Appendix

A Derivations

A.1 Proof of Lemma 1

Use (6) to substitute for w_s in the financial sector's first-order condition and then take the derivative with respect to the transfer T_0:.

$$\frac{d^2 f(K_0, s_0)}{ds_0^2} \frac{ds_0}{dT_0} p_{solv} + w_s \frac{dp_{solv}}{dT_0} - c''(s_0) \frac{ds_0}{dT_0} = 0$$

$$\frac{ds_0}{dT_0} = -w_s \frac{dp_{solv}}{dT_0} / \left(\frac{d^2 f(K_0, s_0)}{ds_0^2} p_{solv} - c''(s_0) \right) \tag{A.1.1}$$

Since $dp_{solv}/dT_0 = p(\underline{A}_1)$, this term is positive so long as $\underline{A}_1$ is in the support of $\tilde{A}_1$ and the transfer increases the probability of solvency by decreasing the solvency threshold $\underline{A}_1$. Hence the numerator of the right hand side in the second line is negative. That the denominator is also negative follows from the concavity of f and the convexity of c. This establishes that the right side is positive and hence $ds_0/dT_0 > 0$.

A.2 A Candidate for $V(K)$ based on $f(K, s)$

Consider the frictionless counterpart to our setting, with $p_{solv} = 1$. In a dynamic setting, the expression for V would reflect the value of future production of the non-financial sector as a function of its future capital, K. For simplicity, consider one extra period of output. The case of more than one future period should be similar as it is the sum of multiple one-period output. The output of the additional period is given by $\max_s f(K, s)$. It is natural then to let

$$V(K) = \max_s f(K, s) - w_s s$$

with w_s determined by the financial sector's first-order condition. With $f(K, s) = \alpha K^{1-\vartheta} s^{\vartheta}$, this implies that

$$V(K) = (1 - \vartheta)\alpha K^{1-\vartheta} s^{*\vartheta}$$

where s^* is the optimal choice of s.

Let $c(s) = \frac{1}{m} s^m$ for $m \geq 2$. Then the first-order condition of the financial sector implies that $w_s = s^{m-1}$ and the first-order condition of the non-financial sector implies that:

$$\vartheta \alpha K^{1-\vartheta} s^{\vartheta - 1} = w_s = s^{m-1}$$

Solving for s^*, substituting into the expression above for $V(K)$, and simplifying gives:

$$s^* = (\vartheta\alpha)^{\frac{1}{m-\vartheta}} K^{\frac{1-\vartheta}{m-\vartheta}}$$
$$V(K) = (1-\vartheta)\alpha^{\frac{m}{m-\vartheta}} K^{\gamma} \quad \text{where} \quad \gamma = \frac{(1-\vartheta)}{1-\frac{\vartheta}{m}}$$

Hence, $V(K)$ has the power form that is used in the paper. Moreover,for $m \geq 2$ (which is assumed), $\gamma < 1$.

A.3 Properties of Expected Tax Revenue: $\mathcal{T}$

For the assumed parametric forms, we obtained the following results:

$$\mathcal{T} = \theta_0 \gamma^{\frac{\gamma}{1-\gamma}} (1-\theta_0)^{\frac{\gamma}{1-\gamma}}$$
$$\frac{d\mathcal{T}}{d\theta_0} = \gamma^{\frac{\gamma}{1-\gamma}}(1-\theta_0)^{\frac{\gamma}{1-\gamma}} - \theta_0 \frac{\gamma}{1-\gamma}\gamma^{\frac{\gamma}{1-\gamma}}(1-\theta_0)^{\frac{\gamma}{1-\gamma}-1} = \frac{\mathcal{T}}{\theta}\left(1 - \frac{\gamma}{1-\gamma}\frac{\theta_0}{1-\theta_0}\right)$$
$$\frac{d^2\mathcal{T}}{d\theta_0^2} = -2\frac{\gamma}{1-\gamma}\gamma^{\frac{\gamma}{1-\gamma}}(1-\theta_0)^{\frac{\gamma}{1-\gamma}-1} + \frac{\theta_0}{1-\theta_0}\left(\frac{\gamma}{1-\gamma} - 1\right)\frac{\gamma}{1-\gamma}\gamma^{\frac{\gamma}{1-\gamma}}(1-\theta_0)^{\frac{\gamma}{1-\gamma}-1}$$

The second line shows that $d\mathcal{T}/d\theta_0 > 0$ on $[0, \theta_0^{max})$ and $d\mathcal{T}/d\theta_0 < 0$ on $(\theta_0^{max}, 1)$ where θ_0^{max} solves: $\frac{\gamma}{1-\gamma}\frac{\theta_0^{max}}{1-\theta_0^{max}} = 1$. It is zero at θ^{max} and at 1 (where $\mathcal{T} = 0$).

The third line implies that $d^2\mathcal{T}/d\theta_0^2 < 0$ on $[0, \theta_0^{max}]$ so $\mathcal{T}$ is *increasing* but *concave* on this region. To see this, note that the third line can be rewritten as:

$$\frac{d^2\mathcal{T}}{d\theta_0^2} = \left(-2 + \frac{\gamma}{1-\gamma}\frac{\theta_0}{1-\theta_0} - \frac{\theta_0}{1-\theta_0}\right)\frac{\gamma}{1-\gamma}\gamma^{\frac{\gamma}{1-\gamma}}(1-\theta_0)^{\frac{\gamma}{1-\gamma}-1}$$

We know that $-1 + \frac{\gamma}{1-\gamma}\frac{\theta_0}{1-\theta_0} < 0$ on $[0, \theta_0^{max}]$ and so, on this region, the leading term in parenthesis is negative. Since the remaining terms are positive, $d^2\mathcal{T}/d\theta_0^2 < 0$ in this region.

A.4 The Government's First-Order Condition

From (3) we obtain the following first order condition of the government for the tax rate, θ_0:

$$\left[\frac{\partial f(K_0, s_0)}{\partial s_0} - c'(s_0)\right]\frac{ds_0}{dT_0}\frac{dT_0}{d\mathcal{T}}\frac{d\mathcal{T}}{d\theta_0} + [V'(K_1) - 1]\frac{dK_1}{d\theta_0} = 0 \tag{A.4.1}$$

Note that the derivatives of s_0 and $\mathcal{T}$ here are total derivatives, since the government's choices are subject to the equilibrium choices of the financial and non-financial sectors.

As shown above, $d\mathcal{T}/d\theta_0$ is positive and decreasing (towards zero), but remains positive, on $[0, \theta_0^{max}]$. Therefore, the mapping from tax level (θ_0) to the marginal rate of transformation of taxes into tax revenue ($d\mathcal{T}/d\theta_0$), is invertible on this region. A high tax rate corresponds to a low marginal rate of transformation of taxes into tax revenue and vice versa. Note that the optimal tax rate must be in the region $[0, \theta_0^{max}]$, since any further increase in θ_0 beyond θ_0^{max} reduces tax revenue and investment. Hence, we can limit the consideration of the optimal tax rate to this region. Since $d\mathcal{T}/d\theta_0$ is positive and the mapping from θ_0 to $\mathcal{T}$ is invertible in this region, we can instead consider the government's first order condition with respect to $\mathcal{T}$, which turns out to be more intuitive for analyzing the government's problem. Dividing (A.4.1) through by $d\mathcal{T}/d\theta_0$, and rewriting $(dK_1/d\theta_0)/(d\mathcal{T}/d\theta_0) = dK_1/d\mathcal{T}$ we obtain this alternative first-order condition:

$$\left[\frac{\partial f(K_0, s_0)}{\partial s_0} - c'(s_0)\right]\frac{ds_0}{dT_0} + [V'(K_1) - 1]\frac{dK_1}{d\mathcal{T}} = 0 \tag{A.4.2}$$

where the term $dT_0/d\mathcal{T}$, which equals 1 under a no-default government policy, is omitted from the expression.

A.5 Under-Investment Loss Due to Taxes

We want to obtain an expression for the second term in (8), the transfer version of the government's first-order condition:

$$\frac{[V'(K_1) - 1]\frac{dK_1}{d\theta_0}}{\frac{d\mathcal{T}}{d\theta_0}}$$

The first-order condition for investment of the non-financial sector, (7), and the parametric form for V imply that:

$$\begin{aligned} V'(K_1) - 1 &= \theta_0 V'(K_1) \\ &= \theta_0 \gamma K^{\gamma-1} \end{aligned}$$

Substituting in the parametric form also gives:

$$\frac{dK_1}{d\theta_0} = \frac{1}{1-\theta_0}\frac{1}{\gamma-1}K_1$$

Moreover, from (7) we can solve for the equilibrium K_1 as a function of θ_0:

$$K_1 = \gamma^{\frac{1}{1-\gamma}}(1-\theta_0)^{\frac{1}{1-\gamma}}$$

We can obtain the numerator to our fraction of interest by multiplying the expressions for the two terms together:

$$
\begin{aligned}
[V'(K_1)-1]\frac{dK_1}{d\theta_0} &= \frac{\theta_0\gamma}{(1-\theta_0)(\gamma-1)}K^{\gamma} \\
&= \frac{\theta_0}{1-\theta_0}\frac{\gamma}{\gamma-1}\gamma^{\frac{\gamma}{1-\gamma}}(1-\theta_0)^{\frac{\gamma}{1-\gamma}} \\
&= \frac{\mathcal{T}}{\theta_0}\frac{\theta_0}{1-\theta_0}\frac{\gamma}{\gamma-1}
\end{aligned}
$$

where the second line follows by substituting in the expression for K_0 and the third line follows by substituting in the expression for $\mathcal{T}$. Appendix A.3 derives $d\mathcal{T}/d\theta_0$. Dividing the expression for the numerator by the expression for $d\mathcal{T}/d\theta_0$ shows that the marginal loss per transfer is given by:

$$
\frac{d\mathcal{L}}{d\mathcal{T}} = \frac{[V'(K_1)-1]\frac{dK_1}{d\theta_0}}{\frac{d\mathcal{T}}{d\theta_0}} = \frac{-\frac{\theta_0}{1-\theta_0}\frac{\gamma}{1-\gamma}}{1-\frac{\theta_0}{1-\theta_0}\frac{\gamma}{1-\gamma}}
$$

From this it is clear that $d\mathcal{L}/d\mathcal{T} \to -\infty$ as $\theta_0 \to \theta^{max}$ (since at θ^{max} the denominator is 0). Additionally, we have:

$$
\frac{d^2\mathcal{L}}{d\mathcal{T}^2} = \frac{d^2\mathcal{L}}{d\theta_0 d\mathcal{T}}\frac{d\theta_0}{d\mathcal{T}} < 0 \qquad \text{for} \quad \theta_0 \in [0, \theta^{max}) \quad .
$$

Hence, the marginal loss to the economy is increasing in magnitude (getting worse) as the tax rate increases up to θ^{max} and expected tax revenue rises to $\mathcal{T}^{max}$. In other words, marginal tax revenues becomes increasingly expensive to raise as the marginal loss to the economy from underinvestment rises in the tax rate/level of tax revenues.

A.6 Proof of Proposition 1

Substituting (6) into (5) and solving, we obtain the equilibrium level of s_0 (note that we refer to the *equilibrium* level of s_0 also as s_0, an abuse of notation intended to reduce clutter):

$$
s_0 = \left(\frac{\vartheta\alpha}{\beta}\right)^{\frac{1}{m-\vartheta}} K_0^{\frac{1-\vartheta}{m-\vartheta}} p_{solv}^{\frac{1}{m-\vartheta}}
$$

Now substitute this into the expression for $d\mathcal{G}/d\mathcal{T}$ to get:

$$
\frac{d\mathcal{G}}{d\mathcal{T}} = \frac{\partial f(K_0, s_0)}{\partial s}(1-p_{solv})\frac{ds_0}{dT_0} = \frac{1}{m-\vartheta}\left(\vartheta\alpha K_0^{1-\vartheta}\right)^{\frac{m}{m-\vartheta}}\beta^{\frac{-\vartheta}{m-\vartheta}}p_{solv}^{\frac{\vartheta}{m-\vartheta}-1}(1-p_{solv})\frac{dp_{solv}}{dT_0}
$$

Taking derivative again with respect to $\mathcal{T}$ shows that:

$$\begin{aligned}\frac{d^2\mathcal{G}}{d\mathcal{T}^2} \propto \quad & \left(\frac{\vartheta}{m-\vartheta}-1\right) p_{solv}^{\frac{\vartheta}{m-\vartheta}-2}(1-p_{solv})\frac{dp_{solv}}{dT_0} \\ & - p_{solv}^{\frac{\vartheta}{m-\vartheta}-1}\left(\frac{dp_{solv}}{dT_0}\right)^2 \quad + p_{solv}^{\frac{\vartheta}{m-\vartheta}-1}(1-p_{solv})\frac{d^2p_{solv}}{dT_0^2}\end{aligned}$$

where $dT_0/d\mathcal{T} = 1$ is omitted. Since the second term in the above expression is always negative, a sufficient condition to ensure that $d^2\mathcal{G}/d\mathcal{T}^2 < 0$ is to ensure that the first and third terms in the above expression are non-positive. The condition: $m - 2\vartheta \geq 0$ ensures that the first term is non-positive. The third term is negative if the slope of the probability density of $\tilde{A}_1$ at $\underline{A}_1$ is non-positive. Letting $\tilde{A}_1$ take a uniform distribution sets this term to zero.[23]

Since we have shown that both $\mathcal{G}$ and $\mathcal{L}$ are concave in $\mathcal{T}$, the government's problem is concave in $\mathcal{T}$. Furthermore, the optimum tax revenue, $\hat{\mathcal{T}}$, must correspond to a tax rate $\hat{\theta} < \theta^{max}$, because the first-order condition is *negative* at θ^{max}. To see that this is the case, note that $d\mathcal{L}/d\mathcal{T} \to \infty$ as $\theta \to \theta^{max}$ while $d\mathcal{G}/d\mathcal{T}$ is finite for $p_{solv} > 0$.

A.6.1 Impact of L_1 and N_D on $\mathcal{T}$

Let $x = L_1$ or N_D. Rewriting (8) using the gain and loss notation as $d\mathcal{G}/d\mathcal{T} + d\mathcal{L}/d\mathcal{T} = 0$ and then taking the derivative with respect to x gives:

$$\frac{d^2\mathcal{G}}{dxd\mathcal{T}} + \frac{d^2\mathcal{L}}{dxd\mathcal{T}} = 0 \tag{A.6.1}$$

Using the Implicit Function Theorem, the two terms on the right side evaluate to the following:

$$\begin{aligned}\frac{d^2\mathcal{G}}{dxd\mathcal{T}} &= \frac{d}{dp_{solv}}\left(\frac{d\mathcal{G}}{d\mathcal{T}}\right)\left\{\frac{\partial p_{solv}}{\partial T_0}\left(\frac{\partial T_0}{\partial \mathcal{T}}\frac{d\mathcal{T}}{dx} + \frac{\partial T_0}{\partial x}\right) + \frac{\partial p_{solv}}{\partial x}\right\} \\ \frac{d^2\mathcal{L}}{dxd\mathcal{T}} &= \frac{d^2\mathcal{L}}{d\mathcal{T}^2}\frac{d\mathcal{T}}{dx}\end{aligned}$$

Substituting into (A.6.1) and combining the terms multiplying $d\mathcal{T}/dx$ yields:

$$\frac{d\mathcal{T}}{dx}\left[\frac{d}{dp_{solv}}\left(\frac{d\mathcal{G}}{d\mathcal{T}}\right)\frac{\partial p_{solv}}{\partial T_0}\frac{\partial T_0}{\partial \mathcal{T}} + \frac{d^2\mathcal{L}}{d\mathcal{T}^2}\right] = -\frac{d}{dp_{solv}}\left(\frac{d\mathcal{G}}{d\mathcal{T}}\right)\left\{\frac{\partial p_{solv}}{\partial T_0}\frac{\partial T_0}{\partial x} + \frac{\partial p_{solv}}{\partial x}\right\} \tag{A.6.2}$$

[23] Using an exponential distribution would also be sufficient. For the log-normal distribution, this term will be negative for a range of values below a cutoff.

Note for the left-hand side term in parenthesis:

$$\frac{d}{dp_{solv}}\left(\frac{d\mathcal{G}}{d\mathcal{T}}\right)\frac{\partial p_{solv}}{\partial T_0}\frac{\partial T_0}{\partial \mathcal{T}}+\frac{d^2\mathcal{L}}{d\mathcal{T}^2}= \quad \frac{d^2\mathcal{G}}{d\mathcal{T}^2}+\frac{d^2\mathcal{L}}{d\mathcal{T}^2} \quad <0$$

For $x = N_D$:

$$\frac{\partial p_{solv}}{\partial T_0}\frac{\partial T_0}{\partial x}+\frac{\partial p_{solv}}{\partial x}=\frac{\partial p_{solv}}{\partial T_0}(k_A-1)<0$$

since $\partial T_0/\partial N_D = -1$ and $\partial p_{solv}/\partial N_D = (\partial p_{solv}/\partial T_0)k_A$.
For $x = L_1$:

$$\frac{\partial p_{solv}}{\partial T_0}\frac{\partial T_0}{\partial x}=0 \quad \text{and} \quad \frac{\partial p_{solv}}{\partial x}<0$$

so for either value of x, the term in braces on the right side is negative. Finally, the intermediate steps in the proof of the concavity of G in $\mathcal{T}$ show that

$$\frac{d}{dp_{solv}}\left(\frac{d\mathcal{G}}{d\mathcal{T}}\right)<0$$

Combining these results shows that $d\mathcal{T}/dx > 0$ for $x = L_1$ or N_D.

A.6.2 Impact of N_D on T_0

To show how T_0 changes with N_D, begin by using the result above for $\mathcal{T}$. In particular, letting $x = N_D$ in (A.6.2) and simplifying the right-side expression using $\frac{\partial p_{solv}}{\partial T_0}\frac{\partial T_0}{\partial x}+\frac{\partial p_{solv}}{\partial x}=\frac{\partial p_{solv}}{\partial T_0}(k_A-1)$ and $d^2\mathcal{G}/(dT_0 d\mathcal{T}) = d^2\mathcal{G}/d\mathcal{T}^2$ gives:

$$\frac{d\mathcal{T}}{dN_D}\left[\frac{d^2\mathcal{G}}{d\mathcal{T}^2}+\frac{d^2\mathcal{L}}{d\mathcal{T}^2}\right]=(1-k_A)\frac{d^2\mathcal{G}}{d\mathcal{T}^2}$$

$$\frac{d\mathcal{T}}{dN_D}=\frac{(1-k_A)\frac{d^2G}{d\mathcal{T}^2}}{\frac{d^2\mathcal{G}}{d\mathcal{T}^2}+\frac{d^2\mathcal{L}}{d\mathcal{T}^2}} \quad \Rightarrow \quad 0<\frac{d\mathcal{T}}{dN_D}<1-k_A$$

Since $T_0 = \mathcal{T} - N_D$,

$$\frac{dT_0}{dN_D}=\frac{d\mathcal{T}}{dN_D}-1 \quad \Rightarrow \quad -1<\frac{dT_0}{dN_d}<-k_A$$

Moreover, this shows that $T_0 + k_A N_D$, the *gross* transfer to the financial sector, is *decreasing* in N_D.

A.6.3 Impact of Factor Share on $\mathcal{T}$

Next we examine the effect of the factor share of financial services on $\mathcal{T}$, while holding constant total output. To that end, we consider the impact of a change in ϑ while simultaneously adjusting α (the level of productivity) to keep output constant. Let $D(\cdot)$ be the following differential with respect to $d\vartheta$ and $d\alpha$

$$Dg = \frac{dg}{d\vartheta}d\vartheta + \frac{dg}{d\alpha}d\alpha$$

where the derivatives are taken holding $\mathcal{T}$ constant but include the change caused by $ds_0/d\vartheta$ and $ds_0/d\alpha$. Now let $d\alpha$ be set to keep total output constant, e.g., $Df = 0$, where f is equilibrium output. This implies $d\alpha = -(df/d\vartheta)/(df/d\alpha)d\vartheta$, which gives:

$$\frac{Dg}{d\vartheta} = \frac{dg}{d\vartheta} - \frac{dg}{d\alpha}\left(\frac{df/d\vartheta}{df/d\alpha}\right)$$

Hence, to find the impact of ϑ on $\mathcal{T}$ while holding output constant, we analyze $D\mathcal{T}/d\vartheta$. Applying this differentiation operator to the first-order condition for $\mathcal{T}$ and collecting terms gives:

$$\left(\frac{d^2\mathcal{G}}{d\mathcal{T}^2} + \frac{d^2\mathcal{L}}{d\mathcal{T}^2}\right)\frac{D\mathcal{T}}{d\vartheta} + \frac{D}{d\vartheta}\frac{dG}{d\mathcal{T}} + \frac{D}{d\vartheta}\frac{d\mathcal{L}}{d\mathcal{T}} = 0 \tag{A.6.3}$$

Note that the application of the D operator is linear as it is simply a sum of two derivatives. Furthermore,

$$\frac{D}{d\vartheta}\left(\frac{d\mathcal{G}}{d\mathcal{T}}\right) > 0$$
$$\frac{D}{d\vartheta}\left(\frac{d\mathcal{L}}{d\mathcal{T}}\right) = 0$$

The first line is proved below, while the second line follows directly since $d\mathcal{L}/d\mathcal{T}$ is not a function of ϑ or α. Using the second-order condition, it follows that $D\mathcal{T}/d\vartheta > 0$.

To prove the first line from above, note that the sign of this term in question is equal to the sign of $D(\partial f/\partial s_0 \times ds_0/dT_0)/d\vartheta$. This follows from the expression for $d\mathcal{G}/d\mathcal{T}$ and that p_{solv} does not depend on ϑ or α. Substituting (6) into (5) and using the functional form of $c(s_0)$ shows that

$$\text{sgn}\left(\frac{D}{d\vartheta}\frac{\partial f}{\partial s_0}\right) = \text{sgn}\left((m-1)s_0^{m-2}\frac{Ds_0}{d\vartheta}\right)$$

Since $m > 1$, this last term equals $\text{sgn}(Ds_0/d\vartheta)$. To find $\text{sgn}(Ds_0/d\vartheta)$, substitute (6) into (5), multiply both sides of the resulting expression by s_0, and substitute in the functional forms of f and $c(s_0)$ to obtain:

$$\vartheta f(K_0, s_0) p_{solv} = \beta s_0^m \quad .$$

Applying the D operator to both sides of this expression gives:

$$\begin{aligned}\frac{D(\vartheta f(K_0, s_0) p_{solv})}{d\vartheta} &= f(K_0, s_0) p_{solv} \\ \frac{D(\beta s_0^m)}{d\vartheta} &= m s_0^{m-1} \frac{D s_0}{d\vartheta}\end{aligned}$$

Since the right-hand side of the first line is positive, so must be the right-hand side of the second line, showing that $Ds_0/d\vartheta > 0$ and hence, $\text{sgn}\,(D(\partial f/\partial s_0)/d\vartheta) > 0$.

It remains to find $\text{sgn}\,(D(ds_0/dT_0)/d\vartheta)$, which can be found using (A.1.1). Using similar steps to those immediately above, it can be shown that if $m \leq 2$ then $\text{sgn}\,(D(d^2 f/s_0^2)/d\vartheta) \geq 0$. Moreover, direct differentiation and $Ds_0/d\vartheta > 0$ show that if $m \leq 2$ then $\text{sgn}(c''(s_0)) < 0$. It is then straightforward to show that $\text{sgn}\,(D(ds_0/dT_0)/d\vartheta) > 0$.

A.7 Proof of Lemma 4

The derivative of the government's objective with respect to N_T is given by:

$$\frac{d\mathcal{G}}{dT_0} \frac{dT_0}{dN_T}$$

When $N_T + N_D \geq \mathcal{T}$ (Region 2), then $T_0 = N_T P_0 = \frac{N_T}{N_T + N_D} \mathcal{T}$ and

$$\frac{dT_0}{dN_T} = P_0 + N_T \frac{dP_0}{dN_T} = P_0 \left(\frac{N_D}{N_T + N_D} \right) \quad .$$

Therefore $dT_0/dN_T > 0$ if $N_D > 0$. Moreover, this implies that $N_T \to \infty$ is optimal in the default region. Alternatively, if $N_D = 0$, then increasing N_T into the default region provides no benefit but does incur the loss of D.

When $N_T \to \infty$, then $T_0 = \mathcal{T}$, as pre-existing bondholders are completely diluted. Note that $T_0 = \mathcal{T}$ is the same situation as if N_D were set to 0. Conditional on this, the government's problem reduces to the same problem it faces in Region 1. Therefore, to determine if default is optimal, the government needs to compare this optimum-cum-default-loss, $W_1 \big| N_D = 0 - D$, with the maximum from region 1, W_1. Since the optimum within the default region can be found by setting $N_D = 0$, Appendix A.6.2 shows that the transfer will be bigger conditional on default. By Appendix A.1 this implies the equilibrium provision of financial services is greater.

A.8 Proof of Proposition 2

As Lemma 4 indicates, the tradeoff involved in default is the deadweight cost D, versus the larger transfer and reduced taxes made possible by diluting pre-existing debt. The net benefit of this tradeoff can be written as follows:

$$\int_{\hat{T_0}^{no\_def}}^{\hat{T_0}^{def}-kN_D} \frac{d\,\mathcal{G}}{dT_0} dT_0 + \int_{\hat{\mathcal{T}}^{no\_def}}^{\hat{\mathcal{T}}^{def}} \frac{d\mathcal{L}}{d\mathcal{T}} d\mathcal{T} - D \tag{A.8.1}$$

where the first integral is the gain due to increasing the (gross) transfer, while the second integral is the reduction in underinvestment loss due to reducing tax revenue. Note that $d\mathcal{G}/dT_0$ here is evaluated at the no-default values. If (A.8.1) is positive, it is optimal for the sovereign to choose default, while if it is negative then no-default is optimal.

To prove point (1), take the derivative of (A.8.1) with respect to L_1 and simplify the resulting expression to obtain:

$$\int_{\hat{T_0}^{no\_def}}^{\hat{T_0}^{def}-k_A N_D} \frac{d}{dL_1}\left(\frac{d\mathcal{G}}{dT_0}\right) > 0$$

The intermediate steps in Appendix A.4 show that the derivative in the integrand is positive. As shown in Appendix A.6.2, the *gross* transfer is decreasing in N_D, so $T_0^{def} > k_A N_D + T_0^{no\_def}$ and hence the integral is positive.

To prove the statement for N_D, take the derivative of (A.8.1) with respect to N_D. Simplifying the derivative at the upper integration boundary gives $-k_A d\mathcal{G}/dT_0\big|_{\hat{T_0}^{def}-k_A N_D}$ while from the lower boundary we get we get $d\mathcal{G}/dT_0\big|_{\hat{T_0}^{no\_def}}$. The remaining part of the derivative is:

$$\begin{aligned}\int_{\hat{T_0}^{no\_def}}^{\hat{T_0}^{def}-k_A N_D} \frac{d}{dN_D}\left(\frac{d\mathcal{G}}{dT_0}\right) &= k_A \int_{\hat{T_0}^{no\_def}}^{\hat{T_0}^{def}-k_A N_D} \frac{d}{dT_0}\left(\frac{d\mathcal{G}}{dT_0}\right)\\ &= k_A\left(\frac{d\mathcal{G}}{dT_0}\bigg|_{\hat{T_0}^{def}-K_A N_D} - \frac{d\mathcal{G}}{dT_0}\bigg|_{\hat{T_0}^{no\_def}}\right)\end{aligned}$$

Combining the three parts of the derivatives gives: $(1-k_A)d\mathcal{G}/dT_0\big|_{\hat{T_0}^{no\_def}} > 0$. To show that the benefit of defaulting is convex in N_D, take a second derivative to obtain: $(1-k_A)d^2\mathcal{G}/dT_0^2\big|_{\hat{T_0}^{no\_def}} dT_0^{no\_def}/dN_D > 0$.

To prove the statement for factor share, apply the operator $D/d\vartheta$ (defined in Appendix A.6.3) to (A.8.1) and again simplify to get:

$$\int_{\hat{T_0}^{no\_def}}^{\hat{T_0}^{def}-k_A N_D} \frac{D}{d\vartheta}\left(\frac{d\mathcal{G}}{dT_0}\right) > 0$$

The integrand is positive as shown in Appendix A.6.3, so again the integral is positive.

Finally, taking the derivative with respect to k, we obtain $-(d\mathcal{G}/dT_0)N_D < 0$ at the upper integration boundary and 0 at the lower boundary. In the interior we obtain

$$\int_{\hat{T_0}^{no\_def}}^{\hat{T_0}^{def}-k_A N_D} \frac{d}{dk_A}\left(\frac{d\mathcal{G}}{dT_0}\right) = N_D \int_{\hat{T_0}^{no\_def}}^{\hat{T_0}^{def}-k_A N_D} \frac{d}{dT_0}\left(\frac{d\mathcal{G}}{dT_0}\right) < 0$$

so the derivative is negative.

A.9 Optimal Tax Revenue Under Uncertainty

The first order condition for the government's choice of $\mathcal{T}$ is given by:

$$\frac{d\mathcal{G}}{dT_0}\frac{dT_0}{d\mathcal{T}} + \frac{d\mathcal{L}}{d\mathcal{T}} = 0$$

Whereas under certainty $dT_0/d\mathcal{T}$=1, this is no longer the case. Taking the derivative of T_0 with respect to $\mathcal{T}$ in (11) (while holding H constant) and then using (9) to substitute into the resulting expression gives $dT_0/d\mathcal{T} = P_0H$. Therefore, the first-order condition for $\mathcal{T}$ is:

$$\frac{d\mathcal{G}}{\partial T_0}HP_0 + \frac{d\mathcal{L}}{d\mathcal{T}} = 0 \tag{A.9.1}$$

with T_0 given in (11). The loss due to underinvestment, $\mathcal{L}$, is the same as under certainty. Recall that it is concave, with the magnitude of the marginal loss, $d\mathcal{L}/d\mathcal{T}$, increasing in $\mathcal{T}$. Similarly, $d\mathcal{G}/dT_0$, the gain to the economy from the increased provision of financial services, remains the same with uncertainty and is decreasing in T_0. However, the rate at which T_0 increases in $\mathcal{T}$ is now HP_0 rather than 1. Note that this rate is a constant in $\mathcal{T}$, as P_0 is only a function of H, and is less than 1.[24] Finally, the second-order condition for $\mathcal{T}$ holds

$$\frac{d^2\mathcal{G}}{\partial T_0^2}(HP_0)^2 + \frac{d^2\mathcal{L}}{d\mathcal{T}^2} < 0$$

as $\mathcal{G}$ and $\mathcal{L}$ are concave and HP_0 is a function only of H.

A.10 Optimal Probability of Default Under Uncertainty

Changing H affects two components of the government's objective. As can be seen from (11), increasing H changes T_0. Unlike the case with $\mathcal{T}$, however, increasing H does not have any effect on investment. Instead, the cost associated with increasing H is that it increases the probability of default, and so also the expected deadweight cost. The first-order condition

[24]To see this, note that $HP_0 = E_0\left[\min\left(H, \tilde{R}_V\right)\right] < E_0[\tilde{R}_V] = 1$.

for H shows this tradeoff:

$$\frac{d\mathcal{G}}{dT_0}\frac{dT_0}{dH} - D\frac{dp_{def}}{dH} = 0 \tag{A.10.1}$$

From (10), it is clear that $dp_{def}/dH > 0$ and we can think of choosing H exactly as choosing the probability of default. The effect on $T_0 = P_0 N_T$ is less immediately clear, since increasing H increases N_T, but decreases P_0. However, (11) shows that $dT_0/dH > 0$. To see this we break up T_0 into two terms based on (11) and consider their derivatives:

$$d\,(\mathcal{T} - \frac{N_D}{H})/dH = \frac{N_D}{H^2} > 0 \tag{A.10.2}$$

$$d\,E_0\left[\min\left(H, \tilde{R}_V\right)\right]/dH = (1 - p_{def}) > 0 \tag{A.10.3}$$

Demonstrating the equivalence in the second line is straightforward, as shown in Appendix A.11. We refer to (A.10.2) as increasing the *dilution* of existing bondholders' claim, since the increase in H reduces the share of tax revenues that goes to the holders of the existing debt, N_D. We refer to (A.10.3) as reducing either the *default buffer* or *precautionary taxation*, since by increasing H, it increases the probability that $\tilde{R}_V < H$, in which case the government defaults. Hence, (A.10.2) and (A.10.3) show that increasing H (while holding $\mathcal{T}$ constant) increases T_0. It immediately follows that $d\mathcal{G}/dH > 0$ and there is a benefit to increasing H. Substituting in for dT_0/dH, the first-order condition becomes:

$$\frac{d\mathcal{G}}{dT_0}\left(\frac{N_D}{H^2}E_0\left[\min\left(H, \tilde{R}_V\right)\right] + (\mathcal{T} - \frac{N_D}{H})(1 - p_{def})\right) - D\frac{dp_{def}}{dH} = 0$$

Appendix A.11 also shows that as H increases, raising it further becomes decreasingly effective at increasing T_0:

$$\frac{d^2T_0}{dH^2} = \frac{-2N_D}{H^3}\int_0^H x p_{\tilde{R}_V}(x)dx - (\mathcal{T} - \frac{N_D}{H})p_{\tilde{R}_V}(H) < 0$$

where $p_{\tilde{R}_V}(x)$ denotes the probability density of $\tilde{R}_V$ evaluated at x. In other words, T_0 is concave in H. Together with the concavity of $\mathcal{G}$ in T_0, this implies that $\mathcal{G}$ is concave in H, e.g., $d^2\mathcal{G}/dH^2$.[25] The implication is that while increasing H provides a benefit to the government by increasing the transfer through dilution and reduction of precautionary taxation, the marginal benefit is decreasing. Meanwhile, the government bears a cost for increasing H; the resulting increased likelihood of default increases the expected deadweight cost of default.

[25]Note that in the first-order conditions, we have assumed that the government takes into account the (negative) impact of higher H on prices. Thus, we have *NOT* treated the government here as a price-taker. If we instead treat the government as a price-taker, the resulting conditions are simpler: $dT_0/dH = P_0\mathcal{T}$ (as dP_0/dH is omitted due to the price-taking assumption) and the first-order condition is: $d\mathcal{G}/dT_0(P_0\mathcal{T}) - Dd\,p_{def}/dH = 0$. In this case, concavity of $\mathcal{G}$ in H still holds because $\mathcal{G}$ is concave in T_0.

We assume that at the optimal choice of H, $d^2\, p_{def}/d^2 H \geq 0$.

A.11 Uncertainty Calculations

To derive $d\, E_0 \left[\min\left(H, \tilde{R}_V\right)\right]/dH$, rewrite the expectation as:

$$E_0 \left[\min\left(H, \tilde{R}_V\right)\right] = \int_0^H x\, p_{\tilde{R}_V}(x)dx + H \int_H^\infty p_{\tilde{R}_V}(x)dx$$

Now taking the derivative with respect to H, one obtains:

$$\begin{aligned} d\, E_0 \left[\min\left(H, \tilde{R}_V\right)\right]/dH =& H p_{\tilde{R}_V}(H) - H p_{\tilde{R}_V}(H) + \int_H^\infty p_{\tilde{R}_V}(x)dx \\ =& \int_H^\infty p_{\tilde{R}_V}(x)dx \\ =& (1 - p_{def}) \end{aligned}$$

The first line is just the derivative, while the last line follows by definition of p_{def}.

Using this result we have that:

$$\frac{dT_0}{dH} = \frac{N_D}{H^2} E_0 \left[\min\left(H, \tilde{R}_V\right)\right] + \left(\mathcal{T} - \frac{N_D}{H}\right)(1 - p_{def})$$

Substituting in the expression above for $E_0 \left[\min\left(H, \tilde{R}_V\right)\right]$, taking the derivative with respect to T_0, and simplifying gives:

$$\begin{aligned} \frac{d^2T_0}{dH^2} =& \frac{-2N_D}{H^3} \left[\int_0^H x\, p_{\tilde{R}_V}(x)dx + H \int_H^\infty p_{\tilde{R}_V}(x)dx\right] + \frac{N_D}{H^2}(1 - p_{def}) \\ &+ \frac{N_D}{H^2}(1 - p_{def}) - \left(\mathcal{T} - \frac{N_D}{H}\right) p_{\tilde{R}_V}(H) \\ =& \frac{-2N_D}{H^3} \left[\int_0^H x\, p_{\tilde{R}_V}(x)dx\right] - \left(\mathcal{T} - \frac{N_D}{H}\right) p_{\tilde{R}_V}(H) \end{aligned}$$

Since $(\mathcal{T} - N_D/H) = N_T/H > 0$, it is clear that $d^2T_0/dH^2 < 0$.

A.12 Proof of Proposition 3

The starting point are the first-order conditions for $\mathcal{T}$ and for H, given by (A.9.1) and (A.10.1), respectively. Substituting out $\frac{d\mathcal{G}}{dT_0}$ and rearranging gives the relation

$$-\frac{d\mathcal{L}}{d\mathcal{T}}\frac{dT_0}{dH} = HP_0D\frac{dp_{def}}{dH} = 0 \tag{A.12.1}$$

Differentiating with respect to L_1 gives on the left-hand side:

$$-\frac{d^2\mathcal{L}}{d\mathcal{T}^2}\frac{d\mathcal{T}}{dL_1}\frac{dT_0}{dH} - \frac{d\mathcal{L}}{d\mathcal{T}}\frac{d^2T_0}{d\mathcal{T}dH}\frac{d\mathcal{T}}{dL_1} - \frac{d\mathcal{L}}{d\mathcal{T}}\frac{d^2T_0}{dH^2}\frac{dH}{dL_1}$$

and on the right-hand side:

$$(1-p_{def})D\frac{dp_{def}}{dH}\frac{dH}{dL_1} + HP_0D\frac{d^2p_{def}}{dH^2}\frac{dH}{dL_1}$$

Combining the terms in $\frac{d\mathcal{T}}{dL_1}$ gives:

$$\frac{d^2\mathcal{L}}{d\mathcal{T}^2}\frac{dT_0}{dH} - \frac{d\mathcal{L}}{d\mathcal{T}}\frac{d^2T_0}{d\mathcal{T}dH}$$

and it is not difficult to see that each term has a positive sign. Combining the terms in $\frac{dH}{dL_1}$ gives:

$$\frac{d\mathcal{L}}{d\mathcal{T}}\frac{d^2T_0}{dH^2} + (1-p_{def})D\frac{dp_{def}}{dH} + HP_0D\frac{d^2p_{def}}{dH^2}$$

and again each term is positive. Thus, we see that at the optimal values, $\operatorname{sgn}\left(\frac{d\mathcal{T}}{dL_1}\right) = \operatorname{sgn}\left(\frac{dH}{dL_1}\right)$. It remains to show that both of these signs are indeed *positive*.

To that end, let V represent the objective function of the government with the first-order conditions given by (A.9.1) and (A.10.1). Let $X = [\mathcal{T}, H]$ be the vector of the two controls. Then the first order conditions can be written as just $dV/dX = 0$. Differentiating this with respect to L_1 then gives

$$\frac{dV}{dL_1dX} + \frac{d^2V}{dX^2}\frac{dX}{dL_1} = 0 \quad .$$

By assumption, the optimal X is internal and so d^2V/dX^2 is negative definite. Isolating dX/dL_1 then gives

$$\frac{dX}{dL_1} = -\left(\frac{d^2V}{dX^2}\right)^{-1}\frac{dV}{dL_1dX} \quad .$$

Premultiplying by $\frac{dV^T}{dL_1 dX}$ we obtain

$$\frac{dV^T}{dL_1 dX}\frac{dX}{dL_1} = -\frac{dV^T}{dL_1 dX}\left(\frac{d^2V}{dX^2}\right)^{-1}\frac{dV}{dL_1 dX} > 0$$

where the sign follows since the Hessian is negative definite. Since

$$\frac{d^2\mathcal{G}}{dL_1 d\mathcal{T}} > 0$$

it is straightforward to see that $\frac{dV}{dL_1 dX} > 0$, i.e., both terms in the vector are positive. Hence, we must have that $dX/dL_1 > 0$ as well since both terms in this vector are of the same sign. Similar steps prove the result for ϑ.

A.13 Proposition 4

Below we derive the return on financial sector equity, debt, and the sovereign bond. A complication created by the guarantee is that the number of outstanding sovereign bonds is state contingent, since it depends on the realization of $\tilde{A}_1$. Let $N_G(\tilde{A}_1)$ denote the number of new bonds issued towards the guarantee. This means there will also be a different price for sovereign bonds contingent on the realization of $\tilde{A}_1$. Hence, P_0 will now depend on $\tilde{A}_1$, as will T_0. This state-contingency is implicit below but will be omitted to avoid excessive notation.

Assume that $\tilde{A}_1 \sim U[A_{min}, A_{max}]$ and consider two types of shocks. The first is a shock to the value of the risky asset held by the financial sector. This shock changes the mean of $\tilde{A}_1$ by shifting the support of $\tilde{A}_1$ by an amount dA. Thus, $\tilde{A}_1$ remains uniformly distributed with the same dispersion, but a different mean. The second shock affects the sovereign bond price by changing the expected growth rate of future output by dR. For $\tilde{R}_V$ uniformly distributed this corresponds to a dR shift in its support.

From the model we have that the value of financial sector equity is given by

$$E = \int_{\underline{A}_1}^{A_{max}} (\tilde{A}_1 + T_0 - L_1)p(\tilde{A}_1)d\tilde{A}_1$$

where $p(\tilde{A}_1)$ is the uniform probability density. Calculating the change in E induced by a shock dA gives

$$\frac{d\,E}{d\,A} = p_{solv} + \frac{T_0(A_{max}) - T_0(\underline{A}_1)}{A_{max} - A_{min}} = p_{solv} \quad .$$

The second equality follows by the fact that there is no change in the guarantee once $\tilde{A}_1 > \underline{A}_1$ because at this point the financial sector is solvent. Calculating the change in E due to a shock dR gives

$$\frac{d\,E}{d\,R} = \frac{dP_0(\underline{A}_1)}{dR} N_T p_{solv}$$

Note that since there is no change in the guarantee for $\tilde{A}_1 > \underline{A}_1$, the quantity dP_0/dR is the same for any $\tilde{A}_1 > \underline{A}_1$.

Next, we have that the value of financial sector debt is given by

$$D = \int_{\underline{A}_1}^{A_{max}} L_1 p(\tilde{A}_1) d\tilde{A}_1 + \int_{A_{min}}^{\underline{A}_1} (\tilde{A}_1 + T_0) p(\tilde{A}_1) d\tilde{A}_1 + \int_{A_{min}}^{\underline{A}_1} (L_1 - \tilde{A}_1 - T_0) P_0 p(\tilde{A}_1) d\tilde{A}_1$$

The last term gives the value of the guarantee. Differentiating, simplifying, and combining terms gives that the change in D induced by a shock dA is

$$\frac{d\,D}{d\,A} = (1 - p_{solv})(1 - P_0(A_{min})) + \frac{T_0(\underline{A}) - T_0(A_{min})}{A_{max} - A_{min}}(1 - P_0(A_{min}))$$

The change in D due to a shock dR is given by

$$\frac{d\,D}{d\,R} = \int_{A_{min}}^{\underline{A}_1} \frac{dP_0}{dR} N_T (1 - P_0) p(\tilde{A}_1) d\tilde{A}_1 + \int_{A_{min}}^{\underline{A}_1} (L_1 - \tilde{A}_1 - T_0) \frac{dP_0}{dR} p(\tilde{A}_1) d\tilde{A}_1$$

The second term represents the change in value of the existing guarantee due to the change in the sovereign bond price. The first term incorporates both the change in the value of the existing transfer plus the change in the 'amount' of guarantee. That is, if dR is positive, the transfer increases in value by dT_0/dR, but this reduces the amount of guarantee given by the government for each realization by that same amount. This is true for each realization of $\tilde{A}_1$ under the integral sign.

We now approximate these values by by ignoring the state-dependence of P_0 on $\tilde{A}_1$ in the above expressions. This simplifies them to:

$$\frac{dE}{dA} = p_{solv}$$
$$\frac{dE}{dR} \approx \frac{dP_0}{dR} N_T p_{solv}$$

and

$$\frac{dD}{dA} \approx (1 - p_{solv})(1 - P_0)$$
$$\frac{dD}{dR} \approx \frac{dP_0}{dR} N_T (1 - p_{solv})(1 - P_0) + \frac{1}{2}\frac{dP_0}{dR}(1 - p_{solv})(\underline{A}_1 - A_{min})$$

By inspection one can then see that the following relation holds for these approximations:

$$dD \approx \frac{1 - p_{solv}}{p_{solv}}(1 - P_0)\, dE + \frac{1}{2}(1 - p_{solv})(\underline{A}_1 - A_{min})\, dP_0$$

Simple algebra and a substitution then give (12),

$$\frac{d\,D}{D} \approx \frac{(1 - p_{solv})(1 - P_0)}{p_{solv}} \frac{E}{D} \frac{dE}{E} + \frac{(1 - p_{solv})^2 (A_{max} - A_{min})}{2} \frac{P_0}{D} \frac{dP_0}{P_0} \quad .$$