A. Proofs

A.1. Proof of Generalized Proposition 1

In the following, we present the proof of a generalized version of Proposition 1 with different switching costs (δ_{01} and δ_{10}). Proposition 1 as presented in the main body can be obtained by equating the two switching costs ($\delta_{01}=\delta_{10}=\delta$).

Let c be the coupon rate. We first assume that the equity holders may pursue temporary suspension of production in a non-empty set of states (i.e., the product price). The equity values in the operating and idle modes follow:

$$V^f(s;c)=\begin{cases} V^{f,1}=-(K+c)\frac{\gamma}{r}+\frac{\gamma}{q}s+\theta s^{\alpha_2} & \text{for } s>s^f_{10}\\ V^{f,0}=-c\frac{\gamma}{r}+\phi_1 s^{\alpha_1}+\phi_2 s^{\alpha_2} & \text{for } s^f_{0d}\le s\le s^f_{01},\end{cases}$$

where $\alpha_1=\frac{1}{2}-\frac{(r-q)}{\sigma^2}+\sqrt{\left(\frac{1}{2}-\frac{r-q}{\sigma^2}\right)^2+\frac{2r}{\sigma^2}}>1$, and $\alpha_2=\frac{1}{2}-\frac{(r-q)}{\sigma^2}-\sqrt{\left(\frac{1}{2}-\frac{r-q}{\sigma^2}\right)^2+\frac{2r}{\sigma^2}}<0$. To simplify notation, we drop the superscripts f and r in the corresponding proofs. The value matching and smooth pasting conditions at the default barrier are

$$-c\frac{\gamma}{r}+\phi_1 s_{0d}^{\alpha_1}+\phi_2 s_{0d}^{\alpha_2}=0, \tag{A.1}$$

$$\alpha_1\phi_1 s_{0d}^{\alpha_1}+\alpha_2\phi_2 s_{0d}^{\alpha_2}=0. \tag{A.2}$$

The matching and smooth pasting conditions at the production suspension point are:

$$-\frac{\gamma}{r}K+\frac{\gamma}{q}s_{10}+\theta s_{10}^{\alpha_2}+\delta_{10}=\phi_1 s_{10}^{\alpha_1}+\phi_2 s_{10}^{\alpha_2}, \tag{A.3}$$

$$\frac{\gamma}{q}s_{10}+\alpha_2\theta s_{10}^{\alpha_2}=\alpha_1\phi_1 s_{10}^{\alpha_1}+\alpha_2\phi_2 s_{10}^{\alpha_2}, \tag{A.4}$$

and at the restart point we have:

$$-\frac{\gamma}{r}K+\frac{\gamma}{q}s_{01}+\theta s_{01}^{\alpha_2}-\delta_{01}=\phi_1 s_{01}^{\alpha_1}+\phi_2 s_{01}^{\alpha_2}, \tag{A.5}$$

$$\frac{\gamma}{q}s_{01}+\alpha_2\theta s_{01}^{\alpha_2}=\alpha_1\phi_1 s_{01}^{\alpha_1}+\alpha_2\phi_2 s_{01}^{\alpha_2}. \tag{A.6}$$

From (A.1) and (A.2), we have:

$$\phi_1=-\frac{\alpha_2\gamma c}{(\alpha_1-\alpha_2)r}s_{0d}^{-\alpha_1}, \tag{A.7}$$

$$\phi_2=\frac{\alpha_1\gamma c}{(\alpha_1-\alpha_2)r}s_{0d}^{-\alpha_2}. \tag{A.8}$$

From (A.3) and (A.4), we can eliminate θ to obtain:

$$\left(-\frac{\gamma}{r}K+\delta_{10}\right)\alpha_2=(\alpha_2-\alpha_1)\phi_1 s_{10}^{\alpha_1}-\frac{\gamma}{q}(\alpha_2-1)s_{10}. \tag{A.9}$$

Similarly, equations (A.5) and (A.6) lead to:

$$\left(-\frac{\gamma}{r}K-\delta_{01}\right)\alpha_2=(\alpha_2-\alpha_1)\phi_1 s_{01}^{\alpha_1}-\frac{\gamma}{q}(\alpha_2-1)s_{01}. \tag{A.10}$$

Now, define $s_{01}=\kappa_1 s_{10}$. Note that $\kappa_1\ge 1$, i.e., $s_{01}\ge s_{10}$, otherwise the threshold policy is not feasible. Then (A.4) and (A.6) lead to:

$$\phi_1 s_{10}^{\alpha_1}=\kappa_2 s_{10}, \tag{A.11}$$

where

$$\kappa_2 = \frac{\gamma}{\alpha_1 q}\frac{(\kappa_1 - \kappa_1^{\alpha_2})}{(\kappa_1^{\alpha_1} - \kappa_1^{\alpha_2})}.$$

Substituting (A.11) into the right hand side of (A.9) and solving for s_{10} leads to $s_{10} = \frac{\alpha_2(\gamma K/r-\delta_{10})}{(\alpha_1-\alpha_2)\kappa_2-\gamma(1-\alpha_2)/q}$. The value for κ_1 can be solved by taking the ratio of equations (A.9) and (A.10) and substituting $s_{01} = \kappa_1 s_{10}$ together with equation (A.11). This leads to the following expression, the solution of which identifies κ_1:

$$\frac{K\gamma/r + \delta_{01}}{K\gamma/r - \delta_{10}} = \kappa_1^{\alpha_1}\left(\frac{(\alpha_1-\alpha_2)\kappa_2 - \frac{\gamma}{q}(1-\alpha_2)\kappa_1^{1-\alpha_1}}{(\alpha_1-\alpha_2)\kappa_2 - \frac{\gamma}{q}(1-\alpha_2)}\right). \tag{A.12}$$

(Note that if $\delta_{01} = \delta_{10} = 0$, the two systems of equations, (A.3,A.4) and (A.5,A.6) are the same, and the problem degenerates to the case where $s_{01} = s_{10}$, with $\kappa_1 = 1$. Moreover, applying $\kappa_1 = 1$, and use the definitions of κ_2, α_1 and α_2, it is straightforward to show that $s_{10} = K$.)

From the above expression, we see κ_1 and κ_2 are independent of c. Hence, s_{10} and s_{01} are also independent of c. From (A.11), we see that ϕ_1 is also independent of c. Combining (A.7) and (A.11) and rearranging for the default barrier leads to:

$$s_{0d} = \zeta c^{1/\alpha_1}, \tag{A.13}$$

where $\zeta = \left(\frac{-\alpha_2\gamma}{r(\alpha_1-\alpha_2)\kappa_2}\right)^{1/\alpha_1} s_{10}^{1-1/\alpha_1}$ is a constant, independent of the coupon. Finally, we can identify θ from equation (A.4). In particular, we have $\frac{\gamma}{q}s_{10} + \alpha_2\theta s_{10}^{\alpha_2} = \alpha_1\phi_1 s_{10}^{\alpha_1} + \alpha_2\phi_2 s_{10}^{\alpha_2}$. Hence,

$$\theta = \frac{\alpha_1}{\alpha_2}\phi_1 s_{10}^{\alpha_1-\alpha_2} + \phi_2 - \frac{\gamma}{\alpha_2 q}s_{10}^{1-\alpha_2}. \tag{A.14}$$

Now combining (A.8) and (A.13) leads to $\phi_2 = \frac{\alpha_1\gamma}{(\alpha_1-\alpha_2)r}\zeta^{-\alpha_2}c^{1-\alpha_2/\alpha_1}$. Substituting the expression of ϕ_2 and (A.11) in (A.14) leads to:

$$\theta = \eta_1 + \eta_2 c^{1-\alpha_2/\alpha_1}, \tag{A.15}$$

where $\eta_1 = \frac{(\alpha_1\kappa_2-\gamma/q)s_{10}^{1-\alpha_2}}{\alpha_2}$ and $\eta_2 = \frac{\alpha_1\gamma\zeta^{-\alpha_2}}{(\alpha_1-\alpha_2)r}$, which completes the proof. □

A.2. Proof of Corollary 1

To prove Corollary 1, we first prove the following two technical lemmas.

LEMMA EC.1. *When $\delta_{01} = \delta_{10} = \delta$, we have $\frac{\partial\kappa_1}{\partial\delta} > 0$, where κ_1 is defined in equation (A.12).*

Proof: We have from equation (A.12):

$$\frac{K\gamma/r + \delta}{K\gamma/r - \delta} = \kappa_1^{\alpha_1}\left(\frac{(\alpha_1-\alpha_2)\kappa_2 - \frac{\gamma}{q}(1-\alpha_2)\kappa_1^{1-\alpha_1}}{(\alpha_1-\alpha_2)\kappa_2 - \frac{\gamma}{q}(1-\alpha_2)}\right), \tag{A.16}$$

which can be rearranged to:

$$\frac{K\gamma/r + \delta}{K\gamma/r - \delta} = \kappa_1^{\alpha_1}\left(1 + \frac{\frac{\gamma}{q}(1-\alpha_2)\left(1-\kappa_1^{1-\alpha_1}\right)}{(\alpha_1-\alpha_2)\kappa_2 - \frac{\gamma}{q}(1-\alpha_2)}\right). \tag{A.17}$$

Substituting for $\kappa_2 = \frac{\gamma}{\alpha_1 q}\frac{(\kappa_1 - \kappa_1^{\alpha_2})}{(\kappa_1^{\alpha_1} - \kappa_1^{\alpha_2})}$ in the above equation leads to:

$$\frac{K\gamma/r+\delta}{K\gamma/r-\delta} = \kappa_1^{\alpha_1}\left(1+\frac{(1-\alpha_2)(1-\kappa_1^{1-\alpha_1})}{\left(\frac{\alpha_1-\alpha_2}{\alpha_1}\right)\left(1+\frac{1-\kappa_1^{\alpha_1-1}}{\kappa_1^{\alpha_1-1}-\kappa_1^{\alpha_2-1}}\right)-(1-\alpha_2)}\right).$$

Note that RHS increases in κ_1 since $\alpha_1 > 1$ and $\alpha_2 < 0$, so the numerator of the second term in the parenthesis increases in κ_1 and its denominator decreases in κ_1. Since LHS increases in δ and RHS increases in κ_1, κ_1 increases in δ. ☐

Lemma EC.2. *When $\delta_{01} = \delta_{10} = \delta$, $s^f_{0d}(c,\delta) = \zeta \cdot c^{\frac{1}{\alpha_1}}$, defined in (A.13), increases in δ for a fixed c. Equivalently, $\frac{\partial \zeta}{\partial \delta} > 0$.*

Proof: Since $V^{f,0}(s;c)$ decreases in δ and increases in s, for $\delta_1 < \delta_2$, we have $V^{f,0}(s^f_{0d}(\delta_2);c,\delta_2) = 0$ and $V^{f,0}(s^f_{0d}(\delta_2);c,\delta_1) > 0$. Therefore, $s_{0d}(\delta_1) < s_{0d}(\delta_2)$ for $V^{f,0}(s^f_{0d}(\delta_1);c,\delta_1) = 0$. ☐

Proof of Corollary 1: Substituting for κ_2 in the expression for s^f_{10} leads to:

$$s^f_{10} = \frac{(-K\gamma/r+\delta)\alpha_2}{\frac{\gamma}{q}\left(1-\alpha_2-\frac{(\alpha_1-\alpha_2)(\kappa_1-\kappa_1^{\alpha_2})}{\alpha_1(\kappa_1^{\alpha_1}-\kappa_1^{\alpha_2})}\right)}.$$

Since κ_1 increases in δ, $\frac{(\kappa_1-\kappa_1^{\alpha_2})}{(\kappa_1^{\alpha_1}-\kappa_1^{\alpha_2})} = 1 + \frac{1-\kappa_1^{\alpha_1-1}}{\kappa_1^{\alpha_1-1}-\kappa_1^{\alpha_2-1}}$ decreases in κ_1 and hence decreases in δ. Therefore, the denominator of s^f_{10} increases in δ. Moreover, the numerator decreases in δ as $\alpha_2 < 0$. Hence s^f_{10} decreases in δ.

For s^f_{01}, note that the equations (A.3, A.4) and (A.5, A.6), based on which we defined κ_1 and κ_2, are symmetric, except for the sign in the front of the switching cost δ. Therefore, by symmetry, we have s^f_{01} increases in δ.

The property of $s_{0d} = \zeta c^{1/\alpha_1}$ follows directly from Lemma EC.2. ☐

A.3. Proof of Corollary 2

Define $\bar{c} = \sup\left\{c | V^{f,1}(s_{10}(c);c) \geq 0\right\}$. To show that when $c \leq \bar{c}$, the flexible policy is feasible (i.e., $V^{f,1}(s_{10}(c);c) \geq 0$, so that the firm does not default when s is above s_{10}), it suffices to show that $V^{f,1}(s_{10}(c);c)$ decreases in c under the flexible policy. We have:

$$V^{f,1}(s_{10};c) = -(K+c)\tfrac{\gamma}{r} + \tfrac{\gamma}{q}s_{10} + \theta s_{10}^{\alpha_2}.$$

Since we have shown that s_{10} is independent of c,

$$\begin{aligned}\tfrac{\partial V^{f,1}}{\partial c} &= -\tfrac{\gamma}{r} + \tfrac{\partial \theta}{\partial c}s \\ &= -\tfrac{\gamma}{r} + \eta_2\left(\tfrac{\alpha_1-\alpha_2}{\alpha_1}\right)s_{10}^{\alpha_2}c^{-\alpha_2/\alpha_1}.\end{aligned}$$

For the derivative to be negative we need to show that:

$$\tfrac{r}{\gamma}\left(\tfrac{\alpha_1-\alpha_2}{\alpha_1}\right)\eta_2 s_{10}^{\alpha_2} < c^{\alpha_2/\alpha_1}. \tag{A.18}$$

Now, substituting for η_2 we obtain $\frac{r}{\gamma}\left(\frac{\alpha_1-\alpha_2}{\alpha_1}\right)\eta_2 = \zeta^{-\alpha_2}$. Moreover, since $s_{0d} = \zeta c^{1/\alpha_1}$ we have $\zeta^{-\alpha_2} = s_{0d}^{-\alpha_2}c^{\alpha_2/\alpha_1}$ and the left hand side of equation (A.18) becomes $\zeta^{-\alpha_2}s_{10}^{\alpha_2} =$

$(\frac{s_{0d}}{s_{10}})^{-\alpha_2} c^{\alpha_2/\alpha_1}$ and the condition reduces to $(\frac{s_{0d}}{s_{10}})^{-\alpha_2} < 1$, which is true under the flexible policy. Therefore, $\frac{\partial V^{f,1}}{\partial c} < 0$. Hence, when $c < \bar{c}$, $V^{f,1}(s_{10}(c);c) \geq 0$ and the flexible policy is feasible.

The fact that s_{10} and s_{01} are independent of c follows directly from the fact that κ_1 and κ_2 are independent of c. ☐

A.4. Proof of Proposition 2

In the paper, we assume that the firm starts with production at time $t = 0$. For completeness, here, we derive the value function for the case where the firm starts in the idle mode as well, i.e., $V^{r,0}$. The ODEs can be written as:

$$\begin{aligned} \tfrac{1}{2}\sigma^2 s^2 V^{r,1}_{ss} + (r-q)sV^{r,1}_s - rV^{r,1} + ((s-K-c))\gamma &= 0 \text{ for } s > s^r_{1d} \\ \tfrac{1}{2}\sigma^2 s^2 V^{r,0}_{ss} + (r-q)sV^{r,0}_s - rV^{r,0} &= 0 \text{ for } s^r_{0d} < s < s^r_{01}, \end{aligned}$$

where the boundary conditions are:

$$V^{r,0}(s^r_{0d};c) = 0, \tag{A.19}$$
$$V^{r,1}(s^r_{1d};c) = 0, \tag{A.20}$$
$$V^{r,0}(s^r_{01};c) = V^{r,1}(s^r_{01};c) - \delta_{01}, \tag{A.21}$$

with smooth pasting conditions:

$$V^{r,0}_s(s^r_{0d};c) = 0, \tag{A.22}$$
$$V^{r,1}_s(s^r_{1d};c) = 0, \tag{A.23}$$
$$V^{r,1}_s(s^r_{01};c) = V^{r,0}_s(s^r_{01};c). \tag{A.24}$$

We can obtain the solution for the above ODEs as:

$$V^{r,1}(s;c) = -(K+c)\tfrac{\gamma}{r} + \tfrac{\gamma}{q}s + \rho s^{\alpha_2} \text{ for } \quad s \geq s^r_{1d} \tag{A.25}$$
$$V^{r,0}(s;c) = -c\tfrac{\gamma}{r} + G_1 s^{\alpha_1} + G_2 s^{\alpha_2} \text{ for } \quad s \leq s^r_{01} \tag{A.26}$$

where the coefficients ρ, G_1 and G_2 as well as the threshold values are determined by the boundary conditions. Combining (A.20) and (A.23) leads to:

$$\begin{aligned} -(K+c)\frac{\gamma}{r} + \frac{\gamma}{q}s_{1d} + \rho s_{1d}^{\alpha_2} &= 0 \\ \frac{\gamma}{q}s_{1d} + \alpha_2 \rho s_{1d}^{\alpha_2} &= 0. \end{aligned}$$

Then, we obtain: $s_{1d} = \frac{-\alpha_2}{1-\alpha_2}\frac{q}{r}(K+c)$ and $\rho = -\frac{\gamma}{q\alpha_2} s_{1d}^{1-\alpha_2}$.. The remaining four unknowns, s^r_{01} , $s^r_{0d,}$, G_1 and G_2 can be solved from the four equations, (A.19), (A.21), (A.22) and (A.24). ☐

A.5. Proof of Corollary 3

Since $s_{1d} = \frac{-\alpha_2}{1-\alpha_2}\frac{q}{r}(K+c)$ with $\alpha_2 < 0$, s_{1d} clearly decreases in c. Note $V^{r,1}(s;c) = -(K+c)\frac{\gamma}{r} + \frac{\gamma}{q}s - \frac{\gamma}{q\alpha_2} s_{1d}^{1-\alpha_2} s^{\alpha_2}$, which is independent of the switching cost. By definition, the switching cost does not enter the value function under the rigid policy when the firm starts in the production mode. ☐

A.6. Proof of Proposition 3

Since $V^{f,1}(s;c) = -(K+c)\frac{\gamma}{r} + \frac{\gamma}{q}s + \theta s^{\alpha_2}$ and $V^{r,1}(s;c) = -(K+c)\frac{\gamma}{r} + \frac{\gamma}{q}s + \rho s^{\alpha_2}$, to compare these two value functions, we only need to compare θ and ρ. From (A.15), we have $\theta = \eta_1 + \eta_2 c^{1-\frac{\alpha_2}{\alpha_1}}$. Note that $\eta_2 > 0$ since $\alpha_2 < 0$ and $\alpha_1 > 0$.

Now, $\rho = -\frac{\gamma}{q\alpha_2}s_{1d}^{1-\alpha_2} > 0$. Since $s_{1d} = \frac{-\alpha_2}{1-\alpha_2}\frac{q}{r}(K+c)$, ρ can be written as $M_1(K+c)^{1-\alpha_2}$ with $M_1 = -\frac{\gamma}{q\alpha_2}(\frac{-\alpha_2 q}{(1-\alpha_2)r})^{1-\alpha_2} > 0$. Note that when $c = 0$, there is no debt, and the flexible policy leads to higher equity value ($\theta(0) > \rho(0)$). The power terms in θ and ρ are $1-\frac{\alpha_2}{\alpha_1}$ and $1-\alpha_2$ respectively. Since $\alpha_1 > 1$ and $\alpha_2 < 0$, $1-\frac{\alpha_2}{\alpha_1} < 1-\alpha_2$. Hence, when c is sufficiently large, $\rho > \theta$. Moreover,

$$\frac{\partial\theta/\partial c}{\partial\rho/\partial c} = \frac{(1-\alpha_2/\alpha_1)\eta_2 c^{-\alpha_2/\alpha_1}}{(1-\alpha_2)M_1(K+c)^{-\alpha_2}} = \frac{(1-\alpha_2/\alpha_1)\eta_2 c^{\alpha_2-\alpha_2/\alpha_1}}{(1-\alpha_2)M_1(K/c+1)^{-\alpha_2}}.$$

Since $\alpha_1 > 1$ and $\alpha_2 < 0$, $\alpha_2 - \frac{\alpha_2}{\alpha_1} < 0$. Hence, the denominator increases in c, whereas the numerator decreases in c, and thus $\frac{\partial\theta/\partial c}{\partial\rho/\partial c}$ decreases in c. We use this monotonicity property to show that $\theta(c)$ and $\rho(c)$ can only cross once.

We consider two cases: $0 < \frac{\partial\theta}{\partial c}|_{c=0} < \frac{\partial\rho}{\partial c}|_{c=0}$ or $\frac{\partial\theta}{\partial c}|_{c=0} \geq \frac{\partial\rho}{\partial c}|_{c=0} > 0$. If $\frac{\partial\theta}{\partial c}|_{c=0} < \frac{\partial\rho}{\partial c}|_{c=0}$, since $\frac{\partial\theta/\partial c}{\partial\rho/\partial c}$ decreases in c, $\frac{\partial\theta}{\partial c}|_c < \frac{\partial\rho}{\partial c}|_c$ for all $c > 0$. Therefore, θ and ρ can cross at most once.

On the other hand, if $\frac{\partial\theta}{\partial c}|_{c=0} \geq \frac{\partial\rho}{\partial c}|_{c=0}$, since $\theta(0) > \rho(0)$ but $\theta(c) < \rho(c)$ as c is sufficiently large, there must exist a set (with measure >0) such that $\frac{\partial\theta}{\partial c}|_c < \frac{\partial\rho}{\partial c}|_c$ in this set. Now, based on the monotonicity property of $\frac{\partial\theta/\partial c}{\partial\rho/\partial c}$ in c, there must exist a threshold c_1 such that when $c < c_1$, $\frac{\partial\theta}{\partial c}|_c > \frac{\partial\rho}{\partial c}|_c$; and when $c > c_1$, $\frac{\partial\theta}{\partial c}|_c < \frac{\partial\rho}{\partial c}|_c$. In this case, since $\theta(0) > \rho(0)$, $\theta(c) > \rho(c)$ for $c \leq c_1$. Therefore, θ and ρ can cross only once, which is in the interval of $[c_1, \infty]$.

The facts that θ and ρ can cross only once in c for both cases and that $\theta > \rho$ when c is small and $\theta < \rho$ when c is large assert that the result of this proposition is true, and we define the crossing point as $\hat{c}(\delta)$.

We next show the monotonicity property of $\hat{c}(\delta)$. Suppose there are two values δ_1 and δ_2 with $\delta_1 < \delta_2$. Define

$$\begin{aligned} c_1 &= \hat{c}(\delta_1) = \sup\{c|V^{f,1}(s;c,\delta_1) > V^{r,1}(s;c)\} = \sup\{c|\theta(c,\delta_1) > \rho(c)\} \\ c_2 &= \hat{c}(\delta_2) = \sup\{c|V^{f,1}(s;c,\delta_2) > V^{r,1}(s;c)\} = \sup\{c|\theta(c,\delta_2) > \rho(c)\}. \end{aligned}$$

Now, we prove this proposition by contradiction. Assume $c_1 < c_2$. Then, based on the above definition of c_1, we have $V^{f,1}(s = s_{10}(\delta_1); c_2, \delta_1) < V^{r,1}(s = s_{10}(\delta_1); c_2)$, i.e.,

$$(-(K+c)\tfrac{\gamma}{r} + \tfrac{\gamma}{q}s + \theta(c_2,\delta_1)\cdot s^{\alpha_2})|_{s=s_{10}(\delta_1)} < -(K+c)\frac{\gamma}{r} + \frac{\gamma}{q}s + \rho(c_2)s^{\alpha_2}|_{s=s_{10}(\delta_1)},$$

which implies $\theta(c_2,\delta_1) < \rho(c_2)$.

Moreover, we know $V^{f,1}(s = s_{10}(\delta_1); c_2, \delta_2) < V^{f,1}(s = s_{10}(\delta_1); c_2, \delta_1)$, i.e.,

$$(-(K+c)\tfrac{\gamma}{r} + \tfrac{\gamma}{q}s + \theta(c_2,\delta_2)\cdot s^{\alpha_2})|_{s=s_{10}(\delta_1)} < (-(K+c)\frac{\gamma}{r} + \frac{\gamma}{q}s + \theta(c_2,\delta_1)\cdot s^{\alpha_2})|_{s=s_{10}(\delta_1)},$$

which implies $\theta(c_2,\delta_2) < \theta(c_2,\delta_1)$ for $\delta_1 < \delta_2$.

The above assets that $\theta(c_2,\delta_2) < \rho(c_2)$, then $V^{f,1}(s;c_2,\delta_2) < V^{r,1}(s;c_2)$ for any s, which contradicts with the definition of c_2. □

A.7. Proof of Corollary 4

Since $\hat{c}$ decreases in δ, we can define $\hat{\delta}(c)$ as the inverse function of $\hat{c}(\delta)$. The threshold and monotonicity results follow directly. □

A.8. Proof of Lemma 1

For the flexible policy we have:

$$\begin{aligned} V^{f,1}(s_0;c) &= -(K+c)\tfrac{\gamma}{r} + \tfrac{\gamma}{q}s + \theta s_0^{\alpha_2} \\ &= -(K+c)\tfrac{\gamma}{r} + \tfrac{\gamma}{q}s + (\eta_1 + \eta_2 c^{1-\alpha_2/\alpha_1})s^{\alpha_2}, \\ B^{f,1}(s_0;c) &= \tfrac{c}{r}(1-(\tfrac{s}{s_{0d}})^{\alpha_2}) = \tfrac{c}{r}(1-(\tfrac{s}{\zeta})^{\alpha_2}c^{-\alpha_2/\alpha_1}), \end{aligned}$$

based on the assumption that the firm loses its entire value upon default (i.e., $b=1$). To establish the optimal coupon we differentiate these two equations leading to:

$$\begin{aligned} \tfrac{\partial V^{f,1}}{\partial c} &= -\tfrac{\gamma}{r} + (\tfrac{\alpha_1-\alpha_2}{\alpha_1})\eta_2 s^{\alpha_2}c^{-\alpha_2/\alpha_1}, \\ \tfrac{\partial B^{f,1}}{\partial c} &= \tfrac{1}{r} - \tfrac{1}{r}(\tfrac{\alpha_1-\alpha_2}{\alpha_1})s^{\alpha_2}\zeta^{-\alpha_2}c^{-\alpha_2/\alpha_1}. \end{aligned}$$

Therefore, the first order condition for firm value maximization is:

$$\left(\tfrac{1-\gamma}{r}\right) + \left(\tfrac{\alpha_1-\alpha_2}{\alpha_1}\right)s^{\alpha_2}\left(\eta_2 - \tfrac{\zeta^{-\alpha_2}}{r}\right)c^{-\alpha_2/\alpha_1} = 0.$$

Substituting $\eta_2 = \frac{\alpha_1\gamma\zeta^{-\alpha_2}}{(\alpha_1-\alpha_2)r}$ into the above expression leads to:

$$\left(\tfrac{1-\gamma}{r}\right) + \left(\tfrac{\alpha_1-\alpha_2}{\alpha_1 r}\right)s^{\alpha_2}\zeta^{-\alpha_2}\left(\tfrac{\alpha_1\gamma-(\alpha_1-\alpha_2)}{\alpha_1-\alpha_2}\right)c^{-\alpha_2/\alpha_1} = 0, \tag{A.27}$$

from which we obtain the optimal unconstrained coupon:

$$c^* = \left(\tfrac{s_0}{\zeta}\right)^{\alpha_1}\left(\tfrac{\alpha_1(1-\gamma)-\alpha_2}{\alpha_1(1-\gamma)}\right)^{\alpha_1/\alpha_2}.$$

Using (A.13), we have:

$$s_{0d}^* = \left(\frac{(1-\gamma)\alpha_1-\alpha_2}{(1-\gamma)\alpha_1}\right)^{\frac{1}{\alpha_2}} s_0. \tag{A.28}$$

Now consider the rigid case. We have:

$$\begin{aligned} V^{r,1}(s;c) &= -(K+c)\frac{\gamma}{r} + \frac{\gamma}{q}s - \frac{\gamma}{q\alpha_2}s_{1d}^{1-\alpha_2}\cdot s^{\alpha_2} \text{ for } s \geq s_{1d} \\ B^{r,1}(s;c) &= \tfrac{c}{r}\left(1-\left(\tfrac{s}{s_{1d}}\right)^{\alpha_2}\right). \end{aligned}$$

Differentiating the two terms yields:

$$\begin{aligned} \tfrac{dV^{r,1}}{dc} &= -\tfrac{\gamma}{r} - \left(\tfrac{\gamma}{q\alpha_2}\right)(s)^{\alpha_2}\,s_{1d}^{-\alpha_2}(1-\alpha_2)\cdot\frac{-\alpha_2}{1-\alpha_2}\frac{q}{r} = -\tfrac{\gamma}{r} + \frac{\gamma}{r}\cdot(\frac{s}{s_{1d}})^{\alpha_2}, \\ \tfrac{dB^{r,1}}{dc} &= \frac{1}{r}\left(1-\left(\tfrac{s}{s_{1d}}\right)^{\alpha_2}\right) + \tfrac{\alpha_2 c}{r}s^{\alpha_2}s_{1d}^{-\alpha_2-1}\cdot\frac{-\alpha_2}{1-\alpha_2}\frac{q}{r}. \end{aligned}$$

Therefore, the first order condition for firm value maximization is:

$$\frac{1}{r}(1-\gamma) - \frac{1}{r}(1-\gamma)\cdot\left(\frac{s_0}{\frac{-\alpha_2}{1-\alpha_2}\frac{q}{r}(K+c)}\right)^{\alpha_2} + \tfrac{\alpha_2}{r}\cdot\frac{c}{K+c}\cdot\left(\frac{s_0}{\frac{-\alpha_2}{1-\alpha_2}\frac{q}{r}(K+c)}\right)^{\alpha_2} = 0,$$

which can be simplified to: $(1-\frac{\alpha_2 c}{(1-\gamma)(K+c)})\left(\frac{s_0(1-\alpha_2)r}{-\alpha_2 q(K+c)}\right)^{\alpha_2} - 1 = 0.$ □

A.9. Proof of Theorem 1

Theorem 1, as presented in the main body, focuses on the setting with identical switching costs. The theorem can be generalized with different switching costs when one of them is fixed or proportional to the other; the proof is available upon request. There are two important conclusions in Theorem 1. The first is the threshold policy that there are two thresholds $\underline{\delta}$ and $\bar{\delta}$ dividing the domain of switching cost into three regions such that the optimal coupon rate takes value of c_f^*, $\hat{c}$ and c_r^* respectively. Correspondingly, the equity holders use the flexible policy in $[0,\bar{\delta}]$ and then adopt the rigid policy in $[\bar{\delta},\infty]$. The second is the non-monotonic property of the optimal coupon rate c^*. It first decreases in $[0,\bar{\delta}]$ and then jumps up to a flat level in $[\bar{\delta},\infty]$. As the proof of threshold property requires some of the non-monotonicity results, we will prove the non-monotoncity result first. The monotonicity of $\hat{c}(\delta)$ shown in Proposition 3, together with Lemmas EC.3, and EC.4 provided below will be used to prove the non-mononicity result with respect to δ in Theorem 1, and Lemmas EC.5, EC.6, and EC.7 provided below will be used to prove the threshold policy in the theorem.

We denote $\mathcal{A}^f$ as the value of the firm when operating using flexiblity, and $\mathcal{A}^r$ as the value for the firm when operating without flexiblity.

Lemma EC.3. *$c_f^*(\delta)$ decreases in δ.*

Proof: Let $\mathcal{A}$ denote the sum of the equity value and the debt value. To show this lemma, we first assume that $\frac{\partial^2\mathcal{A}^f}{\partial c\partial\delta} < 0$, which implies that $\frac{\partial\mathcal{A}^f}{\partial c}$ decreases in δ. Now suppose $\delta_1 < \delta_2$. Then $\frac{\partial\mathcal{A}^f}{\partial c}|_{\delta=\delta_2,c=c_f^*(\delta_1)} < \frac{\partial\mathcal{A}^f}{\partial c}|_{\delta=\delta_1,c=c_f^*(\delta_1)} = 0$. Taking the derivative of (A.27), we have $\frac{\partial^2\mathcal{A}^f}{\partial c^2} < 0$. Given that $\mathcal{A}^f$ is concave in c and the fact that $\frac{\partial\mathcal{A}^f}{\partial c}|_{\delta=\delta_2,c=c_f^*(\delta_1)} < \frac{\partial\mathcal{A}^f}{\partial c}|_{\delta=\delta_1,c=c_f^*(\delta_1)} = 0$, we can conclude $c_f^*(\delta_2) < c_f^*(\delta_1)$, i.e., $c_f^*(\delta)$ decreases in δ.

In the following, we show the above assumption ($\frac{\partial^2\mathcal{A}^f}{\partial c\partial\delta} < 0$) is true. Since $\frac{\partial^2\mathcal{A}^f}{\partial c\partial\delta} = \frac{\partial^2 V^{f,1}}{\partial c\partial\delta} + \frac{\partial^2 B^{f,1}}{\partial c\partial\delta}$, we compute the two terms separately.

$$\frac{\partial^2 V^{f,1}}{\partial c\partial\delta} = \frac{\partial^2\theta}{\partial c\partial\delta}s_0^{\alpha_2} = \frac{\partial^2(\frac{\alpha_1}{\alpha_2}\phi_1 s_{10}^{\alpha_1-\alpha_2} + \phi_2 - \frac{\gamma}{\alpha_2 q}s_{10}^{1-\alpha_2})}{\partial c\partial\delta}s_0^{\alpha_2} = \frac{\partial^2\phi_2}{\partial c\partial\delta}s_0^{\alpha_2}.$$

The last equality follows from the fact that ϕ_1 and s_{10} are independent of c (shown in the proof of Proposition 1). Since $\phi_2 = \frac{\alpha_1\gamma c}{(\alpha_1-\alpha_2)r}(s_{0d}^f)^{-\alpha_2} = \frac{\alpha_1}{\alpha_1-\alpha_2}\cdot\frac{c\gamma}{r}\cdot(\zeta c^{\frac{1}{\alpha_1}})^{-\alpha_2}$,

$$\frac{\partial^2 V^{f,1}}{\partial c\partial\delta} = \frac{\partial^2\phi_2}{\partial c\partial\delta}s_0^{\alpha_2} = \frac{\alpha_1}{\alpha_1-\alpha_2}\cdot\frac{\gamma}{r}\cdot\frac{\partial\zeta^{-\alpha_2}}{\partial\delta}\cdot\frac{\partial c^{1-\frac{\alpha_2}{\alpha_1}}}{\partial c}.$$

$$\frac{\partial^2 B^{f,1}}{\partial c\partial\delta} = \frac{\partial^2(\frac{c}{r}(1-(\frac{s_0}{s_{0d}})^{\alpha_2}))}{\partial c\partial\delta} = -\frac{\partial^2\frac{c}{r}(s_{0d})^{-\alpha_2}}{\partial c\partial\delta} = -\frac{1}{r}\cdot\frac{\partial\zeta^{-\alpha_2}}{\partial\delta}\cdot\frac{\partial c^{1-\frac{\alpha_2}{\alpha_1}}}{\partial c}.$$

Since $\frac{\alpha_1\gamma}{\alpha_1-\alpha_2} < 1$ and $\frac{\partial\zeta}{\partial\delta} > 0$ (Lemma EC.2), we have $\frac{\partial^2 V^{f,1}}{\partial c\partial\delta} + \frac{\partial^2 B^{f,1}}{\partial c\partial\delta} < 0$. □

Lemma EC.4. $c_r^* \geq c^*(0)$.

Proof: see Theorem 1 in Ritchken and Wu (2021). □

The flow to prove the threshold policy in Theorem 1 is as follows:

1. For the proof of the threshold $\underline{\delta}$, we first define it to be $\underline{\delta} = \sup\{\delta | V^{f,1}(s_{10}(\delta); c_f^*(\delta), \delta) \geq V^{r,1}(s_{10}(\delta); c_f^*(\delta))\}$ and we show that for all $\delta < \underline{\delta}$, $V^{f,1}(s_{10}(\delta); c_f^*(\delta), \delta) \geq V^{r,1}(s_{10}(\delta); c_f^*(\delta))$. We prove this result by iteration. To do so, we use a small step size $\hat{\Delta}$ that is independent of δ. In the first step, we show that $V^{f,1}(s_{10}(\delta); c_f^*(\delta), \delta) < V^{r,1}(s_{10}(\delta); c_f^*(\delta))$ holds on $[\underline{\delta} - \hat{\Delta}, \underline{\delta}]$, since $V^{f,1}(s_{10}(\delta - \Delta); c_f^*(\delta - \Delta), \delta - \Delta) > V^{f,1}(s_{10}(\delta - \Delta); c_f^*(\delta), \delta)$ for any $\Delta < \hat{\Delta}$ (Lemma EC.5) and $V^{r,1}(s; c_f^*(\delta))$ increases in δ (Lemma EC.6). We then conduct this iteratively and extend the result to $[0, \underline{\delta}]$ using a finite covering.

2. For the second threshold $\bar{\delta}$, we define $\bar{\delta} = \inf\{\delta > \underline{\delta} | \mathcal{A}^f(s; \hat{c}(\delta), \delta) \leq \mathcal{A}^r(s; c_r^*)\}$. Then we show that for any $\delta > \bar{\delta}$, $\mathcal{A}^f(s; \hat{c}(\delta), \delta) \leq \mathcal{A}^r(s; c_r^*)$.

Lemma EC.5. *There exists a $\hat{\Delta}$ that is independent of δ such that for any $\Delta < \hat{\Delta}$, $V^{f,1}(s_{10}(\delta - \Delta); c_f^*(\delta - \Delta), \delta - \Delta) > V^{f,1}(s_{10}(\delta - \Delta); c_f^*(\delta), \delta)$.*

Proof: First, for any $s > s_{10}(\delta)$, $V^{f,1}(s; c_f^*(\delta), \delta)$ can be written as:

$$V^{f,1}(s; c_f^*(\delta), \delta) = \int_{s_{0d}}^{s_{10}(\delta)} \frac{\partial V^{f,0}(s; c_f^*(\delta), \delta)}{\partial s} ds + \int_{s_{10}(\delta)}^{s} \frac{\partial V^{f,1}(s; c_f^*(\delta), \delta)}{\partial s} ds + \delta.$$

Recall from (A.28) that s_{0d} is a constant independent of δ under the flexible policy with $c = c_f^*$. Moreover, the equity value at s_{0d} is zero.

Hence, for any $\Delta > 0$, we have

$$\begin{aligned} & V^{f,1}(s_{10}(\delta - \Delta); c_f^*(\delta - \Delta), \delta - \Delta) - V^{f,1}(s_{10}(\delta - \Delta); c_f^*(\delta), \delta) \\ = & \int_{s_{0d}}^{s_{10}(\delta)} \frac{\partial V^{f,0}(s; c_f^*(\delta - \Delta), \delta - \Delta)}{\partial s} - \frac{\partial V^{f,0}(s; c_f^*(\delta), \delta)}{\partial s} ds \\ & + \int_{s_{10}(\delta)}^{s_{10}(\delta - \Delta)} \frac{\partial V^{f,0}(s; c_f^*(\delta - \Delta), \delta - \Delta)}{\partial s} - \frac{\partial V^{f,1}(s; c_f^*(\delta), \delta)}{\partial s} ds + \Delta. \end{aligned}$$

Note that for $s \in [s_{10}(\delta), s_{10}(\delta - \Delta)]$, if the switching cost equals $\delta - \Delta$, the production will be temporarily suspended, whereas, if the switching cost equals δ, the production will continue. Since $\frac{\partial V^{f,0}(s; c_f^*(\delta_2 - \Delta), \delta_2 - \Delta)}{\partial s} > \frac{\partial V^{f,0}(s; c_f^*(\delta), \delta)}{\partial s}$, the first integral is positive. Therefore, the above expression is bounded below by $\int_{s_{10}(\delta)}^{s_{10}(\delta - \Delta)} \frac{\partial V^{f,1}(s; c_f^*(\delta - \Delta), \delta - \Delta)}{\partial s} - \frac{\partial V^{f,0}(s; c_f^*(\delta), \delta)}{\partial s} ds + \Delta$. For the expression of $s_{10} = \frac{(-K\frac{\gamma}{r} + \delta) \cdot \alpha_2}{\frac{\gamma}{q}\left(1 - \alpha_2 - \frac{(\alpha_1 - \alpha_2)(1 - \kappa_1^{\alpha_2})}{\alpha_1(\kappa_1^{\alpha_1} - \kappa_1^{\alpha_2})}\right)}$ and $0 < \frac{(1 - \kappa_1^{\alpha_2})}{(\kappa_1^{\alpha_1} - \kappa_1 \alpha_2)} < \frac{-\alpha_2}{\alpha_1 - \alpha_2}$, we have $\frac{-q\alpha_2\Delta}{\gamma(1 - \alpha_2)} > s_{10}(\delta_2 - \Delta) - s_{10}(\delta_2) > \frac{-q\alpha_2\Delta}{\gamma\left(1 - \alpha_2 - \frac{\alpha_2}{\alpha_1}\right)}$. Note that the absolute value of the integrand is bounded by the following relationship:

$$\begin{aligned} \frac{\partial V^{f,1}(s; c_f^*(\delta - \Delta), \delta - \Delta)}{\partial s} - \frac{\partial V^{f,1}(s; c_f^*(\delta), \delta)}{\partial s} &< \frac{\partial V^{f,0}(s; c_f^*(\delta - \Delta), \delta - \Delta)}{\partial s} - \frac{\partial V^{f,1}(s; c_f^*(\delta), \delta)}{\partial s} \\ \frac{\partial V^{f,0}(s; c_f^*(\delta - \Delta), \delta - \Delta)}{\partial s} - \frac{\partial V^{f,0}(s; c_f^*(\delta), \delta)}{\partial s} &> \frac{\partial V^{f,0}(s; c_f^*(\delta - \Delta), \delta - \Delta)}{\partial s} - \frac{\partial V^{f,1}(s; c_f^*(\delta), \delta)}{\partial s}. \end{aligned}$$

Below, we show that the left hand sides of the above two inequalities converge to zero uniformly for all δ as $\Delta \to 0$.

$$\frac{\partial V^{f,0}(s; c_f^*(\delta - \Delta), \delta - \Delta)}{\partial s} - \frac{\partial V^{f,0}(s; c_f^*(\delta), \delta)}{\partial s}$$

$$
\begin{aligned}
&= \alpha_1\Big(\phi_1(c_f^*(\delta-\Delta)) - \phi_1(c_f^*(\delta))\Big)\cdot s^{\alpha_2-1} + \alpha_2\Big(\phi_2(c_f^*(\delta-\Delta)) - \phi_2(c_f^*(\delta))\Big)s^{\alpha_2-1}\\
&= \frac{\alpha_1\alpha_2\gamma}{r(\alpha_1-\alpha_2)}\big(-s_{0d}^{-\alpha_1}\cdot s^{\alpha_1-1} - s_{0d}^{-\alpha_2}\cdot s^{\alpha_2-1}\big)\cdot\big(c_f^*(\delta-\Delta)-c_f^*(\delta)\big),
\end{aligned}
$$

and

$$
\begin{aligned}
&\frac{\partial V^{f,1}(s;c_f^*(\delta-\Delta),\delta-\Delta)}{\partial s} - \frac{\partial V^{f,1}(s;c_f^*(\delta),\delta)}{\partial s}\\
&= \alpha_2\big(\theta(c_f^*(\delta-\Delta),\delta-\Delta) - \theta(c_f^*(\delta-\Delta),\delta-\Delta)\big)s_{0d}^{-\alpha_2}\cdot s^{\alpha_2-1}\\
&= \alpha_2\Bigg(\frac{\alpha_1}{\alpha_2}\cdot\frac{-\alpha_2\big(c_f^*(\delta-\Delta)-c_f^*(\delta)\big)\gamma}{r(\alpha_1-\alpha_2)}\big(s_{10}(\delta-\Delta)\big)^{\alpha_1-\alpha_2}s_{0d}^{-\alpha_1}\\
&+ \frac{\alpha_1}{\alpha_2}\cdot\frac{-\alpha_2 c_f^*(\delta)\gamma}{r(\alpha_1-\alpha_2)}\Big(\big(s_{10}(\delta-\Delta)\big)^{\alpha_1-\alpha_2} - \big(s_{10}(\delta)\big)^{\alpha_1-\alpha_2}\Big)s_{0d}^{-\alpha_1}\\
&+ \frac{\alpha_1\big(c_f^*(\delta-\Delta)-c_f^*(\delta)\big)\gamma}{r(\alpha_1-\alpha_2)}s_{0d}^{-\alpha_2} - \frac{\gamma}{\alpha_2 q}\big(s_{10}(\delta-\Delta)-s_{10}(\delta)\big)\Bigg)s_{0d}^{-\alpha_2}\cdot s^{\alpha_2-1}.
\end{aligned}
$$

Clearly, to show the above two expressions uniformly converge to 0, we only need to show that $\big|c_f^*(\delta-\Delta)-c_f^*(\delta)\big|$ and $\big|s_{10}(\delta-\Delta)-s_{10}(\delta)\big|$ uniformly converge to zero, as $\Delta\to 0$, for all δ. As $\frac{-q\alpha_2\Delta}{\gamma(1-\alpha_2)} < s_{10}(\delta-\Delta)-s_{10}(\delta) < \frac{-q\alpha_2\Delta}{\gamma(1-\alpha_2-\frac{-\alpha_2}{\alpha_1})}$, the uniform convergence of $\big|s_{10}(\delta_2-\Delta)-s_{10}(\delta_2)\big|$ follows directly. For c_f^*, we have:

$$
\begin{aligned}
\frac{\partial c_f^*}{\partial\delta} &= \frac{\partial}{\partial\delta}s_{0d}\cdot\big(s_{10}(\delta)\big)^{1-\alpha_1}\left(\frac{r(\alpha_1-\alpha_2)\cdot\kappa_2(\delta)}{-\alpha_2\gamma}\right)\\
&= s_{0d}^{\alpha_1}\cdot\frac{r(\alpha_1-\alpha_2)}{-\alpha_2\gamma}\cdot\left[(1-\alpha_1)\cdot\big(s_{10}(\delta)\big)^{-\alpha_1}\frac{\partial s_{10}(\delta)}{\partial\delta}\kappa_2(\delta) + \big(s_{10}(\delta)\big)^{1-\alpha_1}\frac{\partial\kappa_2(\delta)}{\partial\delta}\right].
\end{aligned}
$$

Note from equation (A.9), we have: $(-K\frac{\gamma}{r}+\delta)\cdot\alpha_2 + \frac{\gamma}{q}(\alpha_2-1)s_{10} = (\alpha_2-\alpha_1)\phi_1 s_{10}^{\alpha_1}$, and thus, $\alpha_2 + \frac{\gamma}{q}(\alpha_2-1)\frac{\partial s_{10}}{\partial\delta} = (\alpha_2-\alpha_1)\frac{\partial\kappa_2 s_{10}}{\partial\delta}$. Simplifying the above equation, we have

$$
\frac{\partial\kappa_2}{\partial\delta} = -\frac{1}{(\alpha_1-\alpha_2)s_{10}}\cdot\left[(\alpha_1-\alpha_2)\kappa_2 - \frac{\gamma}{q}(1-\alpha_2)\right]\cdot\frac{\partial s_{10}}{\partial\delta} - \frac{\alpha_2}{(\alpha_1-\alpha_2)s_{10}}.
$$

Since $\frac{\partial s_{10}}{\partial\delta}$ is bounded in $[\frac{-\alpha_2}{\frac{\gamma}{q}(1-\alpha_2-\frac{-\alpha_2)}{\alpha_1})}, \frac{-\alpha_2}{\frac{\gamma}{q}(1-\alpha_2)}]$, $\kappa_2 = \frac{(1-\kappa_1^{\alpha_2})}{\alpha_1(\kappa_1^{\alpha_1}-\kappa_1^{\alpha_2})}\cdot\frac{\gamma}{q} \in [0, \frac{-\alpha_2}{\alpha_1(\alpha_1-\alpha_2)}\cdot\frac{\gamma}{q}]$, and $s_{10}\in[s_{0d},K]$, $\frac{\partial\kappa_2}{\partial\delta}$ is also bounded uniformly, independent of δ. Hence, $\frac{\partial c_f^*}{\partial\delta}$ is bounded uniformly. Hence, $\big|c_f^*(\delta-\Delta)-c_f^*(\delta)\big|\to 0$ uniformly as $\Delta\to 0$.

Consequently, $|\frac{\partial V^{f,0}(s;c_f^*(\delta-\Delta),\delta-\Delta)}{\partial s} - \frac{\partial V^{f,1}(s;c_f^*(\delta),\delta)}{\partial s}|\to 0$ uniformly for any δ as $\Delta\to 0$. Therefore,

$$
\begin{aligned}
&\int_{s_{10}(\delta)}^{s_{10}(\delta-\Delta)} \frac{\partial V^{f,1}(s;c_f^*(\delta-\Delta),\delta-\Delta)}{\partial s} - \frac{\partial V^{f,0}(s;c_f^*(\delta),\delta)}{\partial s}ds + \Delta\\
&>\Delta\big(1-\max_s\left|\frac{\partial V^{f,1}(s;c_f^*(\delta-\Delta),\delta-\Delta)}{\partial s} - \frac{\partial V^{f,0}(s;c_f^*(\delta),\delta)}{\partial s}\right|\cdot\frac{-\alpha_2 q}{\gamma(1-\alpha_2)}\big).
\end{aligned}
$$

Since $\max_s \left| \frac{\partial V^{f,1}(s;c_f^*(\delta-\Delta),\delta-\Delta)}{\partial s} - \frac{\partial V^{f,0}(s;c_f^*(\delta),\delta)}{\partial s} \right| \to 0$ uniformly for any δ, there must exist a $\hat{\Delta}$ that is independent of δ such that $1 - \max_s \left| \frac{\partial V^{f,1}(s;c_f^*(\delta-\hat{\Delta}),\delta-\hat{\Delta})}{\partial s} - \frac{\partial V^{f,0}(s;c_f^*(\delta),\delta)}{\partial s} \right| \cdot \frac{-\alpha_2}{\frac{\gamma}{q}(1-\alpha_2\ )} > 0$. Therefore, we have shown that there exists a $\hat{\Delta}$ such that for any $\Delta < \hat{\Delta}$, $V^{f,1}(s_{10}(\delta-\Delta); c_f^*(\delta-\Delta), \delta-\Delta) > V^{f,1}(s_{10}(\delta-\Delta); c_f^*(\delta), \delta)$. □

Lemma EC.6. $\frac{\partial V^{r,1}}{\partial c} \le 0$. *And since $c_f^*(\delta)$ decreases in δ (Lemma EC.3), $V^{r,1}(s; c_f^*(\delta))$ increases in δ.*

Proof: Note $V^{r,1} = -(K+c)\frac{\gamma}{r} + \frac{\gamma}{q}s + \rho s^{\alpha_2}$ where $\rho = -\frac{\gamma}{q\alpha_2} s_{1d}^{1-\alpha_2} = -\frac{\gamma}{\alpha_2 q}(\frac{-\alpha_2 q}{(1-\alpha_2)r})^{1-\alpha_2}(K+c)^{1-\alpha_2}$. Taking derivative wrt c leads to:

$$\begin{aligned}\frac{\partial V^{r,1}}{\partial c} &= -\frac{\gamma}{r} - \frac{\gamma}{q\alpha_2}(1-\alpha_2)(\frac{-\alpha_2 q}{(1-\alpha_2)r})^{1-\alpha_2}(K+c)^{-\alpha_2} \\ &= -\frac{\gamma}{r}(1-(\frac{s_0}{s_{1d}^r(c)})^{\alpha_2}) < 0. \end{aligned} \tag{A.29}$$

The last step is true since $s_0 > s_{1d}^r(c)$; otherwise, it would be trivial since the firm would default on date 0 and the equity value would be equal to 0. □

Lemma EC.7. *If $\hat{c}(\delta) < c_f^*(\delta)$, $\mathcal{A}^f(s; \hat{c}(\delta), \delta)$ decreases in δ.*

Proof: Note $\frac{d\mathcal{A}^f}{d\delta} = \frac{\partial \mathcal{A}^f}{\partial c} \cdot \frac{\partial \hat{c}}{\partial \delta} + \frac{\partial \mathcal{A}^f}{\partial \delta}$. It is clear that $\frac{\partial \mathcal{A}^f}{\partial \delta} < 0$, and we have shown $\frac{\partial \hat{c}}{\partial \delta} < 0$ (Proposition 3). Moreover, since $\hat{c} < c_f^*$, by convexity of $\mathcal{A}^f$,we have $\frac{\partial \mathcal{A}^f}{\partial c}|_{c=\hat{c}} > 0$. Therefore, $\frac{d\mathcal{A}^f(s;\hat{c}(\delta),\delta)}{d\delta} < 0$. □

Now, we prove Theorem 1 based on the above technical lemmas.

Proof of Theorem 1: From the proof of Proposition 3, we know that the comparison between $V^{f,1}(s; c_f^*(\delta), \delta)$ and $V^{r,1}(s; c_f^*(\delta))$ is independent of s. Therefore, if we can show $V^{f,1}(s_{10}; c_f^*(\delta), \delta) \ge V^{r,1}(s_{10}; c_f^*(\delta))$, the inequality holds for all s. Towards this goal, define $\underline{\delta} = \sup\{\delta | V^{f,1}(s_{10}(\delta); c_f^*(\delta), \delta) \ge V^{r,1}(s_{10}(\delta); c_f^*(\delta))\}$. Since $V^{f,1}((s_{10}(\delta); c_f^*(\delta), \delta)$ and $V^{r,1}(s_{10}(\delta); c_f^*(\delta))$ are continuous in δ, $V^{f,1}(s_{10}(\underline{\delta}); c_f^*(\underline{\delta}), \underline{\delta}) = V^{r,1}(s_{10}(\underline{\delta}); c_f^*(\underline{\delta}))$. Hence, we need to show that for any $0 < \delta < \underline{\delta}$, $V^{f,1}(s_{10}(\delta); c_f^*(\delta), \delta) \ge V^{r,1}(s_{10}(\delta); c_f^*(\delta))$.

From Lemma EC.5, there exists a $\hat{\Delta}$ that is independent of δ, s.t., for any $\Delta < \hat{\Delta}$, $V^{f,1}(s_{10}(\underline{\delta}-\Delta); c_f^*(\underline{\delta}-\Delta), \underline{\delta}-\Delta) > V^{f,1}(s_{10}(\underline{\delta}-\Delta); c_f^*(\underline{\delta}), \underline{\delta})$. From Lemma EC.6, we have $V^{r,1}(s_{10}(\underline{\delta}-\Delta); c_f^*(\underline{\delta}-\Delta)) < V^{r,1}(s_{10}(\underline{\delta}-\Delta); c_f^*(\underline{\delta}))$. Therefore, $V^{f,1}(s_{10}(\underline{\delta}-\Delta); c_f^*(\underline{\delta}-\Delta), \underline{\delta}-\Delta) > V^{r,1}(s_{10}(\underline{\delta}-\Delta); c_f^*(\underline{\delta}-\Delta))$.

Now repeat the process, with the same $\hat{\Delta}$. For any $\Delta < \hat{\Delta}$, $V^{f,1}(s_{10}(\underline{\delta}-2\Delta); c_f^*(\underline{\delta}-2\Delta), \underline{\delta}-2\Delta) > V^{f,1}(s_{10}(\underline{\delta}-2\Delta); c_f^*(\underline{\delta}-\Delta), \underline{\delta}-\Delta)$. And from Lemma EC.6, we have $V^{r,1}(s_{10}(\underline{\delta}-2\Delta); c_f^*(\underline{\delta}-2\Delta)) < V^{r,1}(s_{10}(\underline{\delta}-2\Delta); c_f^*(\underline{\delta}-\Delta))$. Therefore, $V^{f,1}(s_{10}(\underline{\delta}-2\Delta); c_f^*(\underline{\delta}-2\Delta), \underline{\delta}-2\Delta) > V^{r,1}(s_{10}(\underline{\delta}-2\Delta); c_f^*(\underline{\delta}-2\Delta))$.

Since $\hat{\Delta}$ is a constant that is independent of δ, we can repeat this process of subtracting Δ from δ until it hits 0 (i.e., $s_{10}(\delta)$ hits $s_{10}(0) = K$). In other words, we find a finite covering in $[0, \underline{\delta}]$ such that we have $V^{f,1}(s_{10}(\delta); c_f^*(\delta), \delta) > V^{r,1}(s_{10}(\delta); c_f^*(\delta))$ for all $\delta \in [0, \underline{\delta}]$, and the equity holders choose the flexible policy. In other words, $\hat{c}(\delta) > \underline{\delta}$. Therefore, we only need to show that from the firm's perspective, $\mathcal{A}^f(s; c_f^*(\delta), \delta) > \mathcal{A}^r(s; c_r^*)$, which is true because the unconstrained flexibility always has a non-negative value. Therefore, when $\delta < \underline{\delta}$, it is optimal for the firm to set $c = c_f^*$ and the equity holders will operate the firm with flexibility.

Next, let $\bar{\delta} = \inf\{\delta > \underline{\delta} | \mathcal{A}^f(s;\hat{c}(\delta),\delta) \leq \mathcal{A}^r(s;c_r^*)\}$. Note that the firm chooses between $\mathcal{A}^f(s;\hat{c}(\delta),\delta)$ and $\mathcal{A}^r(s;c_r^*)$ when deciding its financial choice. Clearly, when $\delta \to \infty$, the flexibility will never been used. It follows from Lemma EC.7 that $\mathcal{A}^f(s;\hat{c}(\delta),\delta)$ decreases in δ and $\mathcal{A}^r(s;c_r^*)$ is independent of δ, and thus there is a single crossing point $\bar{\delta}$ of $\mathcal{A}^f(s;\hat{c}(\delta),\delta)$ and $\mathcal{A}^r(s;c)$.

By the definition of $\hat{c}$, $c_f^*(\underline{\delta}) = \hat{c}(\underline{\delta})$. Therefore, by the monotonicity of $\hat{c}(\delta)$ shown in Proposition 3 and Lemma EC.3, $c^*(\delta)$ decreases in δ on $[0,\underline{\delta}]$. Based on Lemma EC.4, $c^*(0) = c_f^*(0) < c_r^*(\delta)$. Hence, $c_r^* > c^*(\delta)$, for any $\delta < \bar{\delta}$. In sum, we have shown that c^* decreases in δ in $[0,\bar{\delta}]$ and then jumps up to a flat level c_r^*. And by the definition of $\hat{c}(\delta)$, $c^*(\delta)$ leads to flexible operations on $[0,\bar{\delta}]$ and to rigid operations when $\delta > \bar{\delta}$. Thus, this completes the proof for Theorem 1. □

A.10. First Best Optimal Coupon

In the following, we show the optimal coupon rate of the first-best solution.

THEOREM EC.1. *There exist two thresholds $\underline{\delta}^{FB}$ and $\bar{\delta}^{FB}$ with respect to the switching cost such that the optimal coupon rate of the first-best solution follows:*

$$c_{FB}^*(\delta) = \begin{cases} c_f^* & \text{if } \delta < \underline{\delta}^{FB} \\ \bar{c} & \text{if } \underline{\delta}^{FB} \leq \delta \leq \bar{\delta}^{FB} \\ c_r^* & \text{if } \delta \geq \bar{\delta}^{FB} \end{cases}. \tag{A.30}$$

Proof: Let $\underline{\delta}^{FB} = \sup\{\delta | V^{f,1}(s_{10}(\delta);c_f(\delta),\delta) \geq 0)\}$. To show the threshold property for $\underline{\delta}^{FB}$, we show that for any $\delta < \underline{\delta}^{FB}$, $V^{f,1}(s_{10}(\delta);c_f^*(\delta),\delta) \geq 0$. We apply Lemma EC.5 and follow similar steps in the proof of Theorem 1. First, there exists a $\hat{\Delta}$ that is independent of δ, s.t., for any $\Delta < \hat{\Delta}, V^{f,1}(s_{10}(\underline{\delta}^{FB} - \Delta);c_f^*(\underline{\delta}^{FB} - \Delta),\underline{\delta}^{FB} - \Delta) > V^{f,1}(s_{10}(\underline{\delta}^{FB} - \Delta);c_f^*(\underline{\delta}^{FB}),\underline{\delta}^{FB})$. Since $V^{f,1}(s_{10}(\underline{\delta}^{FB} - \Delta);c_f^*(\underline{\delta}^{FB}),\underline{\delta}^{FB}) > V^{f,1}(s_{10}(\underline{\delta}^{FB});c_f^*(\underline{\delta}^{FB}),\underline{\delta}^{FB}) \geq 0$, $V^{f,1}(s_{10}(\underline{\delta}^{FB} - \Delta);c_f^*(\underline{\delta}^{FB} - \Delta),\underline{\delta}^{FB} - \Delta) \geq 0$. Given that $\hat{\Delta}$ is a constant that is independent of δ, we can repeat this process of subtracting Δ from δ in finite times until it hits 0. That is, we can find a finite covering, where $V^{f,1}(s_{10}(\delta);c_f^*(\delta),\delta) \geq 0$. Therefore, this inequality holds for any $\delta \in [0,\underline{\delta}^{FB}]$.

Next, let $\bar{\delta}^{FB} = \inf\{\delta > \underline{\delta}^{SB} | \mathcal{A}^f(s;\bar{c}(\delta),\delta) \leq \mathcal{A}^r(s;c_r^*)\}$. To show the threshold property for $\bar{\delta}^{FB}$, we need to show that for any $\delta > \bar{\delta}^{FB}$, $\mathcal{A}^f(s;\bar{c}(\delta),\delta) \leq \mathcal{A}^r(s;c_r^*)$. Since $\mathcal{A}^r(s;c_r^*)$ is independent of δ, we only need to show $\mathcal{A}^f(s;\bar{c}(\delta),\delta)$ decreases in δ. Taking derivative of $\mathcal{A}^f(s;\bar{c}(\delta),\delta)$, $\frac{d\mathcal{A}^f(s;\bar{c}(\delta),\delta)}{d\delta} = \frac{\partial\mathcal{A}^f(s;\bar{c}(\delta),\delta)}{\partial\delta} + \frac{\partial\mathcal{A}^f(s;\bar{c}(\delta),\delta)}{\partial c} \cdot \frac{\partial\bar{c}(\delta)}{\partial\delta}$. Note that $\frac{\partial\mathcal{A}^f}{\partial\delta} < 0$, $\frac{\partial\mathcal{A}^f(s;c,\delta)}{\partial c}|_{c=\bar{c}(\delta)} > 0$, and $\bar{c}(\delta) < c_f^*$ for $\delta > \underline{\delta}^{FB}$, hence we only need to show $\frac{\partial\bar{c}(\delta)}{\partial\delta} < 0$. Since $V(s_{10}(\delta);\bar{c}(\delta),\delta) = 0$, $\frac{dV(s_{10}(\delta);\bar{c}(\delta),\delta)}{d\delta} = 0$. $\frac{dV(s_{10}(\delta);\bar{c}(\delta),\delta)}{d\delta} = \frac{\partial V^{f,1}}{\partial s} \cdot \frac{\partial s_{10}(\delta)}{\partial\delta} + \frac{\partial V^{f,1}}{\partial c} \cdot \frac{\partial\bar{c}(\delta)}{\partial\delta} + \frac{\partial V^{f,1}}{\partial\delta}$. Note $\frac{\partial V^{f,1}}{\partial s} \geq 0$ and $\frac{\partial V^{f,1}}{\partial\delta} \leq 0$, and we also have $\frac{\partial s_{10}(\delta)}{\partial\delta} < 0$ (Corollary 1) and $\frac{\partial V^{f,1}}{\partial c} < 0$ (shown in the proof of Corollary 2). Hence, $\frac{\partial\bar{c}(\delta)}{\partial\delta} < 0$. Therefore, $\mathcal{A}^f(s;\bar{c}(\delta),\delta)$ decreases in δ and there is a single crossing between $\mathcal{A}^f(s;\bar{c}(\delta),\delta)$ and $\mathcal{A}^r(s;c_r^*)$, which completes the proof. □

A.11. Optimal investment timing

Denote the value of the growth option as $G(s;c) = \psi \cdot s_0^{\alpha_1}$, where ψ is to be determined later by the continuity of the value function and the optimality of the investment. The investment point U and the debt level are determined jointly to maximize the firm's value. As shown in Proposition 3, when the coupon rate $c \le \hat{c}$, the firm will be operated under the flexible policy. The value matching and smooth pasting conditions lead to:

$$\psi^f \cdot U_f^{\alpha_1} = -(K+c)\tfrac{\gamma}{r} + \tfrac{\gamma}{q}U_f + \theta U_f^{\alpha_2} + \tfrac{c}{r}(1-(\tfrac{U_f}{s_{0d}})^{\alpha_2}) - I, \tag{A.31}$$

$$\alpha_1\psi^f \cdot U_f^{\alpha_1} = \tfrac{\gamma}{q}U_f + \theta\alpha_2 U_f^{\alpha_2} - \alpha_2\tfrac{c}{r}(\tfrac{U_f}{s_{0d}})^{\alpha_2}. \tag{A.32}$$

Therefore, (A.31) $*\alpha_1-$ (A.32) results in:

$$-\alpha_1\left((K+c)\tfrac{\gamma}{r} - \tfrac{c}{r} + I_1\right) + \tfrac{\gamma}{q}(\alpha_1-1)U_f + \left(\theta - \tfrac{c}{r\cdot s_{0d}^{\alpha_2}}\right)\cdot(\alpha_1-\alpha_2)(U_f)^{\alpha_2} = 0.$$

Using the value matching condition, the coefficient can be solved by:

$$\psi^f = -\big((K+c)\tfrac{\gamma}{r} + I - \frac{c}{r}\big)U_f^{-\alpha_1} + \tfrac{\gamma}{q}U_f^{1-\alpha_1} + \big(\theta - \tfrac{c}{r}(\tfrac{1}{s_{0d}})^{\alpha_2}\big)U_f^{\alpha_2-\alpha_1}.$$

When the coupon level is $c > \hat{c}$, the value matching and smoothing pasting conditions lead to:

$$\psi \cdot U_r^{\alpha_1} = -(K+c)\tfrac{\gamma}{r} + \tfrac{\gamma}{q}U_r - \frac{\gamma}{q\alpha_2}s_{1d}^{1-\alpha_2}U_r^{\alpha_2} + \tfrac{c}{r}(1-(\tfrac{U_r}{s_{1d}})^{\alpha_2}) - I$$

$$\alpha_1\psi \cdot U_r^{\alpha_1} = \tfrac{\gamma}{q}U_r - \frac{\gamma}{q}\cdot\left(\frac{U_r}{s_{1d}}\right)^{\alpha_2}\cdot s_{1d} - \tfrac{c}{r}\alpha_2(\tfrac{U_r}{s_{1d}})^{\alpha_2} = \tfrac{\gamma}{q}U_r - \left(\frac{\gamma}{q}\cdot\left(\frac{1}{s_{1d}}\right)^{\alpha_2-1} + \alpha_2\tfrac{c}{r}(\tfrac{1}{s_{1d}})^{\alpha_2}\right)U_r^{\alpha_2}.$$

U^r can be solved by:

$$-\alpha_1\left((K+c)\tfrac{\gamma}{r} - \frac{c}{r} + I\right) + (\alpha_1-1)\tfrac{\gamma}{q}U_r - (\alpha_1-\alpha_2)\left(\frac{1}{\alpha_2}\frac{\gamma}{q}\cdot s_{1d} + \frac{c}{r}\right)\cdot\left(\frac{U_r}{s_{1d}}\right)^{\alpha_2} = 0 \tag{A.33}$$

and ψ can be solved by:

$$\psi^r = -\big((K+c)\tfrac{\gamma}{r} + I - \frac{c}{r}\big)U_r^{-\alpha_1} + \tfrac{\gamma}{q}U_r^{1-\alpha_1} + \left(\rho - \frac{c}{r}(\tfrac{1}{s_{1d}})^{\alpha_2}\right)U_r^{\alpha_2-\alpha_1}. \tag{A.34}$$

B. Empirical Analysis and Robustness Check

B.1. Empirical Test on the Interplay of Production Flexibility and Financial Leverage in the Iron and Steel Industry

As discussed in the introduction and literature review sections, most empirical studies on production flexibility and financial leverage use indirect measures of flexibility that face multiple problems. Indirect measures can only capture realized flexibility, but its existence rather than its realization influences financing decisions. Moreover, indirect measures often are influenced by accounting rules and may proxy for other factors. For example, proxies typically do not differentiate between operational flexibility and operational leverage. Operational leverage is defined as the proportion of the fixed production cost to the total cost (Van Horne 1977; Ferri and Jones 1979; Huffman 1983; Dugan et al. 1994; Novy-Marx 2011). The fixed cost is to be paid regardless of production quantity and needs to be paid continuously as a rate per unit time. It plays a different role from switching costs that we use as a measure of flexibility. One of the differences is that operational leverage is not directly linked to risk shifting since the fixed cost is an ongoing cost. In contrast, operational flexibility refers to the ease of altering production quantities, and since the timing of switching is controlled by the equity holders, risk shifting problems are encountered. Endogeneity issues are more of a concern when indirect measures are used.

To avoid these problems, Reinartz and Schmid (2016) obtain direct measures on flexibility based on properties at the plant level of energy utility firms. Their tests involve firms in the energy utility industry, which is heavily regulated. It is unclear whether these results extend to other sectors. In this supplement, we conduct a similar analysis for the iron and steel industry. We choose this industry because there are two well developed technologies that have very different degrees of flexibility and make up the bulk of production: the *Basic Oxygen Furnace* (BOF) process and the *Electric Arc Furnace* (EAF) process. The steel facilities using the BOF process are typically called integrated mills. They make steel from scratch, by melting iron ore in a blast furnace to produce liquid iron, removing the slag, and then combining the liquid iron with scrap metal in a Basic Oxygen Furnace, where pure oxygen is blown in at supersonic velocities. The BOF process relies on large volumes to achieve efficiency and is inflexible. In contrast, the EAF process is a more flexible modern production method. It uses steel scrap and pig iron as raw materials and uses electricity to melt them to make new steel. The steel mills using the EAF process are typically called mini mills. They produce steel in smaller volumes and can stop and restart production relatively quickly. In the United States, approximately 75% of the steel is made using the EAF process, and 25% is made using BOF. [7]

To develop the empirical test, we focus on the U.S. listed steel manufacturers. We first obtain their financial and operational data from the COMPUSTAT quarterly database. Let i be the index for firms, t for fiscal years, and q for fiscal quarters. We denote $SALES_{itq}$ as the sales revenue, $COGS_{itq}$ as the cost of goods sold, TA_{itq} as the total assets, $MKVLT_{itq}$ as the market value of equity, $DLTT_{itq}$ as the long-term debt, DLC_{itq} as the short-term debt, $OIBDPQ_{itq}$ as Operating Income Before Depreciation, and $PPENT_{itq}$ as total Property Plant and Equipment.

Second, to construct a direct measure of production flexibility, we utilize the recent Greenhouse Gas (GHG) dataset in the Envirofacts database obtained from the EPA website.[8] Starting from 2010, the Greenhouse Gas Reporting Program (GHGRP) requires firms to provide their greenhouse gas data from large emission sources in the United States. This dataset can be used to track and compare facilities' greenhouse gas emissions. Facilities are required to submit annual reports if greenhouse gas emissions from covered sources exceed 25,000 metric tons of carbon dioxide, CO_2, per year. In the iron and steel industry, about 40 percent of the carbon used in steelmaking leaves the process as carbon dioxide gas. The required reports not only provide the aggregated level of emissions for each applicable facility, but also the type of emitting facility. Specifically, each facility reports the emission volume of CO_2 and other greenhouse gases by specific furnace type (BOF or EAF). This provides us with detailed information at the facility level of the type of steel-making technology that firms in this industry are using.

We retrieve the information from Greenhouse Gas Customized Search for the iron and steel industry. Each facility is required to report its highest level US parent company. We combine the facilities according to their parent company and then manually match them with those in the quarterly COMPUSTAT dataset that we extract for the iron and steel industry (SIC code from 3200-3299) from 2010 to 2019, which leads to 403 matched quarterly observations. Table EC.1 displays the total amounts of emissions

---

[7] https://www.conklinmetal.com/steelmaking-101/

[8] The data is available throught https://www.epa.gov/enviro/greenhouse-gas-customized-search

	Total Emissions		Matched Emissions		Matched Percentage	
Year	EAF	BOF	EAF	BOF	EAF	BOF
2010	3.418	4.034	2.721	3.855	79.60%	95.58%
2011	4.096	4.268	3.192	4.039	77.92%	94.64%
2012	4.179	4.165	3.319	4.165	79.41%	100.00%
2013	4.396	4.215	3.509	4.215	79.82%	100.00%
2014	4.595	4.082	3.709	4.082	80.73%	100.00%
2015	4.387	3.466	3.526	3.466	80.38%	100.00%
2016	5.156	3.253	4.054	3.253	78.63%	100.00%
2017	5.048	3.478	4.112	3.478	81.47%	100.00%
2018	6.081	3.420	4.775	3.420	78.54%	100.00%
2019	5.600	3.320	4.452	3.320	79.50%	100.00%
Total	4.696	3.770	3.737	3.729	79.59%	98.92%

Table EC.1 The total emissions (in millions of metric tons) and the portions matched to COMPUSTAT Quarterly

reported in the GHG dataset and the amounts of emissions from the firms that we are able to match with the COMPUSTAT dataset. Most of the major steel makers in the US are public firms and are included in our analysis.

For the measure of financial leverage, we consider both market and book leverages defined as: $MLV_{itq} = \frac{DLC_{itq}+DLTT_{itq}}{DLC_{itq}+DLTT_{itq}+MKVLT_{itq}}$ and $BLV_{itq} = \frac{DLC_{itq}+DLTT_{itq}}{TA_{itq}}$, where the numerator is the total debt. The market leverage is generally more volatile than the book leverage since the former contains the market value of equity. Both measures are widely used in the literature to reflect the amount of debt used by a firm to finance its business.

For the measure of production flexibility, we define two variables. The first is the ratio of a firm's emissions contributed by the flexible EAF process to its total emissions in each year:

$$FE_{it} = \frac{\text{total emission of EAF}}{\text{total emission of EAF+total emission of BOF}},$$

which reflects how much a firm relies on the flexible EAF process. It is worth noting that to produce the same amount of steel, the BOF process typically releases more CO_2 than does the EAF process, varying from one to three times. Therefore, in our robustness test, we also alter the definition to place a greater weight on the EAF emission. The second variable is a dummy variable FEI_{it}, that equals 1 if the firm only uses the EAF process and 0 otherwise. We do not separate out the firms that only use the BOF process due to the small number of such firms. Therefore, when $FEI_{it} = 0$, the firm may use the BOF process only, or use both BOF and EAF processes. For such firms with $FEI_{it} = 0$, on average, about 86.7 percent of the emissions are from the BOF process, and only one firm-year observation in our matched dataset has the percentage of emissions from the BOF process below 50 percent.

We use the same control variables as in Reinartz and Schmid (2016) that are standard in the finance literature (Lemmon et al. 2008; Danis et al. 2014). Specifically, profitability is computed as the earnings before interest, taxes, depreciation and amortization, divided by total assets. The market to book value is computed as the sum of the market value of equity and the book value of debt, divided by total assets. Tangibility is computed as the value of property, plant and equipment, divided by total assets. Firm size is defined as the logarithm of net sales (or the logarithm of total assets in robustness tests).

Table EC.2 shows the descriptive statistics for the dependent and independent variables used in our empirical tests. We observe that on average, the firms that use the

Variable	(Mainly depend on) BOF		EAF	
	Mean	Standard Deviation	Mean	Standard Deviation
Book Leverage	0.288	0.122	0.223	0.129
Market Leverage	0.506	0.152	0.247	0.142
Profitability	0.0178	0.0156	0.0278	0.028
Tangibility	0.450	0.036	0.376	0.133
Market to book	0.578	0.164	0.943	0.312
log(AT)	9.768	1.328	7.319	1.727
log(sales)	8.465	1.045	6.084	1.046
Obs	120		283	

Table EC.2 Summary statistics

This table reports the summary statistics of financial and operational characteristics for the two different types of firms.

BOF process are more financially levered than the counterparts in both leverage measures.[9] We also find that the firms in the two groups differ largely in their profitability, tangibility, and market to book value. Those using the inflexible BOF process are generally larger, with lower profitability, higher tangibility, and significantly lower market to book value. Similar observations are made in Reinartz and Schmid (2016) for the power industry, which indicates that it is important to control for these factors.

We test the empirical specification as in Reinartz and Schmid (2016):

$$Lev_{itq} = \lambda_t + \eta_q + \beta_1 \cdot F_{it} + \gamma \cdot X_{itq} + \varepsilon_{itq}, \tag{B.1}$$

where Lev_{itq} is the financial leverage measure (either MLV_{itq} or BLV_{itq}), F_{it} is the flexibility measure (either FE_{it} or FEI_{it}), and X_{itq} is the control variable vector including profitability, tangibility, market to book value, and firm size. All control variables are lagged for one quarter. We control for the fixed year (λ_t) and fixed quarter (η_q) effects.

Table EC.3 reports the regression results. The estimated coefficients of the main model (B.1) with four different combinations among the two leverage measures and the two flexibility measures are provided in columns (1-2) and columns (6-7), respectively. We see that the coefficient for the flexibility variable is always negative under every combination, which is statistically significant except under model specification (7). This finding indicates that the usage of the flexible EAF process is associated with less financial leverage with other important factors controlled, which confirms the presence of a negative association between financial leverage and production flexibility in the iron and steel industry. In particular, based on model specification (2) where the dummy flexibility variable is used, a hypothetical steel maker that uses only the EAF process would have a market leverage about 10.3 percent lower than an otherwise identical firm that significantly relies on the BOF process, which makes a significant difference given that the average market leverage for the EAF-only firms is at 24.7 percent in our dataset. From the other columns in Table EC.3, we see that this result is robust when we use the alternative firm size definition, add a square term FE_{it}^2 into the model, or lag all the independent variables. Moreover, when we add the square term FE_{it}^2 into the model, the associated coefficient is positive, which indicates that the relationship between financial leverage and production flexible can be nonlinear (convex). Therefore, these findings can complement Reinartz and Schmid (2016) in that higher operational flexibility can be associated with less debt, and their relationship might be nonlinear.

---

[9] The big difference between the book leverage and market leverage stems from the lower stock price for the traditional steel makers that depend on BOF. This is consistent with some financial analysts calling some steel firms in this category “value stocks” based on the low market-to-book ratio (https://www.nasdaq.com/articles/5-value-stocks-with-low-price-to-book-ratio-for-big-returns-2017-05-31).

	Market Leverage					Book Leverage				
Model	1	2	R3	R4	R5	6	7	R8	R9	R10
FE	-0.122***	-0.103***	-0.151***	-0.169***	-0.272**	-0.040**	-0.020	-0.065**	-0.060***	-0.437***
or FEI	(0.024)	(0.021)	(0.023)	(0.023)	(0.134)	(0.020)	(0.018)	(0.199)	(0.120)	(0.113)
FE^2					0.101					0.371***
					(0.128)					(0.109)
Profitability	0.669	0.641	0.767	0.705	0.732	0.548	0.592	0.628	0.480	0.580
	(0.475)	(0.473)	(0.486)	(0.456)	(0.458)	(0.391)	(0.385)	(0.412)	(0.387)	(0.392)
Tangibility	0.506***	0.526***	0.455***	0.605***	0.608***	0.296***	0.309***	0.264***	0.321***	0.333***
	(0.059)	(0.058)	(0.061)	(0.056)	(0.058)	(0.050)	(0.049)	(0.051)	(0.049)	(0.048)
Market to Book	-0.130***	-0.132***	-0.111***	-0.084***	-0.093***	0.121***	0.111***	0.140***	0.129***	0.096***
	(0.027)	(0.028)	(0.028)	(0.027)	(0.029)	(0.023)	(0.023)	(0.023)	(0.023)	(0.025)
Size	0.033***	0.032***	0.027***	0.026***	0.028***	0.025***	0.027***	0.019***	0.022***	0.028***
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.003)	(0.003)	(0.004)
Year Fixed Eff.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Qtr Fixed Eff.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R^2	0.616	0.614	0.596	0.660	0.661	0.413	0.412	0.380	0.412	0.430

Table EC.3 Preliminary analysis and baseline model regression

This table reports the coefficients of the following regression models:

1. Models 1–2: $MLV_{itq} = \eta_q + \lambda_t + \beta_1 \cdot FE_{it} + \gamma \cdot X_{itq} + \varepsilon_{it}$ and $MLV_{itq} = \eta_q + \lambda_t + \beta_1 \cdot FEI_{it} + \gamma \cdot X_{itq} + \varepsilon_{it}$
2. Model 3 is the same as Model 1, except that the firm size is defined as the logarithm of total assets. This is a replication of the model in Reinartz and Schmid (2016).
3. Model 4 is the same as Model 1, except that all the independent variables are lagged for one year. Model 5 adds a square term FE_{it}^2 as an independent variable.
4. Models 6-10 are the same as Models 1-5 except that the dependent variable is changed to BLV_{itq}.

In the above models, X represents the control variable vector including profitability, tangibility, market to book value, and firm size. The sample includes all the firm-quarter observations from 2010 to 2019. ***significant at 0.01. ** significant at 0.05. * significant at 0.1. P-values are two-sided. Standard errors are reported in parentheses below the parameter estimates.