Internet Appendix for
"Investment Efficiency and Product Market Competition"

April 2016

This Internet Appendix Contains:

A Proof of Proposition 1

Since $q_i^e(s_i) = a_i^e + b_i^e \delta_i (s_i - \mu)$, it follows from Eq. (4) that

(A.1)
$$2(\beta + c_i) a_i^e = \mu - \gamma \sum_{\substack{j=1 \\ j \neq i}}^{n} a_j^e,$$

and

(A.2)
$$2(\beta + c_i) b_i^e = \left(1 - \gamma \sum_{\substack{j=1 \\ j \neq i}}^{n} b_j^e \delta_j\right),$$

for all $i = 1, \ldots, n$. Subtracting γa_i^e from both sides of Eq. (A.1) and rearranging terms of the resulting equation yields

(A.3)
$$a_i^e = \frac{1}{2\beta + 2c_i - \gamma}\left(\mu - \gamma \sum_{j=1}^{n} a_j^e\right),$$

for all $i = 1, \ldots, n$. Summing over all i of Eq. (A.3) yields

(A.4)
$$\sum_{j=1}^{n} a_j^e = \left(\sum_{j=1}^{n} \frac{1}{2\beta + 2c_j - \gamma}\right)\left(\mu - \gamma \sum_{j=1}^{n} a_j^e\right).$$

Rearranging terms of Eq. (A.4) yields

(A.5)
$$\sum_{j=1}^{n} a_j^e = \left(\sum_{j=1}^{n} \frac{\mu}{2\beta + 2c_j - \gamma}\right) \Big/ \left(1 + \sum_{j=1}^{n} \frac{\gamma}{2\beta + 2c_j - \gamma}\right).$$

Substituting Eq. (A.5) into Eq. (A.3) yields Eq. (5).

Subtracting $\gamma b_i^e \delta_i$ from both sides of Eq. (A.2) and rearranging terms of the resulting equation yields

$$b_i^e = \left(\frac{1}{2\beta + 2c_i - \gamma\delta_i}\right)\left(1 - \gamma\sum_{j=1}^{n} b_j^e \delta_j\right), \tag{A.6}$$

for all $i = 1, \ldots, n$. Multiplying δ_i to both sides of Eq. (A.6) and summing over all i yields

$$\sum_{j=1}^{n} b_j^e \delta_j = \left(\sum_{j=1}^{n} \frac{\delta_j}{2\beta + 2c_j - \gamma\delta_j}\right)\left(1 - \gamma\sum_{j=1}^{n} b_j^e \delta_j\right). \tag{A.7}$$

Rearranging terms of Eq. (A.7) yields

$$\sum_{j=1}^{n} b_j^e \delta_j = \left(\sum_{j=1}^{n} \frac{\delta_j}{2\beta + 2c_j - \gamma\delta_j}\right) \Big/ \left(1 + \sum_{j=1}^{n} \frac{\gamma\delta_j}{2\beta + 2c_j - \gamma\delta_j}\right). \tag{A.8}$$

Substituting Eq. (A.8) into Eq. (A.6) yields Eq. (6).

B Proof of Proposition 2

Differentiating $R_i(t_j)$ with respect to t_j yields

$$R_i'(t_j) = -\frac{\partial^2 \pi_i(t_1, t_2)}{\partial t_i \partial t_j} \Big/ \frac{\partial^2 \pi_i(t_1, t_2)}{\partial t_i^2}, \tag{A.9}$$

where

$$\frac{\partial^2 \pi_i(t_1, t_2)}{\partial t_i \partial t_j} = -\frac{4(\beta + c_1)(\beta + c_2)(2\beta + 2c_1 - \gamma\delta_1)(2\beta + 2c_2 - \gamma\delta_2)\gamma h}{[4(\beta + c_1)(\beta + c_2) - \gamma^2\delta_1\delta_2]^3 (t_1 + h)^2 (t_2 + h)^2} < 0, \tag{A.10}$$

and

$$\text{(A.11)} \qquad \frac{\partial^2 \pi_i(t_1, t_2)}{\partial t_i^2} = -\frac{2(\beta + c_i)(2\beta + 2c_j - \gamma\delta_j)^2[4(\beta + c_1)(\beta + c_2) - \gamma^2\delta_j]}{[4(\beta + c_1)(\beta + c_2) - \gamma^2\delta_i\delta_j]^3(t_i + h)^3} < 0.$$

Hence, it follows from Eqs. (A.9), (A.10), and (A.11) that $R_i'(t_j) < 0$ for $i = 1$ and 2.

When $t_j = 0$, we can solve Eq. (11) to yield

$$\text{(A.12)} \qquad R_i(0) = \frac{1 - 2h\sqrt{(\beta + c_i)\lambda}}{2\sqrt{(\beta + c_i)\lambda}}.$$

Using Eq. (11), we solve $R_j(\bar{t}_i) = 0$ to yield

$$\text{(A.13)} \qquad \bar{t}_i = \frac{2h(\beta + c_i)[1 - 2h\sqrt{(\beta + c_j)\lambda}]}{\gamma - 2(\beta + c_i)[1 - 2h\sqrt{(\beta + c_j)\lambda}]}.$$

It follows from Eqs. (A.12) and (A.13) with $i = 1$ and condition (12) that $R_1(0) < \bar{t}_1$. Since $c_1 < (=)\, c_2$, condition (12) implies that $\lambda < 1/4(\beta + c_1)h^2$. Hence, Eq. (A.12) with $i = 1$ implies that $R_1(0) > 0$ and thus $\bar{t}_1 > 0$. It then follows from $R_2'(t_i) < 0$ that $R_2(0) > 0$. From Eq. (A.12) with $i = 2$, we have $\lambda < 1/4(\beta + c_2)h^2$, which implies that

$$\lambda < \frac{1}{4(\beta + c_2)h^2}\left[\frac{2\beta + 2c_2 - \gamma}{2\sqrt{(\beta + c_1)(\beta + c_2)} - \gamma}\right]^2,$$

since $c_1 < (=)\, c_2$. It then follows from Eqs. (A.12) and (A.13) with $i = 2$ that $R_2(0) < \bar{t}_2$.

Since $R_i(0) < \bar{t}_i$ and $R_i'(t_j) < 0$ for $i = 1$ and 2, there must exist a unique perfect Bayesian equilibrium pair of precision levels, $t = (t_1^*, t_2^*)$, as depicted in Figure 1. Equating Eq. (11) for $i = 1$

and 2 at $t = (t_1^*, t_2^*)$ yields

$$\text{(A.14)} \qquad \frac{(\beta + c_1)(2\beta + 2c_2 - \gamma\delta_2^*)^2}{(t_1^* + h)^2} = \frac{(\beta + c_2)(2\beta + 2c_1 - \gamma\delta_1^*)^2}{(t_2^* + h)^2}.$$

Solving Eq. (A.14) for t_1^* in terms of t_2^* yields Eq. (13). Note that

$$\text{(A.15)} \qquad \frac{\partial}{\partial c}\left(\frac{2\beta + 2c - \gamma}{\sqrt{\beta + c}}\right) = \frac{2\beta + 2c + \gamma}{2(\beta + c)^{2/3}} > 0.$$

It follows from $c_1 < (=) c_2$ and Eq. (A.15) that $(2\beta + 2c_1 - \gamma)/\sqrt{\beta + c_1} < (=)\ (2\beta + 2c_2 - \gamma)/\sqrt{\beta + c_2}$. Hence, Eq. (13) implies that $t_1^* > (=)\ t_2^*$.

Since α and ε_i are independent, we can use Eqs. (5), (6), and (9) to write Eq. (10) as[1]

$$\text{(A.16)} \qquad D_i = E\left\{\left[\frac{\mu + B_i^*(\alpha - \mu)}{\alpha}\right]^2 + \frac{1}{t_i^*}\left(\frac{B_i^*}{\alpha}\right)^2\right\} - 1,$$

where the coefficient, B_i^*, is given by

$$\text{(A.17)} \qquad B_i^* = \delta_i^*\left(\frac{2\beta + 2c_i - \gamma}{2\beta + 2c_i - \gamma\delta_i^*}\right)\left(1 + \sum_{j=1}^{n}\frac{\gamma}{2\beta + 2c_j - \gamma}\right)\Bigg/\left(1 + \sum_{j=1}^{n}\frac{\gamma\delta_j^*}{2\beta + 2c_j - \gamma\delta_j^*}\right).$$

When $n = 2$, Eq. (A.17) becomes

$$\text{(A.18)} \qquad B_i^* = \frac{\delta_i^*(2\beta + 2c_j - \gamma\delta_j^*)[4(\beta + c_i)(\beta + c_j) - \gamma^2]}{(2\beta + 2c_j - \gamma)[4(\beta + c_i)(\beta + c_j) - \gamma^2\delta_i^*\delta_j^*]}.$$

[1] Ingersoll (1987) shows that two random variables, $\tilde{X}$ and $\tilde{Y}$, are independent if $Cov[A(\tilde{X}), B(\tilde{Y})] = 0$ for all functions, $A(\cdot)$ and $B(\cdot)$.

Eq. (A.18) implies that $B_1^* > B_2^*$ if

(A.19)
$$\frac{\delta_1^*(2\beta + 2c_2 - \gamma\delta_2^*)}{2\beta + 2c_2 - \gamma} > \frac{\delta_2^*(2\beta + 2c_1 - \gamma\delta_1^*)}{2\beta + 2c_1 - \gamma}.$$

Using Eq. (A.14), we can write inequality (A.19) as

(A.20)
$$t_1^* > t_2^*\left(\frac{2\beta + 2c_2 - \gamma}{2\beta + 2c_1 - \gamma}\right)\sqrt{\frac{\beta + c_1}{\beta + c_2}}.$$

It follows from Eq. (13) that inequality (A.20) indeed holds so that $B_1^* > B_2^*$ if $c_1 < c_2$. On the other hand, if $c_1 = c_2$, we have $B_1^* = B_2^*$.

Using Eq. (A.18), we differentiate B_1^* with respect to t_2^* to yield

(A.21)
$$\frac{\partial B_1^*}{\partial t_2^*} = \frac{2(\beta + c_2)[4(\beta + c_1)(\beta + c_2) - \gamma^2]}{(2\beta + 2c_2 - \gamma)[4(\beta + c_1)(\beta + c_2) - \gamma\delta_1^*\delta_2^*]^2} \times\left[2(\beta + c_1)(2\beta + 2c_2 - \gamma\delta_2^*)\frac{\partial \delta_1^*}{\partial t_2^*} - \gamma\delta_1^*(2\beta + 2c_1 - \gamma\delta_1^*)\frac{\partial \delta_2^*}{\partial t_2^*}\right].$$

Note that $\partial\delta_2^*/\partial t_2^* = h/(t_2^* + h)^2 > 0$ and $\partial\delta_1^*/\partial t_2^* = [h/(t_1^* + h)^2] \times \partial t_1^*/\partial t_2^* > (=)\, h/(t_1^* + h)^2$ since $\partial t_1^*/\partial t_2^* > (=)\ 1$ from Eq. (13) if $c_1 < (=)\ c_2$. Hence, the expression inside the squared brackets of the last term on the right-hand side of Eq. (A.21) is greater than (equal to)

(A.22)
$$h\left[\frac{2(\beta + c_1)(2\beta + 2c_2 - \gamma\delta_2^*)}{(t_1^* + h)^2} - \frac{\gamma\delta_1^*(2\beta + 2c_1 - \gamma\delta_1^*)}{(t_2^* + h)^2}\right] = \frac{h(2\beta + 2c_1 - \gamma\delta_1^*)[4(\beta + c_2)(\beta + c_1 - \gamma\delta_1^*) + \gamma^2\delta_1^*\delta_2^*]}{(t_2^* + h)^2(2\beta + 2c_2 - \gamma\delta_2^*)} > 0,$$

where the equality follows from Eq. (A.14) if $c_1 < (=)\ c_2$. Hence, Eqs. (A.21) and (A.22) imply that $\partial B_1^*/\partial t_2^* > 0$. If t_2^* goes to infinity, Eq. (13) implies that t_1^* also goes to infinity. It follows from

Eq. (A.18) with $i = 1$ that $B_1^* = 1$ since $\delta_1^* = \delta_2^* = 1$ in this limiting case. Hence, we conclude that $B_1^* < 1$.

Note that

(A.23)
$$\begin{aligned}&\frac{\partial}{\partial B_i^*}E\left\{\left[\frac{\mu + B_i^*(\alpha-\mu)}{\alpha}\right]^2\right\}\\ =\ & 2E\left\{\left[\frac{\mu + B_i^*(\alpha-\mu)}{\alpha}\right]\left(\frac{\alpha-\mu}{\alpha}\right)\right\}\\ =\ & 2\left[B_i^* + (1-B_i^*)\mu E\left(\frac{1}{\alpha}\right)\right]\left[1-\mu E\left(\frac{1}{\alpha}\right)\right] - 2(1-B_i^*)\mu^2 Var\left(\frac{1}{\alpha}\right) < 0,\end{aligned}$$

where the inequality follows from the fact that $0 < B_i^* < 1$, and Jensen's inequality that $E(1/\alpha) > 1/\mu$. Eqs. (A.16) and (A.23) then imply that

(A.24)
$$D_i > E\left\{\left[\frac{\mu + (\alpha-\mu)}{\alpha}\right]^2 + \frac{1}{t_i^*}\left(\frac{B_i^*}{\alpha}\right)^2\right\} - 1 = \frac{B_i^{*2}}{t_i^*}E\left(\frac{1}{\alpha^2}\right) > 0,$$

where the inequality follows from the fact that $B_i^* < 1$.

When $c_1 < c_2$, we have $B_1^{*2}/t_1^* \leq B_2^{*2}/t_2^*$ if

$$\frac{t_1^*(2\beta + 2c_2 - \gamma\delta_2^*)^2}{(t_1^* + h)^2(2\beta + 2c_2 - \gamma)^2} \leq \frac{t_2^*(2\beta + 2c_1 - \gamma\delta_1^*)^2}{(t_2^* + h)^2(2\beta + 2c_1 - \gamma)^2},$$

which, using Eq. (A.14), reduces to

(A.25)
$$t_1^* \leq t_2^*\left(\frac{2\beta + 2c_2 - \gamma}{2\beta + 2c_1 - \gamma}\right)^2\left(\frac{\beta + c_1}{\beta + c_2}\right).$$

Substituting Eq. (13) into inequality (A.25) yields

$$t_2^*\left(\frac{2\beta+2c_2-\gamma}{2\beta+2c_1-\gamma}\right)\sqrt{\frac{\beta+c_1}{\beta+c_2}} \tag{A.26}$$
$$\geq\ h-\frac{\gamma h}{2\beta+2c_1-\gamma}\left(1-\sqrt{\frac{\beta+c_1}{\beta+c_2}}\right)\left[\left(\frac{2\beta+2c_2-\gamma}{2\beta+2c_1-\gamma}\right)\sqrt{\frac{\beta+c_1}{\beta+c_2}}-1\right]^{-1}.$$

Inequality (A.26) holds if $t_2^* \geq h.$ Hence, a sufficient (but not necessary) condition for $D_1 > D_2$ is that the prior precision level, h, to be sufficiently small.

C Proof of Proposition 3

The perfect Bayesian equilibrium pair of precision levels, $t^* = (t_1^*, t_2^*)$, is characterized by the solution to the following system of equations:

$$\frac{\partial \pi_1(t_1^*, t_2^*)}{\partial t_1}-\lambda=0, \tag{A.27}$$

and

$$t_1^*-t_2^*\left(\frac{2\beta+2c_2-\gamma}{2\beta+2c_1-\gamma}\right)\sqrt{\frac{\beta+c_1}{\beta+c_2}}-2\left(\frac{\beta+c_1}{2\beta+2c_1-\gamma}\right)\left(\sqrt{\frac{\beta+c_2}{\beta+c_1}}-1\right)h=0, \tag{A.28}$$

where Eq. (A.28) is simply Eq. (13). Let

$$H=-\left(\frac{2\beta+2c_2-\gamma}{2\beta+2c_1-\gamma}\right)\sqrt{\frac{\beta+c_1}{\beta+c_2}}\times\frac{\partial^2\pi_1(t_1^*, t_2^*)}{\partial t_1^2}-\frac{\partial^2\pi_1(t_1^*, t_2^*)}{\partial t_1 \partial t_2}>0, \tag{A.29}$$

where $\partial^2\pi_1(t_1^*, t_2^*)/\partial t_1 \partial t_2 < 0$ and $\partial^2\pi_1(t_1^*, t_2^*)/\partial t_1^2 < 0$ are given by Eqs. (A.10) and (A.11) with $i = 1$ and evaluated at $t^* = (t_1^*, t_2^*)$. Totally differentiating Eqs. (A.27) and (A.28) with respect to γ yields

(A.30)
$$\frac{dt_1^*}{d\gamma} = \frac{1}{H}\left[\left(\frac{2\beta + 2c_2 - \gamma}{2\beta + 2c_1 - \gamma}\right)\sqrt{\frac{\beta + c_1}{\beta + c_2}} \times \frac{\partial^2\pi_1(t_1^*, t_2^*)}{\partial t_1 \partial \gamma} - \left(\frac{1}{2\beta + 2c_1 - \gamma}\right)\left(t_1^* - t_2^*\sqrt{\frac{\beta + c_1}{\beta + c_2}}\right) \times \frac{\partial^2\pi_1(t_1^*, t_2^*)}{\partial t_1 \partial t_2}\right],$$

and

(A.31)
$$\frac{dt_2^*}{d\gamma} = \frac{1}{H}\left[\frac{\partial^2\pi_1(t_1^*, t_2^*)}{\partial t_1 \partial \gamma} + \left(\frac{1}{2\beta + 2c_1 - \gamma}\right)\left(t_1^* - t_2^*\sqrt{\frac{\beta + c_1}{\beta + c_2}}\right) \times \frac{\partial^2\pi_1(t_1^*, t_2^*)}{\partial t_1^2}\right],$$

where $H > 0$ is given by Eq. (A.29) and

(A.32)
$$\frac{\partial^2\pi_1(t_1, t_2)}{\partial t_1 \partial \gamma} = \frac{t_2^*}{\gamma(1 - \delta_2^*)} \times \frac{\partial^2\pi_1(t_1, t_2)}{\partial t_1 \partial t_2} < 0.$$

Hence, Eq. (A.31) implies that $dt_2^*/d\gamma < 0$.

Substituting Eq. (A.32) into Eq. (A.30) yields

(A.33)
$$\frac{dt_1^*}{d\gamma} = \frac{1}{H}\left(\frac{2\beta + 2c_2 - \gamma}{2\beta + 2c_1 - \gamma}\right)\sqrt{\frac{\beta + c_1}{\beta + c_2}} \times \frac{\partial^2\pi_1(t_1^*, t_2^*)}{\partial t_1 \partial t_2} \times \left[\frac{t_2^*}{\gamma(1 - \delta_2^*)} - \left(\frac{1}{2\beta + 2c_2 - \gamma}\right)\left(t_1^*\sqrt{\frac{\beta + c_2}{\beta + c_1}} - t_2^*\right)\right].$$

The expression inside the squared brackets on the right-hand side of Eq. (A.33) can be written as

$$
\begin{aligned}
\text{(A.34)} \qquad & \left(\frac{2\beta + 2c_2 - \gamma\delta_2^*}{2\beta + 2c_2 - \gamma}\right)\left[\frac{t_2^*}{\gamma(1-\delta_2^*)} - \left(\frac{t_1^*}{2\beta + 2c_2 - \gamma\delta_2^*}\right)\sqrt{\frac{\beta + c_2}{\beta + c_1}}\right] \\
= \ & \left(\frac{2\beta + 2c_2 - \gamma\delta_2^*}{2\beta + 2c_2 - \gamma}\right)\left[\frac{t_2^*}{\gamma(1-\delta_2^*)} - \frac{\delta_1^*(t_2^* + h)}{2\beta + 2c_1 - \gamma\delta_1^*}\right] \\
= \ & \left[\frac{(2\beta + 2c_2 - \gamma\delta_2^*)t_2^*}{(2\beta + 2c_2 - \gamma)(2\beta + 2c_1 - \gamma\delta_1^*)\gamma(1-\delta_2^*)}\right]\left(2\beta + 2c_1 - \frac{\gamma\delta_1^*}{\delta_2^*}\right),
\end{aligned}
$$

where the first equality follows from Eq. (A.14). Hence, it follows from Eqs. (A.33) and (A.34) that $dt_1^*/d\gamma > 0$ if, and only if, $\delta_1^*/\delta_2^* < 2(\beta + c_1)/\gamma$.

D Competition and Entry

In this section we demonstrate that our results generalize when the number of firms in an industry is endogenized through an entry game. This allows us to demonstrate that competition increases in the number of firms while it decreases with respect to the realized HHI.

We therefore add on an initial stage where each firm chooses simultaneously to pay a fixed entry cost, F, and then subsequently compete with other firms by investing in information production and product market competition. Clearly in this entry game, the number of firms is determined (assuming that F is such that the number of firms must be an integer) by the expression that expected profit is equal to the entry cost. That is, using equation (6) along with the definition of δ^*, evaluated at the optimal precision level, t^*, we obtain

$$
\text{(A.35)} \qquad (\beta + c_i)(a_i^e)^2 + \lambda(t_i^* + h)\frac{t_i^*}{h} = F.
$$

We can show that the left hand side of equation (A.35) is decreasing in the number of firms, n. In fact consider

$$a_i^e = \frac{\mu}{2\beta + 2c_i + (n-1)\gamma},$$

from equation (4). Clearly this is decreasing in n and so the first term on the left hand side of (A.35) is decreasing in n. The second term is decreasing in n as long as t^* is decreasing in n. But from the symmetric Cournot oligopoly case in equation (14), we can see that this is also decreasing in n. Hence the left hand side of (A.35) is decreasing in n and there is an unique solution such that the equilibrium number of firms increases with a lower cost of entry, F. Given that cost of entry is associated with a more competitive environment, we find that a larger number of firms is associated with more competition. Furthermore in the symmetric oligopoly case we have $HHI = 1/n$ we know equivalently that more competition implies a lower HHI.

E Variable List

E.1 Investment Regression Variables

I

Investment is measured by the sum of capital expenditure (Data Item 128), research and development expenditure (Data Item 46), acquisitions (Data Item 129) and sale of property, plant and equipment (Data Item 107) minus amortization and depreciation (Data Item 125) scaled by lagged total assets (Data Item 6).

VP

Growth opportunities of the firm are measured by V over P, where $V = (1-\rho r)BV + \rho(1+r)X -$

$\rho r d$ is the value of assets in place. BV is the book value of assets (Data Item 60), X is operating income after depreciation (Data Item 178), r is the discount rate and $\rho = (\omega/(1-r-\omega))$, d is the annual dividend (Data Item 21). P is the market value of the firm as measured by (Data Item 25) multiplied by (Data Item 199). [2]

LEVERAGE

LEVERAGE is measured by the ratio of the sum of book value of short term debt (Data Item 34) and long term debt (Data Item 9) to the sum of book value of total debt (Data Item 34+ Data Item 9) and book value of equity (Data Item 60).

CASH

CASH is measured by cash and short term investment (Data Item 34) deflated by the total assets (Data Item 6) of the previous year.

AGE

AGE is measured by log of the years the firm has been listed on CRSP.

SIZE

SIZE is measured by the log of total assets (Data Item 6).

RETURN

RETURN is measured by the change of the market value of the firm (Data Item 25 * Data Item 199).

INDUSTRY CLASSIFICATIONS

The industry classifications depend on which competition measure is used. These vary between Compustat three-digit SIC, NAICS and Hoberg-Phillips 10-K text based classifications.

---

[2] We set r=12% and ω=0.62 as in Richardson (2006). ω is the abnormal earnings persistence parameter. In an unreported test, we use book to market instead of V/P to proxy for growth opportunities; our results are unaffected.

E.2 Competition Measures

HHI

The Herfindahl-Hirschman Index of industry j is calculated by

$$HHI_j = \sum_{i=1}^{I} {s_{ij}}^2$$

where s_{ij} is the share of sales of firm i in industry j.

NUM

This is the number of firms in the industry scaled by 1,000.

CENSUS HHI

Four-digit SIC or NAICS code industry level HHI of manufacturing firms by US Census Bureau, available every 5 years starting from 1982.

HP_HHI

TNIC HHI from Hoberg Philips data library. Industry is classified by a similarity score based on 10-K text scanning.

EPCM

The price cost margin is measured by the ratio of operating income before depreciation (Data Item 13) over sales (Data Item 12). EPCM is the price-cost margin minus industry median price-cost margin where industry is classified by three-digit SIC code.

(IMPORT) TARIFF

Four-digit SIC level tariff data from Peter Schott's website.

FREIGHT COST

Four-digit SIC level freight cost data are derived from Peter Schott's website.

TRADE COST

The sum of tariff and freight cost.

E.3 Inefficiency Measures

INF

Absolute value of the residual from the investment regression (17).

E.4 Other Measures

MTB

MTB is calculated as the sum of market value of equity (Data Item 25 * Data Item 199) and book value of debt (Data Item 34+ Data Item 9) divide the book value of asset (Data Item 6).

TANGIBLE

TANGIBLE measures the tangibility of the assets of the firm. It is defined as the ratio of net total property, plant and equipment (Data Item 8) to total assets (Data Item 6).

FCF

Details are provided in the text.

IDVOL

Details are provided in the text.

DOWN_OUTPUT

Downstream output measured by industry GDP. The data is obtained from the website of Bureau of Economic Analysis. `http://www.bea.gov/industry/gdpbyind_data.htm`.

DOWN_SALEG

Average of sales growth of downstream industry firms. Sales is obtained from COMPUSTAT(Data Item 12).

DOWN_MTB

Average of MTB of downstream industry firms. The calculation of MTB is the same as before.

DOWN_ANALYST

Average of long-term growth analysts' forecast of downstream industry firms. The data is obtained from I/B/E/S.

TABLE A.1

Correlation Matrix of Inefficiency and Competition Measures

The two inefficiency measures are constructed as the absolute value of the residual from the two optimal investment regressions respectively when HHI is used as the competition measure. The competition measures are Compustat HHI, the number of firms, U.S. Census HHI, Hoberg-Philipps HHI, the deviation from industry median price-cost margin, tariff, freight cost and trade cost respectively. The p-value is reported below the correlation coefficient.

Panel A. Inefficiency Measures

Correlation	1	2
	INF(HHI)	INF(DOWN)
INF(HHI)	1.00	
INF(DOWN)	0.75	1.00
p-value	0.00	

Panel B. Competition Measures

Correlation	1	2	3	4	5	6	7	8
	HHI	NUM	CENSUS HHI	HP_HHI	EPCM	T	FC	TC
HHI	1							
NUM	-0.48	1						
p-value	0.00							
CENSUS HHI	0.09	0.08	1					
p-value	0.00	0.00						
HP_HHI	0.19	-0.16	-0.13	1				
p-value	0.00	0.00	0.00					
EPCM	0.08	-0.15	-0.01	0.04	1			
p-value	0.00	0.00	0.33	0.00				
TARIFF	0.26	-0.23	-0.14	0.07	0.08	1		
p-value	0.00	0.00	0.00	0.00	0.00			
FREIGHT COST	0.18	-0.40	-0.15	0.09	0.08	0.23	1	
p-value	0.00	0.00	0.00	0.00	0.00	0.00		
TRADE COST	0.27	-0.42	-0.18	0.11	0.11	0.72	0.84	1
p-value	0.00	0.00	0.00	0.00	0.00	0.00	0.00	

TABLE A.2

Optimal Investment Regression: Robustness Competition Measures

The regression is as follows:

$$
\begin{aligned}
I_{i,t} \quad = \quad & \beta_0 + \beta_1 VP_{i,t-1} + \beta_2 \text{LEVERAGE}_{i,t-1} + \beta_3 \text{CASH}_{i,t-1} + \beta_4 \text{SIZE}_{i,t-1} \\
& + \beta_5 \text{RETURN}_{i,t-1} + \beta_6 \text{AGE}_{i,t-1} + \beta_7 I_{i,t-1} + \beta_8 \text{COMPETITION}_{j,t-1} + \epsilon_{i,t}.
\end{aligned}
$$

Table A.2 presents the result of the OLS investment regression. Column 1 uses price-cost margin as the measure of competition. Column 2 uses tariff as the measure of competition. Column 3 uses freight cost as the measure of competition. Column 4 uses trade cost (combination of tariff and freight cost) as the measure of competition. We use the fitted value of this regression as the expected investment as in in Richardson (2006) and the absolute value of residuals as the proxy for inefficiency. The sample covers firm-year data from 1980 to 2012. Firm and year fixed effects are included. Robust t-statistics are reported in parentheses. ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively.

Variables	1 $I_{i,t}$	2 $I_{i,t}$	3 $I_{i,t}$	4 $I_{i,t}$
VP	-0.032***	-0.020***	-0.020***	-0.020***
	(-12.37)	(-3.92)	(-3.91)	(-3.91)
LEVERAGE	-0.159***	-0.176***	-0.176***	-0.176***
	(-18.69)	(-8.34)	(-8.35)	(-8.34)
CASH	0.064***	0.016	0.016	0.016
	(8.00)	(0.93)	(0.93)	(0.93)
SIZE	-0.055***	-0.081***	-0.081***	-0.081***
	(-16.85)	(-9.48)	(-9.50)	(-9.49)
RETURN	0.017***	0.018***	0.018***	0.018***
	(10.62)	(4.84)	(4.85)	(4.85)
AGE	-0.013***	-0.019	-0.019	-0.019
	(-2.95)	(-1.55)	(-1.55)	(-1.55)
$I_{i,t-1}$	-0.001	-0.026	-0.026	-0.026
	(-0.05)	(-0.79)	(-0.79)	(-0.79)
EPCM	-0.004**			
	(-2.07)			
TARIFF		-0.019		
		(-0.13)		
FREIGHT COST			-0.002	
			(-0.04)	
TRADE COST				-0.005
				(-0.10)
Constant	0.402***	0.506***	0.505***	0.505***
	(22.75)	(11.15)	(12.08)	(11.68)
Observations	59,684	14,911	14,911	14,911
Firm Fixed Effects	Y	Y	Y	Y
Year Fixed Effects	Y	Y	Y	Y
Adjusted R^2	0.346	0.399	0.399	0.399

TABLE A.3

Inefficiency and Competition: Robustness Competition Measures

The regression is as follows:

$$\text{INF}_{i,t} = \beta_0 + \beta_1 \text{COMPETITION}_{j,t-1} + \beta' \mathbf{X_{i,t-1}} + \epsilon_{i,t}.$$

Table A.3 reports OLS regression results of INF on COMPETITION (either measured by EPCM, TARIFF, FREIGHT COST or TRADE COST). **X** is a set of firm level control variables. The sample covers firm-year data from 1980 to 2012. Standard Errors are two-way clustered at firm-year level. t-statistics are reported in the parentheses. ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively.

Variables	1 INF	2 INF	3 INF	4 INF	5 INF	6 INF	7 INF	8 INF
EPCM	-0.009***	-0.004***						
	(-9.27)	(-4.11)						
TARIFF			-0.278***	-0.217***				
			(-3.21)	(-2.79)				
FREIGHT COST					-0.168**	-0.057		
					(-2.03)	(-1.23)		
TRADE COST							-0.172***	-0.097**
							(-2.76)	(-2.20)
MTB		0.010***		0.009***		0.009***		0.009***
		(9.67)		(4.47)		(4.61)		(4.48)
LEVERAGE		0.018***		0.025*		0.024*		0.025*
		(4.88)		(1.92)		(1.91)		(1.95)
CASH		0.031***		0.041***		0.043***		0.042***
		(6.32)		(3.99)		(4.18)		(4.07)
SIZE		-0.004***		-0.006***		-0.005***		-0.005***
		(-5.64)		(-3.97)		(-3.99)		(-3.95)
TANGIBLE		0.039***		0.027**		0.031**		0.032***
		(6.90)		(2.28)		(2.57)		(2.69)
AGE		-0.005***		-0.003		-0.003		-0.003
		(-2.97)		(-1.02)		(-1.07)		(-1.06)
Constant	0.077***	0.072***	0.089***	0.079***	0.086***	0.070***	0.091***	0.075***
	(35.30)	(14.13)	(15.27)	(7.76)	(14.90)	(7.69)	(13.36)	(7.56)
Observations	59,684	58,915	14,911	14,804	14,911	14,804	14,911	14,804
Adjusted R^2	0.007	0.029	0.003	0.032	0.002	0.030	0.003	0.031

TABLE A.4

Inefficiency and Competition: Free Cash Flow

The regression is as follows:

$$\text{INF}_{i,t} = \beta_0 + \beta_1 \text{COMPETITION}_{j,t-1} + \beta' \mathbf{X_{i,t-1}} + \epsilon_{i,t}.$$

Table A.4 reports OLS regression results of INF on COMPETITION. **X** is a set of firm level control variables. Firms are grouped by their free cash flows. Firms in Columns 1,3,5 and 7 have low free cash flow while firms in Columns 2,4,6 and 8 have high free cash flow. The sample covers firm-year data from 1980 to 2012. Standard Errors are two-way clustered at firm-year level. t-statistics are reported in the parentheses. ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively.

Variables	1	2	3	4	5	6	7	8
	INF	INF	INF	INF	INF	INF	INF	INF
HHI	-0.016	-0.023***						
	(-1.33)	(-5.28)						
NUM			0.078**	0.146***				
			(2.57)	(4.33)				
CENSUS HHI					-0.153***	-0.389*		
					(-3.50)	(-1.76)		
HP_HHI							-0.030***	-0.020***
							(-2.73)	(-2.68)
MTB	0.014***	0.010***	0.014***	0.009***	0.016***	-0.019	0.013***	0.009***
	(6.40)	(5.19)	(6.39)	(4.75)	(4.39)	(-0.90)	(4.65)	(4.05)
LEVERAGE	0.019***	0.028***	0.020***	0.029***	-0.013	0.043*	0.027***	0.032***
	(3.11)	(3.22)	(3.38)	(3.35)	(-0.73)	(1.71)	(3.21)	(3.45)
CASH	0.016**	0.059***	0.013**	0.052**	0.004	0.561**	0.005	0.057*
	(2.46)	(2.63)	(2.03)	(2.45)	(0.23)	(2.10)	(0.63)	(1.94)
SIZE	-0.004***	-0.006***	-0.004***	-0.006***	-0.014***	-0.015	-0.006***	-0.008***
	(-2.93)	(-4.75)	(-2.88)	(-4.84)	(-4.34)	(-1.60)	(-3.39)	(-4.25)
TANGIBLE	0.041***	0.018***	0.044***	0.024***	-0.052	0.009	0.029**	0.001
	(3.82)	(3.25)	(4.21)	(3.99)	(-1.30)	(0.42)	(2.05)	(0.15)
AGE	-0.002	0.006**	-0.002	0.006**	-0.006	0.002	0.004	0.003
	(-0.79)	(2.45)	(-0.72)	(2.53)	(-0.67)	(0.13)	(1.07)	(1.05)
Constant	0.088***	0.043***	0.079***	0.030***	0.178***	0.140***	0.100***	0.066***
	(8.55)	(7.56)	(7.31)	(4.46)	(5.16)	(3.06)	(8.76)	(12.07)
Observations	24,032	24,058	24,045	24,072	2,122	2,124	15,379	15,418
Group	Low FCF	High FCF	Low FCF	High FCF	Low FCF	High FCF	Low FCF	High FCF
Adjusted R^2	0.030	0.014	0.030	0.015	0.087	0.020	0.025	0.011

TABLE A.5

Inefficiency and Competition: Price Informativeness

The regression is as follows:

$$\text{INF}_{i,t} = \beta_0 + \beta_1 \text{COMPETITION}_{j,t-1} + \beta' \mathbf{X_{i,t-1}} + \epsilon_{i,t}.$$

Table A.5 reports OLS regression results of INF on Competition. **X** is a set of firm level control variables. Firms are grouped by the level of idiosyncratic volatility. Firms in Columns 1,3,5 and 7 have low idiosyncratic volatility while firms in Columns 2,4,6 and 8 have high idiosyncratic volatility. The sample covers firm-year data from 1980 to 2012. Standard Errors are two-way clustered at firm-year level. t-statistics are reported in the parentheses. ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively.

Variables	1 INF	2 INF	3 INF	4 INF	5 INF	6 INF	7 INF	8 INF
HHI	-0.027***	-0.016*						
	(-4.20)	(-1.75)						
NUM			0.126***	0.121***				
			(3.75)	(3.35)				
CENSUS HHI					-0.095**	-0.236**		
					(-2.26)	(-2.42)		
HP_HHI							-0.018***	-0.036***
							(-2.60)	(-3.33)
MTB	0.009***	0.016***	0.009***	0.016***	0.006	0.012	0.008***	0.014***
	(7.73)	(8.96)	(7.62)	(8.95)	(1.59)	(1.58)	(4.93)	(7.86)
LEVERAGE	0.026***	0.025***	0.027***	0.027***	-0.012	-0.001	0.038***	0.036***
	(4.46)	(4.10)	(4.59)	(4.29)	(-1.03)	(-0.11)	(4.68)	(3.41)
CASH	0.031***	0.042***	0.026***	0.036***	0.044***	0.174*	0.024***	0.030**
	(4.90)	(3.48)	(4.62)	(3.28)	(5.68)	(1.92)	(3.52)	(2.20)
SIZE	-0.005***	-0.006***	-0.005***	-0.006***	-0.004*	-0.019*	-0.009***	-0.010***
	(-4.48)	(-5.58)	(-4.51)	(-5.76)	(-1.72)	(-1.92)	(-5.38)	(-4.37)
TANGIBLE	0.022***	0.036***	0.025***	0.040***	0.001	-0.062	0.010	0.009
	(4.00)	(4.32)	(4.57)	(5.16)	(0.04)	(-0.96)	(1.33)	(0.74)
AGE	-0.004**	-0.003	-0.004**	-0.002	-0.010***	0.004	-0.003	0.001
	(-2.49)	(-0.93)	(-2.37)	(-0.65)	(-3.51)	(0.17)	(-1.32)	(0.14)
Constant	0.086***	0.067***	0.075***	0.056***	0.130***	0.154***	0.110***	0.099***
	(14.10)	(8.71)	(13.00)	(6.17)	(5.72)	(5.62)	(16.08)	(10.41)
Observations	28,549	28,582	28,549	28,582	2,474	2,621	14,924	14,952
Group	Low IDVOL	High IDVOL	Low IDVOL	High IDVOL	Low IDVOL	High IDVOL	Low IDVOL	High IDVOL
Adjusted R^2	0.030	0.024	0.031	0.025	0.036	0.013	0.026	0.020

References

Ingersoll, J. E. *Theory of financial decision making*, vol. 3. Rowman & Littlefield Pub Incorporated (1987).

Richardson, S. "Over-investment of Free Cash Flow." *Review of Accounting Studies*, 11 (2006), 159–189.