theta}, \overline{\theta}]$, a fact that we utilize for the rest of our analysis. Suppose this were not true at some $\theta \in (\underline{\theta}, \overline{\theta}]$. By Lemma 2, type θ must be indifferent between its allocation, $\{b(\theta), V(b(\theta))\}$, and its left-limit allocation, call it $\{b(\theta^-), V(b(\theta^-))\}$, where the latter has lower borrowing and higher continuation value: $b(\theta^-) < b(\theta)$ and $V(b(\theta^-)) > V(b(\theta))$. Since the government overweighs borrowing relative to society, it follows that a perturbation that assigns type θ its left-limit allocation strictly increases social welfare. Moreover, given this type's indifference, this perturbation is incentive feasible. Hence, $V(b(\theta))$ must be left-continuous, and we use a similar logic to show that $V(b(\underline{\theta})) = \overline{V}(b(\underline{\theta}))$ must also hold in any optimal rule.

Step 2 rules out incentive provision via locally increasing penalties. We show that an optimal rule cannot prescribe a continuation value $V(b(\theta))$ that is continuously strictly decreasing in θ and strictly interior (i.e. strictly between $\underline{V}(b(\theta))$ and $\overline{V}(b(\theta))$) over an interval $[\theta^L, \theta^H]$. Suppose by contradiction that such an interval exists. By Lemma 2, $b(\theta)$ must be strictly increasing over $[\theta^L, \theta^H]$, and by the generic property in Proposition 1, we can take an interval with either $Q'(\theta) > 0$ or $Q'(\theta) < 0$ for all $\theta \in [\theta^L, \theta^H]$. We then show that there exists an incentive feasible perturbation that strictly increases social welfare. If $Q'(\theta) < 0$, we construct a *flattening perturbation* that rotates the increasing $b(\theta)$ schedule clockwise over $[\theta^L, \theta^H]$, which entails reducing incentives with a counterclockwise rotation of the decreasing $V(b(\theta))$ schedule. This perturbation is socially beneficial because, given $Q'(\theta) < 0$, society prefers to concentrate distortions on lower rather than higher types. If instead $Q'(\theta) > 0$, we construct a *steepening perturbation* that drills a hole in the $b(\theta)$ schedule by making allocations in (θ^L, θ^H) no longer available, which entails increasing incentives by moving interior continuation values towards $V(b(\theta^L))$ or $V(b(\theta^H))$. This perturbation is socially beneficial because, given $Q'(\theta) > 0$, society prefers to concentrate distortions on higher rather than lower types. Figure 2 illustrates the perturbations.

Step 3 completes the proof by ruling out intermediate penalties. Suppose $V(b(\theta))$ was strictly interior at some $\theta \in \Theta$. By the previous steps and Lemma 2, type θ must belong

---

[20] Ambrus and Egorov (2013) provide examples in the setting of Amador, Werning, and Angeletos (2006) in which borrowing is at a corner and, as a result, interior punishments can be optimal.

[21] We conjecture that the claims may also extend to the monetary policy model of Athey, Atkeson, and Kehoe (2005).

[22] Some of the arguments that we use when $Q(\theta)$ is decreasing are similar to those employed by Athey, Atkeson, and Kehoe (2005) in their analysis of optimal inflation rules. Unlike in their work, where rules are perfectly enforced, our arguments take into account the constraints due to limited enforcement.
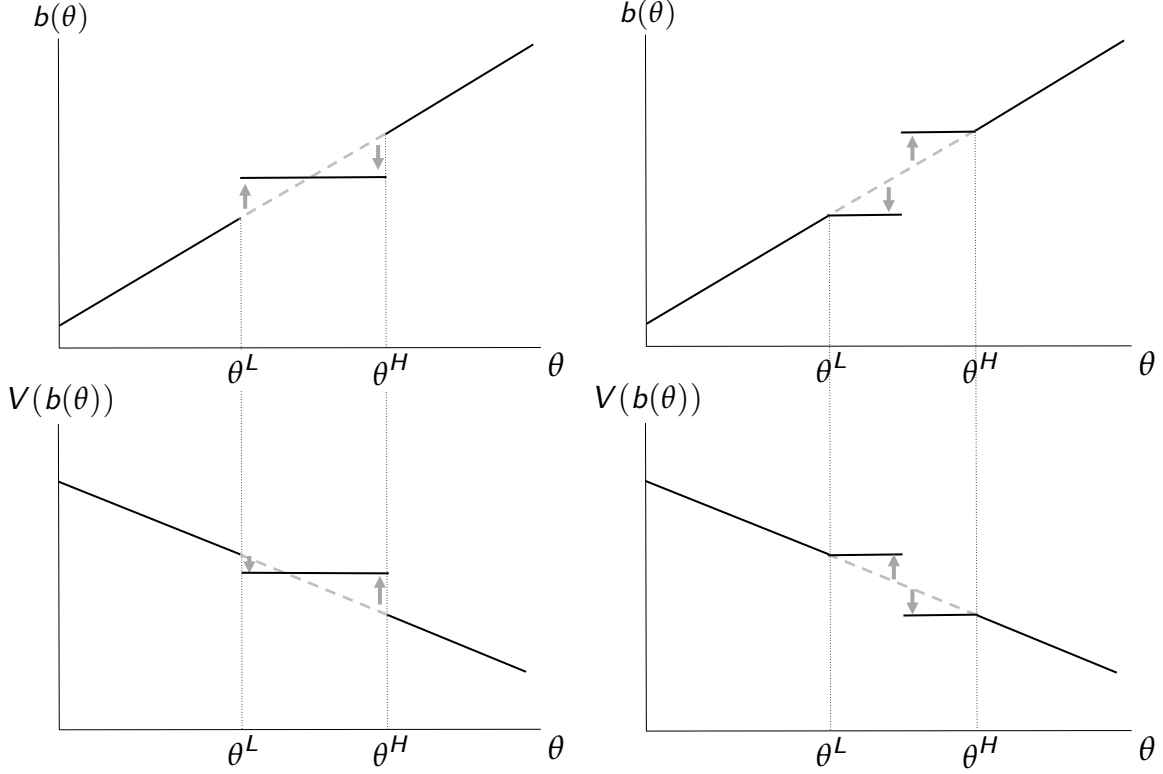
Figure 2: Examples of a flattening perturbation (left panel) and a steepening perturbation (right panel), as used in Step 2 of the proof of Proposition 1.

to a stand-alone segment $(\theta^L, \theta^H]$ such that $b(\theta) = b$ and $V(b(\theta)) = V$ for all $\theta \in (\theta^L, \theta^H]$, $b \in (\underline{b}, \overline{b})$ and $V \in (\underline{V}(b), \overline{V}(b))$ (by assumption), and $b(\theta)$ jumps at each boundary unless $\theta^H = \overline{\theta}$. We then show that there exists an incentive feasible perturbation that changes b and V slightly and strictly increases social welfare. If $\int_{\theta^L}^{\theta^H} Q(\theta)d\theta > \int_{\theta^L}^{\theta^H} Q(\theta^L)d\theta$, we perform a *segment-shifting steepening perturbation* over the interval $(\theta^L, \theta^H]$ as illustrated in Figure 3: we marginally increase b and reduce V so as to leave the government welfare of type θ^H unchanged, thus letting types arbitrarily close to θ^L jump *down* to a lower debt level. This perturbation is socially beneficial because, given $\int_{\theta^L}^{\theta^H} Q(\theta)d\theta > \int_{\theta^L}^{\theta^H} Q(\theta^L)d\theta$, society prefers to concentrate distortions on $(\theta^L, \theta^H]$ compared to θ^L. If instead $\int_{\theta^L}^{\theta^H} Q(\theta)d\theta \leq \int_{\theta^L}^{\theta^H} Q(\theta^L)d\theta$, the generic property in Proposition 1 ensures that $\int_{\theta^L}^{\theta^h} Q(\theta)d\theta < \int_{\theta^L}^{\theta^h} Q(\theta^L)d\theta$ for some $\theta^h \in (\theta^L, \theta^H]$. We then perform a *segment-shifting flattening perturbation* over the interval $(\theta^L, \theta^h]$: we marginally reduce b and increase V so as to leave the government welfare of type θ^h unchanged, thus letting types arbitrarily close to θ^L jump *up* to this allocation. This perturbation is socially beneficial because, given $\int_{\theta^L}^{\theta^h} Q(\theta)d\theta < \int_{\theta^L}^{\theta^h} Q(\theta^L)d\theta$, society prefers to concentrate distortions on θ^L compared to $(\theta^L, \theta^h]$.
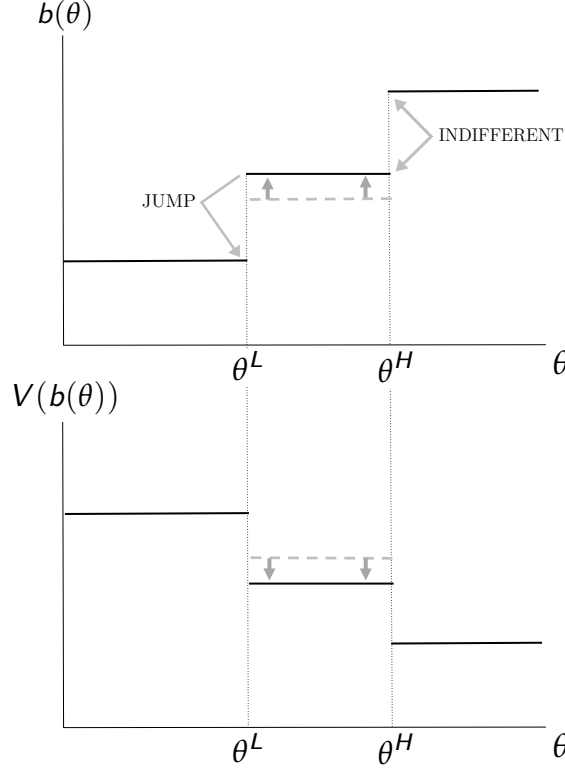
Figure 3: Example of a segment-shifting steepening perturbation, as used in Step 3 of the proof of Proposition 1.

3.3 Monotonic Incentives

To further characterize the schedule of continuation values $V(b(\theta))$, and solve for the optimal fiscal rule in the next subsection, we make the following assumption:

Assumption 1. *There exists $\widehat{\theta} \in \Theta$ such that $Q'(\theta) < 0$ if $\theta < \widehat{\theta}$ and $Q'(\theta) > 0$ if $\theta > \widehat{\theta}$.*

This assumption states that $Q(\theta)$ has a minimum value $Q(\widehat{\theta}) \leq Q(\overline{\theta}) < 0$, being strictly decreasing for $\theta < \widehat{\theta}$ and strictly increasing for $\theta > \widehat{\theta}$. Note that the assumption allows for $Q(\theta)$ to be strictly decreasing or strictly increasing over the whole set Θ; in this case $\widehat{\theta}$ is defined as either the upper bound or the lower bound of the set Θ. Assumption 1 implies the generic property required in Proposition 1, and it holds for a broad range of distribution functions, including uniform, exponential, log-normal, gamma, and beta for a subset of its parameters. This assumption is similar to, but stronger than, the distributional assumption used in Amador, Werning, and Angeletos (2006). We show in Subsection 3.5 that Assumption 1 is necessary for our characterization of optimal rules.

We maintain Assumption 1 for the remainder of our analysis. The next lemma shows that, given the implied shape of $Q(\theta)$, optimal incentives are monotonic.

16

Lemma 3. *If $\{b(\theta), V(b(\theta))\}_{\theta \in \Theta}$ is an optimal rule with $b(\theta) \in (\underline{b}, \overline{b})$ for all $\theta \in \Theta$, then either $V(b(\theta)) = \overline{V}(b(\theta))$ for all $\theta \in \Theta$, or there exists $\theta^{**} \in (\underline{\theta}, \overline{\theta})$ such that $V(b(\theta)) = \overline{V}(b(\theta))$ for all $\theta \in [\underline{\theta}, \theta^{**}]$ and $V(b(\theta)) = \underline{V}(b(\theta))$ for all $\theta \in (\theta^{**}, \overline{\theta}]$.*

The intuition for this result is related to the shape of the $Q(\theta)$ function, which tells us how society wishes to allocate borrowing distortions across types. For types $\theta > \widehat{\theta}$, where $\widehat{\theta}$ is defined in Assumption 1, society prefers to concentrate distortions on relatively high types. This is achieved by using high-powered incentives that punish high levels of borrowing by lowering the continuation value from $\overline{V}(b(\theta))$ to $\underline{V}(b(\theta))$. In contrast, for types $\theta < \widehat{\theta}$, society prefers to concentrate distortions on relatively low types. This is achieved by using flat incentives that keep the continuation value constant at $\overline{V}(b(\theta))$.

The proof of Lemma 3 consists of three steps, which we summarize here. Step 1 shows that any interval of types receiving the maximal penalty $\underline{V}(b(\theta))$ must lie above $\widehat{\theta}$. Suppose by contradiction that this were not true, so that an optimal rule prescribes $V(b(\theta)) = \underline{V}(b(\theta))$ over an interval $[\theta^L, \theta^H]$ with $Q'(\theta) < 0$ for $\theta \in [\theta^L, \theta^H]$. Note that the enforcement constraint requires $b(\theta) = b^p(\theta)$ for all types θ in the interval. We then show that there exists an incentive feasible perturbation that strictly increases social welfare. If $\frac{db^p(\theta)}{d\theta} > 0$ for $\theta \in [\theta^L, \theta^H]$, we perform a flattening perturbation that rotates the borrowing schedule clockwise; by the logic in Step 2 in the proof of Proposition 1 and $Q'(\theta) < 0$, this perturbation is socially beneficial. If instead $b^p(\theta)$ is constant over $[\theta^L, \theta^H]$, we perform a segment-shifting flattening perturbation; by the logic in Step 3 in the proof of Proposition 1 and $Q'(\theta) < 0$, this perturbation is socially beneficial.

Step 2 establishes that incentives are monotonic. We show that if $V(b(\theta^{**})) = \underline{V}(b(\theta^{**}))$ for some type $\theta^{**} \geq \widehat{\theta}$ in an optimal rule, then $V(b(\theta)) = \underline{V}(b(\theta))$ for all types $\theta \geq \theta^{**}$. Suppose by contradiction that $V(b(\theta)) = \overline{V}(b(\theta))$ for some type $\theta > \theta^{**}$. By our left-continuity and bang-bang results, there must exist an interval $(\theta^L, \theta^H]$, $\theta^L \geq \theta^{**}$, such that $V(b(\theta)) = \overline{V}(b(\theta))$ for all $\theta \in (\theta^L, \theta^H]$. We then show that there exists an incentive feasible perturbation that strictly increases social welfare. If $\frac{db(\theta)}{d\theta} > 0$ for $\theta \in (\theta^L, \theta^H]$, we perform a steepening perturbation that drills a hole in the $b(\theta)$ schedule; by the logic in Step 2 in the proof of Proposition 1 and $Q'(\theta) > 0$, this perturbation is socially beneficial. If instead $b(\theta)$ is constant over $(\theta^L, \theta^H]$, we can take a stand-alone segment with constant borrowing lying above $\widehat{\theta}$. Using similar logic as in Step 3 in the proof of Proposition 1 and $Q'(\theta) > 0$, we show that there exists a segment-shifting steepening perturbation that is socially beneficial.

The previous steps, Proposition 1, and Step 1 in the proof of that proposition imply the following: if $V(b(\theta)) = \underline{V}(b(\theta))$ for some type $\theta \in \Theta$, then there exists $\theta^{**} \in [\underline{\theta}, \overline{\theta})$ such that $V(b(\theta)) = \overline{V}(b(\theta))$ for all $\theta \in [\underline{\theta}, \theta^{**}]$ and $V(b(\theta)) = \underline{V}(b(\theta))$ for all $\theta \in (\theta^{**}, \overline{\theta}]$.

Step 3 completes the proof of Lemma 3 by showing that $\theta^{**} > \underline{\theta}$. The idea is simple: if $\theta^{**} = \underline{\theta}$, we can perform a global perturbation that increases the continuation value of types $\theta \in (\underline{\theta}, \overline{\theta}]$ by a constant amount $\Delta > 0$ and assigns type $\underline{\theta}$ the limit allocation to its right. This perturbation keeps the borrowing level of types $\theta \in (\underline{\theta}, \overline{\theta}]$ unchanged and is incentive feasible for $\Delta > 0$ small enough. Moreover, using the representation in (12), we show that the perturbation increases social welfare by $\delta \Delta > 0$.

3.4 Maximally Enforced Deficit Limit

Proposition 1 and Lemma 3 characterize the schedule of continuation values in any optimal rule. To provide a characterization of the borrowing schedule, we require:

Assumption 2. *For all $b \in [\underline{b}, \overline{b}]$, $\overline{V}(b)$ is continuously differentiable and strictly concave.*

This assumption ensures that the government's flexible level of debt conditional on maximal reward, $b^r(\theta)$, is continuous in its type θ. We maintain this assumption for the remainder of our analysis.[23] The next proposition states the main result of the paper:

Proposition 2 (optimal rule)**.** *If $\{b(\theta), V(b(\theta))\}_{\theta \in \Theta}$ is an optimal rule with $b(\theta) \in (\underline{b}, \overline{b})$ for all $\theta \in \Theta$, then it satisfies (9)-(10) for some $\theta^* \in [0, \overline{\theta})$ and finite $\theta^{**} > \max\{\theta^*, \underline{\theta}\}$. Hence, any interior solution is a maximally enforced deficit limit.*

Recall from Lemma 3 that in any optimal rule, either all government types $\theta \in \Theta$ receive the highest continuation value $\overline{V}(b(\theta))$, or the continuation value jumps down from $\overline{V}(b(\theta))$ to $\underline{V}(b(\theta))$ at a point $\theta^{**} \in (\underline{\theta}, \overline{\theta})$. To establish Proposition 2, we take θ^{**} as given and solve for the optimal borrowing allocation above and below this point. If $\theta \in (\theta^{**}, \overline{\theta}]$, the allocation is characterized by the binding enforcement constraint with $\{b(\theta), V(b(\theta))\} = \{b^p(\theta), \underline{V}(b^p(\theta))\}$. If $\theta \in [\underline{\theta}, \theta^{**}]$, we show that the borrowing schedule $b(\theta)$ must be continuous, so the allocation takes the form of bounded discretion. Since a minimum borrowing requirement would reduce social welfare (given the government's bias towards overborrowing), a maximum borrowing limit $b^r(\theta^*)$ is optimal for types in this range. Such a borrowing limit necessarily keeps type θ^{**} indifferent between the allocations $\{b^r(\theta^*), \overline{V}(b^r(\theta^*))\}$ and $\{b^p(\theta^{**}), \underline{V}(b^p(\theta^{**}))\}$.

The proof that $b(\theta)$ is continuous over $[\underline{\theta}, \theta^{**}]$ builds on perturbation arguments similar to those used in the previous subsections. Given the absence of penalty incentives below θ^{**}, a discontinuity would entail a hole in some range $[\theta^L, \theta^H]$. That is, $b(\theta)$ would jump at a point $\theta^M \in (\theta^L, \theta^H)$, such that types $\theta \in (\theta^L, \theta^M)$ borrow at $b^r(\theta^L) < b^r(\theta)$ and

---

[23]This assumption holds in our infinite horizon setting in which $\overline{V}(b)$ is endogenous; see Lemma 4.

types $\theta \in (\theta^M, \theta^H)$ borrow at $b^r(\theta^H) > b^r(\theta)$. However, if $\theta^M < \widehat{\theta}$, an incentive feasible perturbation that slightly closes the hole by raising θ^L (and thus also raising θ^M) would strictly increase social welfare, by similar logic as in Step 2 in the proof of Proposition 1 and $Q'(\theta) < 0$. Moreover, if $\theta^M \geq \widehat{\theta}$, the interval $(\theta^M, \theta^H]$ must belong to a stand-alone segment with constant borrowing.[24] An incentive feasible segment-shifting perturbation would then strictly increase social welfare, by similar logic as in Step 2 in the proof of Lemma 3 and $Q'(\theta) > 0$. It follows that $b(\theta)$ cannot be discontinuous below θ^{**}.

Proposition 2 shows that any interior solution must be a maximally enforced deficit limit, but it is silent on whether this limit is violated along the equilibrium path, namely whether $\theta^{**} < \overline{\theta}$. To address this issue, consider first the problem under perfect enforcement, as in the work of Amador, Werning, and Angeletos (2006). If the enforcement constraint (5) can be ignored, and given Assumption 1 and Assumption 2, it follows from Proposition 2 that the optimal rule solves:

$$\max_{\theta^* \in [0, \overline{\theta}]} \left\{ \int_{\underline{\theta}}^{\theta^*} \left(\theta U(\omega + b^r(\theta)) + \delta \overline{V}(b^r(\theta)) \right) f(\theta) d\theta + \int_{\theta^*}^{\overline{\theta}} \left(\theta U(\omega + b^r(\theta^*)) + \delta \overline{V}(b^r(\theta^*)) \right) f(\theta) d\theta \right\}.$$

Denote the solution by $\theta_e \in [0, \overline{\theta})$. Using the definition of $b^r(\theta)$ and integration by parts, the first-order condition can be shown to be equivalent to[25]

$$\int_{\theta_e}^{\overline{\theta}} Q(\theta) d\theta = 0. \tag{13}$$

The optimal deficit limit under perfect enforcement is such that, on average, the distortion above the limit is zero. Clearly, if this limit can be enforced given $\{\underline{V}(b), \overline{V}(b)\}$, then it is also optimal under limited enforcement:

Corollary 1. *Suppose*

$$\overline{\theta} U(\omega + b^r(\theta_e)) + \beta \delta \overline{V}(b^r(\theta_e)) \geq \overline{\theta} U(\omega + b^p(\overline{\theta})) + \beta \delta \underline{V}(b^p(\overline{\theta})). \tag{14}$$

If $\{b(\theta), V(b(\theta))\}_{\theta \in \Theta}$ is an optimal rule with $b(\theta) \in (\underline{b}, \overline{b})$ for all $\theta \in \Theta$, then it is the perfect-enforcement deficit limit, with $\theta^* = \theta_e$ and $\theta^{**} \geq \overline{\theta}$.

When condition (14) holds, the highest type $\overline{\theta}$, and therefore all types $\theta \in \Theta$, prefer to respect the perfect-enforcement limit $b^r(\theta_e)$ and receive maximal reward rather than

---

[24]Otherwise, by the same logic as in Step 2 in the proof of Lemma 3, social welfare could be increased by drilling a hole in the borrowing allocation above $\widehat{\theta}$.

[25]Note that $Q(\theta) = 1$ for $\theta < \underline{\theta}$.

spend above the limit and receive maximal punishment. The optimal rule under limited enforcement therefore coincides with that under perfect enforcement and features no on-path penalties.

Our interest is in characterizing the optimal rule when condition (14) does not hold, so the perfect-enforcement limit $b^r(\theta_e)$ is not enforceable given $\{\underline{V}(b), \overline{V}(b)\}$. To this end, we can define a unique type θ_c as corresponding to the tightest deficit limit that all types $\theta \in \Theta$ would be willing to respect:

$$\overline{\theta}U(\omega + b^r(\theta_c)) + \beta\delta\overline{V}(b^r(\theta_c)) = \overline{\theta}U(\omega + b^p(\overline{\theta})) + \beta\delta\underline{V}(b^p(\overline{\theta})). \tag{15}$$

Note that $\theta_c \leq \theta_e$ whenever the perfect-enforcement limit is enforceable, and $\theta_c > \theta_e$ otherwise. The value of θ_c increases if $\underline{V}(b)$ is increased for every b, holding all else fixed. Intuitively, if the severity of the penalties that can be imposed on the government is diminished, then the tightest deficit limit that can be enforced becomes less stringent. Moreover, note that the value of θ_c is only a function of preferences (including $\{\underline{V}(b), \overline{V}(b)\}$) and the value of $\overline{\theta}$, and is thus independent of the distribution of types over the support $[\underline{\theta}, \overline{\theta}]$. Using this definition of θ_c, the next proposition provides a necessary and sufficient condition for penalties to be optimally used along the equilibrium path.

Proposition 3 (use of punishment). *If $\{b(\theta), V(b(\theta))\}_{\theta \in \Theta}$ is an optimal rule with $b(\theta) \in (\underline{b}, \overline{b})$ for all $\theta \in \Theta$, then it is the unique such rule. Moreover, if*

$$\int_{\theta_c}^{\overline{\theta}} (Q(\theta) - Q(\overline{\theta}))d\theta \geq 0, \tag{16}$$

this rule a maximally enforced deficit limit with $\theta^* = \max\{\theta_c, \theta_e\}$ and $\theta^{**} \geq \overline{\theta}$. Otherwise, this rule is a maximally enforced deficit limit with $\theta^* \in (\theta_e, \theta_c)$ and $\theta^{**} < \overline{\theta}$.

Whenever the perfect-enforcement limit $b^r(\theta_e)$ is enforceable, i.e. $\theta_c \leq \theta_e$, Assumption 1 guarantees that $\int_{\theta_c}^{\overline{\theta}} Q(\theta)d\theta \geq \int_{\theta_e}^{\overline{\theta}} Q(\theta)d\theta = 0$. Hence, in this case, condition (16) is satisfied and the optimal rule coincides with that under perfect enforcement, as noted in Corollary 1.

If instead the perfect-enforcement limit $b^r(\theta_e)$ is not enforceable, i.e. $\theta_c > \theta_e$, then society faces the following tradeoff. On the one hand, society can raise the value of θ^* to the point that the associated limit $b^r(\theta^*)$ satisfies the enforcement constraint of type $\overline{\theta}$ and thus of all types $\theta \in \Theta$. This option entails setting $\theta^* = \theta_c$ and $\theta^{**} = \overline{\theta}$ and has the benefit of avoiding socially costly penalties along the equilibrium path, albeit at the cost of potentially allowing significant overborrowing within the relaxed deficit limit. On the other hand, society can impose a tighter limit $b^r(\theta^*)$ which does not satisfy the enforcement

constraint of all types. This option sets $\theta^* < \theta_c$ and $\theta^{**} < \bar{\theta}$ and induces higher discipline on types $\theta \leq \theta^{**}$, but at the cost of imposing penalties whenever a shock $\theta > \theta^{**}$ is realized.

Proposition 3 shows that which of these two options is optimal for society depends on whether the inequality in (16) holds or not. The proof of this proposition uses the properties of $Q(\theta)$ to show that there exist a unique threshold θ^* and associated θ^{**} satisfying (10) that optimally resolve the tradeoff between imposing fiscal discipline and avoiding punishments. The intuition for condition (16) is familiar by now: it tells us how society wishes to allocate borrowing distortions, and in particular whether society prefers to concentrate distortions on types $\theta \in [\theta_c, \bar{\theta})$ versus $\bar{\theta}$. A relaxed deficit limit that avoids punishments concentrates distortions on $[\theta_c, \bar{\theta})$, whereas tightening the deficit limit by the use of punishment moves distortions towards $\bar{\theta}$.

Holding fixed the distribution of shocks, condition (16) shows that the use of punishment depends on the ease of enforcement. Recall that θ_c increases if $\underline{V}(b)$ is increased for every b. Moreover, if $\hat{\theta}$ in Assumption 1 is interior, then given this assumption, (16) holds for θ_c low enough but cannot be satisfied if θ_c becomes sufficiently high. Therefore, we find that for any fixed distribution of shocks admitting an interior value of $\hat{\theta}$, we can make penalties sufficiently severe that punishment is not imposed on path, or sufficiently weak that on-path punishment is optimally used. When penalties are weak, the tightest deficit limit that all government types would be willing to respect is too lax, so society prefers to set a tighter limit and let high enough types violate it.

Additionally, holding fixed the support $[\underline{\theta}, \bar{\theta}]$ and the value of θ_c (which, as noted, is determined by preferences), condition (16) tells us how the use of punishment depends on the distribution of shocks. To see this, rewrite (16) as follows (see the proof of Proposition 3 for a derivation):

$$\frac{\mathbb{E}[\theta|\theta \geq \theta_c]}{\theta_c} \geq \frac{1}{\beta} - \left(\frac{\bar{\theta}}{\theta_c} - 1\right) \frac{\bar{\theta} f(\bar{\theta})}{1 - F(\theta_c)} \left(\frac{1}{\beta} - 1\right). \tag{17}$$

For any fixed $[\underline{\theta}, \bar{\theta}]$ and θ_c, we can find a distribution of shocks such that on-path punishment is or is not optimal. Specifically, (17) implies that punishment is not imposed on path if high shocks are likely and expected fiscal needs conditional on the tightest enforceable limit are large, namely if $f(\bar{\theta})/(1-F(\theta_c))$ and $\mathbb{E}[\theta|\theta \geq \theta_c]/\theta_c$ are sufficiently high. This situation arises, for example, under a uniform distribution of shocks. Society in this case benefits from setting a relaxed deficit limit that is never violated, as punishing the government following high shocks would be too costly. In contrast, punishment is optimally imposed on path if $f(\bar{\theta})/(1 - F(\theta_c))$ and $\mathbb{E}[\theta|\theta \geq \theta_c]/\theta_c$ are sufficiently low. This situation arises whenever the perfect-enforcement limit is not enforceable (i.e., $\theta_c > \theta_e$) and $\lim_{\theta \to \bar{\theta}} f(\theta) =$

0, as is true for example under a beta distribution with beta shape parameter greater than one.[26] Society in this case benefits from setting a tight deficit limit that is violated following high enough shocks, as such events are sufficiently rare that the cost of punishing the government is then relatively low.

3.5 Discussion of Distributional Assumption

Our characterization in Subsection 3.4 is derived under Assumption 1, which pins down the shape of the function $Q(\theta)$. Assumption 1 is sufficient to obtain the unique optimality of maximally enforced deficit limits, among interior solutions, under any (bounded and continuous) functions $\underline{V}(b)$ and $\overline{V}(b)$ satisfying Assumption 2. In this section, we explore the necessity of Assumption 1 for our findings.

Definition 2. *Assumption 1 is* weakly violated *if there exist $\theta^L, \theta^H \in \Theta$, $\theta^H > \theta^L$, and $\Delta > 0$ such that (i) $Q'(\theta) \geq 0$ for $\theta \in [\theta^L, \theta^L + \Delta]$ and (ii) $Q'(\theta) \leq 0$ for $\theta \in [\theta^H - \Delta, \theta^H]$. Assumption 1 is* strictly violated *if the inequalities in (i)-(ii) are strict.*

We find that both weak and strict violations of Assumption 1 would affect our results:

Proposition 4 (necessity of distributional assumption)**.** *Consider functions $\underline{V}(b)$ and $\overline{V}(b)$ satisfying Assumption 2. If Assumption 1 is weakly violated, then for any function $\overline{V}(b)$, there exists a function $\underline{V}(b)$ under which not every optimal rule with $b(\theta) \in (\underline{b}, \overline{b})$ for all $\theta \in \Theta$ is a maximally enforced deficit limit. Moreover, if Assumption 1 is strictly violated, then for any function $\overline{V}(b)$, there exists a function $\underline{V}(b)$ under which no such optimal rule is a maximally enforced deficit limit.*

This proposition shows that Assumption 1 is necessary for maximally enforced deficit limits to be uniquely optimal (among interior solutions) given any functions $\{\overline{V}(b), \underline{V}(b)\}$. In this sense, our analysis identifies the minimal structure that guarantees the unique optimality of this class of rules, a class that resembles fiscal rules commonly used in practice in the form of deficit, spending, and debt limits. Note that both weak and strict violations of Assumption 1 are possible even when the generic property in Proposition 1 is satisfied. Under these violations, an optimal rule under limited enforcement would feature bang-bang incentives yet induce an allocation that is not implementable by a deficit limit. We prove the first part of Proposition 4 by construction and the second part by contradiction.[27]

---

[26]Note that for $\lim_{\theta \to \overline{\theta}} f(\theta) = 0$, condition (17) becomes $\frac{\mathbb{E}[\theta | \theta \geq \theta_c]}{\theta_c} \geq \frac{1}{\beta}$. Thus, given θ_c and $\beta < 1$, this condition is violated under distributions of shocks with sufficiently low mass above θ_c.

[27]As noted in Subsection 3.3, Assumption 1 is stronger than the distributional assumption used in Amador, Werning, and Angeletos (2006). Define θ_a as the lowest value such that $\int_{\theta}^{\overline{\theta}} Q(\widetilde{\theta}) d\widetilde{\theta} \leq 0$ for all

4 Self-Enforcing Fiscal Rule

We have characterized the optimal fiscal rule under limited enforcement when the extent of penalties that can be imposed on the government is exogenously given. In this section, we extend our analysis to an infinite horizon environment in which the set of feasible penalties is endogenously determined. To focus our analysis, we assume that there is no external enforcement authority that inflicts sanctions. Instead, the government's policy decision in each period is *self-enforced* by the behavior of future governments.

Specifically, we define a self-enforcing fiscal rule as a perfect public equilibrium of the interaction between successive governments. Such a rule must satisfy a sequence of private information and self-enforcement constraints which are analogous to those in Section 3. In fact, the recursive representation of the optimal rule is equivalent to that in our static environment, where the highest and lowest feasible continuation values given the level of debt, $\overline{V}(b)$ and $\underline{V}(b)$, are now determined by the set of perfect public equilibria of the game. As such, we are able to establish that our results in Section 3 extend to this infinite horizon setting, yielding a characterization of the allocation that corresponds to maximal reward $\overline{V}(b)$. Additionally, we provide a characterization of the allocation that corresponds to maximal punishment $\underline{V}(b)$, and we describe the dynamics of fiscal policy under the optimal self-enforcing rule.

4.1 Infinite Horizon Setting

Consider an infinite horizon setting with periods $t = \{0, 1, \ldots\}$. At the beginning of each period, the government privately observes a shock θ_t, independent and identically distributed (i.i.d.) according to $f(\theta_t)$ and $F(\theta_t)$. We denote by b_t and g_t respectively the government's choices of debt and spending in period t. The government's budget constraint is the dynamic analog of (1), given by

$$g_t = \tau - Rb_{t-1} + b_t, \tag{18}$$

where $\tau > 0$ is the exogenous tax revenue at every date and $R > 1$ is the exogenous gross interest rate on government bonds.

$\theta \geq \theta_a$. Then using our notation and taking $f(\theta)$ to be differentiable, Amador, Werning, and Angeletos (2006) assume $Q'(\theta) \leq 0$ for all $\theta \leq \theta_a$. One can verify that there are distributions $F(\theta)$ satisfying this assumption for which Assumption 1 is strictly violated, and therefore for which there exist continuation value functions such that no optimal rule with interior borrowing is a maximally enforced deficit limit.

Social welfare at the beginning of date t is

$$\sum_{k=0}^{\infty} \delta^k \mathbb{E}\left[\theta_{t+k} U(g_{t+k})\right],$$

whereas the government's welfare following the realization of its type θ_t at date t is

$$\theta_t U(g_t) + \beta \sum_{k=1}^{\infty} \delta^k \mathbb{E}\left[\theta_{t+k} U(g_{t+k})\right]. \tag{19}$$

To make the analysis tractable (as we explain subsequently), we take a social utility of public spending with constant relative risk aversion (CRRA):

$$U(g) = \begin{cases} \dfrac{g^{1-\gamma} - 1}{1 - \gamma} & \text{for } \gamma > 0,\ \gamma \neq 1, \\[2mm] \log(g) & \text{for } \gamma = 1. \end{cases} \tag{20}$$

Following Bernheim, Ray, and Yeltekin (2015), we also require the level of debt in each period t to satisfy

$$b_t \in [\underline{b}(b_{t-1}), \overline{b}(b_{t-1})], \tag{21}$$

where $\underline{b}(b_{t-1}) = \frac{\tau}{R-1} - (1-\nu) R \left(\frac{\tau}{R-1} - b_{t-1}\right)$ and $\overline{b}(b_{t-1}) = \frac{\tau}{R-1} - \nu R \left(\frac{\tau}{R-1} - b_{t-1}\right)$ for some small $\nu \in (0, \frac{1}{2})$. Equation (21) thus defines lower and upper bounds on the fraction of lifetime resources that the government can borrow or save at any time. Note that for $\gamma \geq 1$, (20) yields $U(0) = -\infty$, and therefore (21) is necessary to guarantee that payoffs (and thus punishments) are bounded. We then take $\nu > 0$ small enough that this constraint is otherwise non-binding.[28]

4.2 Equilibrium Definition

We consider the interaction between the governments in each period $t = \{0, 1, \ldots\}$. Our equilibrium concept is perfect public equilibrium. Let $h^{t-1} = \{b_{-1}, b_0, \ldots, b_{t-1}\}$ denote the public history of debt through time $t-1$ and $\mathcal{H}^{t-1}$ the set of all possible such histories. A public strategy for the government in period t is $\sigma_t(h^{t-1}, \theta_t)$, specifying, for each history $h^{t-1} \in \mathcal{H}^{t-1}$ and current government type $\theta_t \in \Theta$, a feasible level of debt $b_t(h^{t-1}, \theta_t)$. Note that given the budget constraint (18), a public history of debt also pins down the public history of spending, and the government's strategy also pins down its choice of spending.

---

[28]See Laibson (1994) for further discussion of the necessity of these bounds in the quasi-hyperbolic model.

Specifically, denoting by $\omega_t(h^{t-1}) \equiv \tau - Rb_{-1}(h^{t-1})$ the available resources at history h^{t-1}, spending given h^{t-1} and θ_t is $g_t(h^{t-1}, \theta_t) = \omega_t(h^{t-1}) + b_t(h^{t-1}, \theta_t)$. We define a perfect public equilibrium as a profile of public strategies $\sigma = (\sigma_t(h^{t-1}, \theta_t))_{t=0}^{\infty}$ such that, for each $t \in \{0, 1, \ldots\}$, $\sigma_t(h^{t-1}, \theta_t)$ maximizes the t-period government's welfare (19) given the continuation strategies $(\sigma_{t+k}(h^{t+k-1}, \theta_{t+k}))_{k=1}^{\infty}$ of all future governments. We henceforth refer to perfect public equilibria as simply equilibria.

Given an equilibrium, let $V_t(h^{t-1})$ denote the continuation value to society at date t starting from a history h^{t-1}. At any (on- or off-path) history h^{t-1}, the continuation value given the equilibrium strategies can be represented recursively:

$$V_t(h^{t-1}) = \mathbb{E}\left[\theta_t U(\omega_t(h^{t-1}) + b_t(h^{t-1}, \theta_t)) + \delta V_{t+1}(h^{t-1}, b_t(h^{t-1}, \theta_t)))\right]. \tag{22}$$

Analogous to constraints (4) and (5) in our static environment, a profile of strategies $(\sigma_t(h^{t-1}, \theta_t))_{t=0}^{\infty}$ constitutes an equilibrium if and only if, for all $t \in \{0, 1 \ldots\}$ and all (on- and off-path) histories h^{t-1}, the following private information and enforcement constraints are satisfied:

$$\theta_t U(\omega_t(h^{t-1}) + b_t(h^{t-1}, \theta_t)) + \beta\delta V_{t+1}(h^{t-1}, b_t(h^{t-1}, \theta_t)) \tag{23}$$
$$\geq \theta_t U(\omega_t(h^{t-1}) + b_t(h^{t-1}, \theta_t')) + \beta\delta V_{t+1}(h^{t-1}, b_t(h^{t-1}, \theta_t')) \ \ \text{for all } \theta_t, \theta_t' \in \Theta$$

and

$$\theta_t U(\omega_t(h^{t-1}) + b_t(h^{t-1}, \theta_t)) + \beta\delta V_{t+1}(h^{t-1}, b_t(h^{t-1}, \theta_t)) \tag{24}$$
$$\geq \theta_t U(\omega_t(h^{t-1}) + b_t^p(h^{t-1}, \theta_t)) + \beta\delta\underline{V}_{t+1}(h^{t-1}, b_t^p(h^{t-1}, \theta_t)) \ \ \text{for all } \theta_t \in \Theta.$$

Here $\underline{V}_{t+1}(h^{t-1}, b_t^p(h^{t-1}, \theta_t))$ denotes the lowest continuation value that can be supported by equilibrium strategies at history $(h^{t-1}, b_t^p(h^{t-1}, \theta_t))$, and $b_t^p(h^{t-1}, \theta_t)$ denotes type θ_t's flexible debt level given maximal punishment following history (h^{t-1}, b_t) for any feasible b_t. That is, analogous to our definition in the static environment, $b_t^p(h^{t-1}, \theta_t) \in \arg\max_{b_t \in [\underline{b}(b_{t-1}), \overline{b}(b_{t-1})]} \{\theta_t U(\omega(h^{t-1}) + b_t) + \beta\delta\underline{V}_{t+1}(h^{t-1}, b_t)\}$.

We examine the equilibrium that maximizes social welfare at date 0; we call this equilibrium the optimal self-enforcing fiscal rule. Let $\underline{V}(b)$ and $\overline{V}(b)$ denote respectively the lowest and highest levels of date-0 social welfare that can be sustained by equilibrium strategies given an initial level of debt $b_{-1} = b$, where we assume $\underline{V}(b) < \overline{V}(b)$.[29] Note that given the repeated nature of the game and the fact that shocks are i.i.d., we can

---

[29]That is, we assume the equilibrium set is not a singleton. Halac and Yared (2017b) analyze a specific case of our model and show that this assumption amounts to the discount factor δ being high enough.

represent initial policies under an optimal self-enforcing rule recursively starting from date 0. That is, as shown in Chade, Prokopovych, and Smith (2008), the recursive structure of the continuation values implies that the techniques of Abreu, Pearce, and Stacchetti (1990) can be applied to our setting. Thus, rather than optimizing over an entire debt sequence, we can assign each type $\theta_0 \in \Theta$ a date-0 level of debt $b_0(\theta_0)$ and continuation value $V_1(b_0(\theta_0))$, where these must satisfy the private information and self-enforcement constraints, and where the continuation value is itself drawn from the set of continuation values $\left[\underline{V}(b_0(\theta_0)), \overline{V}(b_0(\theta_0))\right]$ that satisfy the private information and self-enforcement constraints. Letting $\{b(\theta), V(b(\theta))\}_{\theta \in \Theta} \equiv \{b_0(\theta_0), V_1(b_0(\theta_0))\}_{\theta_0 \in \Theta}$ and $\omega \equiv \omega_0(h^{-1})$, it therefore follows that an optimal self-enforcing rule in the infinite horizon setting solves our static program in (7).

4.3 Optimal Self-Enforcing Fiscal Rule

Define the savings rate at date t, s_t, as the fraction of lifetime resources that are not spent at t:

$$1 - s_t = \frac{\tau - Rb_{t-1} + b_t}{R\tau/(R-1) - Rb_{t-1}}, \tag{25}$$

where, by (21), feasibility requires $s_t \in [\nu, 1-\nu]$. Note that for any period $t \in \{0, 1, \ldots\}$ and public history of debt $h^{t-1} = \{b_{-1}, b_0, \ldots, b_{t-1}\}$, there is a corresponding public history of initial debt and subsequent savings rates, $\widetilde{h}^{t-1} = \{b_{-1}, s_0, \ldots, s_{t-1}\}$. Moreover, a strategy for the government in period t can be equivalently defined as specifying either a debt level $b_t(h^{t-1}, \theta_t)$ for each history h^{t-1} and current type θ_t, or a savings rate $s_t(\widetilde{h}^{t-1}, \theta_t)$ for each history $\widetilde{h}^{t-1}$ and current type θ_t.

In the Online Appendix, we consider the representation of our problem using savings rates and prove the following results. First, given our assumption on preferences in (20), we show that whether or not a profile of savings rate strategies constitutes an equilibrium is independent of the initial level of debt b_{-1}. The reason is that, while social welfare does depend on b_{-1}, the private information and enforcement constraints (23) and (24) are scalable in debt under (20), and therefore only depend on the sequence of savings rates. Second, as a consequence, we establish that the sequences of savings rates that sustain the lowest and highest continuation values $\underline{V}(b)$ and $\overline{V}(b)$ given initial debt b are independent of b. Finally, using this result and (20), we show:

Lemma 4. $\underline{V}(b)$ and $\overline{V}(b)$ are continuously differentiable and strictly concave in b.

Combined with our observations in the previous subsection and our results in Section 3, this lemma implies that the optimal date-0 policies, $\{b(\theta), V(b(\theta))\}_{\theta \in \Theta}$, are characterized

as follows:

Proposition 5 (infinite horizon). *If $\{b(\theta), V(b(\theta))\}_{\theta \in \Theta}$ is an optimal rule in the infinite horizon economy with $b(\theta) \in (\underline{b}(b_{-1}), \overline{b}(b_{-1}))$ for all $\theta \in \Theta$, then it satisfies (9)-(10) for some $\theta^* \in [0, \overline{\theta})$ and finite $\theta^{**} > \max\{\theta^*, \underline{\theta}\}$. Hence, any interior solution for the maximal reward is a maximally enforced deficit limit. Moreover, θ^* and θ^{**} are unique and independent of b_{-1}.*

The first part of the proposition, on the optimality of maximally enforced deficit limits, is a direct application of Proposition 2, taking into account the properties established in Lemma 4. Observe that given preferences as specified in (20), if $\gamma \geq 1$, then the borrowing allocation is always interior. The reason is that the bounds on savings rates are defined as $[\nu, 1 - \nu]$ for ν arbitrarily close to zero, and thus a corner allocation with $s_t = \nu$ or $s_t = 1 - \nu$ would imply arbitrarily low social welfare under $\gamma \geq 1$. Hence, any optimal fiscal rule features interior debt and, as a result, is a maximally enforced deficit limit.

The second part of Proposition 5 shows that the thresholds θ^* and θ^{**} that define an optimal maximally enforced deficit limit are unique and independent of the initial level of debt b_{-1}. Uniqueness follows from the results in Proposition 3. To see why the thresholds do not depend on b_{-1}, recall that we can represent an optimal fiscal rule as a sequence of savings rates, and given (20), such a sequence is independent of initial debt. A simple implementation of this rule is a maximally enforced minimal savings rate s^*, where s^* corresponds to type θ^*'s flexible savings rate under maximal reward (which is independent of the level of debt). That is, under this rule, types $\theta \leq \theta^*$ choose a savings rate weakly higher than s^*, types $\theta \in (\theta^*, \theta^{**}]$ are constrained and choose s^*, and types $\theta > \theta^{**}$ choose a savings rate strictly lower than s^*. Types $\theta \leq \theta^{**}$ are rewarded at date 1 with a continuation value $\overline{V}_1(b_0)$ whereas types $\theta > \theta^{**}$ are punished at date 1 with a continuation value $\underline{V}_1(b_0)$.

Proposition 5 describes the policies that maximize social welfare given an initial level of debt, and thus those that sustain the maximal reward $\overline{V}_1(b_0)$. To complete our characterization, we next proceed to describe the policies that minimize social welfare given an initial level of debt, namely those that sustain the worst punishment $\underline{V}_1(b_0)$.

4.4 Optimal Punishment

In principle, different continuation equilibria could serve as punishment for a government breaking the deficit limit in a given period. In fact, the result in Proposition 5 holds independently of the exact structure of $\underline{V}(\cdot)$. However, the optimal self-enforcing deficit

limit requires that the worst punishment be used, as such a punishment maximally relaxes the constraints of the problem and thus maximizes social welfare.

Given initial debt $b_{-1} = b$, the worst punishment solves an analogous recursive problem to that in (7), with the only difference that social welfare is now minimized rather than maximized. The solution to this problem yields that the date-0 policies $\{b(\theta), V(b(\theta))\}_{\theta \in \Theta}$ corresponding to maximal punishment are characterized as follows:

Proposition 6 (characterization of punishment)**.** *If $\{b(\theta), V(b(\theta))\}_{\theta \in \Theta}$ minimizes social welfare in the infinite horizon economy with $b(\theta) \in (\underline{b}(b_{-1}), \overline{b}(b_{-1}))$ for all $\theta \in \Theta$, then there exist finite $\theta_n^* > \underline{\theta}$ and $\theta_n^{**} \in [\underline{\theta}, \min\{\theta_n^*, \overline{\theta}\})$ such that*

$$\{b(\theta), V(b(\theta))\} = \begin{cases} \{b^r(\theta), \overline{V}(b^r(\theta))\} & \text{if } \theta > \theta_n^*, \\ \{b^r(\theta_n^*), \overline{V}(b^r(\theta_n^*))\} & \text{if } \theta \in [\theta_n^{**}, \theta_n^*], \\ \{b^p(\theta), \underline{V}(b^p(\theta))\} & \text{if } \theta < \theta_n^{**}, \end{cases}$$

where

$$\theta_n^{**} U(\omega + b^r(\theta_n^*)) + \beta\delta\overline{V}(b^r(\theta_n^*)) = \theta_n^{**} U(\omega + b^p(\theta_n^{**})) + \beta\delta\underline{V}(b^p(\theta_n^{**})). \tag{26}$$

Hence, any interior solution for the worst punishment is a maximally enforced surplus limit. Moreover, θ_n^* and θ_n^{**} are unique and independent of b_{-1}.

In the absence of enforcement constraints, the worst punishment would entail forcing all government types in all future periods to choose the highest or lowest borrowing level that is feasible, so as to minimize the value of social welfare. However, such a harsh punishment would not be self-enforcing. Proposition 6 shows that the worst punishment that is self-enforcing takes the form of a maximally enforced surplus limit, associated with a minimum debt level $b^r(\theta_n^*)$.[30] Government types that respect the surplus limit by choosing debt weakly above $b^r(\theta_n^*)$ are maximally rewarded; government types that violate the surplus limit by choosing debt strictly below $b^r(\theta_n^*)$ are maximally punished. Because a positive mass of types $\theta \geq \theta_n^{**}$ respect the limit, the equilibrium transitions back to the optimal deficit limit yielding maximal reward $\overline{V}(\cdot)$ with strictly positive probability. Figure 4 provides an illustration.

A maximally enforced surplus limit minimizes social welfare by incentivizing overborrowing. Intuitively, given a government type θ, there are two ways in which society can reduce welfare: either by inducing too little borrowing or by inducing too much borrowing. Since the government is biased towards overborrowing in the present, the latter relaxes self-enforcement constraints, and it is thus a more efficient means of reducing welfare. As

---

[30]This surplus limit can be shown to be incentive compatible by analogous logic as in Lemma 1.

(g) ....

ems:

)]   elfare given $\overline{V}(b)$ –

d private info θ_t Ifare given $\overline{V}(b)$ −

)$\}_{\theta^t \in \Theta^t}\}_{t=0}^{\infty}$ nd lower min

tives

sufficiently binding

$V(b(\theta))$

extreme

cal rectitude and fiscal profligacy sustain each
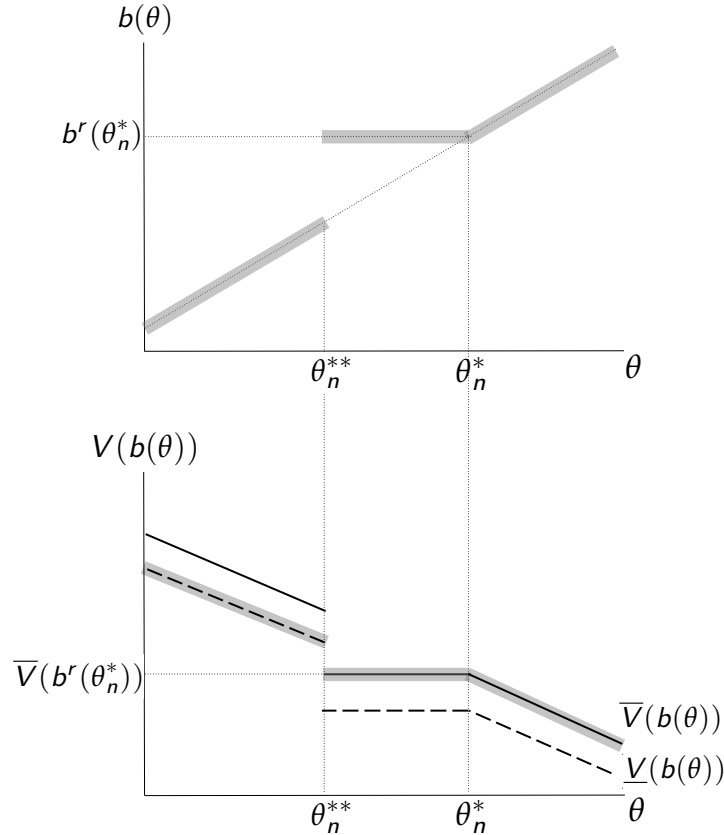
■ Phases of fiscal rectitud

Figure 4: An example of a maximally enforced surplus limit. The thick grey line depicts the borrowing allocation in the top panel and the allocation of continuation values in the bottom panel. The solid and dashed black lines in the bottom panel depict $\overline{V}(b(\theta))$ and $\underline{V}(b(\theta))$ respectively.

a result, in the worst-punishment allocation, all types borrow above the socially optimal level, and in fact weakly above the level preferred by the government.[31] Moreover, consistent with Proposition 1, overborrowing is incentivized by the use of bang-bang continuation values, with the maximal reward for types that respect the surplus limit and the maximal punishment for those that violate it. High-powered incentives allow society to maximize distortions.

The characterization in Proposition 6 follows from analogous arguments to those used to obtain Proposition 2. As noted, the key difference is that instead of maximizing the social welfare function in (12), the worst punishment minimizes this function. As a result, recalling our definition of $\widehat{\theta}$ in Assumption 1, our claims are now reversed: the arguments that apply to types $\theta < \widehat{\theta}$ when maximizing welfare—which make use of the fact that $Q(\theta)$

---

[31] These features are related to the results of Bernheim, Ray, and Yeltekin (2015), who study self-enforcing consumption rules in an environment without private information. The worst punishment in their model takes the form of over-consumption.

in (12) is strictly decreasing in this region—instead apply to types $\theta > \widehat{\theta}$ when minimizing welfare—since $-Q(\theta)$ is strictly decreasing in this region. The same is true with regards to the arguments that apply to types $\theta > \widehat{\theta}$ in the maximization of welfare, which apply to types $\theta < \widehat{\theta}$ in the minimization of welfare. This explains why the worst punishment is a maximally enforced surplus limit while the optimal rule is a maximally enforced deficit limit.

One step in the proof of Proposition 6 that requires additional care is establishing that the optimal surplus limit is indeed respected by a positive mass of government types. That is, we prove that the worst punishment is not an absorbing state in which $\underline{V}(\cdot)$ is sustained by all governments choosing $b^p(\theta)$ for all $\theta \in \Theta$. The proof consists of showing that a surplus limit that is respected by types $\theta \in [\theta_n^{**}, \overline{\theta}]$, for $\theta_n^{**} < \overline{\theta}$, achieves lower social welfare than such an absorbing state in which all types choose their optimum flexibly given a continuation value $\underline{V}(\cdot)$. That is, we show that the social cost of increasing overborrowing outweighs the benefit of increasing the continuation value for high enough types.

The result that any interior solution for the worst punishment is a maximally enforced surplus limit does not rely on the preference structure in (20), and it requires only the assumptions introduced in our static setting to obtain Proposition 2. By assuming that preferences do satisfy (20), we obtain that the thresholds θ_n^* and θ_n^{**} that define an optimal surplus limit are independent of the initial level of debt b_{-1} and thus associated with a maximal savings rate s_n^*. Moreover, when such preferences feature $\gamma \geq 1$, the borrowing allocation is always interior, implying that the worst punishment is always a maximally enforced surplus limit as described in Proposition 6.

4.5 Dynamics

Proposition 5 and Proposition 6 above imply specific dynamics for the economy under the optimal self-enforcing fiscal rule. To begin with, the bang-bang property tells us that, along the equilibrium path, continuation values travel only to extreme points in the feasible set $[\underline{V}(b), \overline{V}(b)]$. The maximal reward and the maximal punishment therefore sustain each other, being jointly determined in equilibrium. These features remind one of the seminal work of Abreu, Pearce, and Stacchetti (1990), who establish the optimality of bang-bang continuation values in a class of repeated games (see Theorem 7, p.1055 in their paper). However, their analysis restricts attention to settings with finite actions and a rich continuous public signal, whereas our model features a continuous action. Thus, while the implications are related, their results do not apply to our environment.

Given the bang-bang property, the dynamics induced by the optimal fiscal rule are

determined by primitives. Specifically, by our results in Proposition 3, the dynamics of the economy depend on whether condition (16) holds or not, which in turn depends on the government's deficit bias and the distribution of shocks. If condition (16) holds (which includes the case where the perfect-enforcement limit is self-enforcing), then the equilibrium implements the same deficit limit in every period, associated with the flexible debt level of a type θ^* under maximal reward $\overline{V}(\cdot)$. In contrast, if condition (16) does not hold, then the economy transitions in and out of the best equilibrium with value $\overline{V}(\cdot)$ and the worst punishment with value $\underline{V}(\cdot)$.

The transitions are stochastic as they depend on the realization of shocks. In particular, starting from the initial period, the government is subject to a deficit limit associated with the flexible debt level of a type θ^*. If the realized shock is $\theta \leq \theta^{**}$ (with θ^{**} given by (10)), the government respects the deficit limit and the best equilibrium restarts in the second period; if the realized shock is $\theta > \theta^{**}$, the government violates the deficit limit and the equilibrium transitions to punishment in the second period. Starting from a punishment period, the government is subject to a surplus limit associated with the flexible debt level of a type θ_n^*. If the realized shock is $\theta \geq \theta_n^{**}$ (with θ_n^{**} given by (26)), the government respects the surplus limit and the equilibrium transitions to the best equilibrium in the next period; if the realized shock is $\theta < \theta_n^{**}$, the government violates the surplus limit and the equilibrium remains in punishment in the next period. Consequently, in resemblance with real-world experiences such as that of Chile mentioned in the Introduction, we find that the economy fluctuates between periods of fiscal responsibility and periods of fiscal irresponsibility, with transitions being triggered by shocks given the optimal self-enforcing fiscal rule.

5 Concluding Remarks

We have studied the optimal design of fiscal rules when enforcement is limited. Under perfect enforcement, the optimal rule is a deficit limit which is never breached in equilibrium. Under limited enforcement, the optimal rule is a maximally enforced deficit limit, which, if violated, leads to the maximal penalty for the government. We established necessary and sufficient conditions under which the optimal deficit limit is violated on path, following bad enough shocks to the economy. We also showed that if penalties are not externally enforced but endogenously imposed by future governments, then punishment takes the form of overspending, and periods of reward and punishment self-enforce each other.

We believe there are potentially interesting directions for future research. For example, one could explore the generality of our result that optimal incentives are bang-bang. As

we discussed, this result does not rely on the presence enforcement constraints and may apply to other models of delegation and repeated adverse selection. It would be useful to understand whether this result is robust to different specifications of continuation values: while joint penalties (as we have assumed) are natural given our focus on fiscal policy, and are also consistent with the formalization of money burning in delegation models, one could consider settings in which penalties are experienced asymmetrically by a principal and an agent. Another possible direction for future work would be to explore environments in which a group of countries or subnational regions are subject to a coordinated fiscal rule. The properties of an optimal common rule would depend on governments' enforcement constraints and on the nature of feasible collective punishments.

Finally, while we have focused on fiscal policy, the insights of this paper may be applied to other settings featuring a commitment-versus-flexibility tradeoff and limited enforcement. For example, consider an individual who suffers from a self-control problem and establishes rules for herself to curb her consumption of a temptation good such as television or alcohol. Such an individual values discipline as well as the flexibility to increase her consumption when highly valuable. Furthermore, any rule the individual imposes on herself must be enforced by penalties that either originate from a third party or are sustained by her own future behavior. We find that the optimal such rule is a consumption threshold, which the individual may violate when her value of the temptation good is high enough. If penalties are endogenous, then they entail over-consumption, and the individual may transition in and out of periods of self-enforcing binging.

References

ABREU, D., D. PEARCE, AND E. STACCHETTI (1990): "Toward a Theory of Discounted Repeated Games with Imperfect Monitoring," *Econometrica*, 58, 1041–63.

ACEMOGLU, D., M. GOLOSOV, AND A. TSYVINSKI (2008): "Political Economy of Mechanisms," *Econometrica*, 76, 619–41.

AGUIAR, M. AND M. AMADOR (2011): "Growth under the Shadow of Expropriation," *Quarterly Journal of Economics*, 126, 651–97.

ALESINA, A. AND G. TABELLINI (1990): "A Positive Theory of Fiscal Deficits and Government Debt," *Review of Economic Studies*, 57, 403–14.

ALFARO, L. AND F. KANCZUK (2016): "Fiscal Rules and Sovereign Default," Working Paper.

ALONSO, R. AND N. MATOUSCHEK (2008): "Optimal Delegation," *Review of Economic Studies*, 75, 259–93.

AMADOR, M. AND K. BAGWELL (2013): "The Theory of Delegation with an Application to Tariff Caps," *Econometrica*, 81, 1541–600.

——— (2016): "Regulating a Monopolist With Uncertain Costs Without Transfers," Working Paper.

AMADOR, M., I. WERNING, AND G.-M. ANGELETOS (2003): "Commitment Vs. Flexibility," NBER Working Paper.

——— (2006): "Commitment Vs. Flexibility," *Econometrica*, 74, 365–396.

AMBRUS, A. AND G. EGOROV (2013): "Comment on 'Commitment vs. Flexibility'," *Econometrica*, 81, 2113–24.

——— (2015): "Delegation and Nonmonetary Incentives," Working Paper.

ATHEY, S., A. ATKESON, AND P. J. KEHOE (2005): "The Optimal Degree of Discretion in Monetary Policy," *Econometrica*, 73, 1431–75.

ATHEY, S., K. BAGWELL, AND C. SANCHIRICO (2004): "Collusion and Price Rigidity," *Review of Economic Studies*, 71, 317–49.

ATKESON, A. AND R. LUCAS (1992): "On Efficient Distribution with Private Information," *Review of Economic Studies*, 59, 427–53.

AZZIMONTI, M. (2011): "Barriers to Investment in Polarized Societies," *American Economic Review*, 101, 2182–204.

AZZIMONTI, M., M. BATTAGLINI, AND S. COATE (2016): "The Costs and Benefits of Balanced Budget Rules: Lessons from a Political Economy Model of Fiscal Policy," *Journal of Public Economics*, 136, 45–61.

BASSETTO, M. AND T. SARGENT (2006): "Politics and Efficiency of Separating Capital and Ordinary Government Budgets," *The Quarterly Journal of Economics*, 121, 1167–1210.

BATTAGLINI, M. AND S. COATE (2008): "A Dynamic Theory of Public Spending, Taxation, and Debt," *American Economic Review*, 98, 201–36.

BERNHEIM, B. D., D. RAY, AND S. YELTEKIN (2015): "Poverty and Self-Control," *Econometrica*, 83, 1877–911.

BISIN, A., A. LIZZERI, AND L. YARIV (2015): "Government Policy with Time Inconsistent Voters," *American Economic Review*, 105, 1711–37.

CAO, D. AND I. WERNING (2017): "Saving and Dissaving with Hyperbolic Discounting," Working Paper.

CASELLI, F., L. EYRAUD, A. HODGE, F. D. KALAN, Y. KIM, S. M. VICTOR LLEDÓ, A. POPESCU, W. H. REUTER, J. REYNAUD, E. TURE, AND P. WINGENDER (2018): "Second-Generation Fiscal Rules: Balancing Simplicity, Flexibility, and Enforceability," International Monetary Fund Staff Discussion Note 18-04.

CHADE, H., P. PROKOPOVYCH, AND L. SMITH (2008): "Repeated Games with Present-Biased Preferences," *Journal of Economic Theory*, 139, 157–175.

DOVIS, A. AND R. KIRPALANI (2017): "Fiscal Rules, Bailouts, and Reputation in Federal Governments," Working Paper.

FARHI, E. AND I. WERNING (2007): "Inequality and Social Discounting," *Journal of Political Economy*, 115, 365–402.

FUCHS, W. AND A. SKRZYPACZ (2015): "Government Interventions in a Dynamic Market with Adverse Selection," *Journal of Economic Theory*, 158, 371–406.

FUDENBERG, D. AND J. TIROLE (1991): *Game Theory*, Cambridge, Mass.: MIT Press.

HALAC, M. AND P. YARED (2014): "Fiscal Rules and Discretion under Persistent Shocks," *Econometrica*, 82, 1557–614.

——— (2017a): "Commitment vs. Flexibility with Costly Verification," Working Paper.

——— (2017b): "Fiscal Rules and Discretion under Self-Enforcement," Working Paper.

——— (2018): "Fiscal Rules and Discretion in a World Economy," *American Economic Review*, 108, 2305–2334.

HATCHONDO, J. C., L. MARTINEZ, AND F. ROCH (2017): "Fiscal Rules and the Sovereign Default Premium," Working Paper.

HOLMSTRÖM, B. (1977): "On Incentives and Control in Organizations," Ph.D. Thesis, Stanford University.

——— (1984): "On the Theory of Delegation," in *Bayesian Models in Economic Theory*, ed. by M. Boyer and R. Kihlstrom, North-Holland; New York.

JACKSON, M. O. AND L. YARIV (2014): "Present Bias and Collective Dynamic Choice in the Lab," *American Economic Review*, 104, 4184–204.

——— (2015): "Collective Dynamic Choice: The Necessity of Time Inconsistency," *American Economic Journal: Microeconomics*, 7, 150–78.

KRUSELL, P., B. KRUSCU, AND A. SMITH, JR. (2010): "Temptation and Taxation," *Econometrica*, 78, 2063–84.

KRUSELL, P. AND J.-V. RIOS-RULL (1999): "On the Size of U.S. Government: Political Economy in the Neoclassical Growth Model," *American Economic Review*, 89, 1156–81.

KRUSELL, P. AND A. SMITH, JR. (2003): "Consumption-Savings Decisions with Quasi-Geometric Discounting," *Econometrica*, 71, 365–75.

LAIBSON, D. (1994): "Self-Control and Saving," Department of Economics, Harvard University.

——— (1997): "Golden Eggs and Hyperbolic Discounting," *Quarterly Journal of Economics*, 112, 443–77.

LIZZERI, A. (1999): "Budget Deficits and Redistributive Politics," *American Economic Review*, 66, 909–28.

LIZZERI, A. AND L. YARIV (2014): "Collective Self-Control," Working Paper.

LLEDÓ, V., S. YOON, X. FANG, S. MBAYE, AND Y. KIM (2017): "Fiscal Rules at a Glance," *International Monetary Fund*, March.

MOSER, C. AND P. O. DE SOUZA E SILVA (2017): "Optimal Paternalistic Savings Policies," Working Paper.

PERSSON, T. AND L. SVENSSON (1989): "Why a Stubborn Conservative Would Run a Deficit: Policy with Time-Inconsistent Preferences," *Quarterly Journal of Economics*, 104, 325–45.

PHELAN, C. (1998): "On the Long Run Implications of Repeated Moral Hazard," *Journal of Economic Theory*, 79, 174–191.

PHELPS, E. AND R. POLLAK (1968): "On Second Best National Savings and Game-Equilibrium Growth," *Review of Economic Studies*, 35, 185–99.

PIGUILLEM, F. AND A. SCHNEIDER (2016): "Coordination, Efficiency and Policy Discretion," Working Paper.

RILEY, J. AND R. ZECKHAUSER (1983): "Optimal Selling Strategies: When to Haggle, When to Hold Firm," *Quarterly Journal of Economics*, 98, 267–89.

SCHAECHTER, A., T. KINDA, N. BUDINA, AND A. WEBER (2012): "Fiscal Rules in Response to the Crisis – Toward the 'Next-Generation' Rules. A New Dataset," IMF Working Paper WP/12/187.

SLEET, C. (2004): "Optimal Taxation with Private Government Information," *Review of Economic Studies*, 71, 1217–39.

SLEET, C. AND S. YELTEKIN (2006): "Credibility and Endogenous Societal Discounting," *Review of Economic Dynamics*, 9, 410–37.

——— (2008): "Politically Credible Social Insurance," *Journal of Monetary Economics*, 55, 129–51.

SONG, Z., K. STORESLETTEN, AND F. ZILIBOTTI (2012): "Rotten Parents and Disciplined Children: A Politico-Economic Theory of Public Expenditure and Debt," *Econometrica*, 80, 2785–803.

THOMAS, J. AND T. WORRALL (1990): "Income Fluctuation and Asymmetric Information: An Example of a Repeated Principal-Agent Problem," *Journal of Economic Theory*, 51, 367–90.

YARED, P. (2010): "Politicians, Taxes and Debt," *Review of Economic Studies*, 77, 806–40.

——— (2018): "Rising Government Debt: Causes and Solutions for a Decades-Old Trend," NBER Working Paper No. 24979.

A   Appendix: Proofs

This Appendix contains the proofs of Lemma 1, Proposition 1, Lemma 3, and Proposition 2. See the Online Appendix for the remaining proofs.

A.1 Preliminaries

We describe two functions that we will use in our proofs.

Lemma 5. *Given $\underline{\theta} \leq \theta^L \leq \theta^M \leq \theta^H \leq \overline{\theta}$, define the functions*

$$S^L(\theta^L, \theta^M) = \int_{\theta^L}^{\theta^M} \left(\beta\theta - \theta^L\right) f(\theta) \, d\theta + (1-\beta) \theta^M f(\theta^M)(\theta^M - \theta^L),$$

$$S^H(\theta^H, \theta^M) = \int_{\theta^M}^{\theta^H} \left(\beta\theta - \theta^H\right) f(\theta) d\theta + (1-\beta) \theta^M f(\theta^M)(\theta^H - \theta^M).$$

Then $S^L(\theta^L, \theta^M) > 0$ if $Q'(\theta) < 0$ for all $\theta \in (\theta^L, \theta^M)$, $S^L(\theta^L, \theta^M) < 0$ if $Q'(\theta) > 0$ for all $\theta \in (\theta^L, \theta^M)$, $S^H(\theta^H, \theta^M) > 0$ if $Q'(\theta) > 0$ for all $\theta \in (\theta^M, \theta^H)$, and $S^H(\theta^H, \theta^M) < 0$ if $Q'(\theta) < 0$ for all $\theta \in (\theta^M, \theta^H)$.

Proof. Consider the claims about $S^L(\theta^L, \theta^M)$. Note that $S^L(\theta, \theta^M)|_{\theta=\theta^M} = 0$, and hence $S^L(\theta^L, \theta^M) = -\int_{\theta_L}^{\theta^M} \frac{dS^L(\theta,\theta^M)}{d\theta} d\theta$. Moreover,

$$\frac{dS^L(\theta, \theta^M)}{d\theta} = -\int_{\theta}^{\theta^M} f(\widetilde{\theta}) d\widetilde{\theta} + (1-\beta) \theta f(\theta) - (1-\beta) \theta^M f(\theta^M),$$

and thus $\frac{dS^L(\theta,\theta^M)}{d\theta}|_{\theta=\theta^M} = 0$. Therefore, $S^L(\theta^L, \theta^M) = \int_{\theta_L}^{\theta^M} \int_{\theta}^{\theta^M} \frac{d^2S^L(\widetilde{\theta},\theta^M)}{d\widetilde{\theta}^2} d\widetilde{\theta} d\theta$, where

$$\frac{d^2 S^L(\theta, \theta^M)}{d\theta^2} = (2 - \beta) f(\theta) + (1-\beta) \theta f'(\theta).$$

Note that $\frac{d^2S^L(\theta,\theta^M)}{d\theta^2} > 0$ if $Q'(\theta) < 0$, $\frac{d^2S^L(\theta,\theta^M)}{d\theta^2} = 0$ if $Q'(\theta) = 0$, and $\frac{d^2S^L(\theta,\theta^M)}{d\theta^2} < 0$ if $Q'(\theta) > 0$. The claims about $S^L(\theta^L, \theta^M)$ follow.

The proof for the claims about $S^H(\theta^H, \theta^M)$ is analogous and thus omitted. $\qquad\square$

A.2 Proof of Lemma 1

We proceed in three steps.

Step 1. Suppose $\theta^* \geq \underline{\theta}$. We show that constraints (4) and (5) are satisfied for types $\theta \in [\underline{\theta}, \theta^*]$.

The claim follows immediately from the fact that all types $\theta \in [\underline{\theta}, \theta^*]$ are assigned their flexible debt levels with maximum continuation value. Thus, given $\theta \in [\underline{\theta}, \theta^*]$, type θ's welfare cannot be increased, and (4) and (5) are trivially satisfied.

Step 2. We show that constraints (4) and (5) are satisfied for types $\theta \in (\theta^*, \theta^{**}]$.

Take first the enforcement constraint (5). We can rewrite this constraint for $\theta \in (\theta^*, \theta^{**}]$ as

$$\theta U(\omega + b^r(\theta^*)) + \beta\delta\overline{V}(b^r(\theta^*)) - \theta U(\omega + b^p(\theta)) - \beta\delta\underline{V}(b^p(\theta)) \geq 0. \tag{27}$$

Differentiating the left-hand side with respect to θ, given θ^* and the definition of $b^p(\theta)$, yields

$$U(\omega + b^r(\theta^*)) - U(\omega + b^p(\theta)), \tag{28}$$

which is weakly decreasing in θ, since $b^p(\theta)$ is nondecreasing. This means that the left-hand side of (27) is weakly concave. Since (27) holds as a strict inequality for $\theta = \theta^*$ and as an equality for $\theta = \theta^{**}$ (by (10)), this weak concavity implies that (27) holds as a strict inequality for all $\theta \in (\theta^*, \theta^{**})$. Thus, the enforcement constraint is satisfied for all $\theta \in (\theta^*, \theta^{**}]$.

Consider next the private information constraint (4). This constraint is trivially satisfied for all $\theta \in (\theta^*, \theta^{**}]$ given $\theta' \in [\theta^*, \theta^{**}]$, since all types $\theta \in [\theta^*, \theta^{**}]$ are assigned the same allocation. We next show that the constraint is also satisfied given $\theta' > \theta^{**}$ and $\theta' < \theta^*$:

Step 2a: We show that (4) is satisfied for all $\theta \in (\theta^*, \theta^{**}]$ given $\theta' > \theta^{**}$. Note that $\{b(\theta'), V(b(\theta'))\} = \{b^p(\theta'), \underline{V}(b^p(\theta'))\}$ for all $\theta' > \theta^{**}$, and by the definition of $b^p(\theta)$,

$$\theta U(\omega + b^p(\theta)) + \beta\delta\underline{V}(b^p(\theta)) \geq \theta U(\omega + b^p(\theta')) + \beta\delta\underline{V}(b^p(\theta'))$$

for all $\theta' \in \Theta$. Thus, the fact that the enforcement constraint (5) is satisfied for all $\theta \in (\theta^*, \theta^{**}]$ implies that (4) is satisfied for all such types given $\theta' > \theta^{**}$.

Step 2b: We show that (4) is satisfied for all $\theta \in (\theta^*, \theta^{**}]$ given $\theta' < \theta^*$. Suppose by contradiction that this is not the case, that is,

$$\theta(U(\omega + b^r(\theta^*)) - U(\omega + b^r(\theta'))) < \beta\delta\left(\overline{V}(b^r(\theta')) - \overline{V}(b^r(\theta^*))\right) \tag{29}$$

for some $\theta \in (\theta^*, \theta^{**}]$ and $\theta' < \theta^*$. Note that by Step 1, (4) holds as a strict inequality for type θ^* given $\theta' < \theta^*$:

$$\theta^*(U(\omega + b^r(\theta^*)) - U(\omega + b^r(\theta'))) > \beta\delta\left(\overline{V}(b^r(\theta')) - \overline{V}(b^r(\theta^*))\right). \tag{30}$$

Combining (29) and (30) yields

$$(\theta^* - \theta)(U(\omega + b^r(\theta^*)) - U(\omega + b^r(\theta'))) > 0,$$

which is a contradiction since $\theta > \theta^*$ and $b^r(\theta') \leq b^r(\theta^*)$. Therefore, (4) is satisfied for all types $\theta \in (\theta^*, \theta^{**}]$ given $\theta' < \theta^*$.

Step 3. Suppose $\theta^{**} < \overline{\theta}$. We show that constraints (4) and (5) are satisfied for types $\theta \in (\theta^{**}, \overline{\theta}]$.

Constraint (5) holds as an equality for all $\theta \in (\theta^{**}, \overline{\theta}]$ and is thus satisfied for all these types. It is also immediate that constraint (4) is satisfied for all $\theta \in (\theta^{**}, \overline{\theta}]$ given $\theta' \in (\theta^{**}, \overline{\theta}]$, since all such types are assigned their flexible debt level with minimum continuation value. Consider next constraint (4) for $\theta \in (\theta^{**}, \overline{\theta}]$ given $\theta' \in [\theta^*, \theta^{**}]$. Note that $\{b(\theta'), V(b(\theta')\} = \{b^r(\theta^*), \overline{V}(b^r(\theta^*))\}$ for all $\theta' \in [\theta^*, \theta^{**}]$. Thus, satisfaction of this constraint is ensured if (27) is violated for $\theta \in (\theta^{**}, \overline{\theta}]$. The latter follows from the fact that, as shown above, the left-hand side of (27) is weakly concave and (27) holds as an equality for $\theta = \theta^{**}$ and a strict inequality for $\theta \in (\theta^*, \theta^{**})$.

Finally, consider constraint (4) for $\theta \in (\theta^{**}, \overline{\theta}]$ given $\theta' < \theta^*$. Since (4) is satisfied given $\theta' \in [\theta^*, \theta^{**}]$, satisfaction of this constraint given $\theta' < \theta^*$ is ensured if

$$\theta(U(\omega + b^r(\theta^*)) - U(\omega + b(\theta'))) \geq \beta\delta\overline{V}(b^r(\theta')) - \overline{V}(b^r(\theta^*))$$

for $\theta \in (\theta^{**}, \overline{\theta}]$. The latter follows from the same logic as in Step 2b above.

A.3 Proof of Proposition 1

Take any solution to program (7) with $b(\theta) \in (\underline{b}, \overline{b})$ for all $\theta \in \Theta$. We proceed in three steps.

Step 1. We show that $V(b(\theta))$ is left-continuous at each $\theta \in (\underline{\theta}, \overline{\theta}]$ and $V(b(\underline{\theta})) = \overline{V}(b(\underline{\theta}))$.

Consider the first claim. Suppose by contradiction that there exists $\theta \in (\underline{\theta}, \overline{\theta}]$ at which $V(b(\theta))$ is not left-continuous. Denote the left limit by $\{b(\theta^-), V(b(\theta^-))\} = \lim_{\theta' \uparrow \theta}\{b(\theta'), V(b(\theta'))\}$. By Lemma 2,

$$0 < \theta\left(U(\omega + b(\theta)) - U(\omega + b(\theta^-))\right) = \beta\delta\left(V(b(\theta^-)) - V(b(\theta))\right).$$

Given $\beta \in (0, 1)$, this implies

$$\theta\left(U(\omega + b(\theta)) - U(\omega + b(\theta^-))\right) < \delta\left(V(b(\theta^-)) - V(b(\theta))\right).$$

It follows that a perturbation that assigns $\{b(\theta^-), V(b(\theta^-))\}$ to type θ is incentive feasible,

strictly increases social welfare from type θ, and does not affect social welfare from types other than θ. Hence, $V(b(\theta))$ must be left-continuous at each $\theta \in (\underline{\theta}, \overline{\theta}]$.

Consider next the second claim. Suppose by contradiction that $V(b(\underline{\theta})) < \overline{V}(b(\underline{\theta}))$. Then we can perform a perturbation where we change $b(\underline{\theta}) \in (\underline{b}, \overline{b})$ by $db(\underline{\theta}) < 0$ arbitrarily small and increase $V(b(\underline{\theta}))$ so as to keep type $\underline{\theta}$ equally well off:

$$-\underline{\theta}U'(\omega + b(\underline{\theta})) + \beta\delta\frac{dV(b(\underline{\theta}))}{db(\underline{\theta})} = 0.$$

This perturbation is incentive feasible and does not affect social welfare from types $\theta \in (\underline{\theta}, \overline{\theta}]$. The change in social welfare from type $\underline{\theta}$ is equal to

$$-\underline{\theta}U'(\omega + b(\underline{\theta})) + \delta\frac{dV(b(\underline{\theta}))}{db(\underline{\theta})} = -\underline{\theta}U'(\omega + b(\underline{\theta}))\left(1 - \frac{1}{\beta}\right) > 0.$$

Hence, the perturbation strictly increases social welfare, yielding a contradiction.

Step 2. We show that $V(b(\theta))$ is a step function over any interval $[\theta^L, \theta^H]$ with $V(b(\theta)) \in (\underline{V}(b(\theta)), \overline{V}(b(\theta)))$ for $\theta \in [\theta^L, \theta^H]$.

By the private information constraints, $V(b(\theta))$ is piecewise continuously differentiable and nonincreasing. Suppose by contradiction that there is an interval $[\theta^L, \theta^H]$ over which $V(b(\theta))$ is continuously strictly decreasing in θ and satisfies $\underline{V}(b(\theta)) < V(b(\theta)) < \overline{V}(b(\theta))$. By Lemma 2, $b(\theta)$ must be continuously strictly increasing over the interval, and without loss we can take an interval over which $b(\theta)$ is continuously differentiable. Moreover, by the generic property in Proposition 1, we can take an interval with either $Q'(\theta) > 0$ or $Q'(\theta) < 0$ for all $\theta \in [\theta^L, \theta^H]$. We consider each possibility in turn.

<u>Case 1:</u> Suppose $Q'(\theta) < 0$ for all $\theta \in [\theta^L, \theta^H]$. We show that there exists an incentive feasible flattening perturbation that rotates the increasing borrowing schedule $b(\theta)$ clockwise over $[\theta^L, \theta^H]$ and strictly increases social welfare. Define

$$\overline{U} = \frac{1}{(\theta^H - \theta^L)} \int_{\theta^L}^{\theta^H} U(\omega + b(\theta))d\theta.$$

For given $\kappa \in [0, 1]$, let $\widetilde{b}(\theta, \kappa)$ be the solution to

$$U(\omega + \widetilde{b}(\theta, \kappa)) = \kappa\overline{U} + (1 - \kappa)U(\omega + b(\theta)), \tag{31}$$

which clearly exists. Define $\widetilde{V}(\widetilde{b}(\theta), \kappa)$ as the solution to

$$\theta U(\omega + \widetilde{b}(\theta, \kappa)) + \beta\delta\widetilde{V}(\widetilde{b}(\theta, \kappa), \kappa)$$
$$= \theta^L U(\omega + b(\theta^L)) + \beta\delta V(b(\theta^L)) + \int_{\theta^L}^{\theta} U(\omega + \widetilde{b}(\widetilde{\theta}, \kappa))d\widetilde{\theta}. \qquad (32)$$

The original allocation corresponds to $\kappa = 0$. We consider a perturbation where we increase κ marginally above zero if and only if $\theta \in [\theta^L, \theta^H]$. Note that differentiating (31) and (32) with respect to κ yields

$$\frac{d\widetilde{b}(\theta, \kappa)}{d\kappa} = \frac{\overline{U} - U(\omega + b(\theta))}{U'(\omega + \widetilde{b}(\theta, \kappa))}, \qquad (33)$$

$$\frac{d\widetilde{b}(\theta, \kappa)}{d\kappa}\theta U'(\omega + \widetilde{b}(\theta, \kappa)) + \beta\delta\frac{d\widetilde{V}(\widetilde{b}(\theta, \kappa), \kappa)}{d\kappa} = \int_{\theta^L}^{\theta} \frac{d\widetilde{b}(\widetilde{\theta}, \kappa)}{d\kappa}U'(\omega + \widetilde{b}(\widetilde{\theta}, \kappa))d\widetilde{\theta}. \qquad (34)$$

Substituting (33) in (34) yields that for a type $\theta \in [\theta^L, \theta^H]$, the change in government welfare from a marginal increase in κ, starting from $\kappa = 0$, is equal to

$$D(\theta) \equiv \int_{\theta^L}^{\theta} \left(\overline{U} - U(\omega + b(\widetilde{\theta}))\right)d\widetilde{\theta}.$$

We begin by showing that the perturbation satisfies constraints (4)-(6). For the private information constraint (4), note that $D(\theta^L) = D(\theta^H) = 0$, so the perturbation leaves the government welfare of types θ^L and θ^H (and that of types $\theta < \theta^L$ and $\theta > \theta^H$) unchanged. Using Lemma 2 and the representation in (11), it then follows from equation (32) and the fact that $\widetilde{b}(\theta, \kappa)$ is nondecreasing that the perturbation satisfies constraint (4) for all $\theta \in \Theta$ and any $\kappa \in [0, 1]$.

To prove that the perturbation satisfies the enforcement constraint (5), we show that the government welfare of types $\theta \in [\theta^L, \theta^H]$ weakly rises when κ increases marginally. Since $D(\theta^L) = D(\theta^H) = 0$, it is sufficient to show that $D(\theta)$ is concave over (θ^L, θ^H) to prove that $D(\theta) \geq 0$ for all θ in this interval. Indeed, we verify:

$$D'(\theta) = \overline{U} - U(\omega + b(\theta)),$$
$$D''(\theta) = -U'(\omega + b(\theta))\frac{db(\theta)}{d\theta} < 0.$$

Lastly, observe that constraint (6) is satisfied for $\kappa > 0$ small enough. This follows from $\underline{V}(b(\theta))$ and $\overline{V}(b(\theta))$ being continuous and the fact that $V(b(\theta)) \in (\underline{V}(b(\theta)), \overline{V}(b(\theta)))$ for $\theta \in [\theta^L, \theta^H]$ in the original allocation.

We next show that the perturbation strictly increases social welfare. Using the representation in (12), the change in social welfare from an increase in κ is equal to

$$\frac{1}{\beta} \int_{\theta^L}^{\theta^H} \frac{d\widetilde{b}(\theta, \kappa)}{d\kappa} U'(\omega + \widetilde{b}(\theta, \kappa)) Q(\theta) d\theta.$$

Substituting with (33) and the expression for $Q(\theta)$ yields that at $\kappa = 0$, this is equal to

$$\frac{1}{\beta} \int_{\theta^L}^{\theta^H} \left(\overline{U} - U\left(\omega + b(\theta)\right)\right) \left(1 - F(\theta) - \theta f(\theta)(1 - \beta)\right) d\theta.$$

This is an integral over the product of two terms. The first term is strictly decreasing in θ since $b(\theta)$ is strictly increasing over $[\theta^L, \theta^H]$. The second term is also strictly decreasing in θ; this follows from $Q'(\theta) < 0$ for all $\theta \in [\theta^L, \theta^H]$. Therefore, these two terms are positively correlated with one another, and thus the change in social welfare is strictly greater than

$$\frac{1}{\beta} \int_{\theta^L}^{\theta^H} \left(\overline{U} - U\left(\omega + b(\theta)\right)\right) d\theta \int_{\theta^L}^{\theta^H} \left(1 - F(\theta) - (1 - \beta)\theta f(\theta)\right) d\theta,$$

which is equal to 0. It follows that the change in social welfare from the perturbation is strictly positive. Hence, if $V(b(\theta))$ is strictly interior and $Q'(\theta) < 0$ over a given interval, then $V(b(\theta))$ must be a step function over the interval.

<u>Case 2:</u> Suppose $Q'(\theta) > 0$ for all $\theta \in [\theta^L, \theta^H]$. Recall that $b(\theta)$ is continuously strictly increasing over $[\theta^L, \theta^H]$. We begin by showing that the enforcement constraint cannot bind for all $\theta \in [\theta^L, \theta^H]$. Suppose by contradiction that it does. Using the representation of government welfare in (11), this implies

$$\int_{\theta}^{\theta^H} (U(\omega + b^p(\widetilde{\theta})) - U(\omega + b(\widetilde{\theta})))d\widetilde{\theta} = 0$$

for all $\theta \in [\theta^L, \theta^H]$, which requires $\{b(\theta), V(b(\theta))\} = \{b^p(\theta), \underline{V}(b^p(\theta))\}$ for all $\theta \in (\theta^L, \theta^H)$. However, this contradicts the assumption that $V(b(\theta)) \in (\underline{V}(b(\theta)), \overline{V}(b(\theta)))$ for all $\theta \in [\theta^L, \theta^H]$. Hence, the enforcement constraint cannot bind for all types in the interval, and without loss we can take an interval with this constraint being slack for all $\theta \in [\theta^L, \theta^H]$.

We next show that there exists a steepening perturbation that is incentive feasible and strictly increases social welfare. Specifically, consider drilling a hole around a type θ^M within $[\theta^L, \theta^H]$ so that we marginally remove the allocation around this type. That is, type θ^M can no longer choose $\{b(\theta^M), V(b(\theta^M))\}$ and is indifferent between jumping to the lower or upper limit of the hole. With some abuse of notation, denote the limits of the hole by θ^L

and θ^H, where the perturbation marginally increases θ^H from θ^M. Since the enforcement constraint is slack for all $\theta \in [\theta^L, \theta^H]$, the perturbation is incentive feasible. The change in social welfare from the perturbation is equal to

$$\int_{\theta^M}^{\theta^H} \frac{db(\theta^H)}{d\theta^H} \left(\theta U'(\omega + b(\theta^H)) + \delta V'(b(\theta^H))\right) f(\theta)d\theta$$

$$+\frac{d\theta^M}{d\theta^H} \left(\theta^M U(\omega + b(\theta^L)) + \delta V(b(\theta^L)) - \theta^M U(\omega + b(\theta^H)) - \delta V(b(\theta^H))\right) f(\theta^M).$$

Note that by the private information constraint for type θ^H,

$$\frac{db(\theta^H)}{d\theta^H} \left(\theta^H U'(\omega + b(\theta^H)) + \beta\delta V'(b(\theta^H))\right) = 0, \tag{35}$$

and by indifference of type θ^M,

$$\theta^M U(\omega + b(\theta^L)) + \beta\delta V(b(\theta^L)) = \theta^M U(\omega + b(\theta^H)) + \beta\delta V(b(\theta^H)). \tag{36}$$

Substituting with these expressions, the change in social welfare is equal to

$$\frac{db(\theta^H)}{d\theta^H} U'(\omega + b(\theta^H)) \int_{\theta^M}^{\theta^H} \left(\theta - \frac{\theta^H}{\beta}\right) f(\theta)d\theta \tag{37}$$

$$+\frac{d\theta^M}{d\theta^H} \theta^M \left(U(\omega + b(\theta^L)) - U(\omega + b(\theta^H))\right) \left(1 - \frac{1}{\beta}\right) f(\theta^M).$$

Differentiating (36) with respect to θ^H and substituting with (35) yields

$$\frac{d\theta^M}{d\theta^H} = \frac{db(\theta^H)}{d\theta^H} U'(\omega + b(\theta^H)) \frac{(\theta^H - \theta^M)}{U(\omega + b(\theta^H)) - U(\omega + b(\theta^L))}.$$

Substituting back into (37) and dividing by $\frac{1}{\beta} \frac{db(\theta^H)}{d\theta^H} U'(\omega + b(\theta^H)) > 0$, we find that the change in social welfare takes the same sign as

$$S^H(\theta^H, \theta^M) = \int_{\theta^M}^{\theta^H} \left(\beta\theta - \theta^H\right) f(\theta)d\theta + (1 - \beta)\theta^M f(\theta^M)(\theta^H - \theta^M).$$

Since $Q'(\theta) > 0$ for all $\theta \in [\theta^M, \theta^H]$, Lemma 5 implies $S^H(\theta^H, \theta^M) > 0$, and thus the perturbation strictly increases social welfare. Hence, if $V(b(\theta))$ is strictly interior and $Q'(\theta) > 0$ over a given interval, then $V(b(\theta))$ must be a step function over the interval.

Step 3. We show that $V(b(\theta)) \in \{\underline{V}(b(\theta)), \overline{V}(b(\theta))\}$ for all $\theta \in \Theta$.

43

Suppose by contradiction that $V(b(\theta)) \in (\underline{V}(b(\theta)), \overline{V}(b(\theta)))$ for some $\theta \in \Theta$. By the previous steps and Lemma 2, type θ belongs to a stand-alone segment $(\theta^L, \theta^H]$, such that $b(\theta) = b$ and $V(b(\theta)) = V$ for all $\theta \in (\theta^L, \theta^H]$, $b \in (\underline{b}, \overline{b})$ and $V \in (\underline{V}(b), \overline{V}(b))$ (by assumption), $b(\theta)$ jumps at θ^L, and $b(\theta)$ jumps at θ^H unless $\theta^H = \overline{\theta}$.

We first show that the enforcement constraint must be slack for all $\theta \in (\theta^L, \theta^H)$. Express the enforcement constraint as the difference between the left-hand and right-hand sides of (5), so that this constraint must be weakly positive and it is equal to zero if it binds. By the private information constraints, the derivative of the enforcement constraint with respect to θ is equal to $U(\omega + b(\theta)) - U(\omega + b^p(\theta))$. Since $b(\theta)$ is constant over $(\theta^L, \theta^H]$ and $b^p(\theta)$ is nondecreasing, it follows that the enforcement constraint is weakly concave over the interval. Then the constraint can only bind at a point $\theta' \in (\theta^L, \theta^H)$ if it binds at all $\theta \in (\theta^L, \theta^H)$. However, by the arguments used in Case 2 in Step 2 above, that would require $b = b^p(\theta)$ and $V = \underline{V}(b)$ for $\theta \in (\theta^L, \theta^H)$, contradicting the assumption that V is strictly interior.

We next show that there exists an incentive feasible perturbation that strictly increases social welfare. We consider segment-shifting perturbations that marginally change the constant borrowing level b and continuation value V. As we describe next, the perturbation that we perform depends on the shape of the function $Q(\theta)$ over $(\theta^L, \theta^H]$:

<u>Case 1:</u> Suppose $\int_{\theta^L}^{\theta^H} Q(\theta^L) d\theta < \int_{\theta^L}^{\theta^H} Q(\theta) d\theta$. Consider a perturbation that marginally changes the borrowing level by $db > 0$ and reduces V in order to keep type θ^H equally well off. This means that $\frac{dV}{db}$ is given by

$$\theta^H U'(\omega + b) + \beta\delta\frac{dV}{db} = 0. \tag{38}$$

Note that for any arbitrarily small $db > 0$, this perturbation makes the lowest types in $(\theta^L, \theta^H]$, arbitrarily close to θ^L, jump either to the allocation of type θ^L or to their flexible allocation under maximal punishment $\{b^p(\theta), \underline{V}(b^p(\theta))\}$, where we let the perturbation introduce the latter. In the limit as db goes to zero, the change in social welfare due to the perturbation is thus equal to[32]

$$\int_{\theta^L}^{\theta^H} \left(\theta U'(\omega + b) + \delta\frac{dV}{db}\right) f(\theta) d\theta \tag{39}$$

$$+\frac{d\theta^L}{db} \left(\theta^L U(\omega + b(\theta^L)) + \delta V(b(\theta^L)) - \theta^L U(\omega + b) - \delta V\right) f(\theta^L),$$

---

[32]The arguments that follow are unchanged if $\{b(\theta^L), V(b(\theta^L))\}$ is replaced with $\{b^p(\theta^L), \underline{V}(b^p(\theta^L))\}$ for the cases where the enforcement constraint binds.

where the following indifference condition holds:

$$\theta^L U(\omega + b) + \beta\delta V = \theta^L U(\omega + b(\theta^L)) + \beta\delta V(b(\theta^L)).$$

To verify that the perturbation is incentive feasible, note that the enforcement constraint is slack for all $\theta \in (\theta^L, \theta^H)$, V is strictly interior, and the government welfare of types θ^L and θ^H remains unchanged with the perturbation. Hence, the perturbation is incentive feasible for db arbitrarily close to zero.

To verify that the perturbation strictly increases social welfare, substitute (38) and the indifference condition of type θ^L into (39) to obtain:

$$U'(\omega+b) \int_{\theta^L}^{\theta^H} \left(\theta - \frac{\theta^H}{\beta}\right) f(\theta)d\theta + \frac{d\theta^L}{db}\theta^L \left(U(\omega + b(\theta^L)) - U(\omega + b)\right)\left(1 - \frac{1}{\beta}\right) f(\theta^L). \quad (40)$$

Differentiating the indifference condition of type θ^L and substituting with (38) yields

$$\frac{d\theta^L}{db} = -U'(\omega + b)\frac{\left(\theta^H - \theta^L\right)}{U(\omega + b(\theta^L)) - U(\omega + b)}.$$

Substituting back into (40) and dividing by $\frac{1}{\beta}U'(\omega + b) > 0$, we find that the change in social welfare takes the same sign as

$$S^H(\theta^H, \theta^L) = \int_{\theta^L}^{\theta^H} \left(\beta\theta - \theta^H\right) f(\theta)d\theta + (1 - \beta)\theta^L f(\theta^L)(\theta^H - \theta^L),$$

which can be rewritten as

$$S^H(\theta^H, \theta^L) = \int_{\theta^L}^{\theta^H}\int_{\theta^L}^{\theta} Q'(\widetilde{\theta})d\widetilde{\theta}d\theta = \int_{\theta^L}^{\theta^H} (Q(\theta) - Q(\theta^L))d\theta.$$

By the assumption that $\int_{\theta^L}^{\theta^H} Q(\theta^L)d\theta < \int_{\theta^L}^{\theta^H} Q(\theta)d\theta$, the above expression is strictly positive. The perturbation therefore strictly increases social welfare, yielding a contradiction.

<u>Case 2:</u> Suppose $\int_{\theta^L}^{\theta^H} Q(\theta^L)d\theta \geq \int_{\theta^L}^{\theta^H} Q(\theta)d\theta$. By the generic property in Proposition 1, there must exist $\theta^h \in (\theta^L, \theta^H]$ such that $\int_{\theta^L}^{\theta^h} Q(\theta^L)d\theta > \int_{\theta^L}^{\theta^h} Q(\theta)d\theta$. Then consider a perturbation where, for $\theta \in (\theta^L, \theta^h]$, we marginally change the borrowing level by $db < 0$ and increase V in order to keep type θ^h equally well off. This perturbation makes types arbitrarily close to θ^L jump up to the allocation of the stand-alone segment. Arguments analogous to those in Case 1 above imply that the perturbation is incentive feasible. Moreover, following analogous steps as in that case yields that the implied change in social

welfare takes the same sign as

$$-S^H(\theta^h, \theta^L) = -\int_{\theta^L}^{\theta^h} \left(\beta\theta - \theta^h\right) f(\theta)d\theta - (1 - \beta)\,\theta^L f(\theta^L)(\theta^h - \theta^L),$$

which can be rewritten as

$$-S^H(\theta^h, \theta^L) = -\int_{\theta^L}^{\theta^h} \int_{\theta^L}^{\theta} Q'(\widetilde{\theta})d\widetilde{\theta}d\theta = -\int_{\theta^L}^{\theta^h} (Q(\theta) - Q(\theta^L))d\theta.$$

By the assumption that $\int_{\theta^L}^{\theta^h} Q(\theta^L)d\theta > \int_{\theta^L}^{\theta^h} Q(\theta)d\theta$, the above expression is strictly positive. The perturbation therefore strictly increases social welfare, yielding a contradiction.

A.4   Proof of Lemma 3

Take any solution to program (7) with $b(\theta) \in (\underline{b}, \overline{b})$ for all $\theta \in \Theta$. We proceed in three steps.

Step 1. We show that if $V(b(\theta^{**})) = \underline{V}(b(\theta^{**}))$, then $\theta^{**} \geq \widehat{\theta}$.

By Proposition 1 and Step 1 in the proof of that proposition, if $V(b(\theta^{**})) = \underline{V}(b(\theta^{**}))$ for some $\theta^{**} \in \Theta$, then $V(b(\theta)) = \underline{V}(b(\theta))$ over an interval $(\theta^L, \theta^H]$ that contains θ^{**}. Take the largest such interval. We establish that $\theta^L \geq \widehat{\theta}$. Suppose by contradiction that $\theta^L < \widehat{\theta}$. Note that the enforcement constraint (5) requires $b(\theta) = b^p(\theta)$ for all $\theta \in (\theta^L, \theta^H]$. There are two cases to consider:

<u>Case 1:</u> Suppose $b^p(\theta)$ is strictly increasing over a subset of $(\theta^L, \theta^H]$ below $\widehat{\theta}$, and without loss take a subset over which $b^p(\theta)$ is continuously differentiable. Then we can perform a flattening perturbation that rotates the borrowing schedule clockwise over this subset, analogous to the perturbation used in Step 2 in the proof of Proposition 1. By the arguments in that step, this perturbation is incentive feasible. In particular, note that since the perturbation weakly increases the government welfare of all types θ in the subset while simultaneously changing their borrowing allocation, it follows from the definition of $b^p(\theta)$ that the perturbation must necessarily increase $V(b(\theta))$ above $\underline{V}(b(\theta))$. Moreover, by $Q'(\theta) < 0$ for all types θ in the subset (by the subset being below $\widehat{\theta}$ and Assumption 1), the perturbation strictly increases social welfare, yielding a contradiction.

<u>Case 2:</u> Suppose $b^p(\theta)$ is constant for $\theta \in (\theta^L, \theta^M]$, where $\theta^M \equiv \min\{\theta^H, \widehat{\theta}\}$. Then we can perform an incentive feasible segment-shifting perturbation analogous to that described in Step 3 in the proof of Proposition 1: for $\theta \in (\theta^L, \theta^M]$, we marginally reduce the constant

borrowing level and increase the constant continuation value so as to keep the government welfare of type θ^M unchanged. Since $Q'(\theta) < 0$ over $(\theta^L, \theta^M]$ implies $\int_{\theta^L}^{\theta^M} Q(\theta^L)d\theta > \int_{\theta^L}^{\theta^M} Q(\theta)d\theta$, this perturbation strictly increases social welfare, yielding a contradiction.

Step 2. We show that if $V(b(\theta^{**})) = \underline{V}(b(\theta^{**}))$, then $V(b(\theta)) = \underline{V}(b(\theta))$ for all $\theta \geq \theta^{**}$.

Suppose by contradiction that $V(b(\theta^{**})) = \underline{V}(b(\theta^{**}))$ for $\theta^{**} \in \Theta$ and $V(b(\theta)) > \underline{V}(b(\theta))$ for some $\theta > \theta^{**}$. By Step 1, $\theta^{**} \geq \widehat{\theta}$. Moreover, by Proposition 1 and Step 1 in the proof of that proposition, there exist $\theta^H > \theta^L \geq \theta^{**}$ such that $V(b(\theta)) = \overline{V}(b(\theta))$ for all $\theta \in (\theta^L, \theta^H]$.

We begin by establishing that $b(\theta) = b$ for all $\theta \in (\theta^L, \theta^H]$ and some $b \in (\underline{b}, \overline{b})$. Suppose by contradiction that the borrowing schedule $b(\theta)$ is strictly increasing at some point $\theta' \in (\theta^L, \theta^H]$. Note that the private information constraint (4) implies $b(\theta) = b^r(\theta)$, and thus a slack enforcement constraint, in the neighborhood of such a type θ'. Then we can perform an incentive feasible steepening perturbation that drills a hole in the $b(\theta)$ schedule in this neighborhood, as that described in Step 2 (Case 2) in the proof of Proposition 1. By the arguments in that step, this perturbation strictly increases social welfare, yielding a contradiction.

We next show that a segment $(\theta^L, \theta^H]$ with $b(\theta) = b$ and $V(b(\theta)) = \overline{V}(b)$ for all $\theta \in (\theta^L, \theta^H]$ and $\theta^L \geq \theta^{**}$ cannot exist. Suppose by contradiction that it does. Take θ^L to be the lowest point weakly above θ^{**} at which $V(b(\theta))$ jumps, and take θ^H to be the lowest point above θ^L at which $V(b(\theta))$ jumps again, or $\theta^H = \overline{\theta}$ if $V(b(\theta))$ does not jump above θ^L. Then $(\theta^L, \theta^H]$ is a stand-alone segment with constant borrowing b and continuation value $\overline{V}(b)$. Note that by arguments analogous to those in Step 3 of the proof of Proposition 1, the enforcement constraint must be slack for all $\theta \in (\theta^L, \theta^H)$. We then show that there exists an incentive feasible segment-shifting perturbation that is socially beneficial. There are three cases to consider:

<u>Case 1:</u> Suppose $\theta^H U(\omega + b) + \beta\delta\overline{V}(b) \leq \theta^H U(\omega + b') + \beta\delta\overline{V}(b')$ for $b' = b + \varepsilon$, $\varepsilon > 0$ arbitrarily small. Then we perform a segment-shifting perturbation as that in Step 3 in the proof of Proposition 1, where we increase b marginally to b' and set $V(b')$ weakly below $\overline{V}(b')$ so as to keep type θ^H's welfare under this allocation unchanged. This perturbation is incentive feasible. Moreover, since $\theta^L \geq \theta^{**}$ and Assumption 1 imply $\int_{\theta^L}^{\theta^H} Q(\theta^L)d\theta < \int_{\theta^L}^{\theta^H} Q(\theta)d\theta$, this perturbation strictly increases social welfare, yielding a contradiction.

<u>Case 2:</u> Suppose $\theta^H U(\omega + b) + \beta\delta\overline{V}(b) > \theta^H U(\omega + b') + \beta\delta\overline{V}(b')$ for $b' = b + \varepsilon$, $\varepsilon > 0$ arbitrarily small, and $\theta^H < \overline{\theta}$. Then we perform a segment-shifting perturbation where we reduce b marginally to $b'' = b - \varepsilon$ and set $V(b'')$ weakly below $\overline{V}(b'')$ so as to keep type θ^L's welfare under this allocation unchanged. This perturbation is incentive feasible. Denote by

$\{b(\theta^h), V(b(\theta^h))\}$ the allocation above θ^H over which type θ^H is initially indifferent. Note that analogous to the perturbation in Step 3 in the proof of Proposition 1, this perturbation makes the highest types in $(\theta^L, \theta^H]$, arbitrarily close to θ^H, jump either to $\{b(\theta^h), V(b(\theta^h))\}$ or to their flexible allocation under maximal punishment $\{b^p(\theta), \underline{V}(b^p(\theta))\}$, where we let the perturbation introduce the latter. In the limit as ε goes to zero, the change in social welfare due to the perturbation is thus equal to[33]

$$\int_{\theta^L}^{\theta^H} \left(-\theta U'(\omega + b) + \delta \frac{dV}{db} \right) f(\theta) \, d\theta \tag{41}$$

$$-\frac{d\theta^H}{db} \left(\theta^H U(\omega + b(\theta^h)) + \delta V(b(\theta^h)) - \theta^H U(\omega + b) - \delta V \right) f(\theta^H),$$

where $V = \overline{V}(b)$, $\frac{dV}{db}$ solves

$$-\theta^L U'(\omega + b) + \beta \delta \frac{dV}{db} = 0, \tag{42}$$

and the following indifference condition holds:

$$\theta^H U(\omega + b) + \beta \delta V = \theta^H U(\omega + b(\theta^h)) + \beta \delta V(b(\theta^h)).$$

To verify that the perturbation strictly increases social welfare, substitute (42) and the indifference condition of type θ^H into (41) to obtain:

$$-U'(\omega + b) \int_{\theta^L}^{\theta^H} \left(\theta - \frac{\theta^L}{\beta} \right) f(\theta) d\theta - \frac{d\theta^H}{db} \theta^H \left(U(\omega + b(\theta^h)) - U(\omega + b) \right) \left(1 - \frac{1}{\beta} \right) f(\theta^H). \tag{43}$$

Differentiating the indifference condition of type θ^H and substituting with (42) yields

$$\frac{d\theta^H}{db} = -U'(\omega + b) \frac{\left(\theta^H - \theta^L \right)}{U(\omega + b(\theta^h)) - U(\omega + b)}.$$

Substituting back into (43) and dividing by $\frac{1}{\beta} U'(\omega + b) > 0$, we find that the change in social welfare takes the same sign as

$$-S^L(\theta^L, \theta^H) = - \left[\int_{\theta^L}^{\theta^H} \left(\beta \theta - \theta^L \right) f(\theta) d\theta + (1 - \beta) \theta^H f(\theta^H)(\theta^H - \theta^L) \right].$$

---

[33]The arguments that follow are unchanged if $\{b(\theta^h), V(b(\theta^h))\}$ is replaced with $\{b^p(\theta^H), \underline{V}(b^p(\theta^H))\}$ for the cases where the enforcement constraint binds.

By $\theta^L \geq \theta^{**}$, Assumption 1, and Step 1 above, $Q'(\theta) > 0$ for all $\theta \in (\theta^L, \theta^H)$. It then follows from Lemma 5 that $S^L(\theta^L, \theta^H) < 0$. Hence, the perturbation strictly increases social welfare, yielding a contradiction.

<u>Case 3:</u> Suppose $\theta^H U(\omega + b) + \beta\delta\overline{V}(b) > \theta^H U(\omega + b') + \beta\delta\overline{V}(b')$ for $b' = b + \varepsilon$, $\varepsilon > 0$ arbitrarily small, and $\theta^H = \overline{\theta}$. Then we perform a segment-shifting perturbation as that in Case 2 above, where we reduce b marginally to $b'' = b - \varepsilon$ and set $V(b'')$ weakly below $\overline{V}(b'')$ so as to keep type θ^L's welfare under this allocation unchanged. This perturbation is incentive feasible. Note that analogous to Case 2, this perturbation makes the highest types in $(\theta^L, \theta^H]$, arbitrarily close to θ^H, either jump to their flexible allocation under maximal punishment $\{b^p(\theta), \underline{V}(b^p(\theta))\}$ or remain with the perturbed allocation. In the former case, the same arguments as in Case 2 apply, yielding that the perturbation strictly increases social welfare by $-S^L(\theta^L, \theta^H) > 0$. In the latter case, those arguments imply that the change in social welfare is equal to

$$ -\int_{\theta^L}^{\theta^H} \left(\beta\theta - \theta^L \right) f(\theta) d\theta > -S^L(\theta^L, \theta^H) > 0. $$

Hence, the perturbation strictly increases social welfare, yielding a contradiction.

Step 3. We show that $V(b(\theta))$ is right-continuous at $\underline{\theta}$.

Suppose by contradiction that this is not the case. Then by the previous steps, Proposition 1, and Step 1 in the proof of Proposition 1, $V(b(\theta)) = \underline{V}(b(\theta))$ for all $\theta \in (\underline{\theta}, \overline{\theta}]$ and $V(b(\theta))$ jumps down at $\underline{\theta}$ from $\overline{V}(b(\underline{\theta}))$. Note that the enforcement constraint (5) implies $b(\theta) = b^p(\theta)$ for all $\theta \in (\underline{\theta}, \overline{\theta}]$, and indifference of $\underline{\theta}$ requires

$$ \underline{\theta}U(\omega + b(\underline{\theta})) + \beta\delta\overline{V}(b(\underline{\theta})) = \lim_{\theta\downarrow\underline{\theta}}\{\underline{\theta}U(\omega + b^p(\theta)) + \beta\delta\underline{V}(b^p(\theta))\}. $$

Take $\Delta \in (0, \min_{\theta\in\Theta}\{\overline{V}(b(\theta)) - \underline{V}(b(\theta))\})$. Then consider a global perturbation that assigns $V(b(\theta)) = \underline{V}(b(\theta)) + \Delta$ to all $\theta \in (\underline{\theta}, \overline{\theta}]$ and assigns type $\underline{\theta}$ the limit allocation to its right. This perturbation keeps the borrowing allocation of types $\theta \in (\underline{\theta}, \overline{\theta}]$ unchanged and is incentive feasible. Moreover, using the representation in (12), the change in social welfare from this perturbation is equal to $\delta\Delta$. Thus, the perturbation strictly increases social welfare, yielding a contradiction.

A.5 Proof of Proposition 2

Take any solution to program (7) with $b(\theta) \in (\underline{b}, \overline{b})$ for all $\theta \in \Theta$. By Lemma 3 and the enforcement constraint (5), there exists $\theta^{**} > \underline{\theta}$ such that $\{b(\theta), V(b(\theta))\} = \{b^p(\theta), \underline{V}(b^p(\theta))\}$ for all $\theta > \theta^{**}$ and $V(b(\theta)) = \overline{V}(b(\theta))$ for all $\theta \leq \theta^{**}$ (where it is possible that $\theta^{**} > \overline{\theta}$). Moreover, since the enforcement constraint holds with equality at θ^{**}, this type's allocation satisfies

$$\theta^{**} U(\omega + b(\theta^{**})) + \beta\delta\overline{V}(b(\theta^{**})) = \theta^{**} U(\omega + b^p(\theta^{**})) + \beta\delta\underline{V}(b^p(\theta^{**})). \tag{44}$$

These results characterize the allocation for types $\theta \geq \theta^{**}$. To characterize the allocation for types $\theta < \theta^{**}$, we proceed in three steps.

Step 1. We show that $b(\theta)$ is continuous over $[\underline{\theta}, \theta^{**}]$.

Recall from Step 1 in the proof of Lemma 3 that $\theta^{**} \geq \widehat{\theta}$. There are two cases to consider:

<u>Case 1:</u> Suppose by contradiction that $b(\theta)$ has a point of discontinuity below $\widehat{\theta}$: there is a type $\theta^M < \widehat{\theta}$ which is indifferent between choosing $\lim_{\theta\uparrow\theta^M} b(\theta)$ and $\lim_{\theta\downarrow\theta^M} b(\theta) > \lim_{\theta\uparrow\theta^M} b(\theta)$. Note that given $V(b(\theta)) = \overline{V}(b(\theta))$ for all $\theta \in [\underline{\theta}, \theta^{**}]$ and $\theta^{**} \geq \widehat{\theta}$, there must be a hole with types $\theta \in [\theta^L, \theta^M)$ bunched at $b^r(\theta^L)$ and types $\theta \in (\theta^M, \theta^H]$ bunched at $b^r(\theta^H)$, for some $\theta^L < \theta^M < \theta^H$. Now consider perturbing the rule by marginally increasing θ^L, in an effort to slightly close the hole. This perturbation leaves the government welfare of types strictly above θ^M unchanged and is incentive feasible. The change in social welfare from the perturbation is equal to

$$\frac{db^r(\theta^L)}{d\theta^L} \int_{\theta^L}^{\theta^M} \left(\theta U'(\omega + b^r(\theta^L)) + \delta\overline{V}'(b^r(\theta^L)) \right) f(\theta)\, d\theta$$

$$+ \frac{d\theta^M}{d\theta^L} \left[\theta^M U(\omega + b^r(\theta^L)) + \delta\overline{V}(b^r(\theta^L)) - \theta^M U(\omega + b^r(\theta^H)) - \delta\overline{V}(b^r(\theta^H)) \right] f(\theta^M).$$

By the definition of $b^r(\theta)$, we have $\theta^L U'(\omega + b^r(\theta^L)) = -\beta\delta\overline{V}'(b^r(\theta^L))$. Moreover, by indifference of type θ^M, we have

$$\theta^M U(\omega + b^r(\theta^L)) + \beta\delta\overline{V}(b^r(\theta^L)) = \theta^M U(\omega + b^r(\theta^H)) + \beta\delta\overline{V}(b^r(\theta^H)). \tag{45}$$

Substituting these into the expression above yields that the change in social welfare due

to the perturbation is equal to

$$\frac{db^r(\theta^L)}{d\theta^L}U'(\omega + b^r(\theta^L)) \int_{\theta^L}^{\theta^M} \left(\theta - \frac{\theta^L}{\beta}\right) f(\theta)\, d\theta \tag{46}$$

$$+\frac{d\theta^M}{d\theta^L}\left(\frac{1}{\beta} - 1\right)\theta^M f(\theta^M)\left(U(\omega + b^r(\theta^H)) - U(\omega + b^r(\theta^L))\right).$$

Note that differentiating the indifference condition (45) with respect to θ^L (and substituting again with $\theta^L U'(\omega + b^r(\theta^L)) = -\beta\delta\overline{V}'(b^r(\theta^L)))$ yields

$$\frac{d\theta^M}{d\theta^L} = \frac{db^r(\theta^L)}{d\theta^L}U'(\omega + b^r(\theta^L))\frac{(\theta^M - \theta^L)}{U(\omega + b^r(\theta^H)) - U(\omega + b^r(\theta^L))}.$$

Substituting this back into (46) and dividing by $\frac{1}{\beta}\frac{db^r(\theta^L)}{d\theta^L}U'(\omega + b^r(\theta^L)) > 0$, we find that the change in social welfare takes the same sign as

$$S^L(\theta^L, \theta^M) = \int_{\theta^L}^{\theta^M} \left(\beta\theta - \theta^L\right) f(\theta)\, d\theta + (1 - \beta)\,\theta^M f(\theta^M)(\theta^M - \theta^L).$$

If $\theta^L \geq \underline{\theta}$, it follows from $\theta^M < \widehat{\theta}$, Assumption 1, and Lemma 5 that $S^L(\theta^L, \theta^M) > 0$. Moreover, if $\theta^L < \underline{\theta}$, then $S^L(\theta^L, \theta^M) > S^L(\underline{\theta}, \theta^M) > 0$. Thus, the perturbation strictly increases social welfare, showing that $b(\theta)$ cannot jump at a point below $\widehat{\theta}$.

<u>Case 2:</u> Suppose by contradiction that $b(\theta)$ is discontinuous at a point $\theta \in [\widehat{\theta}, \theta^{**}]$. Note that since $V(b(\theta)) = \overline{V}(b(\theta))$ for all $\theta \in [\widehat{\theta}, \theta^{**}]$, we can apply the same logic as in Step 2 in the proof of Lemma 3 to show that $\frac{db(\theta)}{d\theta} = 0$ over any continuous interval in $[\widehat{\theta}, \theta^{**}]$. Hence, if $b(\theta)$ jumps at a point $\theta \in [\widehat{\theta}, \theta^{**}]$, then there exists a stand-alone segment $(\theta^L, \theta^H]$ with constant borrowing $b \in (\underline{b}, \overline{b})$ and continuation value $V = \overline{V}(b)$, satisfying $\theta^L \geq \widehat{\theta}$. However, using again the arguments in Step 2 in the proof of Lemma 3, we can then perform an incentive feasible segment-shifting perturbation that strictly increases social welfare. Thus, $b(\theta)$ cannot jump at a point $\theta \in [\widehat{\theta}, \theta^{**}]$.

Step 2. We show that $b(\theta) \leq b^r(\theta)$ for all $\theta \in [\underline{\theta}, \theta^{**}]$.

By Step 1 above, the allocation over $[\underline{\theta}, \theta^{**}]$ must be bounded discretion, with either a minimum borrowing level or a maximum borrowing level or both. We next show that a binding minimum borrowing requirement is strictly suboptimal. Suppose by contradiction that this is not the case, namely there exist $\theta^\star > \underline{\theta}$ and an optimal rule prescribing $\{b(\theta), V(b(\theta))\} = \{b^r(\theta^\star), \overline{V}(b^r(\theta^\star))\}$ for all $\theta \in [\underline{\theta}, \theta^\star]$, where $b^r(\theta) < b^r(\theta^\star)$ for all

$\theta \in [\underline{\theta}, \theta^\star)$. Consider a perturbation where we remove this minimum borrowing requirement, that is, we set $\{b(\theta), V(b(\theta))\} = \{b^r(\theta), \overline{V}(b^r(\theta))\}$ for all $\theta \in [\underline{\theta}, \theta^\star)$. Clearly, this perturbation is incentive feasible, and it keeps the allocation of types $\theta \in [\theta^\star, \overline{\theta}]$, and thus the social welfare from these types, unchanged. The change in social welfare from each type $\theta \in [\underline{\theta}, \theta^\star)$ is equal to

$$\theta U(\omega + b^r(\theta)) + \delta \overline{V}(b^r(\theta)) - \theta U(\omega + b^r(\theta^\star)) - \delta \overline{V}(b^r(\theta^\star)).$$

Note that by the definition of $b^r(\theta)$,

$$\delta \overline{V}(b^r(\theta)) - \delta \overline{V}(b^r(\theta^\star)) \geq \frac{1}{\beta} \left[\theta U(\omega + b^r(\theta^\star)) - \theta U(\omega + b^r(\theta)) \right].$$

Substituting back into the previous expression, we obtain that the change in social welfare from each $\theta \in [\underline{\theta}, \theta^\star)$ is greater than

$$\left(\frac{1}{\beta} - 1 \right) \left[\theta U(\omega + b^r(\theta^\star)) - \theta U(\omega + b^r(\theta)) \right],$$

which is strictly positive. Thus, the perturbation strictly increases social welfare, implying that a binding minimum borrowing requirement is strictly suboptimal.

Step 3. We show that $b(\theta) < b^r(\theta)$ for some $\theta \in \Theta$.

By Step 1 and Step 2, the allocation for types $\theta \in [\underline{\theta}, \theta^{**}]$ is as described in Definition 1 for some $\theta^* \geq 0$. That is, equation (44) necessarily holds for $b(\theta^{**}) = b^r(\theta^*)$. All that remains to be shown is that $\theta^* < \overline{\theta}$. Suppose by contradiction that this is not true, which implies $\{b(\theta), V(b(\theta))\} = \{b^r(\theta), \overline{V}(b^r(\theta))\}$ for all $\theta \in \Theta$. Consider an incentive feasible perturbation that assigns $\{b(\theta), V(b(\theta))\} = \{b^r(\overline{\theta} - \varepsilon), \overline{V}(b^r(\overline{\theta} - \varepsilon))\}$ to all $\theta \in [\overline{\theta} - \varepsilon, \overline{\theta}]$, where $\varepsilon > 0$ is chosen to be small enough as to continue to satisfy the enforcement constraint (5) for all types $\theta \in \Theta$. Using the representation in (12), the change in social welfare from this perturbation is equal to

$$\frac{1}{\beta} \int_{\overline{\theta} - \varepsilon}^{\overline{\theta}} (U(\omega + b^r(\overline{\theta} - \varepsilon)) - U(\omega + b^r(\theta))) Q(\theta) d\theta.$$

For $\varepsilon > 0$ arbitrarily small, $b^r(\overline{\theta} - \varepsilon) < b^r(\theta)$ and $Q(\theta) < 0$ for all $\theta \in (\overline{\theta} - \varepsilon, \overline{\theta})$. Thus, the perturbation strictly increases social welfare.