Online Appendix for
"Fiscal Rules and Discretion under Limited Enforcement"

by Marina Halac and Pierre Yared

This Online Appendix contains the proofs omitted from the paper, namely those for Proposition 3, Proposition 4, Lemma 4, Proposition 5, and Proposition 6.

B Supplementary Proofs

B.1 Proof of Proposition 3

For any given threshold θ', denote by $\rho(\theta')$ the type exceeding θ' at which (10) holds:

$$\rho(\theta')U(\omega + b^r(\theta')) + \beta\delta\overline{V}(b^r(\theta')) = \rho(\theta')U(\omega + b^p(\rho(\theta'))) + \beta\delta\underline{V}(b^p(\rho(\theta'))). \tag{47}$$

Note that given θ', $\rho(\theta') > \theta'$ is uniquely defined. This follows from the same logic as in Step 2 in the proof of Lemma 1. We prove the proposition in five steps.

Step 1. We show that $d\rho(\theta')/d\theta' > 0$.

Implicit differentiation of (47), taking into account the definition of $b^r(\theta')$, yields

$$\frac{d\rho(\theta')}{d\theta'} = \frac{(\rho(\theta') - \theta')U'(\omega + b^r(\theta'))\frac{db^r(\theta')}{d\theta'}}{U(\omega + b^p(\rho(\theta'))) - U(\omega + b^r(\theta'))}. \tag{48}$$

We show that both the numerator and the denominator in (48) are strictly positive. Note that implicit differentiation of the first-order condition defining $b^r(\theta')$ yields

$$\frac{db^r(\theta')}{d\theta'} = -\frac{U'(\omega + b^r(\theta'))}{\theta' U''(\omega + b^r(\theta')) + \beta\delta\overline{V}''(b^r(\theta'))} > 0.$$

This inequality and $\rho(\theta') > \theta'$ imply that the numerator in (48) is strictly positive. Next, note that satisfaction of the enforcement constraint for type θ' implies

$$\theta' U(\omega + b^r(\theta')) + \beta\delta\overline{V}(b^r(\theta')) > \theta' U(\omega + b^p(\rho(\theta'))) + \beta\delta\underline{V}(b^p(\rho(\theta'))), \tag{49}$$

which combined with (47) yields

$$(\rho(\theta') - \theta')U(\omega + b^p(\rho(\theta'))) > (\rho(\theta') - \theta')U(\omega + b^r(\theta')).$$

Given $\rho(\theta') > \theta'$, this inequality implies $b^p(\rho(\theta')) > b^r(\theta')$, which in turn implies that the denominator in (48) is strictly positive. Therefore, we obtain $d\rho(\theta')/d\theta' > 0$.

Step 2. We show that if $\theta_c \leq \theta_e$, then condition (16) holds and the optimal maximally enforced deficit limit is unique and has $\theta^* = \theta_e$ and $\theta^{**} \geq \overline{\theta}$.

As noted in the text, if $\theta_c \leq \theta_e$, then Assumption 1 guarantees that $\int_{\theta_c}^{\overline{\theta}} Q(\theta)d\theta \geq \int_{\theta_e}^{\overline{\theta}} Q(\theta)d\theta = 0$, so condition (16) is satisfied. Since the perfect-enforcement deficit limit, specifying $\theta^* = \theta_e$ and $\theta^{**} \geq \overline{\theta}$, is enforceable in this case, it follows that the optimal rule coincides with this limit. Finally, this rule is unique since, by Assumption 1, there is a unique value of θ_e satisfying (13).

Step 3. We show that if $\theta_c > \theta_e$, then $\theta^* \leq \theta_c$.

Assume $\theta_c > \theta_e$. Suppose by contradiction that an optimal maximally enforced deficit limit features $\theta^* > \theta_c$, which implies $\theta^{**} \geq \overline{\theta}$. Consider a perturbation that reduces θ^* by $\varepsilon > 0$ arbitrarily small. Since in the original rule the enforcement constraint of all types $\theta \in \Theta$ is slack, this perturbation is incentive feasible. Using the representation in (12), the change in social welfare from the perturbation is equal to

$$-\frac{1}{\beta}\int_{\theta^*}^{\overline{\theta}} U'(\omega + b^r(\theta^*))\frac{db^r(\theta^*)}{d\theta^*}Q(\theta)d\theta. \tag{50}$$

Note that Assumption 1 together with (13) imply $\theta_e < \widehat{\theta}$. It then follows from $\theta^* > \theta_c > \theta_e$ and Assumption 1 that $\int_{\theta^*}^{\overline{\theta}} Q(\theta)d\theta < 0$, and therefore, since $db^r(\theta^*)/d\theta^* > 0$ (as established in Step 1), (50) is strictly positive. Hence, the perturbation strictly increases social welfare, implying that $\theta^* > \theta_c$ cannot hold.

Step 4. We show that if $\theta_c > \theta_e$ and condition (16) holds, then the optimal maximally enforced deficit limit is unique and has $\theta^* = \theta_c$ and $\theta^{**} = \overline{\theta}$.

Assume $\theta_c > \theta_e$ and that condition (16) holds. By Step 3, an optimal maximally enforced deficit limit has $\theta^* \leq \theta_c$. Suppose by contradiction that $\theta^* < \theta_c$, which implies $\theta^{**} = \rho(\theta^*) < \overline{\theta}$ for $\rho(\cdot)$ as defined in (47). Consider a perturbation that changes θ^* by some $\varepsilon \gtrless 0$ for $|\varepsilon|$ arbitrarily small, where $\theta^{**} = \rho(\theta^*)$ is also changed to preserve (47). This perturbation is incentive feasible. Using the representation in (12), for this perturbation to not increase social welfare for any arbitrarily small $\varepsilon \gtrless 0$, it must be that the following condition holds:

$$\int_{\theta^*}^{\rho(\theta^*)} U'(\omega + b^r(\theta^*))\frac{db^r(\theta^*)}{d\theta^*}Q(\theta)d\theta$$
$$+\frac{d\rho(\theta^*)}{d\theta^*}\left[U(\omega + b^r(\theta^*)) - U(\omega + b^p(\rho(\theta^*)))\right]Q(\rho(\theta^*)) = 0.$$

Using (48) to substitute for $\frac{d\rho(\theta^*)}{d\theta^*}$ and simplifying terms, we can rewrite this condition as

$$\int_{\theta^*}^{\rho(\theta^*)} (Q(\theta) - Q(\rho(\theta^*)))d\theta = 0. \tag{51}$$

Given Assumption 1, (51) requires $\theta^* < \widehat{\theta} < \rho(\theta^*)$ with

$$Q(\theta^*) > Q(\rho(\theta^*)). \tag{52}$$

Now note that the derivative of the left-hand side of (51) with respect to θ^* is equal to

$$-\left(Q(\theta^*) - Q(\rho(\theta^*))\right) - \int_{\theta^*}^{\rho(\theta^*)} Q'(\rho(\theta^*)))\frac{d\rho(\theta^*)}{d\theta^*}d\theta. \tag{53}$$

By (52), the first term in (53) is strictly negative. Moreover, since $\rho(\theta^*) > \widehat{\theta}$, Assumption 1 implies $Q'(\rho(\theta^*))) > 0$. Given $d\rho(\theta')/d\theta' > 0$ (as established in Step 1), it then follows that the second term in (53) is also strictly negative. Hence, the derivative of the left-hand side of (51) with respect to θ^* is strictly negative. However, using the contradiction assumption that $\theta^* < \theta_c$, condition (51) then requires that the left-hand side of (16) be strictly negative, contradicting the assumption that condition (16) holds. Therefore, there exists a perturbation that changes θ^* by some $\varepsilon \gtrless 0$ which strictly increases social welfare, implying that the unique optimal maximally enforced deficit limit has $\theta^* = \theta_c$ and $\theta^{**} = \overline{\theta}$.

Step 5. We show that if $\theta_c > \theta_e$ and condition (16) does not hold, then the optimal maximally enforced deficit limit is unique and has $\theta^* \in (\theta_e, \theta_c)$ and $\theta^{**} < \overline{\theta}$.

Assume $\theta_c > \theta_e$ and that condition (16) is violated. By Step 3, an optimal maximally enforced deficit limit has $\theta^* \leq \theta_c$. We begin by showing that $\theta^* = \theta_c$ cannot be optimal. Suppose by contradiction that an optimal maximally enforced deficit limit sets $\theta^* = \theta_c$ and thus $\theta^{**} = \rho(\theta_c) = \overline{\theta}$. Consider a perturbation that reduces θ^* by $\varepsilon > 0$ arbitrarily small, where $\theta^{**} = \rho(\theta^*)$ is also changed to preserve (47). This perturbation is incentive feasible. Using the representation in (12), for this perturbation to not increase social welfare for any arbitrarily small $\varepsilon > 0$, it must be that the following condition holds:

$$-\int_{\theta^*}^{\rho(\theta^*)} U'(\omega + b^r(\theta^*))\frac{db^r(\theta^*)}{d\theta^*}Q(\theta)d\theta$$
$$-\frac{d\rho(\theta^*)}{d\theta^*}\left[U(\omega + b^r(\theta^*)) - U(\omega + b^p(\rho(\theta^*)))\right]Q(\rho(\theta^*)) \leq 0.$$

By analogous logic as in Step 4 above, we can rewrite this condition as

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

where we have taken into account that $\theta^* = \theta_c$ and $\theta^{**} = \rho(\theta_c) = \overline{\theta}$. However, this inequality contradicts the assumption that condition (16) does not hold. Therefore, the perturbation strictly increases social welfare, implying that any optimal maximally enforced deficit limit has $\theta^* < \theta_c$ and $\theta^{**} = \rho(\theta^*) < \overline{\theta}$.

We next show that the optimal values of θ^* and $\theta^{**} = \rho(\theta^*)$ are unique with $\theta^* > \theta_e$. By analogous logic as in Step 4 above, the optimal value of θ^* must satisfy (51). As shown in Step 4, the left-hand side of (51) is strictly decreasing in θ^*. This has two implications. First, it implies that there is a unique value of θ^* and associated $\theta^{**} = \rho(\theta^*)$ which solve (51). Second, given (13), Assumption 1, and the fact that the left-hand side of (51) is strictly decreasing in $\rho(\theta^*)$, it implies that if $\theta^* \leq \theta_e$, then the left-hand side of (51) must be strictly positive, a contradiction. Therefore, the unique value of θ^* that solves (51) must satisfy $\theta^* > \theta_e$.

We are left to establish the claim in the text that condition (16) is equivalent to condition (17). Integration by parts of (16) implies

$$\int_{\theta_c}^{\overline{\theta}}(Q(\theta) - Q(\overline{\theta}))d\theta = \int_{\theta_c}^{\overline{\theta}}\left[(\theta - \theta_c)f(\theta) - \theta f(\theta)(1-\beta) + \overline{\theta}f(\overline{\theta})(1-\beta)\right]d\theta.$$

Dividing through by $1 - F(\theta_c)$ and rearranging terms yields (17).

B.2 Proof of Proposition 4

Let $\theta^L, \theta^H \in \Theta$ and $\Delta > 0$ be defined as in Definition 2. We prove the proposition by proving the following three claims.

Claim 1. Suppose Assumption 1 is strictly violated. If a maximally enforced deficit limit with thresholds $\{\theta^*, \theta^{**}\}$ is a solution to (7) for given functions $\{\underline{V}(b), \overline{V}(b)\}$, then $\theta^* \leq \theta^L$ and $\theta^{**} \geq \theta^H$.

Proof of Claim 1. Assume that Assumption 1 is strictly violated. Suppose by contradiction that a maximally enforced deficit limit with $\theta^* > \theta^L$ is a solution to (7). Then analogous to Step 2 (Case 2) in the proof of Proposition 1, consider a perturbation that drills a hole in the borrowing schedule in the range $[\theta^L, \theta^L + \varepsilon]$ for arbitrarily small $\varepsilon > 0$ satisfying $\theta^L + \varepsilon < \min\left\{\theta^*, \theta^L + \Delta\right\}$. This perturbation is incentive feasible. Moreover, since $Q(\theta)$ is strictly increasing in this range, the arguments in Step 2 in the proof of Proposition 1 imply that this perturbation strictly increases social welfare. It follows that $\theta^* > \theta^L$ cannot hold and we must have $\theta^* \leq \theta^L$.

Next, suppose by contradiction that a maximally enforced deficit limit with $\theta^{**} < \theta^H$ is a solution to (7). Then consider types $\theta \in [\theta^H - \varepsilon, \theta^H]$ for arbitrarily small $\varepsilon > 0$ satisfying $\theta^H - \varepsilon > \max\left\{\theta^{**}, \theta^H - \Delta\right\}$. For each such type θ, we have $\{b(\theta), V(b(\theta))\} = \{b^p(\theta), \underline{V}(b^p(\theta))\}$ and $Q'(\theta) < 0$. Thus, this is the same situation as in Step 1 in the proof of Lemma 3, and analogous to that step, we can show that there is a perturbation of the allocation that strictly increases social welfare, yielding a contradiction. It follows that $\theta^{**} < \theta^H$ cannot hold and we must have $\theta^{**} \geq \theta^H$.

Claim 2. Suppose Assumption 1 is strictly violated. For any function $\overline{V}(b)$, there exists a function $\underline{V}(b)$ for which no solution to (7) is a maximally enforced deficit limit.

Proof of Claim 2. Suppose Assumption 1 is strictly violated. Given $\overline{V}(b)$, define $\underline{V}(b) = \overline{V}(b) - m$ for $m > 0$. By Claim 1, if a maximally enforced deficit limit with thresholds $\{\theta^*, \theta^{**}\}$ solves (7), then $\theta^* \leq \theta^L$ and $\theta^{**} \geq \theta^H$. Consider the indifference condition (10) which defines, for any given θ^*, a unique value of $\theta^{**} > \theta^*$. This condition shows that given $\overline{V}(b)$ and $\underline{V}(b) = \overline{V}(b) - m$, the value of $(\theta^{**} - \theta^*)$ is continuous in m and approaches 0 as m goes to 0. It follows that if we take $m > 0$ small enough, then $\theta^* \leq \theta^L < \theta^H \leq \theta^{**}$ cannot hold. Thus, for any given $\overline{V}(b)$, there exists $\underline{V}(b)$ such that a maximally enforced deficit limit is not a solution to (7).

Claim 3. Suppose Assumption 1 is weakly violated. For any function $\overline{V}(b)$, there exists a function $\underline{V}(b)$ for which not every solution to (7) is a maximally enforced deficit limit.

Proof of Claim 3. Suppose Assumption 1 is weakly violated and a maximally enforced deficit limit with thresholds $\{\theta^*, \theta^{**}\}$ is a solution to (7). Then $\{\theta^*, \theta^{**}\}$ satisfy condition (10) and analogous arguments as in the proof of Claim 2 above imply that, given $\overline{V}(b)$, there exists a function $\underline{V}(b)$ such that $\theta^* \leq \theta^L < \theta^H \leq \theta^{**}$ cannot hold. This means that given such continuation value functions, any maximally enforced deficit limit $\{\theta^*, \theta^{**}\}$ solving (7) must have either $\theta^* > \theta^L$ or $\theta^{**} < \theta^H$ (or both). Suppose first that $\theta^* > \theta^L$. Then consider a perturbation as in the proof of Claim 1 above which drills a hole in the borrowing schedule in the range $[\theta^L, \theta^L + \varepsilon]$ for arbitrarily small $\varepsilon > 0$ satisfying $\theta^L + \varepsilon < \min\left\{\theta^*, \theta^L + \Delta\right\}$.

The same arguments as in the proof of Claim 1, given $Q'(\theta) \geq 0$ for $\theta \in [\theta^L, \theta^L + \varepsilon]$, imply that this perturbation weakly increases social welfare. The resulting allocation is therefore a solution to (7), and it is not a maximally enforced deficit limit.

Suppose next that $\theta^{**} < \theta^H$. Then as in the proof of Claim 1 above, consider types $\theta \in [\theta^H - \varepsilon, \theta^H]$ for arbitrarily small $\varepsilon > 0$ satisfying $\theta^H - \varepsilon > \max\{\theta^{**}, \theta^H - \Delta\}$. For each such type θ, we have $\{b(\theta), V(b(\theta))\} = \{b^p(\theta), \underline{V}(b^p(\theta))\}$ and $Q'(\theta) \leq 0$. Thus, we can perturb the allocation of these types as in Step 1 in the proof of Lemma 3 and weakly increase social welfare. The resulting allocation is therefore a solution to (7), and it is not a maximally enforced deficit limit.

B.3 Proof of Lemma 4

To prove this lemma, we consider the representation of our problem using savings rates. Define the savings rate s_t as in equation (25). As explained in the text, a government's choice of debt b_t given b_{t-1} is equivalent to a choice s_t, and the public history h^{t-1} given b_{-1} is equivalent to a public history $\widetilde{h}^{t-1} = \{b_{-1}, s_0, \ldots, s_{t-1}\} \in \mathbb{R}^t$. As such, it is without loss to redefine strategies and payoffs as conditional on $\widetilde{h}^{t-1}$, with $V_t(\widetilde{h}^{t-1})$ and $b_{t-1}(\widetilde{h}^{t-1})$ denoting the continuation value and debt at history $\widetilde{h}^{t-1}$, and $s_t(\widetilde{h}^{t-1}, \theta_t)$ denoting the savings rate at history $(\widetilde{h}^{t-1}, \theta_t)$.

Suppose the preferences in (20) satisfy $\gamma \neq 1$. Let $\widetilde{V}_t(\widetilde{h}^{t-1})$ denote the continuation value normalized by the level of debt starting from a history $\widetilde{h}^{t-1}$, defined as:

$$\widetilde{V}_t(\widetilde{h}^{t-1}) = \frac{V_t(\widetilde{h}^{t-1}) + \mathbb{E}[\theta_t]/\left[(1-\gamma)(1-\delta)\right]}{\left[R\tau/(R-1) - Rb_{t-1}(\widetilde{h}^{t-1})\right]^{1-\gamma}}. \tag{54}$$

Note that dividing both sides of (22) by $\left[R\tau/(R-1) - Rb_{t-1}(\widetilde{h}^{t-1})\right]^{1-\gamma}$ and substituting with (20) yields

$$\widetilde{V}_t(\widetilde{h}^{t-1}) = \mathbb{E}\left[\theta_t \frac{(1 - s_t(\widetilde{h}^{t-1}, \theta_t))^{1-\gamma}}{1-\gamma} + \delta(s_t(\widetilde{h}^{t-1}, \theta_t)R)^{1-\gamma} \widetilde{V}_{t+1}(\widetilde{h}^{t-1}, s_t(\widetilde{h}^{t-1}, \theta_t))\right]. \tag{55}$$

This equation shows that the normalized continuation value at any given history $\widetilde{h}^{t-1}$ is independent of the inherited level of debt $b_{t-1}(\widetilde{h}^{t-1})$ and depends only on the future sequence of saving rates. We next use this observation to show that whether or not a profile of savings rate strategies constitutes an equilibrium is independent of initial debt. Specifically, note that the incentive compatibility constraints (23) and (24) are equivalent to the following constraint:

$$\theta_t U(\omega_t(h^{t-1}) + b_t(h^{t-1},\theta_t)) + \beta\delta V_{t+1}(h^{t-1}, b_t(h^{t-1},\theta_t)) \tag{56}$$
$$\geq \theta_t U(\omega_t(h^{t-1}) + b'_t) + \beta\delta V_{t+1}(h^{t-1}, b'_t) \;\text{ for all } \theta_t \in \Theta \text{ and } b'_t \in [\underline{b}(b_{t-1}), \overline{b}(b_{t-1})].$$

Dividing both sides by $\left[R\tau/(R-1) - Rb_{t-1}(\widetilde{h}^{t-1})\right]^{1-\gamma}$ and substituting with (20), we can rewrite this constraint as

$$\theta_t \frac{(1 - s_t(\widetilde{h}^{t-1},\theta_t))^{1-\gamma}}{1-\gamma} + \beta\delta(s_t(\widetilde{h}^{t-1},\theta_t)R)^{1-\gamma}\widetilde{V}_{t+1}(\widetilde{h}^{t-1}, s_t(\widetilde{h}^{t-1},\theta_t)) \tag{57}$$
$$\geq \theta_t \frac{(1-s'_t)^{1-\gamma}}{1-\gamma} + \beta\delta(s'_t R)^{1-\gamma}\widetilde{V}_{t+1}(\widetilde{h}^{t-1}, s'_t) \;\text{ for all } \theta_t \in \Theta \text{ and } s'_t \in [\nu, 1-\nu].$$

Observe therefore that if a profile of savings rate strategies is an equilibrium starting from some initial debt $b_{-1} = b$—that is, if such a profile satisfies (57) given $b_{-1} = b$—then the same profile is an equilibrium starting from any other initial debt $b_{-1} = b' \neq b$. The reason is that the incentive compatibility constraints in (57) are independent of initial debt and depend only on the sequence of savings rates. Moreover, as a consequence, we obtain that the highest and lowest normalized continuation values at time 0, namely the highest and lowest values of $\widetilde{V}_0(\widetilde{h}^{-1})$, are independent of the initial debt b_{-1}. We can thus represent these values by $\widetilde{\overline{V}}$ and $\widetilde{\underline{V}}$, and, using the definition in (54), we obtain:

$$\begin{aligned} \overline{V}(b) &= \widetilde{\overline{V}}\,[R\tau/(R-1) - Rb]^{1-\gamma} - \mathbb{E}[\theta]/[(1-\gamma)(1-\delta)], \\ \underline{V}(b) &= \widetilde{\underline{V}}\,[R\tau/(R-1) - Rb]^{1-\gamma} - \mathbb{E}[\theta]/[(1-\gamma)(1-\delta)]. \end{aligned}$$

By (55), $\widetilde{\overline{V}} > 0$ and $\widetilde{\underline{V}} > 0$ if $\gamma < 1$, and $\widetilde{\overline{V}} < 0$ and $\widetilde{\underline{V}} < 0$ if $\gamma > 1$. Hence, it follows that $\overline{V}(b)$ and $\underline{V}(b)$ are continuously differentiable, strictly decreasing, and strictly concave in b. This proves the lemma for the case of $\gamma \neq 1$.

Finally, suppose $\gamma = 1$, so that $U(g) = \log(g)$. Then the arguments are immediate, since (57) is independent of the inherited level of debt even without performing a normalization as above. In this case, we define $\widetilde{V}_t(\widetilde{h}^{t-1}) = V_t(\widetilde{h}^{t-1}) - \log(R\tau/(R-1) - Rb_{t-1}(\widetilde{h}^{t-1}))$, so that:

$$\begin{aligned} \overline{V}(b) &= \widetilde{\overline{V}} + \log\left[R\tau/(R-1) - Rb\right], \\ \underline{V}(b) &= \widetilde{\underline{V}} + \log\left[R\tau/(R-1) - Rb\right]. \end{aligned}$$

Hence, it follows again that $\overline{V}(b)$ and $\underline{V}(b)$ are continuously differentiable, strictly decreasing, and strictly concave in b, completing the proof.

B.4 Proof of Proposition 5

Given the results established in the proofs of Proposition 2, Proposition 3, and Lemma 4, the proof of this proposition follows from the arguments in the text.

B.5 Proof of Proposition 6

The proof of this proposition is analogous to the proof of Proposition 2, with the uniqueness of the optimal thresholds θ_n^* and θ_n^{**} following from the same logic as in the proof of Proposition 3, and their independence of b_{-1} following from the results established in the proof of Lemma 4. We therefore describe this proof only briefly here, focusing on the steps that differ from those in the previous proofs.

The worst punishment solves the analog of the program in (7), where we now minimize social welfare as opposed to maximizing it:

$$\min_{\left\{b(\theta)\in[\underline{b},\overline{b}],V(b(\theta))\right\}_{\theta\in\Theta}} \mathbb{E}\big[\theta U(\omega+b(\theta))+\delta V(b(\theta))\big] \tag{58}$$

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

Consider first the proof of Proposition 1. By analogous arguments to those in the first part of Step 1 in that proof, we can establish that $V(b(\theta))$ must be right-continuous at each $\theta \in [\underline{\theta},\overline{\theta})$ in any solution to (58). Moreover, by arguments analogous to those in the second part of Step 1, we can establish that $V(b(\overline{\theta})) = \overline{V}(b(\overline{\theta}))$. Specifically, if $V(b(\overline{\theta})) \in (\underline{V}(b(\overline{\theta})), \overline{V}(b(\overline{\theta})))$, then a perturbation that marginally increases $b(\overline{\theta}) \in (\underline{b}, \overline{b})$ and changes $V(b(\overline{\theta}))$ so as to keep type $\overline{\theta}$'s welfare unchanged is incentive feasible and strictly reduces social welfare. Such a perturbation is also incentive feasible (and welfare reducing) if $V(b(\overline{\theta})) = \underline{V}(b(\overline{\theta}))$, as in this case $b(\overline{\theta}) = b^p(\overline{\theta})$ by the enforcement constraint (5) and thus the perturbation requires setting $V(b(\overline{\theta})) > \underline{V}(b(\overline{\theta}))$. It follows that $V(b(\overline{\theta})) = \overline{V}(b(\overline{\theta}))$ must hold in any solution to (58).

The claims in Step 2 and Step 3 in the proof of Proposition 1 also apply when solving the program in (58). The reason is that perturbations that apply whenever $Q'(\theta) > 0$ in the maximization of social welfare now apply whenever $Q'(\theta) < 0$ in the minimization of social welfare, and vice versa. Hence, the arguments in these steps, together with those in Step 1 just described, imply that continuation values are bang-bang: in any solution to (58), $V(b(\theta)) \in \{\underline{V}(b(\theta)), \overline{V}(b(\theta))\}$ for all $\theta \in [\underline{\theta}, \overline{\theta}]$.

Consider next the proof of Lemma 3. Step 1 and Step 2 in that proof are isomorphic in the sense that the arguments applying to types $\theta < \widehat{\theta}$ in the maximization of social welfare now apply to types $\theta > \widehat{\theta}$ in the minimization of social welfare, and vice versa. Combined with the claims above, these steps thus imply the following: in any solution to (58), there exists $\theta_n^{**} \leq \overline{\theta}$ such that $V(b(\theta)) = \underline{V}(b(\theta))$ for all $\theta \in [\underline{\theta}, \theta_n^{**})$ and $V(b(\theta)) = \overline{V}(b(\theta))$ for all $\theta \in [\theta_n^{**}, \overline{\theta}]$.

The analog of Step 3 in the proof of Lemma 3 then consists of showing that $\theta_n^{**} < \overline{\theta}$. To see why this must be true, suppose by contradiction that $\theta_n^{**} = \overline{\theta}$, namely that $V(b(\theta)) = \underline{V}(b(\theta))$ for all $\theta \in [\underline{\theta}, \overline{\theta})$ and $V(b(\theta))$ jumps at $\overline{\theta}$ to $\overline{V}(b(\overline{\theta}))$. Note that the enforcement constraint (5) implies $b(\theta) = b^p(\theta)$ for all $\theta \in [\underline{\theta}, \overline{\theta})$, and using the representation in (12), social welfare is equal to

$$\frac{1}{\beta}\underline{\theta}U(\omega + b^p(\underline{\theta})) + \delta\underline{V}(b^p(\underline{\theta})) + \frac{1}{\beta}\int_{\underline{\theta}}^{\overline{\theta}} U(\omega + b^p(\theta))Q(\theta)d\theta. \tag{59}$$

Consider a global perturbation in which all types $\theta \in \Theta$ are assigned the allocation corresponding to a maximally enforced surplus limit $\{\theta_n^*, \theta_n^{**}\}$, with $\theta_n^{**} \in (\underline{\theta}, \overline{\theta})$, $Q(\theta_n^{**}) < 0$, and $\theta_n^* \geq \overline{\theta}$ (and with equation (26) being satisfied). Note that this is feasible since $Q(\overline{\theta}) < 0$ and $Q(\cdot)$ is continuous. Using the representation in (59) and taking into account that the perturbation keeps the allocation of types $\theta \in [\underline{\theta}, \theta_n^{**})$ unchanged, we find that the change in social welfare from the perturbation is equal to

$$\frac{1}{\beta}\int_{\theta_n^{**}}^{\overline{\theta}} (U(\omega + b^r(\theta_n^*)) - U(\omega + b^p(\theta))Q(\theta)d\theta. \tag{60}$$

Note that $b^r(\theta_n^*) > b^p(\theta)$ and $Q(\theta) < 0$ for all $\theta \in [\theta_n^{**}, \overline{\theta}]$ (by construction, Assumption 1, and the surplus limit being incentive compatible). Hence, the perturbation strictly reduces social welfare, implying that $\theta_n^{**} < \overline{\theta}$ must hold in any solution to (58).

Given the claims above, the next step in the proof of Proposition 6 is to show that $b(\theta)$ is continuous for $\theta \geq \theta_n^{**}$. Here analogous arguments to those in the proof of Proposition 2 apply. The optimality of a surplus limit that is binding also follows from analogous arguments as in that proof. Note that the optimal surplus limit must satisfy $\theta_n^{**} \geq \underline{\theta}$: otherwise, if $\theta_n^{**} < \underline{\theta}$, then a perturbation that tightens the limit by raising θ_n^{**} to $\underline{\theta}$ is incentive feasible and strictly reduces social welfare. Finally, as noted above, uniqueness of the optimal thresholds θ_n^* and θ_n^{**} follows from analogous logic as in the proof of Proposition 3, and their independence of initial debt follows from the results established in the proof of Lemma 4.