B Online Appendix

B.1 Proof of Lemma 1

We proceed in three steps.

Step 1. Suppose $\theta^* \geq \underline{\theta}$. We show that (3) and (4) 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 no penalty. Thus, given $\theta \in [\underline{\theta}, \theta^*]$, type θ's welfare cannot be increased, and (3) and (4) are trivially satisfied.

Step 2. We show that (3) and (4) are satisfied for types $\theta \in (\theta^*, \theta^{**}]$.

Take first the enforcement constraint (4). We can rewrite it for $\theta \in (\theta^*, \theta^{**}]$ as

$$\theta U(\omega + b^r(\theta^*)) + \beta\delta V(b^r(\theta^*) - \theta U(\omega + b^p(\theta)) - \beta\delta(V(b^p(\theta)) - \overline{P}(b^p(\theta))) \geq 0. \tag{B.1}$$

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

$$U(\omega + b^r(\theta^*)) - U(\omega + b^p(\theta)),$$

which is weakly decreasing in θ, since $b^p(\theta)$ is nondecreasing. This means that the left-hand side of (B.1) is weakly concave. Since (B.1) holds as a strict inequality for $\theta = \theta^*$ and as an equality for $\theta = \theta^{**}$ (by (8)), this weak concavity implies that (B.1) holds as a strict inequality for all $\theta \in (\theta^*, \theta^{**})$. Thus, constraint (4) is satisfied for all $\theta \in (\theta^*, \theta^{**}]$.

Take next the truthtelling constraint (3). 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 (3) is satisfied for all $\theta \in (\theta^*, \theta^{**}]$ given $\theta' > \theta^{**}$. Note that $(b(\theta'), P(\theta')) = (b^p(\theta'), \overline{P}(b^p(\theta')))$ for all $\theta' > \theta^{**}$, and by the definition of $b^p(\theta)$,

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

for all $\theta' \in \Theta$. Thus, the fact that the enforcement constraint (4) is satisfied for all $\theta \in (\theta^*, \theta^{**}]$ implies that (3) is satisfied for all such types given $\theta' > \theta^{**}$.

Step 2b: We show that (3) is satisfied for all $\theta \in (\theta^*, \theta^{**}]$ given $\theta' < \theta^*$. Suppose by contradiction that this is not the case, that is,

$$\theta(U(\omega + b^r(\theta^*)) - U(\omega + b^r(\theta'))) < \beta\delta\left(V(b^r(\theta')) - V(b^r(\theta^*))\right) \tag{B.2}$$

for some $\theta \in (\theta^*, \theta^{**}]$ and $\theta' < \theta^*$. By Step 1, (3) holds for type θ^* given $\theta' < \theta^*$:

$$\theta^*(U(\omega + b^r(\theta^*)) - U(\omega + b^r(\theta'))) \geq \beta\delta\left(V(b^r(\theta')) - V(b^r(\theta^*))\right). \tag{B.3}$$

Combining (B.2) and (B.3) yields

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

which is a contradiction since $\theta > \theta^*$ and $b^r(\theta') \leq b^r(\theta^*)$. The claim follows.

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

Constraint (4) is satisfied as an equality for all $\theta \in (\theta^{**}, \overline{\theta}]$. It is immediate that constraint (3) 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 maximum penalty. Consider next constraint (3) for $\theta \in (\theta^{**}, \overline{\theta}]$ given $\theta' \in [\theta^*, \theta^{**}]$. Note that $(b(\theta'), P(\theta')) = (b^r(\theta^*), 0)$ for all $\theta' \in [\theta^*, \theta^{**}]$. Thus, satisfaction of this constraint is ensured if (B.1) is violated for $\theta \in (\theta^{**}, \overline{\theta}]$. The latter is true since, as shown above, the left-hand side of (B.1) is weakly concave and (B.1) holds as an equality for $\theta = \theta^{**}$ and a strict inequality for $\theta \in (\theta^*, \theta^{**})$.

Finally, consider constraint (3) for $\theta \in (\theta^{**}, \overline{\theta}]$ given $\theta' < \theta^*$. Since (3) is satisfied given $\theta' \in [\theta^*, \theta^{**}]$, satisfaction of this constraint given $\theta' < \theta^*$ is ensured if

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

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

B.2 Proof of Corollary 1

Consider optimal rules with $b(\theta) \in (\underline{b}, \overline{b})$ for all $\theta \in \Theta$. We proceed in four steps.

Step 1. We show that an optimal maximally enforced deficit limit solves:

$$\max_{\theta^*,\theta^{**}} \left\{ \int_0^{\theta^*} U(\omega + b^r(\theta))Q(\theta)d\theta + \int_{\theta^*}^{\theta^{**}} U(\omega + b^r(\theta^*))Q(\theta)d\theta + \int_{\theta^{**}}^{\overline{\theta}} U(\omega + b^p(\theta))Q(\theta)d\theta \right\} \tag{B.4}$$

$$\text{subject to (8),}$$

where $Q(\theta) = 1$ for $\theta < \underline{\theta}$ and, by convention, the last integral equals zero if $\theta^{**} \geq \overline{\theta}$.

By the arguments in the text, social welfare can be written as

$$\frac{1}{\beta}\underline{\theta}U(\omega + b(\underline{\theta})) + \delta\left(V(b(\underline{\theta})) - P(\underline{\theta})\right) + \frac{1}{\beta}\int_{\underline{\theta}}^{\overline{\theta}} U(\omega + b(\theta))Q(\theta)d\theta,$$

which in turn can be rewritten as

$$\lim_{\underline{\theta}'\downarrow 0} \frac{1}{\beta}\underline{\theta}'U(\omega + b(\underline{\theta}')) + \delta\left(V(b(\underline{\theta}')) - P(\underline{\theta}')\right) + \frac{1}{\beta}\int_0^{\overline{\theta}} U(\omega + b(\theta))Q(\theta)d\theta,$$

where $Q(\theta) = 1$ for $\theta < \underline{\theta}$. Hence, social welfare under a maximally enforced deficit limit can be represented as

$$\lim_{\underline{\theta}'\downarrow 0} \frac{1}{\beta}\underline{\theta}'U(\omega + b^r(\underline{\theta}')) + \delta\left(V(b^r(\underline{\theta}')) - P(\underline{\theta}')\right) \tag{B.5}$$

$$+\frac{1}{\beta}\int_0^{\theta^*} U(\omega + b^r(\theta))Q(\theta)d\theta + \frac{1}{\beta}\int_{\theta^*}^{\theta^{**}} U(\omega + b^r(\theta^*))Q(\theta)d\theta + \frac{1}{\beta}\int_{\theta^{**}}^{\overline{\theta}} U(\omega + b^p(\theta))Q(\theta)d\theta.$$

Since the first term in (B.5) is independent of the choice of $\theta^* > 0$ and $\theta^{**} > \theta^*$, and since the constant $\frac{1}{\beta}$ multiplies all other terms, the objective in (B.4) is equivalent to (B.5).

Step 2. Consider the following relaxed program:

$$\max_{\theta^*} \left\{ \int_0^{\theta^*} U(\omega + b^r(\theta))Q(\theta)d\theta + \int_{\theta^*}^{\overline{\theta}} U(\omega + b^r(\theta^*))Q(\theta)d\theta \right\}.$$

We show that any solution to this program yields strictly higher social welfare than any solution to program (B.4) with $\theta^{**} < \overline{\theta}$.

Take any solution $\{\theta^*, \theta^{**}\}$ to program (B.4) with $\theta^{**} < \overline{\theta}$. To prove the claim, it suffices to show that social welfare strictly increases if we change the allocation of types

$\theta \in [\theta^{**}, \overline{\theta}]$ from $(b(\theta), P(\theta)) = (b^p(\theta), \overline{P}(b^p(\theta)))$ to $(b(\theta), P(\theta)) = (b^r(\theta^*), 0)$. To prove this, note first that by Step 1 in the proof of Proposition 2, the solution $\{\theta^*, \theta^{**}\}$ to program (B.4) has $\theta^{**} \geq \widehat{\theta}$. Hence, by Assumption 1, $Q(\theta) < 0$ for all $\theta \in [\theta^{**}, \overline{\theta}]$. Given the representation in (B.4), the claim then follows if $b^r(\theta^*) < b^p(\theta)$ for all $\theta \in [\theta^{**}, \overline{\theta}]$. We show next that this inequality holds. Given the solution $\{\theta^*, \theta^{**}\}$, the following conditions hold for all $\theta \in [\theta^{**}, \overline{\theta}]$:

$$\theta U(\omega + b^r(\theta^*)) + \beta\delta V(b^r(\theta^*)) \leq \theta U(\omega + b^p(\theta)) + \beta\delta(V(b^p(\theta)) - \overline{P}(b^p(\theta)))$$

and

$$\theta^* U(\omega + b^r(\theta^*)) + \beta\delta V(b^r(\theta^*)) > \theta^* U(\omega + b^p(\theta)) + \beta\delta(V(b^p(\theta)) - \overline{P}(b^p(\theta))).$$

Combining these two inequalities yields

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

which implies $b^p(\theta) > b^r(\theta^*)$ for all $\theta \in [\theta^{**}, \overline{\theta}]$.

Step 3. We show that the solution to the relaxed program in Step 2 is $\theta^* = \theta_e$, where $\theta_e \in [0, \overline{\theta})$ is uniquely defined by (11). Moreover, if $\theta^* = \theta_e$ satisfies constraint (8) for some $\theta^{**} \geq \overline{\theta}$, then these values correspond to the unique solution to program (B.4).

To prove the first claim, consider the first-order condition of the relaxed program in Step 2:

$$\frac{db^r(\theta^*)}{d\theta^*} U'(\omega + b^r(\theta^*)) \int_{\theta^*}^{\overline{\theta}} Q(\theta) d\theta = 0.$$

Since $\frac{db^r(\theta^*)}{d\theta^*} > 0$ and $U'(\omega + b^r(\theta^*)) > 0$, this condition requires that the integral be equal to 0. Hence, by the definition in (11), we obtain $\theta^* = \theta_e$. Note that this value is uniquely defined since, by Assumption 1, $\int_{\theta^*}^{\overline{\theta}} Q(\theta) d\theta = 0$ requires $\theta^* < \widehat{\theta}$ and $Q(\theta^*) > 0$, and hence $\int_{\theta^*}^{\overline{\theta}} Q(\theta) d\theta$ is strictly decreasing in θ^*. Since $\int_{\theta^*}^{\overline{\theta}} Q(\theta) d\theta$ is strictly positive for $\theta^* = \varepsilon$ and strictly negative for $\theta^* = \overline{\theta} - \varepsilon$ for sufficiently small $\varepsilon > 0$,[30] it follows that a

[30] To see that $\int_{\varepsilon}^{\overline{\theta}} Q(\theta) d\theta > 0$ for ε sufficiently small, note that using integration by parts yields

$$\int_{\varepsilon}^{\overline{\theta}} Q(\theta) d\theta = -(1 - F(\varepsilon))\varepsilon + \int_{\varepsilon}^{\overline{\theta}} f(\theta)\theta d\theta - \int_{\varepsilon}^{\overline{\theta}} f(\theta)\theta(1-\beta) d\theta,$$

which approaches $\beta\mathbb{E}[\theta] > 0$ as ε goes to 0.

unique interior $\theta_e \in (0, \overline{\theta})$ exists and is the unique optimum.

To prove the second claim, note that if constraint (8) holds under $\theta^* = \theta_e$ and some $\theta^{**} \geq \overline{\theta}$, then such a deficit limit $\{\theta_e, \theta^{**}\}$ is feasible in program (B.4). Moreover, since this deficit limit yields the same social welfare as the relaxed program, it follows from Step 2 and the above claim that it yields strictly higher social welfare than any other feasible deficit limit and is thus the unique solution to program (B.4).

Step 4. We show that if (12) holds, then the solution to (B.4) has $\theta^* = \theta_e$ and $\theta^{**} \geq \overline{\theta}$.

The claim follows from Step 3 and the fact that if (12) holds, then constraint (8) is satisfied under $\theta^* = \theta_e$ and some $\theta^{**} \geq \overline{\theta}$.

B.3 Proof of Proposition 4

For any given threshold θ', denote by $\rho(\theta')$ the type exceeding θ' at which (8) holds:

$$\rho(\theta')U(\omega+b^r(\theta'))+\beta\delta V(b^r(\theta'))=\rho(\theta')U(\omega+b^p(\rho(\theta')))+\beta\delta(V(b^p(\rho(\theta')))-\overline{P}(b^p(\rho(\theta')))). \tag{B.6}$$

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 this proposition in five steps.

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

Implicit differentiation of (B.6), taking into account the definition of $b^r(\theta')$, yields

$$\frac{d\rho(\theta')}{d\theta'} = \frac{(\rho(\theta')-\theta')U'(\omega+b^r(\theta'))\frac{db^r(\theta')}{d\theta'}}{U(\omega+b^p(\rho(\theta')))-U(\omega+b^r(\theta'))}. \tag{B.7}$$

Note that since $\frac{db^r(\theta')}{d\theta'} > 0$ and $\rho(\theta') > \theta'$, the numerator in (B.7) is strictly positive. Moreover, by the arguments in Step 2 of the proof of Corollary 1, we have $b^p(\rho(\theta')) > b^r(\theta')$, which implies that the denominator is also strictly positive. Thus, we obtain $\frac{d\rho(\theta')}{d\theta'} > 0$.

Step 2. We show that if $\theta_c \leq \theta_e$, then condition (14) 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 $\theta_c \leq \theta_e$, 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 (14) is satisfied. The claim then follows from Corollary 1.

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. The change in social welfare, using the representation in (B.4), is

$$-\int_{\theta^*}^{\overline{\theta}} \frac{db^r(\theta^*)}{d\theta^*} U'(\omega + b^r(\theta^*))Q(\theta)d\theta. \tag{B.8}$$

Assumption 1 together with (11) 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 thus, since $\frac{db^r(\theta^*)}{d\theta^*} > 0$, (B.8) 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 (14) holds, then the optimal maximally enforced deficit limit is unique and has $\theta^* = \theta_c$ and $\theta^{**} = \overline{\theta}$.

Assume that $\theta_c > \theta_e$ and condition (14) 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 (B.6). Consider a perturbation that changes θ^* by some $\varepsilon \gtrless 0$ for $|\varepsilon|$ arbitrarily small, where $\theta^{**} = \rho(\theta^*)$ is also changed to preserve (B.6). This perturbation is incentive feasible. Using the representation in (B.4), for this perturbation to not increase social welfare for any arbitrarily small $\varepsilon \gtrless 0$, we must have

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

Using (B.7) to substitute for $\frac{d\rho(\theta^*)}{d\theta^*}$ and simplifying terms, we can rewrite this condition as

$$\int_{\theta^*}^{\rho(\theta^*)} (Q(\theta) - Q(\rho(\theta^*)))d\theta = 0. \tag{B.9}$$

Given Assumption 1, (B.9) requires $\theta^* < \widehat{\theta} < \rho(\theta^*)$ with

$$Q(\theta^*) > Q(\rho(\theta^*)). \tag{B.10}$$

Now note that the derivative of the left-hand side of (B.9) with respect to θ^* is equal to

$$-(Q(\theta^*)-Q(\rho(\theta^*)))-\int_{\theta^*}^{\rho(\theta^*)}Q'(\rho(\theta^*)))\frac{d\rho(\theta^*)}{d\theta^*}d\theta. \quad \text{(B.11)}$$

By (B.10), the first term is strictly negative. Moreover, since $\rho(\theta^*) > \widehat{\theta}$, Assumption 1 implies $Q'(\rho(\theta^*))) > 0$. Given $\frac{d\rho(\theta')}{d\theta'} > 0$ (as established in Step 1), it then follows that the second term in (B.11) is also strictly negative. Hence, the derivative of the left-hand side of (B.9) with respect to θ^* is strictly negative. However, using the contradiction assumption that $\theta^* < \theta_c$, condition (B.9) then requires that the left-hand side of (14) be strictly negative, contradicting the assumption that condition (14) 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 (14) 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 that $\theta_c > \theta_e$ and condition (14) 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 (B.6). This perturbation is incentive feasible. Using the representation in (B.4), for this perturbation to not increase social welfare for any arbitrarily small $\varepsilon > 0$, we must have

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

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

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

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 (14) 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 (B.9). As shown in Step 4, the left-hand side of (B.9) 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 (B.9). Second, given (11), Assumption 1, and the fact that the left-hand side of (B.9) is strictly decreasing in $\rho(\theta^*)$, it implies that if $\theta^* \leq \theta_e$, then the left-hand side of (B.9) must be strictly positive, a contradiction. Therefore, the unique value of θ^* that solves (B.9) must satisfy $\theta^* > \theta_e$.

B.4 Proof of Proposition 5

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 $\{\theta^*, \theta^{**}\}$ is a solution to (6) for given functions $V(b), \overline{P}(b)$, then $\theta^* \leq \theta^L$ and $\theta^{**} \geq \theta^H$.

Proof of Claim 1. Suppose Assumption 1 is strictly violated. Suppose by contradiction that a maximally enforced deficit limit with $\theta^* > \theta^L$ is a solution to (6). 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 $\theta^L + \varepsilon < \min\{\theta^*, \theta^L + \Delta\}$. 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, yielding a contradiction.

Next, suppose by contradiction that a maximally enforced deficit limit with $\theta^{**} < \theta^H$ is a solution to (6). Then 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 $(b(\theta), P(\theta)) = (b^p(\theta), \overline{P}(b^p(\theta)))$ and $Q'(\theta) < 0$. Thus, this is the same situation as in Step 1 in the proof of Proposition 2. Analogous to that step, we can show that there is an incentive feasible perturbation that strictly increases social welfare, yielding a contradiction.

Claim 2. Suppose Assumption 1 is strictly violated. For any function $V(b)$, there exists a function $\overline{P}(b)$ such that no solution to (6) is a maximally enforced deficit limit.

Proof of Claim 2. Suppose Assumption 1 is strictly violated. Given $V(b)$, define $\overline{P}(b) = P$ for $P > 0$. By Claim 1, if a maximally enforced deficit limit $\{\theta^*, \theta^{**}\}$ solves (6), then $\theta^* \leq \theta^L$ and $\theta^{**} \geq \theta^H$. Consider the indifference condition (8) which defines, for any given θ^*, a unique value of $\theta^{**} > \theta^*$. This condition shows that given $V(b)$ and $\overline{P}(b) = P$, the value of $(\theta^{**} - \theta^{*})$ is continuous in P and approaches 0 as P goes to 0. Hence, if we take $P > 0$ small enough, then $\theta^{*} \leq \theta^{L} < \theta^{H} \leq \theta^{**}$ cannot hold. The claim follows.

Claim 3. Suppose Assumption 1 is weakly violated. For any function $V(b)$, there exists a function $\overline{P}(b)$ such that not every solution to (6) is a maximally enforced deficit limit.

Proof of Claim 3. Suppose Assumption 1 is weakly violated and a maximally enforced deficit limit $\{\theta^{*}, \theta^{**}\}$ is a solution to (6). Then $\{\theta^{*}, \theta^{**}\}$ satisfy condition (8) and analogous arguments as in the proof of Claim 2 above imply that, given $V(b)$, there exists a function $\overline{P}(b)$ such that $\theta^{*} \leq \theta^{L} < \theta^{H} \leq \theta^{**}$ cannot hold. This means that given such functions, any maximally enforced deficit limit $\{\theta^{*}, \theta^{**}\}$ solving (6) 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 (6), 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 $\theta^{H} - \varepsilon > \max\left\{\theta^{**}, \theta^{H} - \Delta\right\}$. For each such type θ, we have $(b(\theta), P(\theta)) = (b^{p}(\theta), \overline{P}(b^{p}(\theta)))$ and $Q'(\theta) \leq 0$. Thus, we can perturb the allocation of these types as in Step 1 in the proof of Proposition 2 and weakly increase social welfare. The resulting allocation is therefore a solution to (6), and it is not a maximally enforced deficit limit.

B.5 Proof of Proposition 6

We prove each part of the proposition in order.

Part 1. Suppose the enforcement constraint binds under $\overline{P}(b)$. Then for $k = 0$ we have

$$\overline{\theta}U(\omega + b^{r}(\theta_{e})) + \beta\delta V(b^{r}(\theta_{e})) < \overline{\theta}U(\omega + b^{p}(\overline{\theta})) + \beta\delta(V(b^{p}(\overline{\theta})) - \overline{P}(b^{p}(\overline{\theta})) - k). \tag{B.12}$$

Observe that there exists a finite value $k' > 0$ such that the right-hand side of (B.12) equals the left-hand side under $k = k'$. If $k \in [0, k')$, the inequality in (B.12) is preserved and the enforcement constraint continues to bind under $\overline{P}(b) + k$. If instead $k \geq k'$, this inequality no longer holds and the enforcement constraint does not bind under $\overline{P}(b) + k$.

Part 2. Suppose the enforcement constraint binds and on-path penalties are optimal under $\overline{P}(b)$. By analogous arguments as in the proof of Part 1 above, there exists a finite $k''' > 0$ such that the enforcement constraint under $\overline{P}(b)+k$ binds if $k \in [0,k''')$ and does not bind if $k \geq k'''$. To complete the proof, take $k \in [0,k''')$ and define $\theta_c(k)$ as the solution to

$$\overline{\theta}U(\omega+b^r(\theta_c(k)))+\beta\delta V(b^r(\theta_c(k))) = \overline{\theta}U(\omega+b^p(\overline{\theta}))+\beta\delta(V(b^p(\overline{\theta}))-\overline{P}(b^p(\overline{\theta}))-k). \tag{B.13}$$

The value of $\theta_c(k)$ corresponds to the value of θ_c defined in (13) as a function of the additional penalty $k \in [0,k''')$. We show that $\theta_c(k)$ is strictly decreasing. Implicit differentiation of (B.13) yields

$$\frac{d\theta_c(k)}{dk} = -\frac{\beta\delta}{(\overline{\theta}-\theta_c(k))\frac{db^r(\theta_c(k))}{d\theta}U'(\omega+b^r(\theta_c(k)))} < 0, \tag{B.14}$$

where we have used the fact that $\theta_c(k)U'(\omega+b^r(\theta_c(k))) = -\beta\delta V'(b^r(\theta_c(k)))$. Since on-path penalties are optimal under $k=0$, Proposition 4 implies

$$\int_{\theta_c(0)}^{\overline{\theta}}(Q(\theta)-Q(\overline{\theta}))d\theta < 0. \tag{B.15}$$

By the definition of k''', the value of $\theta_c(k)$ approaches θ_e from above as k approaches k'''. Given the definition of θ_e in (11) and the fact that $Q(\overline{\theta}) < 0$, it follows that

$$\int_{\theta_c(k''')}^{\overline{\theta}}(Q(\theta)-Q(\overline{\theta}))d\theta > 0. \tag{B.16}$$

Equations (B.15) and (B.16) imply that there exists $k'' \in (0,k''')$ satisfying

$$\int_{\theta_c(k'')}^{\overline{\theta}}(Q(\theta)-Q(\overline{\theta}))d\theta = 0. \tag{B.17}$$

Note that k'' is unique: the derivative of the left-hand side of (B.17) with respect to k is

$$-\frac{d\theta_c(k'')}{dk}(Q(\theta_c(k''))-Q(\overline{\theta})) > 0,$$

where the inequality follows from the fact that $\frac{d\theta_c(k'')}{dk} < 0$ (by (B.14)) and $Q(\theta_c(k'')) > Q(\overline{\theta})$ (by (B.17) and Assumption 1). Therefore, we obtain $\int_{\theta_c(k)}^{\overline{\theta}}(Q(\theta)-Q(\overline{\theta}))d\theta < 0$ if

$k \in [0, k'')$ and $\int_{\theta_c(k)}^{\overline{\theta}} (Q(\theta) - Q(\overline{\theta}))d\theta > 0$ if $k \in (k'', k''')$. By Proposition 4, it follows that on-path penalties are optimal if $k \in [0, k'')$ and suboptimal if $k \in [k'', k''')$.

B.6 Proof of Proposition 7

We prove each part of the proposition in order.

Part 1. There are two cases to consider.

Case 1: Suppose that on-path penalties are suboptimal. By Proposition 4, the optimal rule sets $\theta^* = \theta_c(k)$ for $\theta_c(k)$ defined in (B.13) in the proof of Proposition 6. Since $\theta_c(k)$ is strictly decreasing in k by (B.14), it follows that θ^* strictly decreases (increases) when $\overline{P}(b)$ is shifted to $\overline{P}(b) + k$ for $k > 0$ ($k < 0$).

Case 2: Suppose that on-path penalties are optimal. We prove the result for the case of a positive penalty shift. The proof of the negative-shift case is analogous and thus omitted.

Given a penalty shift k, define $\rho^k(\theta)$ as the unique solution to

$$\rho^k(\theta)U(\omega + b^r(\theta)) + \beta\delta V(b^r(\theta)) \\ = \rho^k(\theta)U(\omega + b^p(\rho^k(\theta))) + \beta\delta(V(b^p(\rho^k(\theta))) - \overline{P}(b^p(\rho^k(\theta))) - k).$$

Observe that $\rho^k(\theta)$ corresponds to the value of θ^{**} that satisfies the indifference condition (8) given $\theta = \theta^*$ and the penalty shift k, and for $k = 0$ it corresponds to $\rho(\theta^*)$ defined in the proof of Proposition 4. It follows from Step 1 in that proof that $\rho^k(\theta)$ is strictly increasing in θ. Moreover, by implicit differentiation,

$$\frac{d\rho^k(\theta)}{dk} = -\frac{\beta\delta}{U(\omega + b^r(\theta)) - U(\omega + b^p(\rho^k(\theta)))} > 0,$$

where we have used the fact that $b^p(\rho^k(\theta)) > b^r(\theta)$, as implied by the arguments in Step 2 of the proof of Corollary 1.

Consider the optimal deficit limit $\{\theta^*, \theta^{**}\}$ under $\overline{P}(b)$ and denote by $\{\theta^{*k}, \theta^{**k}\}$ the optimal deficit limit under $\overline{P}(b) + k$. Since the enforcement constraint binds, we have $\theta^{**} = \rho(\theta^*)$ and $\theta^{**k} = \rho^k(\theta^{*k})$. By Step 4 in the proof of Proposition 4, the following first-order conditions uniquely define θ^* and θ^{*k}:

$$\int_{\theta^*}^{\rho(\theta^*)} (Q(\theta) - Q(\rho(\theta^*)))d\theta = 0, \tag{B.18}$$

$$\int_{\theta^{*k}}^{\rho^k(\theta^{*k})} (Q(\theta) - Q(\rho^k(\theta^{*k})))d\theta = 0. \tag{B.19}$$

By Assumption 1, these conditions require that $\theta^* < \widehat{\theta} < \rho(\theta^*)$ and $\theta^{*k} < \widehat{\theta} < \rho^k(\theta^{*k})$ and that $Q(\theta^*) > Q(\rho(\theta^*))$ and $Q(\theta^{*k}) > Q(\rho^k(\theta^{*k}))$.

Suppose by contradiction that $\theta^* \leq \theta^{*k}$ for some $k > 0$. Then given Assumption 1, conditions (B.18) and (B.19), and the fact that $\rho^k(\theta)$ is strictly increasing in θ and k, we must have

$$\theta^* \leq \theta^{*k} < \widehat{\theta} < \rho(\theta^*) < \rho^k(\theta^{*k}) \tag{B.20}$$

and

$$Q(\theta^*) \geq Q(\theta^{*k}) > Q(\rho^k(\theta^{*k})) > Q(\rho(\theta^*)). \tag{B.21}$$

Note that by the arguments in Step 4 in the proof of Proposition 4, the function

$$\int_{\theta^L}^{\theta^H} (Q(\theta) - Q(\theta^H))d\theta$$

is strictly decreasing in θ^L and in θ^H for any θ^L and θ^H satisfying $Q(\theta^L) > Q(\theta^H)$ and $\theta^H > \widehat{\theta}$. However, combined with conditions (B.20) and (B.21), this implies

$$\begin{aligned} \int_{\theta^*}^{\rho(\theta^*)} (Q(\theta) - Q(\rho(\theta^*)))d\theta &\geq \int_{\theta^{*k}}^{\rho(\theta^*)} (Q(\theta) - Q(\rho(\theta^*)))d\theta \\ &> \int_{\theta^{*k}}^{\rho^k(\theta^{*k})} (Q(\theta) - Q(\rho^k(\theta^{*k})))d\theta, \end{aligned}$$

which cannot hold simultaneously with equations (B.18) and (B.19). Therefore, it follows that $\theta^* > \theta^{*k}$ for all $k > 0$.

Part 2. We prove the result for the case of a positive penalty shift. The proof of the negative-shift case is analogous and thus omitted.

Suppose by contradiction that $\theta^{**} = \rho(\theta^*) \geq \theta^{**k} = \rho^k(\theta^{*k})$ for some $k > 0$. Since

$\theta^{*k} < \theta^*$ by Part 1, it follows by analogous reasoning as in the proof of Part 1 that

$$\begin{aligned}\int_{\theta^*}^{\rho(\theta^*)}(Q(\theta)-Q(\rho(\theta^*)))d\theta &< \int_{\theta^{*k}}^{\rho(\theta^*)}(Q(\theta)-Q(\rho(\theta^*)))d\theta \\ &\leq \int_{\theta^{*k}}^{\rho^k(\theta^{*k})}(Q(\theta)-Q(\rho^k(\theta^{*k})))d\theta.\end{aligned}$$

However, this cannot hold simultaneously with equations (B.18) and (B.19). Therefore, it follows that $\theta^{**} < \theta^{**k}$ for all $k > 0$.

B.7 Proof of Proposition 8

We prove each part of the proposition in order.

Part 1. Suppose that on-path penalties are suboptimal under $f(\theta)$. By Proposition 4, the following condition holds:

$$\int_{\theta_c}^{\overline{\theta}}(Q(\theta)-Q(\overline{\theta}))d\theta \geq 0. \tag{B.22}$$

Consider a Q-decreasing perturbation that yields $\widetilde{f}(\theta)$ over $\widetilde{\Theta} = \Theta$. Observe that the value of θ_c defined in (13) does not vary with the perturbation since $\overline{\theta} = \widetilde{\overline{\theta}}$. Suppose by contradiction that on-path penalties are optimal under $\widetilde{f}(\theta)$. By Proposition 4, this implies

$$\int_{\theta_c}^{\overline{\theta}}(\widetilde{Q}(\theta)-\widetilde{Q}(\overline{\theta}))d\theta < 0. \tag{B.23}$$

Combining (B.22) and (B.23) yields

$$\int_{\theta_c}^{\overline{\theta}}(\widetilde{Q}(\overline{\theta})-Q(\overline{\theta}))d\theta > \int_{\theta_c}^{\overline{\theta}}(\widetilde{Q}(\theta)-Q(\theta))d\theta. \tag{B.24}$$

However, since the perturbation is Q-decreasing and support-preserving, it necessarily admits

$$\widetilde{Q}(\overline{\theta})-Q(\overline{\theta}) < \widetilde{Q}(\theta)-Q(\theta)$$

for all $\theta \leq \overline{\theta}$. For $\theta \in [\underline{\theta}, \overline{\theta}]$, this inequality follows by the definition of Q-decreasing. For $\theta < \underline{\theta}$, the inequality follows from the fact that $\widetilde{Q}(\theta) = Q(\theta) = 1$ for all $\theta < \underline{\theta}$ and $Q(\overline{\theta}) \geq \widetilde{Q}(\overline{\theta})$, where the latter follows from the fact that $\widetilde{f}(\overline{\theta}) \geq f(\overline{\theta})$ in a support-preserving Q-decreasing perturbation.[31] Hence, we obtain that (B.24) cannot hold, which yields a

[31] See fn. 25.

contradiction and proves that on-path penalties are suboptimal under $\widetilde{f}(\theta)$.

Part 2. Suppose that on-path penalties are optimal under $f(\theta)$. By Proposition 4, the following condition holds:

$$\int_{\theta_c}^{\overline{\theta}} (Q(\theta) - Q(\overline{\theta}))d\theta < 0.$$

Consider a Q-increasing perturbation that yields $\widetilde{f}(\theta)$ over $\widetilde{\Theta} = \Theta$. Suppose by contradiction that on-path penalties are suboptimal under $\widetilde{f}(\theta)$. By Proposition 4, this implies

$$\int_{\theta_c}^{\overline{\theta}} (\widetilde{Q}(\theta) - \widetilde{Q}(\overline{\theta}))d\theta \geq 0.$$

Analogous arguments as in the proof of Part 1 imply that these two inequalities cannot simultaneously hold under a support-preserving, Q-increasing perturbation. We thus obtain a contradiction, which proves that on-path penalties are optimal under $\widetilde{f}(\theta)$.

B.8 Proof of Proposition 9

Denote by $\{\widetilde{\theta}^*, \widetilde{\theta}^{**}\}$ the optimal deficit limit under $\widetilde{f}(\theta)$. Observe that given the binding enforcement constraint, $\widetilde{\theta}^{**} = \rho(\widetilde{\theta}^*)$ for $\rho(\cdot)$ defined in Step 1 of the proof of Proposition 4. We prove each part of the proposition in order.

Part 1. Suppose that on-path penalties are suboptimal. By Proposition 4, the optimal deficit limits under $f(\theta)$ and $\widetilde{f}(\theta)$ set $\theta^* = \theta_c$ and $\widetilde{\theta}^* = \widetilde{\theta}_c$ respectively, where $\widetilde{\theta}_c = \theta_c$ if $\overline{\theta} = \widetilde{\overline{\theta}}$ (since θ_c and $\widetilde{\theta}_c$ are defined by (13)). To complete the proof, it is thus sufficient to prove that $\widetilde{\theta}_c$ strictly increases in $\widetilde{\overline{\theta}}$. Note that $\widetilde{\overline{\theta}} = \rho(\widetilde{\theta}_c)$, where $\rho(\cdot)$ (defined in Step 1 of the proof of Proposition 4) is strictly increasing. It thus follows that $\widetilde{\theta}_c = \rho^{-1}(\widetilde{\overline{\theta}})$ is strictly increasing in $\widetilde{\overline{\theta}}$.

Part 2. We prove the result for the case of a Q-increasing perturbation. The proof for the case of a Q-decreasing perturbation is analogous and thus omitted.

Suppose that on-path penalties are optimal. By Step 4 in the proof of Proposition 4, the following two first-order conditions uniquely define θ^* and $\widetilde{\theta}^*$:

$$\int_{\theta^*}^{\rho(\theta^*)} (Q(\theta) - Q(\rho(\theta^*)))d\theta = 0, \tag{B.25}$$

$$\int_{\widetilde{\theta}^*}^{\rho(\widetilde{\theta}^*)} (\widetilde{Q}(\theta) - \widetilde{Q}(\rho(\widetilde{\theta}^*)))d\theta = 0. \tag{B.26}$$

By Assumption 1, these conditions require that $\theta^* < \widehat{\theta} < \rho(\theta^*)$ and $\widetilde{\theta}^* < \widetilde{\widehat{\theta}} < \rho(\widetilde{\theta}^*)$, where $\widetilde{\widehat{\theta}}$ corresponds to the analog of $\widehat{\theta}$ under the perturbed distribution. Moreover, we must have that $Q(\theta^*) > Q(\rho(\theta^*))$ and $\widetilde{Q}(\widetilde{\theta}^*) > \widetilde{Q}(\rho(\widetilde{\theta}^*))$.

Suppose that $\widetilde{f}(\theta)$ is the result of a Q-increasing perturbation satisfying the conditions in the proposition. Suppose by contradiction that $\widetilde{\theta}^* \geq \theta^*$. It then follows that

$$\theta^* \leq \widetilde{\theta}^* < \widetilde{\widehat{\theta}} < \rho(\widetilde{\theta}^*) \text{ and } \widehat{\theta} < \rho(\theta^*) \leq \rho(\widetilde{\theta}^*) \tag{B.27}$$

and

$$\widetilde{Q}(\theta^*) \geq \widetilde{Q}(\widetilde{\theta}^*) > \widetilde{Q}(\rho(\widetilde{\theta}^*)), \tag{B.28}$$

where we observe that $\widetilde{Q}(\theta)$ is well defined at all $\theta \leq \widetilde{\overline{\theta}}$ and thus at θ^* and $\rho(\theta^*)$. Since the perturbation is Q-increasing, we can show that

$$\int_{\theta^*}^{\rho(\theta^*)} (Q(\theta) - Q(\rho(\theta^*)))d\theta > \int_{\theta^*}^{\rho(\theta^*)} (\widetilde{Q}(\theta) - \widetilde{Q}(\rho(\theta^*)))d\theta. \tag{B.29}$$

The inequality follows from the fact that $\widetilde{Q}(\theta) - Q(\theta) < \widetilde{Q}(\rho(\theta^*)) - Q(\rho(\theta^*))$ for all $\theta \in (\max\{\underline{\theta}, \widetilde{\underline{\theta}}\}, \rho(\theta^*))$ with $\theta^* \geq \max\{\underline{\theta}, \widetilde{\underline{\theta}}\}$. Moreover, by arguments analogous to those in the proof of Part 1 of Proposition 7, and appealing to (B.27) and (B.28), we obtain

$$\begin{aligned}\int_{\theta^*}^{\rho(\theta^*)} (\widetilde{Q}(\theta) - \widetilde{Q}(\rho(\theta^*)))d\theta &\geq \int_{\theta^*}^{\rho(\widetilde{\theta}^*)} (\widetilde{Q}(\theta) - \widetilde{Q}(\rho(\widetilde{\theta}^*)))d\theta \\ &\geq \int_{\widetilde{\theta}^*}^{\rho(\widetilde{\theta}^*)} (\widetilde{Q}(\theta) - \widetilde{Q}(\rho(\widetilde{\theta}^*)))d\theta.\end{aligned} \tag{B.30}$$

However, combining (B.29) and (B.30) yields

$$\int_{\theta^*}^{\rho(\theta^*)} (Q(\theta) - Q(\rho(\theta^*)))d\theta > \int_{\widetilde{\theta}^*}^{\rho(\widetilde{\theta}^*)} (\widetilde{Q}(\theta) - \widetilde{Q}(\rho(\widetilde{\theta}^*)))d\theta,$$

which cannot hold simultaneously with equations (B.25) and (B.26). Therefore, it follows that $\widetilde{\theta}^* < \theta^*$.