Online Appendix:
Signaling Safety

Roni Michaely, Stefano Rossi, and Michael Weber

Not for Publication

1. Theoretical Appendices

In this Section we present our theoretical proofs. Appendix 1.1 solves the signaling model of Section VA. Appendix 1.2 states and proves the main comparative statics results of the signaling model. Appendix 1.3 presents an example with the log production function. Appendix 1.4 verifies that the six assumptions given by Riley (1979) for signaling games hold in our framework. Henceforth we refer to the best separating equilibrium outcome discussed in the text as the "Riley outcome". Appendix 1.5 verifies that the assumptions of Theorem 1, Theorem 2 and Corollary of Mailath (1987) hold for our signaling model, which implies that the Riley outcome is the unique separating equilibrium of our model. Appendix 1.6 verifies that the assumptions of Theorem 1 of Esö and Schummer (2009) hold for our signaling model, which implies that the Riley outome is the unique equilibrium that survives the "credible deviations" refinement (Esö and Schummer (2009); see also Cho and Sobel (1990) and Ramey (1996)). Appendix 1.7 states Theorem 1, Theorem 2 and Corollary of Mailath (1987). Appendix 1.8 states Theorem 1 of Esö and Schummer (2009). Appendix 1.9 considers the baseline setting, both with a general Arrow-Pratt certainty equivalent formulation and in the CARA special case. Appendix 1.10 considers the additive formulation of the agency model. Appendix 1.11 studies the multiplicative formulation of the agency model. Appendix 1.12 proves the comparative statics of signaling and agency models relative to investment opportunities. Appendix 1.13 considers estimates of our production function and compares them with the literature.

1.1. Solving the Signaling Model

Assume we can associate to each level of variance σ^2 a level of dividends D_1 that solves the optimization problem of the manager. We write this correspondence as $\sigma^2(D_1)$. If $\sigma^2(D_1)$ is single-valued and if the market is rational, we get the following condition,

$$V^s(D_1) = V^h(\sigma^2(D_1), D_1) = V^h(\sigma^2, D_1).$$

We then obtain

$$\begin{aligned} V^s(-\sigma^2(D_1), D_1) &= D_1 + R \cdot f(\omega_1 + Y - D_1 - \frac{a}{2}\sigma^2(D_1)), \\ V^h(-\sigma^2, D_1) &= D_1 + R \cdot f(\omega_1 + Y - D_1 - \frac{a}{2}\sigma^2). \end{aligned}$$

Valuation schedules that satisfy the conditions above and solve the maximization problem of the manager are termed "informationally consistent price functions" (Riley (1979)). To find the Pareto-dominant schedule, we start from the boundary condition that the worst firm that has the highest variance, that is, σ^2_{max}, will choose the same optimal dividend D_1 as it would in the full-information case, so that

$$
\begin{aligned}
1 - R \cdot f'(\omega_1 + Y - D_1^\star - \frac{a}{2}\sigma^2) &= 0 \\
\sigma^2(D_1{}^\star) &= \sigma^2_{max}.
\end{aligned}
$$

Because $V^s(D_1) = V^h(\sigma^2(D_1), D_1) = V^h(\sigma^2, D_1)$, the first-order condition is

$$kV^h_{-\sigma^2}(-\sigma^2(D_1), D_1)\frac{\partial(-\sigma^2)}{\partial D_1} + kV^h_d(-\sigma^2(D_1), D_1) + (1-k)V^h_d(-\sigma^2, D_1) = 0.$$

Given $\sigma^2(D_1) = \sigma^2$, the first-order condition is equivalent to the condition

$$kV^h_{-\sigma^2}(-\sigma^2(D_1), D_1)\frac{\partial(-\sigma^2)}{\partial D_1} + V^h_d(-\sigma^2, D_1) = 0;$$

that is,

$$1 - R \cdot f'(\omega_1 + Y - D_1 - \frac{a}{2}\sigma^2) - \frac{ka}{2} \cdot R \cdot f'(\omega_1 + Y - D_1 - \frac{a}{2}\sigma^2(D_1)) \cdot \frac{\partial \sigma^2(D_1)}{\partial D_1} = 0.$$

Then the ordinary differential equation (ODE) together with the boundary condition above uniquely determine the schedule. The worst firm type with the highest variance, σ^2_{max}, sets dividends $D_1^\star$ as in the first-best, full-information case. As variance decreases, firms pay more dividends and forego more investment opportunities. Therefore, relative to the first-best case with full information, the signaling equilibrium features excessive dividend payment and under-investment. Figure 2 illustrates the equilibrium.

We can establish the relevant solution informally by checking the second-order conditions for a maximum of the optimization problem of the manager

$$\frac{\partial}{\partial D_1}\left[kV^h_{-\sigma^2}(-\sigma^2(D_1), D_1)\frac{\partial(-\sigma^2)}{\partial D_1} + kV^h_d(-\sigma^2(D_1), D_1) + (1-k)V^h_d(-\sigma^2, D_1)\right] < 0.$$

Substituting the first-order condition leads to a simple condition guaranteeing a maximum,

$$-V^h_{d\sigma^2}(\sigma^2, D_1)\frac{\partial \sigma^2}{\partial D_1} < 0.$$

Because $V^h_{d\sigma^2}(\sigma^2, D_1) = \frac{a}{2} \cdot R \cdot f''(\omega_1 + Y - D_1 - \frac{a}{2}\sigma^2) < 0$, a maximum occurs if and only if $\frac{\partial \sigma^2}{\partial D_1} < 0$. The relevant solution must lie on the right-hand side of the red line in Figure 2 where higher dividends are associated with a lower variance of cash flow. Thus, only dividends that exceed $D_1^\star$ are optimal. This solution is the Riley equilibrium outcome. In the Appendix, we show this equilibrium is the unique separating equilibrium of our game, by applying the results of Mailath (1987), and we show it is the unique equilibrium that survives standard refinement concepts for this class of games (Esö and

Schummer (2009); see also Ramey (1996) and Cho and Sobel (1990)).

In this model, dividends are a signal to the market about the cash-flow volatility. Because managers care about short-term investors, they would like to signal that their cash flows have low volatility and therefore higher value. For this signal to be credible, it must be costly. To prevent imitation and thus generate a separating equilibrium, the signal must be costlier for low types than for high types. This conclusion follows from the concavity of the production function, because riskier firms have more to lose in terms of foregone investment if they pay a larger dividend in an attempt to imitate safer firms.

1.2. Proof of Comparative Statics

Here we state and prove the main comparative statics.

Prediction 1 (signaling - time series). The dividend changes should be followed by changes in future cash flow volatility in the opposite direction, i.e. $\frac{\partial \sigma^2(D_1)}{\partial D_1} < 0$

Proof. The proof is immediately given in the analysis of the schedules in the main text.

Combining the FOC and SOC of the manager's optimization problem we get a simple condition guaranteeing a maximum

$$-V^h_{d\sigma^2}(\sigma^2, D_1)\frac{\partial \sigma^2}{\partial D_1} < 0 \tag{A.1}$$

With $V^h_{d\sigma^2}(\sigma^2, D_1) = \frac{a}{2}Rf''(\omega_1 + Y - D_1 - \frac{a}{2}\sigma^2) < 0$, a maximum occurs if and only if $\frac{\partial \sigma^2}{\partial D_1} < 0$.

■

Prediction 2 (signaling - cross-section). Following a dividend increase (re. decrease), there's a larger decrease (re. increase) in cash flow volatility for firms with smaller (re. larger) current earnings, i.e. $\frac{\partial^2 \sigma(D_1)}{\partial D_1 \partial Y} > 0$

Proof. Recall the FOC,

$$1 - R \cdot f'(\omega_1 + Y - D_1 - \frac{a}{2}\sigma^2) - \frac{ka}{2} \cdot R \cdot f'(\omega_1 + Y - D_1 - \frac{a}{2}\sigma^2(D_1)) \cdot \frac{\partial \sigma^2(D_1)}{\partial D_1} = 0, \tag{A.2}$$

we get

$$\frac{\partial \sigma^2(D_1)}{\partial D_1} = \frac{1 - R \cdot f'(\omega_1 + Y - D_1 - \frac{a}{2}\sigma^2)}{\frac{ka}{2} \cdot R \cdot f'(\omega_1 + Y - D_1 - \frac{a}{2}\sigma^2(D_1))} \tag{A.3}$$

Then

$$
\begin{aligned}
\frac{\partial^2\sigma^2(D_1)}{\partial D_1 \partial Y} &= \frac{\partial\left(\dfrac{1 - R\cdot f'(\omega_1 + Y - D_1 - \frac{a}{2}\sigma^2)}{\frac{ka}{2}\cdot R\cdot f'(\omega_1+Y-D_1-\frac{a}{2}\sigma^2(D_1))}\right)}{\partial Y} \\
&= -\frac{2f''(\omega_1 + Y - D_1 - \frac{a}{2}\sigma^2(D_1))}{k\cdot a\cdot R\cdot \left[f'(\omega_1+Y-D_1-\frac{a}{2}\sigma^2(D_1)\right]^2} > 0
\end{aligned}
$$

because $f'' < 0$. ■

1.3. An Example

For this example, define $f(\omega_1 + Y - D_1 - \frac{a}{2}\cdot\sigma^2) = \ln\left(\omega_1 + Y - D_1 - \frac{a}{2}\cdot\sigma^2\right)$ and assume $R = 1$. The ODE (i.e. FOC)

$$
1 - f'(\omega_1 + Y - D_1 - \frac{a}{2}\sigma^2(D_1)) - \frac{ka}{2}\cdot f'(\omega_1 + Y - D_1 - \frac{a}{2}\sigma^2(D_1))\cdot\frac{\partial\sigma^2\left(D_1\right)}{\partial D_1} = 0 \tag{A.4}
$$

becomes

$$
\frac{\partial\sigma^2\left(D_1\right)}{\partial D_1} = \frac{2(\omega_1 + Y - D_1 - 1 - \frac{a}{2}\sigma^2)}{k\cdot a}. \tag{A.5}
$$

Together with the boundary condition which says the worst type chooses the dividend such that $\sigma^2(D_1^\star) = \sigma^2_{max}$, we get the solution to this problem,

$$
\sigma^2\left(D_1\right) = \frac{2(\omega_1 + Y - 1 - \frac{a}{2}\sigma^2)D_1 - D_1^2 + a\cdot k\cdot\sigma^2_{max} + D_1^{\star 2} - 2(\omega_1 + Y - 1 - \frac{a}{2}\sigma^2)D_1^\star}{k\cdot a} \tag{A.6}
$$

where $D_1 \geq D_1^\star$. It is then immediate to check that $\dfrac{\partial\sigma^2(D_1)}{\partial D_1} < 0$ and $\dfrac{\partial^2\sigma^2(D_1)}{\partial D_1\partial Y} > 0$.

1.4. Proof of Riley (1979) conditions

This Section shows that our signaling model of Sections IIB-IID satisfies the Riley (1979) conditions for signaling games. **Proof.** Let

$$
\begin{aligned}
W &= k\cdot V^s + (1-k)\cdot V^h \\
V^s &= D_1 + R\cdot f\left(\omega_1 + Y - D_1 - \frac{a}{2}\cdot\sigma^2(D_1)\right) \\
V^h &= D_1 + R\cdot f\left(\omega_1 + Y - D_1 - \frac{a}{2}\cdot\sigma^2\right)
\end{aligned}
$$

so that

$$
W = D_1 + k\cdot R\cdot f\left(\omega_1 + Y - D_1 - \frac{a}{2}\cdot\sigma^2(D_1)\right) + (1-k)\cdot R\cdot f\left(\omega_1 + Y - D_1 - \frac{a}{2}\cdot\sigma^2\right).
$$

Riley (1979) assumptions:

A1. The unobservable attribute, σ^2, is distributed on $\left[\sigma^2_{min}, \sigma^2_{max}\right]$ according to a strictly increasing distribution function

A2. The functions $W(\cdot)$, $V^h(\cdot)$ are infinitely differentiable in all variables

A3. $\frac{\partial W}{\partial V^s} > 0$

A4. $V^h\left(-\sigma^2, D_1\right) > 0$; $\frac{\partial V^h\left(-\sigma^2, D_1\right)}{\partial(-\sigma^2)} > 0$

A5. $\frac{\partial}{\partial(-\sigma^2)}\left(\frac{-\frac{\partial W}{\partial D_1}}{\frac{\partial W}{\partial V^s}}\right) < 0$

A6. $W\left(-\sigma^2; D_1, V^h\left(-\sigma^2, D_1\right)\right)$ has a unique maximum over D_1.

Assumptions **A1-A4** are immediate.

Condition **A5** is also known as the "single crossing condition" of signaling games and is that $\frac{\partial}{\partial(-\sigma^2)}\left(\frac{-\frac{\partial W}{\partial D_1}}{\frac{\partial W}{\partial V^s}}\right) < 0.$

$\frac{\partial W}{\partial D_1} = 1 - R \cdot f'\left(\omega_1 + Y - D_1 - \frac{a}{2} \cdot \sigma^2\right)$

$\frac{\partial W}{\partial V^s} = k.$

Hence:

$$\begin{aligned} \frac{\partial}{\partial\left(-\sigma^2\right)}\left(\frac{-\frac{\partial W}{\partial D_1}}{\frac{\partial W}{\partial V^s}}\right) &= \frac{\partial}{\partial\left(-\sigma^2\right)}\left(\frac{-1 + R \cdot f'\left(\omega_1 + Y - D_1 - \frac{a}{2} \cdot \sigma^2\right)}{k}\right) \\ &= \frac{a \cdot R \cdot f''\left(\omega_1 + Y - D_1 - \frac{a}{2} \cdot \sigma^2\right)}{2k} < 0 \end{aligned}$$

because $f''(\cdot) < 0$.

Condition **A6** requires that $V^h\left(-\sigma^2, D_1\right)$ has a unique maximum over D_1, which it does at the point $D_1^\star$ such that

$$R \cdot f'\left(\omega_1 + Y - D_1 - \frac{a}{2} \cdot \sigma^2\right) = 1 \tag{A.7}$$

with the S.O.C. $R \cdot f''\left(\omega_1 + Y - D_1 - \frac{a}{2} \cdot \sigma^2\right) < 0$ satisfied. ■

Because the Riley conditions are satisfied, we refer to the separating equilibrium of Section IIC in the main text as the "Riley equilibrium" and to the separating equilibrium outcome of Section IIC in the main text as the "Riley outcome".

1.5. Uniqueness of the Separating Equilibrium

This section shows that the Riley equilibrium is the unique separating equilibrium of our model.

According to Theorem 1, Theorem 2 and Corollary in Mailath (1987) (see Appendix D), if the payoff function satisfies Mailath (1987)'s conditions (1)-(5) and the single crossing condition (7), together with the initial value condition (6), then the Riley equilibrium is the unique separating equilibrium solution.

To begin with, in the dividend framework, the set of possible types is the interval $\left[\sigma^2_{min}, \sigma^2_{max}\right] \subset \mathbb{R}$ and the set of possible actions is $\mathbb{R}$. Let $\tau^{-1}(D_1) = \sigma^2(D_1)$ where $\tau : \left[\sigma^2_{min}, \sigma^2_{max}\right] \to \mathbb{R}$ is the proposed equilibrium one-to-one strategy.

Recall that

$$\begin{aligned} W &= k \cdot V^s + (1-k) \cdot V^h \\ V^s &= D_1 + R \cdot f\left(\omega_1 + Y - D_1 - \frac{a}{2} \cdot \sigma^2(D_1)\right) \\ V^h &= D_1 + R \cdot f\left(\omega_1 + Y - D_1 - \frac{a}{2} \cdot \sigma^2\right) \end{aligned}$$

so the expected payoff function is

$$\begin{aligned} W\left(-\sigma^2, -\sigma^2(D_1), D_1\right) &= D_1 + k \cdot R \cdot f\left(\omega_1 + Y - D_1 - \frac{a}{2} \cdot \sigma^2(D_1)\right) \\ &\quad +(1-k) \cdot R \cdot f\left(\omega_1 + Y - D_1 - \frac{a}{2} \cdot \sigma^2\right) \end{aligned}$$

As we already shown $\sigma^2(D_1)$ ($i.e., \tau^{-1}(D_1)$) solves the optimization problem, it satisfies *incentive compatibility*:

$(IC) \quad \tau\left(\sigma^2\right) \in \underset{D_1 \in \tau\left(\left[\sigma^2_{min}, \sigma^2_{max}\right]\right)}{\operatorname{argmax}} W\left(-\sigma^2, -\tau^{-1}\left(D_1\right), D_1\right), \ \forall \sigma^2 \in \left[\sigma^2_{min}, \sigma^2_{max}\right]$

Mailath (1987)'s regularity conditions on the payoff function W,

(1) $W\left(-\sigma^2, -\sigma^2(D_1), D_1\right)$ is C^2 on $\left[\sigma^2_{min}, \sigma^2_{max}\right]^2 \times \mathbb{R}$ (smoothness)

(2) W_2 never equals zero, and so is either positive or negative (belief monotonicity)

(3) W_{13} never equals zero, and so is either positive or negative (type monotonicity)

(4) $W_3\left(-\sigma^2, -\sigma^2, D_1\right) = 0$ has a unique solution in D_1, denoted $\phi\left(\sigma^2\right)$, which maximizes $W\left(-\sigma^2, -\sigma^2, D_1\right)$, and $W_{33}\left(-\sigma^2, -\sigma^2, \phi\left(\sigma^2\right)\right) < 0$ ("strict" quasi-concavity)

(5) there exists $k > 0$ such that for all $\left(-\sigma^2, D_1\right) \in \left[\sigma^2_{min}, \sigma^2_{max}\right] \times \mathbb{R}$, $W_{33}\left(-\sigma^2, -\sigma^2, D_1\right) \geq 0 \Rightarrow \mid W_3\left(-\sigma^2, -\sigma^2, D_1\right) \mid > k$ (boundedness)

The other two conditions which play a role in what follows are

(6) $\tau\left(\sigma^2_w\right) = \phi\left(\sigma^2_w\right)$, where $\sigma^2_w = \sigma^2_{max}$ if $W_2 > 0$ and σ^2_{min} if $W_2 < 0$ (initial value)

(7) $\frac{W_3\left(-\sigma^2, -\sigma^2(D_1), D_1\right)}{W_2\left(-\sigma^2, -\sigma^2(D_1), D_1\right)}$ is a strictly monotonic function of $-\sigma^2$ (single crossing).

Condition (1) is satisfied because it is obvious that $W\left(-\sigma^2, -\sigma^2(D_1), D_1\right)$ is C^2 on $\left[\sigma^2_{min}, \sigma^2_{max}\right]^2 \times \mathbb{R}$. Condition (2) is satisfied because W_2 is always negative and will never be zero.

$$W_2 = \frac{\partial W}{\partial(-\sigma^2(D_1))} = \frac{k \cdot a \cdot R}{2} f'\left(\omega_1 + Y - D_1 - \frac{a}{2} \cdot \sigma^2(D_1)\right) > 0 \tag{A.8}$$

since $f'(\cdot) > 0$.

Condition (3) is satisfied because $W_{13} < 0$ is always negative and will never equal zero.

$$\begin{aligned} W_{13} &= \frac{\frac{\partial W}{\partial(-\sigma^2)}}{\partial D_1} \\ &= \frac{\frac{(1-k)\cdot a\cdot R}{2} f'\left(\omega_1 + Y - D_1 - \frac{a}{2}\cdot\sigma^2\right)}{\partial D_1} \\ &= \frac{-(1-k)\cdot a\cdot R}{2} f''\left(\omega_1 + Y - D_1 - \frac{a}{2}\cdot\sigma^2\right) > 0 \end{aligned}$$

since $f''(\cdot) < 0$.

Condition (4) is satisfied because $f'(\cdot)$ is monotonic with $f''(\cdot) < 0$.

$$\begin{aligned} W_3\left(-\sigma^2, -\sigma^2, D_1\right) &= 0 \\ &\Longleftrightarrow \\ 1 - R\cdot f'\left(\omega_1 + Y - D_1 - \frac{a}{2}\cdot\sigma^2\right) &= 0 \end{aligned}$$

Since $f'(\cdot) > 0$ is monotonic, $W_3\left(-\sigma^2, -\sigma^2, D_1\right) = 0$ has a unique solution in D_1, denoted $\phi(\sigma^2)$. It is easy to show that $\phi(\sigma^2)$ also maximizes $W\left(-\sigma^2, -\sigma^2, D_1\right)$.

$W\left(-\sigma^2, -\sigma^2, D_1\right) = D_1 + R\cdot f\left(\omega_1 + Y - D_1 - \frac{a}{2}\cdot\sigma^2\right)$

To find the optimal D_1 that maximizes $W\left(-\sigma^2, -\sigma^2, D_1\right)$, the F.O.C. is $W_3\left(-\sigma^2, -\sigma^2, D_1\right) = 0$ which is already shown above and the S.O.C. is $W_{33}\left(-\sigma^2, -\sigma^2, D_1\right) < 0$ which is shown below,

$$\begin{aligned} W_{33}\left(-\sigma^2, -\sigma^2, D_1\right) &= \frac{\partial W_3}{\partial D_1} \\ &= \frac{1 - R\cdot f'\left(\omega_1 + Y - D_1 - \frac{a}{2}\cdot\sigma^2\right)}{\partial D_1} \\ &= R\cdot f''\left(\omega_1 + Y - D_1 - \frac{a}{2}\cdot\sigma^2\right) < 0 \end{aligned}$$

since $f''(\cdot) < 0$.

Condition (5) is satisfied because if for all $\left(-\sigma^2, D_1\right) \in \left[\sigma^2_{min}, \sigma^2_{max}\right] \times \mathbb{R}$, $W_{33}\left(-\sigma^2, -\sigma^2, D_1\right) \geq 0$ then there exist some $k > 0$ such that $\mid W_3\left(-\sigma^2, -\sigma^2, D_1\right) \mid > k$.

$$\begin{aligned} W_{33}\left(-\sigma^2, -\sigma^2, D_1\right) &\geq 0 \\ &\Longleftrightarrow \\ R\cdot f''\left(\omega_1 + Y - D_1 - \frac{a}{2}\cdot\sigma^2\right) &\geq 0 \end{aligned}$$

Thus, to maximize the expected payoff function, the manager will never choose $D_1^\star$ where $D_1^\star$ is the solution of $1 - f'\left(\omega_1 + Y - D_1 - \frac{a}{2}\cdot\sigma^2\right) = 0$. The reason is that $D_1^\star(\sigma^2)$ will minimize the expected utility payoff function instead of maximizing it.

$$W_3 = 1 - R \cdot f'\left(\omega_1 + Y - D_1 - \frac{a}{2} \cdot \sigma^2\right) \neq 1 - R \cdot f'\left(\omega_1 + Y - D_1^\star - \frac{a}{2} \cdot \sigma^2\right) = 0 \qquad \text{(A.9)}$$

Thus, $\mid W_3\left(-\sigma^2, -\sigma^2, D_1\right) \mid > 0$. It means we can always find some $k > 0$ such that $\mid W_3\left(-\sigma^2, -\sigma^2, D_1\right) \mid > k$.

The next step is to show that both the initial value condition and the single crossing condition hold.

Condition (6) holds because $W_2 < 0$, in the solution proposed the worst-type firm behaves as if it is in the full information case in equilibrium, i.e. $\tau(\sigma^2_{max}) = \phi(\sigma^2_{max})$.

Condition (7) holds because

$$\begin{aligned}
W_3\left(-\sigma^2, -\sigma^2(D_1), D_1\right) &= 1 - k \cdot R \cdot f'\left(\omega_1 + Y - D_1 - \frac{a}{2} \cdot \sigma^2(D_1)\right) \\
&\quad + \frac{k \cdot a \cdot R}{2} f'\left(\omega_1 + Y - D_1 - \frac{a}{2} \cdot \sigma^2(D_1)\right) \frac{\partial(-\sigma^2(D_1))}{\partial D_1} \\
&\quad - (1-k) \cdot R \cdot f'\left(\omega_1 + Y - D_1 - \frac{a}{2} \cdot \sigma^2\right) \\
W_2\left(-\sigma^2, -\sigma^2(D_1), D_1\right) &= \frac{\partial W}{\partial(-\sigma^2(D_1))} \\
&= \frac{k \cdot a}{2} f'\left(\omega_1 + Y - D_1 - \frac{a}{2} \cdot \sigma^2(D_1)\right) \\
\text{Then, } \frac{\partial \dfrac{W_3\left(-\sigma^2, -\sigma^2(D_1), D_1\right)}{W_2\left(-\sigma^2, -\sigma^2(D_1), D_1\right)}}{\partial(-\sigma^2)} &= -\frac{a}{2} \cdot \frac{2(1-k)}{k \cdot a} \frac{f''\left(\omega_1 + Y - D_1 - \frac{a}{2} \cdot \sigma^2(D_1)\right)}{f'\left(\omega_1 + Y - D_1 - \frac{a}{2} \cdot \sigma^2(D_1)\right)} \\
&= -\frac{1-k}{k} \cdot \frac{f''\left(\omega_1 + Y - D_1 - \frac{a}{2} \cdot \sigma^2(D_1)\right)}{f'\left(\omega_1 + Y - D_1 - \frac{a}{2} \cdot \sigma^2(D_1)\right)} > 0
\end{aligned}$$

Thus $\dfrac{W_3\left(-\sigma^2, -\sigma^2(D_1), D_1\right)}{W_2\left(-\sigma^2, -\sigma^2(D_1), D_1\right)}$ is a strictly increasing function of $(-\sigma^2)$.

Since Mailath (1987)'s conditions (1)-(7) are satisfied, the Riley equilibrium is the unique separating equilibrium of our model.

1.6. Equilibrium Refinement and Uniqueness

In this Section we want to show that our game belongs to the class of monotonic signaling games discussed in Section 3 of Esö and Schummer (2009), (see also Cho and Sobel (1990) and Ramey (1996)) and thus we can apply theorem 1 of Esö and Schummer (2009) to show that in this game the Riley equilibrium (i.e., the unique separating equilibrium as per above) is also the unique equilibrium that is immune to Credible Deviations.

First, let's check the 5 assumptions of Esö and Schummer (2009), A1 to A5, one by one. The firm (Sender) with variance σ^2 (type) decides to pay D_1 (signal). The investors (receivers) in the market buy the share of the firm at price $V^s(D_1)$ in the belief that the dividend D_1 reflect the value of the firm as a function of the unobserved variance, which can be denoted as $\sigma^2(D_1)$.

A1. $W(-\sigma^2, D_1, V^s(D_1))$ is strictly increasing in $V^s(D_1)$ for all $(-\sigma^2, D_1)$. In order to avoid solutions involving arbitrarily large messages and actions we assume that $\lim_{D_1\to\infty} W(-\sigma^2, D_1, V^s(D_1)) = -\infty$.

Proof.

$$\frac{\partial W(-\sigma^2, D_1, V^s(D_1))}{\partial V^s(D_1)} = k > 0$$

■

A2. Assume that $V^s(D_1)$ is such that, for any type σ^2 and message D_1, the Receiver has a unique best response, i.e. that $BR(-\sigma^2, D_1)$ is a singleton.

Proof. Since the investors (Receivers) act as price takers, they purchase the shares of the firm at the price $V^s(D_1)$. Their best response $BR(-\sigma^2, D_1)$ is a singleton $\{V^s(D_1) : V^s(D_1) = V^h(D_1)\}$. ■

A3. Assume that $V^s(D_1)$ is strictly increasing in $(-\sigma^2(D_1)$ for all (D_1, V^s).

Proof.

$$\frac{\partial V^s}{\partial(-\sigma^2(D_1))} = \frac{a}{2} R f'(\omega_1 + Y - D_1 - \frac{a}{2}\sigma^2(D_1)) > 0.$$

In particular,

$$\frac{\partial V^h}{\partial(-\sigma^2)} = \frac{a}{2} R f'(\omega_1 + Y - D_1 - \frac{a}{2}\sigma^2) > 0.$$

■

Together with monotonicity, A3 captures the idea that the manager (Sender) wants to induce the investors (Receivers) to buy the firm at a larger price by trying to convince them that the firm type is better (its variance is lower).

A4. Assume the game satisfies the central assumption in Spencian signaling games, the single crossing condition, that $-(\partial W/\partial D_1)/(\partial W/\partial(V^s(D_1)))$ is strictly decreasing in $-\sigma^2$.

Proof. According to the proof of Riley (1979)'s assumption A5 (see Appendix A), this assumption obviously holds. ■

A5. Assume that $W(-\sigma^2, D_1, V^h(D_1))$ is strictly quasi-concave in D_1.

Proof. Similar with the proof of Mailath (1987)'s condition (4) (see Appendix B), this assumption obviously holds. In detail,

$$\begin{aligned} W(-\sigma^2, D_1, V^h(D_1)) &= V^h(D_1) \\ &= D_1 + R \cdot f\left(\omega_1 + Y - D_1 - \frac{a}{2}\cdot\sigma^2\right). \end{aligned}$$

We have

$$\frac{\partial W(-\sigma^2, D_1, V^h(D_1))}{\partial D_1} = 1 - R \cdot f'\left(\omega_1 + Y - D_1 - \frac{a}{2}\cdot\sigma^2\right) = 0$$

has a unique solution and

$$\frac{\partial^2 W(-\sigma^2, D_1, V^h(D_1))}{\partial D_1^2} = R \cdot f''\left(\omega_1 + Y - D_1 - \frac{a}{2} \cdot \sigma^2\right) < 0.$$

Thus, $W(-\sigma^2, D_1, V^h(D_1))$ is strictly quasi-concave in D_1. ■

Thus, our game satisfies the assumptions of monotonic signaling games discussed in Esö and Schummer (2009). As a result, the Riley outcome is the unique equilibrium outcome that is immune to Credible Deviations.

1.7. Mailath (1987)

This Section states results in Mailath (1987) that are used above.

Suppose the set of possible types is the interval $[m, M] \subset \mathbb{R}$ and the set of possible actions is $\mathbb{R}$. If $\tau : [m, M] \to \mathbb{R}$ is an equilibrium one-to-one strategy for the informed agent, then when he chooses $y \in \tau([m, M])$ the uninformed agents infer his type is $\tau^{-1}(y)$. Thus, his expected payoff is $U\left(\alpha, \tau^{-1}(y), y\right)$. Furthermore, τ is an optimal strategy for the informed agent, so that $\tau(\alpha)$ maximizes the expected payoff. So, for τ to be a separating equilibrium strategy it must be one-to-one and satisfy *incentive compatibility* (IC):

(IC) $\quad \tau(\alpha) \in \underset{y \in \tau([m,M])}{\operatorname{argmax}} U\left(\alpha, \tau^{-1}(y), y\right), \ \forall \alpha \in [m, M]$

If $U\left(\alpha, \tau^{-1}(y), y\right)$ has no other maximizer for $y \in \tau([m, M])$ for all $\alpha \in [m, M]$, then τ satisfies strict incentive compatibility (SIC), i.e.,

(SIC) $\quad \tau(\alpha) = \underset{y \in \tau([m,M])}{\operatorname{argmax}} U\left(\alpha, \tau^{-1}(y), y\right) \ \forall \alpha \in [m, M].$

The regularity conditions on U are (where subscripts denote partial derivatives):

(1) $U(\alpha, \hat{\alpha}, y)$ is C^2 on $[m, M]^2 \times \mathbb{R}$ (smoothness)

(2) U_2 never equals zero, and so is either positive or negative (belief monotonicity)

(3) U_{13} never equals zero, and so is either positive or negative (type monotonicity)

(4) $U_3(\alpha, \alpha, y) = 0$ has a unique solution in y, denoted $\phi(\alpha)$, which maximizes $U(\alpha, \alpha, y)$, and $U_{33}(\alpha, \alpha, \phi(\alpha)) < 0$ ("strict" quasi-concavity)

(5) there exists $k > 0$ such that for all $(\alpha, y) \in [m, M] \times \mathbb{R}$ $U_{33}(\alpha, \alpha, y) \geq 0 \Rightarrow |U_3(\alpha, \alpha, y)| > k$ (boundedness)

The other two conditions which play a role in what follows are

(6) $\tau(\alpha^w) = \phi(\alpha^w)$, where $\alpha^w = M$ if $U_2 < 0$ and m if $U_2 > 0$ (initial value)

(7) $\frac{U_3(\alpha, \hat{\alpha}, y)}{U_2(\alpha, \hat{\alpha}, y)}$ is a strictly monotonic function of α (single crossing)

Theorem 1. *Suppose (1) - (5) are satisfied and $\tau : [m, M] \to \mathbb{R}$ is one-to-one and satisfies incentive compatibility. Then τ has at most one discontinuity on $[m, M]$, and where it is continuous on (m,M), it is differentiable and satisfies (DE)* $\quad \dfrac{d\tau}{d\alpha} = \dfrac{-U_2(\alpha, \alpha, \tau)}{U_3(\alpha, \alpha, \tau)}$.

Furthermore, if τ is discontinuous at a point, α' say, then τ is strictly increasing on one of $[m, \alpha')$ or $(\alpha', M]$ and strictly decreasing on the other, and the jump at α' is of the same sign as U_{13}.

Theorem 2. *Suppose, in addition, that either the initial value condition or the single crossing condition for $(\hat{\alpha}, y)$ in the graph of τ is satisfied. Then τ is strictly monotonic on (m, M) and hence continuous and satisfies the differential equation (DE) there. If the initial value condition is satisfied, then in fact τ is continuous on $[m, M]$ and $\frac{d\tau}{d\alpha}$ has the same sign as U_{13}.*

The following corollary shows that incentive compatibility and the initial value condition together imply uniqueness. Let $\tilde{\tau}$ denote the unique solution to the following restricted initial value problem: (DE), $\tau(\alpha^w) = \phi(\alpha^w)$ and $(d\tau/d\alpha)U_{13} > 0$.

Corollary: suppose (1)-(5) are satisfied and the initial value condition holds. If τ satisfies incentive compatibility, then $\tau = \tilde{\tau}$.

1.8. Esö and Schummer (2009)

This Section states results in Esö and Schummer (2009) that are used above.

Define the Sender-Receiver game which is denoted by the tuple (Θ, π, u_S, u_R). The Sender has private information that is summarized by his type $\theta \in \Theta = \{\theta_1,\ \theta_2, \cdots,\ \theta_n\} \subset \mathbb{R}$, where $\theta_1 < \theta_2 < \cdots < \theta_n$. The commonly known prior probability that the Sender's type is θ is $\pi(\theta)$. Upon realizing his type, the Sender chooses a message $m \in \mathbb{R}_+$. A strategy for the Sender is a function $M : \Theta \to \mathbb{R}_+$. The Sender and Receiver receive respective payoffs of $u_S(\theta, m, a)$ and $u_R(\theta, m, a)$, which are both continuously differentiable in (m, a).

The Receiver's (posterior) beliefs upon receiving the Sender's message is a function $\mu : \mathbb{R}_+ \to \Delta(\Theta)$, where $\Delta(\Theta)$ refers to the set of probability distributions on Θ. For any message $m \in \mathbb{R}_+$ and any fixed (posterior belief) distribution $\tilde{\pi} \in \Delta(\Theta)$, denote the Receiver's best responses to m (given $\tilde{\pi}$) by $BR(\tilde{\pi}, m) \equiv \text{argmax}_{a \in \mathbb{R}} \mathbb{E}\left[u_R(\theta, m, a) \mid \tilde{\pi}\right]$.

Formalizing Credible Deviations

Definition 1. *(Vulnerability to a Credible Deviation) Given an equilibrium (M, A, μ), we say that an out-of-equilibrium message $m \in \mathbb{R}_+ \backslash M(\Theta)$ is a Credible Deviation if the following condition holds for exact one (non-empty) set of types $C \subseteq \Theta$.*

$$C = \{\theta \in \Theta : u_S^\star(\theta) < \min_{a \in BR(C,m)} u_S(\theta, m, a)\} \tag{A.10}$$

We call C the (unique) Credible Deviators' Club for message m. If such a message exits, the equilibrium is Vulnerable to a Credible Deviation.

Monotonic Signaling Games and the Uniqueness of the Equilibrium

Following Cho and Sobel (1990) and Ramey (1996), monotonic signaling games are defined as follows,

A1. $u_S(\theta, m, a)$ is strictly increasing in a for all (θ, m). One can think of a as some sort of compensation for the Sender. In order to avoid solutions involving arbitrarily large messages and actions we assume that $\lim_{m\to\infty} u_S(\theta, m, a) = -\infty$.

A2. Assume that u_R is such that, for any type θ and message m, the Receiver has a unique best response, i.e. that $BR(\theta, m)$ is a singleton. We denote this action as $\{\beta(\theta, m)\} \equiv BR(\theta, m)$ and $\beta(\theta, m)$ is uniformly bounded from above.

A3. Assume that $BR(\tilde{\pi}, m)$ is greater for beliefs that are greater in the first-order stochastic sense, and in particular, $\beta(\theta, m)$ is strictly increasing in θ for all (m, a) (Cho and Sobel 1990, p. 392).

A4. Assume the game satisfies the central assumption in Spencian signaling games, the single crossing condition, that $-(\partial u_S/\partial m)/(\partial u_S/\partial a)$ is strictly decreasing in θ.

A5. Assume that $u_S(\theta, m, \beta(\theta, m))$ is strictly quasi-concave in m.

An additional piece of notation simplifies the exposition. For any θ and m, let $\hat{a}(\theta, m)$ be the action to satisfy,

$$u_S(\theta, m, \hat{a}(\theta, m)) = u_S^{\star}(\theta) \tag{A.11}$$

if such an action exists, and denote $\hat{a}(\theta, m) = \infty$ otherwise. This action by the Receiver would give Sender-type θ his equilibrium payoff after sending m. If such an action exists, it is unique by monotonicity.

Lemma 3 *If an equilibrium (M, A, μ) is not Vulnerable to Credible Deviations, it is a separating equilibrium - no two types send the same message.*

Lemma 4 *Any equilibrium whose outcome is different from the Riley outcome is Vulnerable to Credible Deviations.*

Theorem *The Riley outcome is the unique equilibrium outcome that is not Vulnerable to Credible Deviations.*

1.9. Baseline Setting: General Case and CARA Special Case

In the baseline setting, the manager chooses the dividend payment to maximize

$$\begin{aligned}
&\max_{D_1} \quad D_1 + \mathbb{E}\left[Y_2\right] \\
&\text{subject to} \\
&Y_2 = R \cdot f\left(I_1\right) + \nu \\
&D_1 \leq \omega_1,
\end{aligned}$$

which implies, assuming for illustration that the second constraint is slack, $D_1 < \omega_1$,

$$\max_{D_1} \quad D_1 + R \cdot f\left(\omega_1 + Y - D_1 - \frac{a}{2} \cdot \sigma^2\right),$$

where $\mathbb{E}\left[Y_2\right] = \mathbb{E}\left[R \cdot f(I_1) + \nu\right] = R \cdot f\left(\omega_1 + Y - D_1 - \frac{a}{2} \cdot \sigma^2\right)$ and a is the certainty equivalent coefficient in the sense of Arrow-Pratt. We begin by assuming that the Arrow-Pratt coefficient is scale-invariant, i.e., $a\left(I_1^*\right) \equiv a$, for clarity of illustration, which is the case for exponential production functions. Later in this section we analyze the general case.

To understand this formulation, note that in our framework, randomness in Y reduces the expected profits if the function $f\left(.\right)$ is concave, in which case the firm is essentially risk averse with respect to fluctuations in Y, in the precise sense that $\mathbb{E}[f(Y)] < f(\mathbb{E}[Y])$, that is, Jensen's inequality.[1] The first-order condition is $1 - R \cdot f'\left(\omega_1 + Y - D_1 - \frac{a}{2} \cdot \sigma^2\right) \geq 0$.

Prediction 1 (baseline). The following result is straightforward:

$$\frac{\partial \sigma^2}{\partial D_1} = -\frac{2}{a} < 0.$$

Larger dividends should be associated with subsequent lower cash-flow volatility. Because managers pay dividends before the cash flows are realized, managers take into account, in a certainty-equivalence sense, that paying higher dividends will increase the probability of foregoing future investment opportunities, as the (expected) volatility of cash flows increases. This prediction follows from the precautionary savings motive, because in our setting lower dividends directly translate in higher cash balances available for future investment.[2]

This stylized model already delivers the main hypothesis of our paper; that is, dividend changes should be followed by changes in cash-flow volatility in the opposite direction. We now consider the more general case, $a = a(I_1^*)$, and for illustration we maintain $R = 1$. Recall the manager's maximization problem is,

$$\begin{aligned} \max_{D_1} \quad & D_1 + \mathbb{E}[Y_2] \\ & s.t. \\ & Y_2 = f(I_1) + \nu \\ & D_1 \leq \omega_1 \end{aligned}$$

We can rewrite $D_1 + \mathbb{E}(Y_2)$ as

$$\begin{aligned} D_1 + \mathbb{E}(Y_2) &= D_1 + \mathbb{E}(f(I_1) + \nu) \\ &= D_1 + \mathbb{E}(f(\omega_1 + Y_1 - D_1)) \qquad (\text{since } \mathbb{E}(\nu) = 0) \\ &= D_1 + \mathbb{E}(f(\omega_1 + Y + \nu - D_1)) \qquad (\text{since } \mathbb{E}(Y_1) = f(I_0) = Y) \end{aligned}$$

Let $f(\omega_1 + Y - D_1 - \frac{a}{2}\sigma^2) = \mathbb{E}[f(\omega_1 + Y + \nu - D_1)]$. By first order Taylor expansion of the left-hand side (LHS),

$$f(\omega_1 + Y - D_1 - \frac{a}{2}\sigma^2) \quad \approx \quad f(\omega_1 + Y - D_1) + f'(\omega_1 + Y - D_1)(-\frac{a}{2}\sigma^2)$$

[1]This insight exactly parallels the one in Froot, Scharfstein and Stein (1993) about conditions under which risk management increases firm value. See also Rampini and Viswanathan (2013).

[2]Together with Jensen's inequality, this equivalence between cash balances and the negative of cash payouts implies that lower cash-flow volatility will eventually result in higher investment and cash flows at Time 2. In a richer model with a wedge between cash holdings and cash payouts, this prediction would no longer necessarily hold.

By second order Taylor expansion of the right-hand side (RHS),

$$\begin{aligned} f(\omega_1 + Y + \nu - D_1) &= f(\omega_1 + Y - D_1 + \nu) \\ &\approx f(\omega_1 + Y - D_1) + f'(\omega_1 + Y - D_1)\nu + \frac{f''(\omega_1 + Y - D_1)}{2}\nu^2 \end{aligned}$$

Taking expectation in both sides, obtain

$$\begin{aligned} \mathbb{E}\left[f(\omega_1 + Y + \nu - D_1)\right] &\approx f(\omega_1 + Y - D_1) + f'(\omega_1 + Y - D_1)\mathbb{E}(\nu) + \frac{f''(\omega_1 + Y - D_1)}{2}\mathbb{E}(\nu^2) \\ &= f(\omega_1 + Y - D_1) + \frac{f''(\omega_1 + Y - D_1)}{2}\mathbb{E}(\nu^2) \quad (\text{since } \mathbb{E}(\nu) = 0) \\ &= f(\omega_1 + Y - D_1) + \frac{f''(\omega_1 + Y - D_1)}{2}\sigma^2 \end{aligned}$$

Comparing Taylor expansions of LHS and RHS, we obtain $a(Y, D_1) = -\frac{f''(\omega_1 + Y - D_1)}{f'(\omega_1 + Y - D_1)}$. It is not only a function of Y, but also a function of D_1. Therefore, $D_1 + \mathbb{E}(Y_2)$ can be expressed as $D_1 + f(\omega_1 + Y - D_1 - \frac{a(Y, D_1)}{2}\sigma^2)$. The F.O.C. of the problem is

$$1 - f'(\omega_1 + Y - D_1 - \frac{a(Y, D_1)}{2}\sigma^2) - \frac{\sigma^2}{2}f'(\omega_1 + Y - D_1 - \frac{a(Y, D_1)}{2}\sigma^2)\frac{\partial a(Y, D_1)}{\partial D_1} = 0$$

We can then show that:
$\frac{\partial \sigma^2}{\partial D_1} < 0$ if

$$\frac{a(Y, D_1)f''(\omega_1 + Y - D_1 - \frac{a(Y, D_1)}{2}\sigma^2)}{2f'(\omega_1 + Y - D_1 - \frac{a(Y, D_1)}{2}\sigma^2)} - \frac{1}{\sigma^2}\left[1 - f'(\omega_1 + Y - D_1 - \frac{a(Y, D_1)}{2}\sigma^2)\right] < 0.$$

Proof. Let $G(D_1, \sigma^2) = 1 - f'(\omega_1 + Y - D_1 - \frac{a(Y, D_1)}{2}\sigma^2) - \frac{\sigma^2}{2}f'(\omega_1 + Y - D_1 - \frac{a(Y, D_1)}{2}\sigma^2)\frac{\partial a(Y, D_1)}{\partial D_1} = 0$,

$$\begin{aligned} \frac{\partial \sigma^2}{\partial D_1} &= -\frac{\frac{\partial G}{\partial D_1}}{\frac{\partial G}{\partial \sigma^2}} \\ &= -\frac{f''(\omega_1 + Y - D_1 - \frac{a(Y, D_1)}{2}\sigma^2)\left[1 + \frac{\sigma^2}{2}\frac{\partial a(Y, D_1)}{\partial D_1}\right]^2 - \frac{\sigma^2}{2}f'(\omega_1 + Y - D_1 - \frac{a(Y, D_1)}{2}\sigma^2)\frac{\partial^2 a(Y, D_1)}{\partial D_1^2}}{f''(\omega_1 + Y - D_1 - \frac{a(Y, D_1)}{2}\sigma^2)\left[1 + \frac{\sigma^2}{2}\frac{\partial a(Y, D_1)}{\partial D_1}\right]\frac{a(Y, D_1)}{2} - \frac{1}{2}f'(\omega_1 + Y - D_1 - \frac{a(Y, D_1)}{2}\sigma^2)\frac{\partial a(Y, D_1)}{\partial D_1}} \end{aligned}$$

Recalling the F.O.C.,

$$1 - f'(\omega_1 + Y - D_1 - \frac{a(Y, D_1)}{2}\sigma^2) - \frac{\sigma^2}{2}f'(\omega_1 + Y - D_1 - \frac{a(Y, D_1)}{2}\sigma^2)\frac{\partial a(Y, D_1)}{\partial D_1} = 0, \qquad \text{(A.12)}$$

we have $\dfrac{\partial a(Y,D_1)}{\partial D_1} = \dfrac{2}{\sigma^2}\left[\dfrac{1}{f'(\omega_1+Y-D_1^\star-\frac{a(Y,D_1)}{2}\sigma^2)}-1\right]$. The S.O.C. is

$$f''(\omega_1+Y-D_1-\frac{a(Y,D_1)}{2}\sigma^2)\left[1+\frac{\sigma^2}{2}\frac{\partial a(Y,D_1)}{\partial D_1}\right]^2-\frac{1}{2}f'(\omega_1+Y-D_1-\frac{a(Y,D_1)}{2}\sigma^2)\frac{\partial^2 a(Y,D_1)}{\partial D_1^2}<0$$

Thus,

$$\begin{aligned}
\frac{\partial^2 a(Y,D_1)}{\partial D_1^2} &> \frac{\partial(\frac{2}{\sigma^2}\left[\frac{1}{f'(\omega_1+Y-D_1-\frac{a(Y,D_1)}{2}\sigma^2)}-1\right])}{\partial D_1}\\
&= \frac{2}{\sigma^2}\frac{f''(\omega_1+Y-D_1-\frac{a(Y,D_1)}{2}\sigma^2)\left[1+\frac{\sigma^2}{2}\frac{\partial a(Y,D_1)}{\partial D_1}\right]}{f'(\omega_1+Y-D_1-\frac{a(Y,D_1)}{2}\sigma^2)^2}\\
&= \frac{2}{\sigma^2}\frac{f''(\omega_1+Y-D_1-\frac{a(Y,D_1)}{2}\sigma^2)}{f'(\omega_1+Y-D_1-\frac{a(Y,D_1)}{2}\sigma^2)^3}
\end{aligned}$$

Thus, if $f''\dfrac{a(Y,D_1)}{2f'}-\dfrac{f'}{2}\dfrac{\partial a(Y,D_1)}{\partial D_1}<0$ (i.e., $f''\dfrac{a(Y,D_1)}{2f'}-\dfrac{1}{\sigma^2}\left[1-f'\right]<0$),

$$\frac{\partial\sigma^2}{\partial D_1}<-\frac{\frac{f''}{f'^2}-\frac{\sigma^2}{2}f'\frac{2f''}{\sigma^2 f'^3}}{f''\frac{a(Y,D_1)}{2f'}-\frac{f'}{2}\frac{\partial a(Y,D_1)}{\partial D_1}}\Rightarrow\frac{\partial\sigma^2}{\partial D_1}<0$$

where $f'=f'(\omega_1+Y-D_1-\frac{a(Y,D_1)}{2}\sigma^2)$, $f''=f''(\omega_1+Y-D_1-\frac{a(Y,D_1)}{2}\sigma^2)$ and $f'''=f'''(\omega_1+Y-D_1-\frac{a(Y,D_1)}{2}\sigma^2)$. ■

We can now examine the cross derivative, $\dfrac{\partial^2\sigma^2}{\partial D_1\partial Y}$, and see that in general its sign is ambiguous.

$$\frac{\partial^2\sigma^2}{\partial D_1\partial Y} = \frac{\partial(\frac{\partial\sigma^2}{\partial D_1})}{\partial Y}$$

According to the proof above,

$$\frac{\partial\sigma^2}{\partial D_1} = \frac{-f''\left[1+\frac{\sigma^2}{2}\frac{\partial a(Y,D_1)}{\partial D_1}\right]^2+\frac{\sigma^2}{2}f'\frac{\partial^2 a(Y,D_1)}{\partial D_1^2}}{f''\left[1+\frac{\sigma^2}{2}\frac{\partial a(Y,D_1)}{\partial D_1}\right]\frac{a(Y,D_1)}{2}-\frac{1}{2}f'\frac{\partial a(Y,D_1)}{\partial D_1}}$$

where $f'=f'(\omega_1+Y-D_1-\frac{a(Y,D_1)}{2}\sigma^2)$, $f''=f''(\omega_1+Y-D_1-\frac{a(Y,D_1)}{2}\sigma^2)$ and $f'''=f'''(\omega_1+Y-$

$D_1 - \frac{a(Y,D_1)}{2}\sigma^2)$. Let A denote the numerator of $\frac{\partial \sigma^2}{\partial D_1}$ and B denote the denominator of $\frac{\partial \sigma^2}{\partial D_1}$, then

$$\frac{\partial^2 \sigma^2}{\partial D_1 \partial Y} = \frac{A'B - AB'}{B^2}$$

where

$$\begin{aligned} A'B &= \{\left[\frac{\sigma^2 f''}{2}\frac{\partial^2 a}{\partial D_1^2} - f'''(1+\frac{\sigma^2}{2}\frac{\partial a}{\partial D_1})^2\right]\left[1+\frac{\sigma^2}{2}\frac{\partial a}{\partial D_1}\right] + \frac{\sigma^2}{2}f'\frac{\partial(\frac{\partial^2 a}{\partial D_1^2})}{\partial Y} - 2f''\left[1+\frac{\sigma^2}{2}\frac{\partial a}{\partial D_1}\right]\frac{\sigma^2}{2}\frac{\partial^2 a}{\partial D_1 \partial Y}\} \\ &\times\{f''\left[1+\frac{\sigma^2}{2}\frac{\partial a}{\partial D_1}\right]\frac{a}{2} - \frac{1}{2}f'\frac{\partial a}{\partial D_1}\} \end{aligned}$$

and

$$\begin{aligned} AB' &= \{-f''\left[1+\frac{\sigma^2}{2}\frac{\partial a(Y,D_1)}{\partial D_1}\right]^2 + \frac{\sigma^2}{2}f'\frac{\partial^2 a(Y,D_1)}{\partial D_1^2}\} \\ &\times\{f'''(1+\frac{\sigma^2}{2}\frac{\partial a}{\partial D_1})^2\frac{a}{2} - \frac{1}{2}f''\frac{\partial a}{\partial D_1}(1+\frac{\sigma^2}{2}\frac{\partial a}{\partial D_1}) + \frac{a}{2}f''\frac{\sigma^2}{2}\frac{\partial^2 a}{\partial D_1 \partial Y} - \frac{f'}{2}\frac{\partial^2 a}{\partial D_1 \partial Y}\} \end{aligned}$$

Special Case

Assume $f(x) = A - e^{-kx}$ with constants $A > 0$ and $k > 0$. To begin with,

$$\begin{aligned} f'(x) &= -(-ke^{-kx}) = ke^{-kx} \\ f''(x) &= -k^2 e^{-kx} \end{aligned}$$

In this case, $a = \dfrac{-f''(\omega_1 + Y - D_1)}{f'(\omega_1 + Y - D_1)} = k$ is constant.
The F.O.C. will be

$$1 - f'(\omega_1 + Y - D_1 - \frac{a}{2}\sigma^2) = 0$$

Prediction 1 (Special Case).
$\dfrac{\partial \sigma^2}{\partial D_1} < 0$

Proof. Let $J = 1 - f'(\omega_1 + Y - D_1 - \frac{a}{2}\sigma^2) = 0$,

$$\frac{\partial \sigma^2}{\partial D_1} = -\frac{\frac{\partial J}{\partial D_1}}{\frac{\partial J}{\partial \sigma^2}} = -\frac{f''(\omega_1 + Y - D_1 - \frac{a}{2}\sigma^2)}{f''(\omega_1 + Y - D_1 - \frac{a}{2}\sigma^2)\frac{a}{2}} = -\frac{2}{a} < 0$$

Because $a = k > 0$. ■

Prediction 2 (Special Case).
$\frac{\partial^2 \sigma^2}{\partial D_1 \partial Y} = 0$
Proof: by inspection.

1.10. Agency: Additive Formulation

We report here the proofs of the statements in Section VB in the main text. Under an additive agency formulation, the maximization program is:

$$\begin{aligned} &\max_{D_1} D_1 + \mathbb{E}[Y_2] - c(D_1), \\ &\text{subject to} \\ &Y_2 = R \cdot f(I_1) + \nu \\ &0 \leq d_1 \leq 1. \end{aligned}$$

The first-order condition is $1 - R \cdot f'\left(\omega_1 + Y - D_1 - \frac{a}{2} \cdot \sigma^2\right) - c'(D_1) \geq 0$. Therefore, we obtain:

Prediction i **(additive agency - time series)**.

$$\frac{\partial \sigma^2}{\partial D_1} = -\frac{2}{a} + \frac{2 \cdot c''(D_1)}{a \cdot R \cdot \left[f''\left(\omega_1 + Y - D_1 - \frac{a}{2} \cdot \sigma^2\right)\right]} < 0.$$

As in the baseline setting and in the signaling model, higher dividends should correlate with lower future cash-flow volatility. Two effects are at play. First, as in the baseline setting, lower future cash-flow volatility implies a higher income available for paying dividends, holding investment opportunities fixed. Second, lower future cash-flow volatility enables managers to more easily extract private benefits (re. incur lower agency costs) and pay more dividends, again holding investment fixed.

Now, however, the *larger* the current earnings, the larger the reduction in cash-flow volatility should be following the same dollar of dividend paid:

Prediction ii **(additive agency - cross-section)**.

$$\frac{\partial^2 \sigma^2}{\partial D_1 \partial Y} = -\frac{2 \cdot c''(D_1) \cdot f'''\left(\omega_1 + Y - D_1 - \frac{a}{2} \cdot \sigma^2\right)}{a \cdot R \cdot \left[f''\left(\omega_1 + Y - D_1 - \frac{a}{2} \cdot \sigma^2\right)\right]^2} < 0.$$

1.11. Agency: Multiplicative Formulation

We assume the manager faces a private (agency) cost $c(d_1)$ from paying as dividends a fraction d_1 of cash flows in the first period and the remaining cash in the second period, where the function c is convex, that is, $c' > 0$ and $c'' > 0$. In this setting, the manager chooses the dividend payment to maximize

$$\begin{aligned}
&\max_{d_1} \omega_1 \left[d_1 - c(d_1)\right] + \mathbb{E}\left[Y_2\right]\left[1 - c(d_1)\right] \\
&\text{subject to} \\
&Y_2 = R \cdot f(I_1) + \nu \\
&0 \leq d_1 \leq 1.
\end{aligned}$$

The first-order condition is: $\omega_1 [1 - c'(d_1)] - \omega_1 \cdot R \cdot f'\left((1 - d_1)\omega_1 + Y - \frac{a}{2} \cdot \sigma^2\right)[1 - c(d_1)] - c'(d_1) \cdot R \cdot f\left((1 - d_1)\omega_1 + Y - \frac{a}{2} \cdot \sigma^2\right) \geq 0$ or for brevity $\omega_1(1 - c') - \omega_1 \cdot R \cdot f'(1 - c) - c' \cdot R \cdot f \geq 0$. Therefore, we obtain:

Prediction i **(multiplicative agency - time series)**.

$$\frac{\partial \sigma^2}{\partial d_1} = -\frac{2\omega_1}{a} + \frac{2 \cdot c'' \cdot (\omega_1 + f)}{a \cdot R \cdot [\omega_1 \cdot f''(1 - c) + f'c']} - \frac{2 \cdot \omega_1 \cdot f' \cdot c'}{a \cdot [\omega_1 \cdot f''(1 - c) + f'c']} \gtrless 0$$

To illustrate, assume that $\omega_1 \cdot f''(1 - c) + f'c' < 0$ (the discussion for the reverse case in which $\omega_1 \cdot f''(1 - c) + f'c' > 0$ is largely symmetric).

Three effects are at play. First, as in the baseline setting, lower future cash flow volatility implies a higher income available for paying dividends, holding investment opportunities fixed. Second, lower future cash-flow volatility enables managers to more easily extract private benefits (re. incur lower agency costs) and pay more dividends, again holding investment fixed. These effects are as in the additive agency model, and predict a negative correlation between σ^2 and d_1. Now, however, a third effect also arises, as lower future cash flow volatility, implying a higher future income, also allows managers to divert a larger share of the cash flows. This third effect is captured by the term $f'c'$. As a result, the sign of the comparative static exercise cannot be uniquely determined and the prediction is therefore ambiguous.

We can now compute the cross-derivative:

Prediction ii **(multiplicative agency - cross-section)**.

$$\frac{\partial^2 \sigma^2}{\partial d_1 \partial Y} = \frac{-\omega_1^2 \cdot f''' + c'' \cdot f' + 2\omega_1 \cdot f''c'}{R \cdot \left\{\frac{a}{2} \cdot \left[\omega_1 \cdot f''(1-c) + f'c'\right]\right\}^2} \lesseqgtr 0$$

As in the additive agency model, larger current earnings make extracting more private benefits (re. incur lower agency costs) easier. Now, a further effect also arises, namely, higher current earnings also allow managers to divert larger cash flows. This latter effect is captured by the terms $f'c'$. As a result, this cross-sectional prediction is also ambiguous.

1.12. Investment Opportunities

To model investment opportunities we follow Johnson et al. (2000) and Choe et al. (1993) and assume that the production function, f, is pre-multiplied by a positive parameter R representing investment opportunities. The maximization program of the additive agency model then becomes:

$$\max_{D_1} D_1 + \mathbb{E}\left[Y_2\right] - c\left(D_1\right),$$

subject to

$$Y_2 = R \cdot f\left(I_1\right) + \nu$$

$$D_1 \leq \omega_1.$$

We establish:

Prediction A.1 (additive agency - investment opportunities).

$$\frac{\partial^2 \sigma^2}{\partial D_1 \partial R} = \frac{-2 \cdot c''\left(D_1\right) \cdot f''\left(\omega_1 + Y - D_1 - \frac{a}{2} \cdot \sigma^2\right)}{a \cdot \left[R \cdot f''\left(\omega_1 + Y - D_1 - \frac{a}{2} \cdot \sigma^2\right)\right]^2} > 0$$

In the additive agency model, the decline in cash flow volatility following dividend increases should be more pronounced for firms with *smaller* investment opportunities. The reason is that, if lower cash flow volatility facilitates extraction of private benefits, then high investment opportunities mute this effect because they increase the cost of extracting private benefits relative to engaging in efficient investment.

Conversely, the maximization program of the multiplicative agency model becomes

$$\begin{aligned}&\max_{d_1} \omega_1 \left[d_1 - c\left(d_1\right)\right] + \mathbb{E}\left[Y_2\right]\left[1 - c\left(d_1\right)\right]\\&\text{subject to}\\&Y_2 = R \cdot f\left(I_1\right) + \nu\\&0 \leq d_1 \leq 1.\end{aligned}$$

We obtain:

Prediction A.2 (multiplicative agency - investment opportunities).

$$\frac{\partial^2 \sigma^2}{\partial d_1 \partial R} = \frac{-c'' \cdot \omega_1 \cdot \frac{a}{2} \cdot \left[\omega_1 \cdot f''\left(1-c\right) + f'c'\right]}{\left\{\frac{a}{2} \cdot R \cdot \left[\omega_1 \cdot f''\left(1-c\right) + f'c'\right]\right\}^2} \gtrless 0$$

Relative to the additive agency model, a new effect arises, as higher investment opportunities also allow managers to divert larger cash flows, an effect captured by the terms $f'c'$. As a result, the sign of this comparative statics depends on the term in parentheses. If $\left[\omega_1 \cdot f''\left(1-c\right) + f'c'\right] < 0$ then the multiplicative agency model generates the same prediction as the additive agency model (see above). If instead $\left[\omega_1 \cdot f''\left(1-c\right) + f'c'\right] > 0$ then the multiplicative agency model generates the same prediction as the signaling model (see below).

Finally, the maximization program of the signaling model becomes

$$\begin{aligned}\max_{\{D_1\}} \quad & W_1 = kV_1^s + (1-k)V_1^h\\&\text{subject to}\\&Y_2 = R \cdot f\left(I_1\right) + \nu\\&D_1 \leq \omega_1,\end{aligned}$$

We establish:

Prediction S.1 (signaling - investment opportunities).

$$\frac{\partial^2 \sigma^2}{\partial D_1 \partial R} = \frac{-f'\left(\omega_1 + Y - D_1 - \frac{a}{2} \cdot \sigma^2\right)}{\left[R \cdot f''\left(\omega_1 + Y - D_1 - \frac{a}{2} \cdot \sigma^2\right)\right]^2} < 0$$

In the signaling model, the decline in cash flow volatility following dividend increases should be more pronounced for firms with higher investment opportunities. The intuition is that the scope of using dividends to signal future declines in cash flow volatility is magnified when investment opportunities are larger.

1.13. Curvature of the Production Function

Under the assumption of a (negative) exponential production function, our firm is akin to a risk averse agent with CARA utility and its curvature is fully described by the Arrow-Pratt coefficient of absolute risk aversion (or certainty equivalence), a. As a result, our estimates of equation (17), particularly of $\hat{\beta} = -0.04$ in column 2 of Table 13, implies a coefficient of $\hat{a} = 0.08$. In this section, we discuss how this estimate compares with estimates in the literature. First, we note that much of the quantitative corporate finance literature does not study a setting in which firms are "essentially risk averse" as in ours, and assumes a Cobb-Douglas production function, $y_t = k^{\gamma} \cdot l^{(1-\gamma)}$. The literature then typically calibrates $\hat{\gamma} = 0.33$ based on macroeconomic estimates of the capital share in production functions.

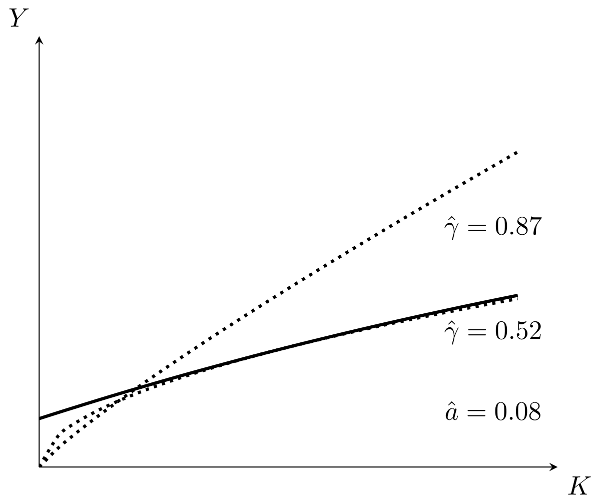
Closer to our focus, Froot, Scharfstein and Stein (1993) present a static model of a firm with a concave production function and stress that such firm can be understood as being effectively risk averse with respect to fluctuations in cash flows, thereby providing a microfoundation for corporate risk management and the use of derivative contracts. Recently, Rampini and Viswanathan (2010) and Rampini and Viswanathan (2013) develop this basic intuition in dynamic models of risk management, capital structure and corporate financing. Li, Whited and Wu (2016) conduct a full-fledged structural estimation of the Rampini and Viswanathan (2013) model, and estimate the curvature of the production function. In their setting, the concavity of the production function together with financial constraints (and particularly collateral constraints) induce curvature in the value function, which is what they estimate in the data. In this sense, therefore, the firm's risk aversion in the Li, Whited and Wu (2016) and Rampini and Viswanathan (2013) papers is an induced property of the value function stemming from the whole theoretical framework rather than a single parameter as in our CARA setting.

Specifically, Li, Whited and Wu (2016) consider a production function, $y_t = z_t k_t^{\gamma}$, and in their baseline setting (i.e., their Table 1) they estimate the curvature parameter $\hat{\gamma}$ to be in the range $[0.52, 0.88]$, where the estimate $\hat{\gamma} = 0.52$ indicates a "high curvature / steep production function / high risk aversion" (in the sense discussed above) and results from realistic tax rates between 0.1 and 0.2, while $\hat{\gamma} = 0.88$ indicates a "flat production function" with "low curvature / low risk aversion" and results from a benchmark setting with zero tax rates. Furthermore, in a series of experiments they report values of their curvature parameter between 0.52 and 0.61 in their model with non-state contingent debt (Table 4).

How do our estimates compare with the Li, Whited and Wu (2016) estimates? Once more, a direct comparison is not possible, because in our setting risk aversion is captured by a single CARA parameter, while in their setting risk aversion is a feature of the concavity of the production function together with collateral constraints. Notwithstanding the above caveats and qualifications, one can nevertheless gauge a visual comparison of our estimate with those of Li, Whited and Wu (2016), and we report them in the Figure A.1. From this exercise, it appears that our estimate of $\hat{a} = 0.08$ has a curvature that is locally "close" to that of Li, Whited and Wu (2016)s upper bound of $\hat{\gamma} = 0.52$.

Fig. A.1 Estimates of the curvature of the production function

This figure plots the implied curvature of the production function in our model (solid line) with the implied curvature of the production function in Li, Whited and Wu (2016) (dotted lines). Li, Whited and Wu (2016) consider a production function, $Y = z \cdot K^{\gamma}$, and estimate $\hat{\gamma}$ to be in the range $[0.52, 0.88]$. We consider a negative exponential function $Y = w \cdot exp(-aK) + v$ and estimate $\hat{a} = 0.08$. For illustration, in the plot below we use $z = 1$, $w = -4.88$, and $v = 5.61$.



	Event firms				Matched firms			
	Nobs	Median	Mean	Std	Nobs	Median	Mean	Std
BM	6,777	0.70	0.92	0.79	6,777	0.80	0.98	0.73
Size	6,777	19.66	19.74	1.93	6,777	18.98	19.20	1.94
Leverage	6,777	0.41	0.43	0.27	6,777	0.39	0.41	0.22
Age	6,777	12.75	15.96	9.93	6,777	12.50	14.44	9.03

Table A.1 Descriptive statistics matching

This table reports descriptive statistics for the event firms and the observationally-similar firms matched on propensity scores. BM is the book-to-market-ratio, Size is the log market cap, Lev is financial leverage and age is the time since the firm has been on CRSP. Our sample period is 1964 until 2013.

	$\Delta Div > 0$ (1)	Initiation (2)	Pooled (3)	$\Delta Div < 0$ (4)	Omission (5)	Pooled (6)
$(\mathbb{Var}(\eta\_cf_{t:t+1} - \mathbb{Var}(\eta\_cf_{t-1:t})/mean(\mathbb{Var}(\eta\_cf_{t-5:t}))$	−9.00%	−17.04%	−11.75%	5.34%	11.58%	7.50%
	(−3.61)	(−5.21)	(−5.91)	(1.97)	(2.87)	(3.33)
$(\mathbb{Var}(\eta\_cf_{t:t+2} - \mathbb{Var}(\eta\_cf_{t-2:t})/mean(\mathbb{Var}(\eta\_cf_{t-5:t}))$	−14.29%	−21.97%	−16.91%	9.36%	11.91%	10.40%
	(−7.96)	(−8.33)	(−11.37)	(4.64)	(4.24)	(6.36)
$(\mathbb{Var}(\eta\_cf_{t:t+3} - \mathbb{Var}(\eta\_cf_{t-3:t})/mean(\mathbb{Var}(\eta\_cf_{t-5:t}))$	−15.98%	−21.63%	−17.71%	8.43%	11.93%	9.63%
	(−10.42)	(−9.21)	(−13.79)	(4.93)	(4.79)	(6.82)
$(\mathbb{Var}(\eta\_cf_{t:t+4} - \mathbb{Var}(\eta\_cf_{t-4:t})/mean(\mathbb{Var}(\eta\_cf_{t-5:t}))$	−16.12%	−23.58%	−18.48%	9.10%	8.14%	8.64%
	(−11.44)	(−10.96)	(−15.65)	(5.71)	(3.47)	(6.54)
$(\mathbb{Var}(\eta\_cf_{t:t+5} - \mathbb{Var}(\eta\_cf_{t-5:t})/mean(\mathbb{Var}(\eta\_cf_{t-5:t}))$	−14.86%	−20.01%	−16.43%	7.29%	6.09%	6.89%
	(−9.65)	(−4.94)	(−16.07)	(4.38)	(2.42)	(6.19)
Nobs	2,441	1,069	3,510	2,461	1,233	3,694

Table A.2 Scaled change in variance of cash-flow news around dividend events: buildup

This table reports changes in cash-flow news around dividend events using the methodology of Vuolteenaho (2002) that we describe in Section 2 We study the build-up of the change in the variance of cash-flow news after the dividend event over time when we expand the event window. Our sample period is 1964 until 2013.

	$\Delta Div > 0$			$\Delta Div < 0$		
	High constraints	Low constraints	Δ	High constraints	Low constraints	Δ
	(1)	(2)	(3)	(4)	(5)	(6)
	Panel A. Δ Scaled variance cash-flow news: $\Delta\,\mathbb{Var}(\eta\_cf)/mean(\mathbb{Var}(\eta\_cf))$					
	−16.98%	−11.58%	5.20%	9.10%	6.98%	−3.11%
	(−5.98)	(−4.56)	(6.78)	(3.07)	(2.56)	(−3.21)
Nobs	813	814		820	820	
	Panel B. Cumulative returns					
	1.02%	0.60%	−0.37%	−0.84%	−0.63%	0.23%
	(5.34)	(4.42)	(−7.13)	(−4.50)	(−3.25)	(4.48)
Nobs	813	814		820	820	

Table A.3 Sample split by financial constraints: scaled change in variance of cash-flow news and announcement returns around dividend events

This table reports the average change in the variance of cash-flow news scaled by the average variance of cash-flow news before the event ($\Delta\,\mathbb{Var}(\eta\_cf)/mean(\mathbb{Var}(\eta\_cf))$) using the methodology of Vuolteenaho (2002) that we describe in Section 2 in Panel A and announcement returns in Panel B. The table splits dividend events by the Hadlock-Pierce Index using the terciles as cutoff excluding the middle tercile. Announcement returns are cumulative returns in a three-day window bracketing the dividend event. We bootstrap the difference between large and small financial constraints. Our sample period is 1964 until 2013.

	Before (1)	After (3)
Initiations	17.82%	21.54%
Dividend Increases	22.38%	21.31%
Omissions	20.62%	21.29%
Dividend Decreases	21.56%	21.34%

Table A.4 Correlation between variance of cash-flow news and stock return volatility

This table reports correlations of the variance of cash-flow news using the methodology of Vuolteenaho (2002) that we describe in Section 2 and stock-return volatility which we calculate using five years of daily data around dividend events. The sample period is from 1964 until 2013.

	$\Delta Div > 0$	Initiation	Pooled	$\Delta Div < 0$	Omission	Pooled
	(1)	(2)	(3)	(4)	(5)	(6)
			Panel A. Δ Annualized return variance			
	18.42%	18.16%	18.31%	–6.42%	5.43%	-4.24%
	(8.39)	(5.24)	(9.82)	(–2.77)	(0.93)	(-1.95)
			Panel B. Δ Scaled annualized return variance			
	3.76%	3.70%	3.73%	–1.31%	1.11%	-0.86%
	(8.39)	(5.24)	(9.82)	(–2.77)	(0.93)	(-1.95)
Nobs	2,441	1,069	3,510	2,461	1,233	3,694

Table A.5 Scaled change in stock return volatility around dividend events

This table reports changes in annualized stock return volatility around dividend events using using five years of daily data before and after the dividend event. Panel A reports the average change in annualized stock return volatility and Panel B reports the average change in annualized stock return volatility scaled by the average stock return volatility before the event. Our sample period is 1964 until 2013.

	$\Delta Div > 0$ (1)	Initiation (2)	Pooled (3)	$\Delta Div < 0$ (4)	Omission (5)	Pooled (6)
	0.72%	2.37%	1.22%	-0.70%	-8.68%	-3.38%
	(7.69)	(11.00)	(13.11)	(-6.11)	(-29.77)	(-24.37)
Nobs	2,441	1,069	3,510	2,461	1,233	3,694

Table A.6 Announcement returns

This table reports three-day cumulative returns on dividend event days for a sample period from 1964 until 2013.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
ΔDiv	−0.26	−0.24	−0.37	−0.13	−0.15	−0.14	−0.25	−0.11
	(−5.55)	(−5.29)	(−5.94)	(−2.76)	(−4.92)	(−4.64)	(−5.04)	(−2.90)
eps		−0.17	−0.12	−0.22		(−0.14)	(−0.10)	(−0.16)
		(−1.59)	(−1.91)	(−4.41)		(−1.43)	(−1.77)	(−3.21)
$\Delta Div \times eps$			0.24	0.17			0.20	0.16
			3.13	3.18			2.67	2.87
$\mathbb{Var}_{t-1}(\eta\_cf)$				−93.65				−80.83
				(−29.17)				(−32.08)
Age				0.00				0.00
				(1.23)				(0.24)
Book-to-market				81.87				112.71
				(0.93)				(1.81)
Leverage				0.16				0.14
				(1.36)				(1.33)
Size				0.04				0.01
				(2.57)				(0.87)
Constant	0.03	0.12	0.08	0.09				
	(0.45)	(1.24)	(1.04)	(0.29)				
Year FE					X	X	X	X
Industry FE					X	X	X	X
R2	2.06%	2.89%	3.89%	39.16%	30.60%	31.15%	31.80%	52.24%
Nobs	3,127	3,127	3,127	3,127	3,127	3,127	3,127	3,127

Table A.7 Regression of changes in variance of cash-flow news around dividend events: initial variance

This table reports estimates from the following specification:

$$\Delta\, \mathbb{Var}(\eta\_cf_{it}) = \alpha + \beta_1 \cdot \Delta D_{it} + \beta_2 \cdot eps_{it} + \beta_3 \cdot \Delta D_{it} \cdot eps_{it} + \delta \cdot X_{it} + \varepsilon_{it}.$$

We regress changes in the scaled variance of cash-flow news around dividend events using the methodology of Vuolteenaho (2002) that we describe in Section 2 of firm i at event t, $\Delta\, \mathbb{Var}(\eta\_cf_{it})$, on the dividend change, ΔD_{it}, earnings per share, eps_{it}, the interaction between the two, and additional covariates, X_{it}, with t-statistics in parentheses. Additional covariates include firm age, size, book-to-market, and financial leverage. We add year and industry fixed effects at the Fama and French 17 industry level whenever indicated. We cluster standard errors at the dividend-quarter level. Our sample period is 1964 until 2013.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
ΔDiv	−0.26	−0.24	−0.37	−0.35	−0.15	−0.14	−0.25	−0.23
	(−5.55)	(−5.31)	(−5.94)	(−6.05)	(−4.92)	(−4.66)	(−5.01)	(−5.03)
eps		−0.17	−0.12	−0.18		−0.14	−0.10	−0.23
		(−1.56)	(−1.87)	(−2.76)		(−1.41)	(−1.75)	(−1.78)
$\Delta Div \times eps$			0.24	0.21			0.19	0.18
			(3.12)	(3.20)			(2.64)	(2.54)
Age				0.00				0.00
				(1.44)				(1.21)
Book-to-Market				34.02				134.39
				(0.40)				(2.44)
Leverage				−0.35				−0.14
				(−2.58)				(−1.12)
Size				0.05				0.02
				(3.35)				(1.23)
Cash				−0.00				−0.00
				(−2.35)				(−1.40)
Constant	0.03	0.12	0.08	−0.86				
	(0.45)	(1.22)	(1.01)	(−2.75)				
Year FE					X	X	X	X
Industry FE					X	X	X	X
R2	2.06%	2.89%	3.89%	5.24%	30.60%	31.15%	31.80%	32.27%
Nobs	3,127	3,127	3,127	3,127	3,127	3,127	3,127	3,127

Table A.8 Regression of changes in variance of cash-flow news around dividend events

This table reports estimates from the following specification:

$$\Delta \mathbb{V}\mathrm{ar}(\eta\_cf_{it}) = \alpha + \beta_1 \cdot \Delta D_{it} + \beta_2 \cdot eps_{it} + \beta_3 \cdot \Delta D_{it} \cdot eps_{it} + \delta \cdot X_{it} + \varepsilon_{it}.$$

We regress changes in the scaled variance of cash-flow news around dividend events using the methodology of Vuolteenaho (2002) that we describe in Section 2 of firm i at event t, $\Delta \mathbb{V}\mathrm{ar}(\eta\_cf_{it})$, on the dividend change, ΔD_{it}, earnings per share, eps_{it}, the interaction between the two, and additional covariates, X_{it}, with t-statistics in parentheses. Additional covariates include firm age, size, book-to-market, financial leverage, and cash. We add year and industry fixed effects at the Fama and French 17 industry level whenever indicated. We cluster standard errors at the dividend-quarter level. Our sample period is from 1964 until 2013.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
ΔDiv	−0.26	−0.27	−0.26	−0.25	−0.15	−0.16	−0.15	−0.14
	(−5.55)	(−5.62)	(−5.47)	(−5.71)	(−4.92)	(−5.11)	(−4.82)	(−4.90)
CF		−0.09	−0.11	−0.09		−0.06	−0.08	−0.08
		(−2.54)	(−3.53)	(−2.87)		(−2.47)	(−3.17)	(−3.20)
$\Delta Div \times CF$			0.02	0.02			0.02	0.02
			(3.05)	(2.91)			(2.40)	(2.62)
Age				0.00				0.00
				(1.38)				(1.23)
Book-to-market				57.49				150.01
				(0.62)				(2.55)
Leverage				−0.33				−0.11
				(−2.43)				(−0.95)
Size				0.03				0.00
				(1.92)				(−0.16)
Constant	0.03	0.03	0.02	−0.53				
	(0.45)	(0.34)	(0.32)	(−1.76)				
Year FE					X	X	X	X
Industry FE					X	X	X	X
R2	2.06%	2.61%	2.75%	3.52%	30.60%	30.80%	30.94%	31.45%
Nobs	3,127	3,127	3,127	3,127	3,127	3,127	3,127	3,127

Table A.9 Regression of changes in variance of cash-flow news around dividend events

This table reports estimates from the following specification:

$$\Delta\, \mathbb{Var}(\eta\_cf_{it}) = \alpha + \beta_1 \cdot \Delta D_{it} + \beta_2 \cdot CF_{it} + \beta_3 \cdot \Delta D_{it} \cdot CF_{it} + \delta \cdot X_{it} + \varepsilon_{it}.$$

We regress changes in the scaled variance of cash-flow news around dividend events using the methodology of Vuolteenaho (2002) that we describe in Section 2 of firm i at event t, $\Delta\, \mathbb{Var}(\eta\_cf_{it})$, on the dividend change, ΔD_{it}, cash flow, CF_{it}, the interaction between the two, and additional covariates, X_{it}, with t-statistics in parentheses. Additional covariates include firm age, size, book-to-market, financial leverage, and cash. We add year and industry fixed effects at the Fama and French 17 industry level whenever indicated. We cluster standard errors at the dividend-quarter level. Our sample period is from 1964 until 2013.

	$\Delta Div > 0$			$\Delta Div < 0$		
	Large vol	Small vol	Δ	Large vol	Small vol	Δ
	(1)	(2)	(3)	(4)	(5)	(6)
	Panel A. Δ Scaled variance cash-flow news: $\Delta\,\mathbb{V}\mathrm{ar}(\eta\_cf)/mean(\mathbb{V}\mathrm{ar}(\eta\_cf))$					
	−16.49%	−13.13%	−3.74%	9.61%	3.79%	3.45%
	(−7.23)	(−6.39)	(−6.27)	(4.28)	(1.55)	(4.64)
Nobs	1,262	1,179		1,482	979	

	$\Delta Div > 0$			$\Delta Div < 0$		
	Large vol	Small vol	Δ	Large vol	Small vol	Δ
	(1)	(2)	(3)	(4)	(5)	(6)
	Panel B. Announcement returns					
	0.93%	0.48%	0.44%	−0.93%	−0.31%	−0.73%
	(5.97)	(5.10)	(11.15)	(−5.44)	(−2.85)	(16.34)
Nobs	1,262	1,179		1,482	979	

Table A.10 Scaled change in variance of cash-flow news and announcement returns around dividend events: total vol

This table reports the average change in the variance of cash-flow news scaled by the average variance of cash-flow news before the event ($\Delta\,\mathbb{V}\mathrm{ar}(\eta\_cf)/mean(\mathbb{V}\mathrm{ar}(\eta\_cf))$) using the methodology of Vuolteenaho (2002) that we describe in Section 2 In Panel A and announcement returns in Panel B. The table splits firms by their ex-ante total stock return volatility. Specifically, we first calculate a firm's ex-ante total volatility on a four-quarter rolling basis using daily data. We then assign a firm into the large total volatility sample if it had a volatility above the 30% percentile of firm volatility in the respective Fama and French 17 industry in the quarter before the dividend event. Announcement returns are cumulative returns in a three-day window bracketing the dividend event. We bootstrap the difference between large and small changes. Our sample period is from 1964 until 2013.