Investment Efficiency and Product Market Competition[*]

Neal M. Stoughton
WU-Vienna University of Economics and Business
neal.stoughton@wu.ac.at

Kit Pong Wong
University of Hong Kong
kitpongwong@hku.hk

Long Yi
Hong Kong Baptist University
jasonyi@hkbu.edu.hk

April 2016

Abstract

Does more competition lead to more information production and greater investment efficiency? This question is largely unexplored in the finance literature. This paper provides both a model and a series of extensive empirical tests. The model features a two-stage Bayesian game in differentiated products market competition. We find that competition causes firms to acquire less information and investments become more inefficient relative to a first best case with the same market structure. Empirically the panel regression analysis provides strong support for the theory and shows that investment is more efficient in concentrated industries.

I. Introduction

Ever since the seminal economic treatise of Joseph Schumpeter (Schumpeter (1962)), economists have been interested in understanding the role of the competitive environment in promoting innovative activity. The theory of the economics of information has arguably played a leading role in recent times in explaining firm and financial market behavior. Clearly information acquisition has a lot to do with successful innovation, which not only benefits firm stakeholders but society in general.

One important area that has seen considerable research activity is that of firm disclosure policies. Since the seminal articles of Verrecchia (1983) and Darrough and Stoughton (1990) there have been many studies that investigate the connection between market structure, potential entry and whether firms disclose useful information in public markets. These studies universally assume that the firm already possesses the information it considers disclosing.

This paper, by contrast, considers a strategic game of information acquisition where market structure plays a major role in determining the nature of information acquired by firms in the first place. With an intensely competitive environment, it seems intuitive that firms will be reluctant to voluntarily disclose information. Therefore if information is to enter prices and quantities sold in product markets it would appear that this necessitates larger numbers of firms all acquiring information. But when incentives for information collection are considered, the payoff from a single firm expending resources on information collection activities is mitigated when there are more firms outstanding. This provides a countervailing effect. The central question we investigate is whether a more concentrated or a more competitive environment is better from the standpoint of economic efficiency when it comes to investment.

In developing any testable theory, one critical complication is that information production activities are generally unobservable. This is because it is often difficult to measure these costs directly; they are intermingled with other costs and further some costs are nonpecuniary and therefore only observed at the managerial level. Motivated by this challenge, our paper develops a theory which connects information production activities with investment costs of firms, and thereby develops a predictive relation which we are then able to bring to the data in the form of testable hypotheses.

Specifically we build a simple model of information production where a set of n firms acquires information that is relevant to consumer demand. The information acquisition phase is the first of a two-stage Bayesian game of imperfect information which culminates in investment and production decisions. Because the market is characterized by differentiated products, we can consider the full range of outcomes from perfect competition to monopolies. Moreover, competition can be defined by the number of firms in the industry or the market shares as in the Hirfindahl-Hirschman Index. Performance in our model is measured by the deviation between the investment costs of firms based on their outputs using their own information as compared to that of the full information case that would be obtained if a large number of firms pooled their private signals. Therefore an important feature of our model is that the first best benchmark is also a function of industry concentration levels. Our most important results are that (1) firms overinvest in terms of investment expenditure as compared to the full information case; (2) investment efficiency is *increasing* with lower degrees of competition and (3) larger firms exhibit greater efficiencies than do small firms, when firm size is aligned with lower unit costs of production. As a result the impact of competitive structure on information is more pronounced for small as compared with large firms. Hence the theoretical part of the paper confirms the Schumpeterian viewpoint that industry concentration is beneficial in terms of investment efficiency.

The intuition for our main result concerns the way in which firms utilize signals to form their beliefs about rivals' actions in the product market equilibrium and the consequences for the payoff to investment in information production. Consider first what happens with respect to a fixed level of information precision. After observation of a signal, a given firm forms expectations of other firms' signals that are equal to this signal. With increasing degrees of competition, the product market reaction function with respect to its own signal has greater sensitivity than in a less competitive environment. Further, in the noncooperative game of selecting signal precisions, information signals are strategic substitutes. This means that with higher levels of competition the payoff to investment in precision is lower. Putting these two insights together we find that with more competition firms invest less in obtaining precise signals. Nevertheless their production activities in equilibrium react more strongly to whatever signal is observed. With the full vector of production activities, this means there is a greater deviation of industry output from a full information outcome where every firm observed perfect information. When costs are a convex function of production levels, the expected deviation from the full information outcome is greater with more competition. Hence there is greater inefficiency with competition, or equivalently lower inefficiency with industry concentration.

In the second part of the paper we test the model empirically, using panel regression techniques based on data from the period 1980--2012. In this case the investment inefficiency measure is obtained by using an empirical model of optimal investment due to Richardson (2006). We find that competition is positively related to investment inefficiency when competition is either measured by the empirical Hirfindahl-Hirschman Index or the number of firms in an industry. There are also nonlinear effects as well. We also confirm strong support for our hypothesis that larger firms are less negatively affected by competition. Additionally we consider an alternative empirical specification

4

for investment inefficiency that is based on downstream industry variables. This approach avoids using firm-level information which itself might be inefficient. We then perform a series of robustness checks on the data. First, endogeneity in the degree of competition is checked for by using other variables that are less likely to be influenced by the investment level of the firm. Second we utilize an alternative method based on the Lerner index of price cost margin. In all cases, our results are strongly maintained. Last, we explore alternative explanations for our empirical results. In this respect we find that private information and the agency costs of free cash flows are not responsible for our results.

In a closely related paper, Peress (2010) examines a noisy rational expectations model in which investors are endowed with private signals about technology shocks to firms that operate under monopolistic competition. Investors in firms with more market power trade more aggressively using their private information than investors in more competitive industries. As a result, more of this information is incorporated into stock market prices for firms with greater degrees of market power. The improved stock price informativeness then helps allocate capital more efficiently across firms with more market power, thereby enhancing investment efficiency. Although our predictions are similar, they arise for different reasons as we assume that firms can directly access information on their own, rather than having to infer it from stock prices. In other words, Peress does not have information gathering by the managers which is the focus of this paper. A recent related paper is that of Foucault and Frésard (2016).

Another relevant paper, Hoberg and Phillips (2010b) shows that analysts are more biased in competitive industries and that investment and stock prices are less efficient following high stock prices in competitive industries as too many firms invest. Hence, similar to this paper, it is more difficult for firms to gather information in competitive industries and to effectively coordinate their

actions. By contrast it is easier with concentration because analyst bias is not present. The major difference to our paper is that the externality of producing too much is internalized only in concentrated industries, while in our paper the amount of information gathered by all firms is rationally anticipated and accounted for in all competitive environments.

The consequences of product market competition for innovation by firms have been extensively addressed. For instance Vives (2008) considers a variety of models and models innovation by cost reduction. He finds qualified support for more competition driving innovation activities. As far as corporate finance decisions are concerned there are a number of dimensions along which competitive forces have been investigated. MacKay and Phillips (2005) consider financial leverage decisions and illustrate how concentration affects the linkage to a firms natural hedge, the capital/labor ratio. Morellec, Nikolov, and Zucchi (2013) develop a more general model including cash and financing constraints.

In the area of shareholder returns, Nickell (1996) uses a sample of UK firms and argues that competition improves corporate performance using a production function approach. He finds that evidence is somewhat inconclusive, however. Hou and Robinson (2006) find a positive stock return relationship to competition. Giroud and Mueller (2010) examine the effect of the adoption of state-level anti-takeover laws, which help to insulate managers from external discipline, on firms in more or less competitive industries. They find that the passage of anti-takeover legislation leads to lower operating performance of firms in non-competitive industries, while firms in competitive industries exhibit no such effect. Giroud and Mueller (2011) find that good governance, as measured by the number of anti-takeover provisions in a firm's charter, is positively associated with firm value only for firms operating in non-competitive industries. These findings support the idea that product market competition acts as a substitute for corporate governance.

When it comes to investigating investment activity, Sundaram, John, and John (1996) utilize an event study concerning R&D changes and link this to a measure of competition which is based on a measure of competitive strategy from the cost functions. Akdoğu and MacKay (2008) focus on investments related to growth options. They examine sensitivities to the Q-ratios for different industry concentrations and find support for the real options approach. Their paper is more concerned with levels and timing of investment, while we focus on efficiency measures.

Our model structure is based on existing work of a number of authors. Vives (1988) is the seminal work on modeling information production in a model of product market competition. Building on this model Hwang (1993) contains a duopoly model with perfect product substitutes, where signals are given. Ali, Klasa, and Yeung (2014) study financial disclosure policy using industry concentration measures and find that in concentrated industries disclosures are more limited and such firms have a propensity to utilize private placements instead of public offerings. This paper, as does ours utilizes the census data on industry concentrations to avoid the bias as pointed out in Ali, Klasa, and Yeung (2009) that occurs using Compustat data for industry concentration because private firms are omitted.

We set up our model in Section II., where n firms are assumed to compete in a differentiated product oligopoly. The model is analyzed for Cournot competition over firm outputs. In an initial stage firms acquire costly information about a common demand parameter for the products. Firms play a noncooperative game over their information precision levels. Performance is measured by looking at the deviation of cost or investment, from its value compared to a theoretical perfect information outcome, in which signals are pooled perfectly across firms. We solve the model analytically for two firms and then extend the results using Monte Carlo simulation techniques to the general case. We develop hypotheses concerning the level of investment distortion as a function of compe-

7

tition measured both by the number of firms as well as by the Hirfindahl-Hirschman index. Our results are taken to the data in Section III.. We perform panel regressions with controls relating distortions from efficient investment levels to industry concentration. We also conduct robustness checks with respect to endogeneity and alternative industry concentration measures. Finally we test some alternate hypotheses. Conclusions appear in Section IV..

II. Model

In this section, we develop an illustrative model that allows us to draw empirical predictions about the relation between investment efficiency, product market competition, and firm size. The model includes multiple firms, differentiated products, information acquisition and product market competition. The game proceeds in two stages. In the first stage, firms collect information about a common product demand characteristic. We refer to this as the information production stage as it models innovation as an information acquisition process which may confer a relative advantage in the second stage, which is the product market competition stage. In this second stage, firms compete in a Cournot oligopoly given differentiated products.

In order to derive our empirical predictions, we further define an investment efficiency variable which is equal to the expected investment cost of production given the endogenous information acquisition *vis a vis* the full information outcome. This investment efficiency variable is then related to different measures of competition. We consider three measures of enhanced competition: (i) a decrease in the degree of product differentiation, (ii) an increase in the number of competitors, and (iii) the Hirfindahl-Hirschman index of market shares.

A.   Model Setup

We begin with a reduced form model of product market competition, representing the second stage of the problem. We then work backward to analyze the information production stage in the beginning. Finally we relate this to the competitive environment in the form of an endogenous entry decision.

Suppose there are n firms, each producing a quantity, q_i, at a price, p_i, so that the quantity vector is $q = (q_1, ..., q_n)$ and the price vector is $p = (p_1, ..., p_n)$. We describe consumer preferences in terms of three parameters. First, let α represent a common component of demand across all products. Second, let β represent the symmetric level of diminishing benefits consumers receive for each product. Finally, let γ represent the degree of differentiation in products. Specifically, we model the inverse demand function, $p_i(q)$, as

$$(1) \qquad\qquad p_i = \alpha - \beta q_i - \gamma \sum_{\substack{j=1 \\ j \neq i}}^{n} q_j,$$

for all $i = 1, ..., n$. These inverse demand functions can be derived from quadratic utility functions of consumers as in Singh and Vives (1984). In this case, we assume that $0 < \gamma < \beta$. Indeed, $\gamma > 0$ implies that the goods are substitutes, which is a reasonable assumption for goods produced by firms competing in the same industry. In the limiting case that $\gamma = \beta$, the n goods become perfect substitutes and are homogeneous. Hence, γ measures the degree of product differentiation such that an increase in γ renders a reduction in product differentiation. Alternatively, we can interpret γ as a competition intensity parameter: the closer γ to β, the closer substitutes the n goods and the more intense the competition in the product market.

Each firm faces a quadratic cost function for output: $C_i(q_i) = c_i q_i^2$, where $c_i > 0$ is a constant cost coefficient. To model the impact of the size of the firm, we assume that firms with large sizes (as indicated by the level of their capital stock) have lower cost coefficients, i.e., c_i is *decreasing* with the size of the firm.[1]

Moving back one stage to the beginning, we now allow firms to acquire information about the common demand term, α, prior to selecting their quantities or prices in the product market. The parameter, α, which represents consumer preferences for the differentiated product, is not known ex ante. The n firms, however, have a common prior that α has a mean of $\mu > 0$ and a precision (the inverse of variance of this prior) of $h > 0$. Firm i can exert non-monetary effort to acquire information about α. This effort expenditure leads to a signal, $s_i = \alpha + \varepsilon_i$, that is privately observed by the firm, where ε_i is independent of α, and is distributed with a mean of zero and a precision of $t_i > 0$. We assume that the n signals, $s_1, ..., s_n$, are mutually independent conditional on α. We assume further that the posterior estimate, $E(\alpha|s_i)$, is affine in s_i, where $E(\cdot|s_i)$ is the expectation operator conditional on the private signal, s_i. Specifically suppose that $E(\alpha|s_i) = E(s_j|s_i) = \mu + \delta_i(s_i - \mu)$, where $\delta_i = t_i/(t_i + h) > 0$. Information acquisition is endogenous in that each firm can select a precision level, t_i, of its signal by incurring a non-pecuniary effort cost, λt_i, where $\lambda > 0$ is the constant marginal cost of information production.[2] We assume that firms cannot credibly share

---

[1]This cost function could be derived for instance from a more primitive environment, viz., firm i produces q_i units of good i according to the Cobb-Douglas production function, $q_i = \sqrt{K_i I_i}$, where $K_i > 0$ is firm i's fixed assets (capital) that determine the size of the firm, and $I_i \geq 0$ is a variable input chosen by the firm at a constant per-unit cost, $r > 0$. Firm i's cost function as such is given by $C_i(q_i) = r I_i = c_i q_i^2$, where $c_i = r/K_i$ is a cost parameter that decreases with the size of the firm.

[2]A typical example of such affine information structure is for α and ε_i to be jointly normally distributed. The signal, s_i, is a single draw from the normal distribution, $s_i = \alpha + \varepsilon_i$. See Vives (1988) for other distributions that also define affine information structures.

information truthfully. In the disclosure literature Gal-Or (1986) shows that Cournot competition and private information about demand leads to non-disclosure while Bertrand competition and private information about costs also leads to non-disclosure. These results are extended in Darrough (1993) where firms do not necessarily make precommitments to a voluntary disclosure policy. Since our model is one of private information about demand we utilize a Cournot equilibrium in product markets to ensure that firms do not have incentives to make their private information public.[3]

To summarize, the model setup is a two-stage game. In the first stage, firm i acquires information by selecting a precision level, t_i, of its signal at a variable effort cost, λt_i. The n firms select their precision levels simultaneously and non-cooperatively. In the second stage of a Cournot oligopoly, firm i chooses its output level, q_i, conditional on its own signal, s_i, without knowing the precision levels acquired and the actual signals received by all other firms. The n firms as such compete in a Bayesian-Cournot fashion in the product market under incomplete information.

B.   Solution

We solve the perfect Bayesian equilibrium of our two-stage game by using backward induction. Firms are risk neutral and maximize expected profit with respect to their own private information. Consider first the product market competition stage given an n-tuple of precision levels, $t = (t_1, ..., t_n)$. In the case of a Cournot oligopoly, firm i chooses its output level, q_i, to maximize

---

[3]In an earlier version of the paper we also analyzed an associated Bertrand equilibrium with strategic decisions about prices and found that the basic results are preserved.

its expected profit conditional on its private information, s_i:

(2)
$$\max_{q_i} E\left\{ \left[\alpha - \beta q_i - \gamma \sum_{\substack{j=1 \\ j \neq i}}^{n} q_j(s_j) \right] q_i - c_i q_i^2 \,\middle|\, s_i \right\}.$$

Following Vives (1988), we seek an equilibrium in which outputs are linear in signals. Solving the first-order condition for (2) with $q_j(s_j) = a_j + b_j \delta_j(s_j - \mu)$ for all $j \neq i$ yields

(3)
$$q_i(s_i) = \frac{1}{2(\beta + c_i)}\left(\mu - \gamma \sum_{\substack{j=1 \\ j \neq i}}^{n} a_j \right) + \frac{\delta_i}{2(\beta + c_i)}\left(1 - \gamma \sum_{\substack{j=1 \\ j \neq i}}^{n} b_j \delta_j \right)(s_i - \mu),$$

where we have used the fact that $E(\alpha|s_i) = E(s_j|s_i) = \mu + \delta_i(s_i - \mu)$.

In the Bayesian-Cournot equilibrium, firm i's output level is $q_i^e(s_i) = a_i^e + b_i^e \delta_i(s_i - \mu)$ for all $i = 1, ..., n$. Solving equation (3) for the coefficients, a_i^e and b_i^e, yields our first proposition.

Proposition 1. *Given an n-tuple of precision levels, $t = (t_1, ..., t_n)$, the Bayesian-Cournot equilibrium output level of firm i conditional on the private signal, s_i, is given by $q_i^e(s_i) = a_i^e + b_i^e \delta_i(s_i - \mu)$, where*

(4)
$$a_i^e = \left(\frac{\mu}{2\beta + 2c_i - \gamma} \right) \middle/ \left(1 + \sum_{j=1}^{n} \frac{\gamma}{2\beta + 2c_j - \gamma} \right),$$

and

(5)
$$b_i^e = \left(\frac{1}{2\beta + 2c_i - \gamma \delta_i} \right) \middle/ \left(1 + \sum_{j=1}^{n} \frac{\gamma \delta_j}{2\beta + 2c_j - \gamma \delta_j} \right),$$

for all $i = 1, ..., n$.

Proof. See Internet Appendix A. □

Proposition 1 shows that the equilibrium output level of firm i prior to receiving the signal, s_i, is given by a_i^e, which is larger if the firm is bigger, i.e., $a_i^e > a_j^e$ if $c_i < c_j$. When the signal is favorable (unfavorable) in that $s_i > (<) \mu$, the firm adjusts its equilibrium output level upward (downward) from a_i^e to $a_i^e + b_i^e \delta_i (s_i - \mu)$. It is evident from Eq. (5) that $\partial b_i^e / \partial c_i < 0$ and $\partial b_i^e / \partial c_j > 0$ for all $j \neq i$. Therefore, when two firms have identical precisions of information, the amount of output adjustment to a given signal is larger for the firm that is bigger. When the information precisions are different, it is ambiguous which firm is more responsive to the signal, unless the bigger firm has more precise information.

We now turn to the first stage of information acquisition. For an n-tuple of precision levels, $t^* = (t_1^*, ..., t_n^*)$, to be a Nash equilibrium in the case of a Cournot oligopoly, firm i's equilibrium output level must be given by $q_i^e(s_i) = a_i^e + b_i^{e*}\delta_i^*(s_i - \mu)$, where b_i^{e*} is given by equation (5) evaluated at t^*, and none of the n firms have any incentives to deviate from the proposed precision strategic profile. Suppose that firm i deviates from its equilibrium precision level, t_i^*, to t_i. Since such a deviation is unobservable, there are no reactions made by all other firms and their equilibrium output levels remain $q_j(s_j) = a_j^e + b_j^{e*}\delta_j^*(s_j - \mu)$ for all $j \neq i$. It then follows from equation (3) with $a_j = a_j^e$ and $b_j\delta_j = b_j^{e*}\delta_j^*$ for all $j \neq i$ that firm i's optimal output level becomes $q_i(s_i) = a_i^e + b_i^{e*}\delta_i(s_i - \mu)$ after the deviation. Since the conditional expected profit of firm i is equal to $(\beta + c_i)q_i(s_i)^2$, the unconditional expected profit of firm i is given by

(6) $$\pi_i(t_i | t^*) = (\beta + c_i)E[q_i(s_i)^2] = (\beta + c_i)\left[(a_i^e)^2 + (b_i^{e*})^2 \frac{\delta_i}{h}\right].$$

It is evident from equation (6) that an increase in t_i (i.e., an increase in δ_i) raises firm i's uncon-ditional expected profit. Since information acquisition is costly, firm i does not want to deviate from t_i^* if, and only if the increase in net profit is equal to the constant marginal cost of information production, i.e.,

$$(7) \qquad \left. \frac{\partial \pi_i(t_i|t^*)}{\partial t_i} \right|_{t_i=t_i^*} = (\beta + c_i)\left(\frac{b_i^{e*}}{t_i^* + h}\right)^2 = \lambda,$$

where we have used equation (6). Solving equation (7) for all $i = 1, ..., n$ simultaneously yields the perfect Bayesian equilibrium n-tuple of precision levels, $t^* = (t_1^*, ..., t_n^*)$. From Proposition 1, firm i's equilibrium output level conditional on the private signal, s_i, is then given by $q_i^*(s_i) = a_i^e + b_i^{e*}\delta_i^*(s_i - \mu)$, where a_i^e is given by equation (4) and b_i^{e*} is given by equation (5) evaluated at t^*.

This completes the derivation of the solution to the problem of endogenous information pro-duction. Before analyzing the solution, we derive the implications for firm investment policies as compared to a benchmark in which firms could conceivably approach the full information optimum by sharing information reliably.

C. Investment Inefficiency

We define investment inefficiency as the expectation of the proportional difference between the investment cost under incomplete information and that under full information. When the n firms in the Cournot oligopoly observe the true value of α in the full-information scenario, we have $\delta_i = 1$ and $s_i = \alpha$ for all $i = 1, ..., n$. It then follows from Proposition 1 that firm i's equilibrium output

level under full information is given by

$$(8) \qquad q_i^{FI}(\alpha) = \left(\frac{\alpha}{2\beta + 2c_i - \gamma}\right) \Big/ \left(1 + \sum_{j=1}^{n} \frac{\gamma}{2\beta + 2c_j - \gamma}\right),$$

for all $i = 1, ..., n$. Our ex-ante measure of investment inefficiency is given by

$$(9) \qquad D_i = E\left\{\frac{C_i[q_i^*(s_i)] - C_i[q_i^{FI}(\alpha)]}{C_i[q_i^{FI}(\alpha)]}\right\} = E\left[\frac{q_i^*(s_i)^2 - q_i^{FI}(\alpha)^2}{q_i^{FI}(\alpha)^2}\right],$$

where the second equality follows from $C_i(q_i) = c_i q_i^2$. Depending on the realized values of α and

the n signals, the ex-post difference between the investment cost under incomplete information

and that under full information can be either positive or negative. However the *ex ante* difference

can be signed unambiguously. It is important to note that this measure of inefficiency is relative to

a full information benchmark *with the same competitive market structure*. Hence we have isolated

the inefficiency only with respect to information and can look at how this varies with competi-

tion/concentration.

D.   The Duopoly Case

In this subsection, we analytically solve the model for the case when $n = 2$ under Cournot

competition. The duopoly case allows us to provide intuition for the main driving forces of the

model and of the comparative statics.

In the duopoly case, equation (7) becomes

$$(10) \qquad \frac{\partial \pi_i(t_i|t^*)}{\partial t_i}\bigg|_{t_i = t_i^*} = \frac{(\beta + c_i)(2\beta + 2c_j - \gamma \delta_j^*)^2}{[4(\beta + c_1)(\beta + c_2) - \gamma^2 \delta_1^* \delta_2^*]^2 (t_i^* + h)^2} = \lambda,$$

FIGURE 1
Perfect Bayesian Equilibrium when Firm 1 is Larger than Firm 2

$R_i(t_j)$ is the best response function of firm i for $i = 1$ and 2, where $c_1 < c_2$. The intersection of the

two best response functions gives rise to the perfect Bayesian equilibrium pair of precision levels,

$t^* = (t_1^*, t_2^*)$, such that $t_1^* > t_2^*$.
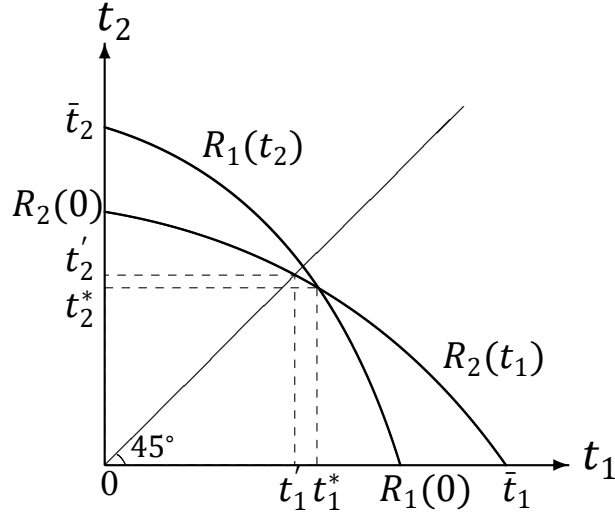
for $i, j = 1$ and 2, and $i \neq j$. Let $R_i(t_j)$ be the best response function of firm i, i.e., $R_i(t_j)$ is the

solution of t_i that solves equation (10) when $t_j^* = t_j$. It is easily verified that the best response

function is downward-sloping so that the precision levels are strategic substitutes.[4] To ensure the

existence of the perfect Bayesian equilibrium, it follows from $R_i'(t_j) < 0$ that $R_i(0)$ has to be pos-

itive, or else firm i never acquires any information. Denote $\bar{t}_i$ as the maximum precision level of

firm i at which firm j optimally chooses not to acquire any information, i.e., $R_j(\bar{t}_i) = 0$. For the

perfect Bayesian equilibrium to be stable, we require that $R_i(0) < \bar{t}_i$. Figure 1 depicts the two best

response functions of firms 1 and 2 in the case wherein firm 1 is larger than firm 2, i.e., $c_1 < c_2$.

The intersection of the two reaction functions gives rise to the unique perfect Bayesian equilibrium,

which is analytically characterized in the following proposition.

---

[4]See the derivation in Internet Appendix B.

Proposition 2. *Suppose that firm 1 is larger than firm 2, $c_1 \leq c_2$. If the marginal effort cost of information production, λ, is sufficiently small such that*

$$(11) \qquad \lambda < \frac{1}{4(\beta + c_1)h^2} \left[\frac{2\beta + 2c_1 - \gamma}{2\sqrt{(\beta + c_1)(\beta + c_2)} - \gamma} \right]^2,$$

there exists a unique perfect Bayesian equilibrium pair of precision levels, $t^* = (t_1^*, t_2^*)$, in which

$$(12) \qquad t_1^* = t_2^* \left(\frac{2\beta + 2c_2 - \gamma}{2\beta + 2c_1 - \gamma} \right) \sqrt{\frac{\beta + c_1}{\beta + c_2}} + 2\left(\frac{\beta + c_1}{2\beta + 2c_1 - \gamma} \right) \left(\sqrt{\frac{\beta + c_2}{\beta + c_1}} - 1 \right) h \geq t_2^*,$$

i.e., firm 1 acquires more information than firm 2. Furthermore, the ex-ante measures of investment inefficiency are positive, and satisfy $D_2 > D_1 \geq 0$, so that the small firm is more inefficient than the large firm.

Proof. See Internet Appendix B. □

Proposition 2 shows that the large firm acquires more information than the small firm in equilibrium. To see the intuition, we refer to Figure 1. Consider the pair of precision levels, $t' = (t_1', t_2')$, on the 45° degree line, where $t_1' = t_2'$. Using equation (10), it is easily verified that

$$(13) \qquad \left. \frac{\partial \pi_1(t_1 | t')}{\partial t_1} \right|_{t_1 = t_1'} > \left. \frac{\partial \pi_2(t_2 | t')}{\partial t_2} \right|_{t_2 = t_2'} = \lambda.$$

Equation (13) implies that the marginal benefit of more precise information is higher for the large firm than for the small firm at $t' = (t_1', t_2')$. The large firm as such is induced to raise the precision level above t_1'. Given the precision levels of the two firms are strategic substitutes, the small firm reacts by setting its precision level below t_2'. This results in the unique perfect Bayesian equilibrium

pair of precision levels, $t^* = (t_1^*, t_2^*)$, where the reaction functions cross, as is shown in Figure 1.

Proposition 2 further shows that the ex-ante measures of investment inefficiency are positive so that the two firms over-invest on average. Under mild conditions that ensure the equilibrium precision levels exceed the prior precision level, h, we show in Internet Appendix B that the large firm is less inefficient than the small firm, i.e., $D_1 < D_2$. This is mainly attributable to the fact that the large firm acquires more reliable information than the small firm.

To obtain the intuition for our results in Proposition 2 regarding inefficiency, recall that we define this as the extra cost relative to the full information benchmark. Since large firms engage in more information production, their signals are more precise and hence closer to the full information outcome. Hence the inefficiency of investment is lower for large firms. The reason why inefficiency is positive using our cost-based measure comes from the convexity of the cost function. Since signals are not infinitely precise, production levels deviate from the full information benchmark. This deviation coupled with convexity of the cost function implies that the average ex ante deviation is positive.

Next we turn to the effect of product market characteristics on our results. First consider the degree of product market differentiation, γ. Proposition 3 derives the result.

Proposition 3. *Consider a Cournot duopoly. Suppose that firm 1 is at least as large as firm 2, i.e., $c_1 \le c_2$. In the perfect Bayesian equilibrium, firm 2 always acquires less information when the degree of product differentiation, γ, increases, i.e., $dt_2^*/d\gamma < 0$. Firm 1 acquires less information when γ increases if, and only if, the ratio, δ_1^*/δ_2^*, is less than the threshold, $2(\beta + c_1)/\gamma$, which exceeds two.*

Proof. See Internet Appendix C. □

18

To gain more insight into the duopoly case with different firm sizes, we perform the following numerical analysis. For the symmetric demand system in equation (1), we set the demand slope, β, equal to unity, and let the degree of product differentiation, γ, take on values in the unit interval, $[0, 1]$. The demand intercept, α, is assumed to be normally distributed with mean, $\mu = 50$, and precision, $h = 0.1$.[5] The large and small firms have cost parameters, $c_1 = 0.5$ and $c_2 = 1$, respectively. The marginal cost of information production is fixed at $\lambda = 0.05$.

The simulation proceeds in two steps. We first solve for the perfect Bayesian equilibrium pair of the precision levels, (t_1^*, t_2^*). In the second step, we obtain a random draw of α from the normal distribution, $N(50, 10)$, a random draw of the noise, ε_1, from $N(0, 1/t_1^*)$, and a random draw of the noise, ε_2, from $N(0, 1/t_2^*)$. We then compute the ex-post difference between the investment cost under incomplete information and that under full information for each firm. We repeat the second step 1 million times.[6] The sample mean of the 1 million ex-post differences is the estimated measure of investment inefficiency, D_i, for firm i.

Figure 2 depicts the results for different degrees of product differentiation, γ. Graph A shows that the equilibrium precision levels are decreasing when the product becomes less differentiated, i.e., when γ increases. Graph B shows that the large firm is less inefficient than the small firm, i.e., $D_1 < D_2$, for all $\gamma \in [0, 1]$. Furthermore, investment inefficiency is more severe when γ increases. If we interpret γ as a competition intensity parameter, Figure 2 suggests that firms acquire less information, and thereby become more inefficient, when competition in the product market is

---

[5]We also use the specification of uniform distributions. The qualitative results are the same.

[6]It is well-known that the Monte Carlo simulation converges at a rate equal to the square root of the simulation frequency. Since we use a simulation frequency equal to one million, the convergence rate is only 0.001. However, the values of D are of similar order. As such the simulated curves that depict D are not smooth. If we increase the simulation frequency to 100 million, the curves would be smoother, but this has no qualitative impact.

more intense. The effect of competition is stronger on the small firm as evidenced from the steeper slope for the small firm when γ varies.

FIGURE 2
Cournot Duopoly Case

The two sub figures plot the relationship between information acquisition, investment inefficiency and competition for the two firm case with Cournot competition. Two Firms $c_1 = 0.5, c_2 = 1$ Parameters used: $\mu = 50, h = 0.1, \beta = 1, \lambda = 0.05$.
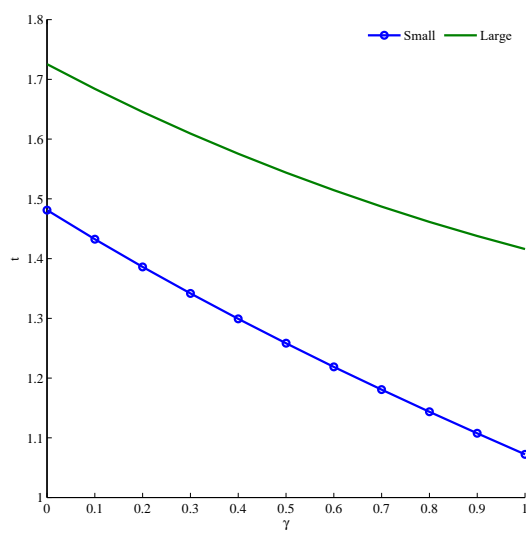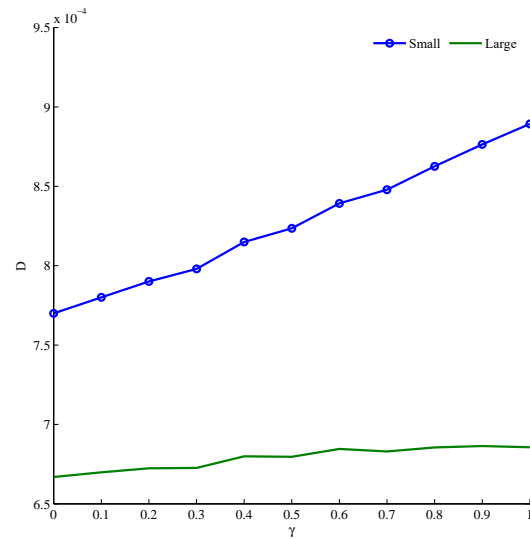
Graph A. Relation between γ and t_i   *Graph B. Relation between γ and D_i*

E.   Cournot Oligopoly

In the duopoly case, we conclude that investment inefficiency is more severe for the small firm than for the large firm. We also find that more intense competition in the product market, as measured by an increase in γ, renders greater investment inefficiency. In this subsection, we examine the robustness of these results in the n-firm case. We also use numerical simulations to illustrate the result.

Suppose first that the n firms are symmetric in that $c_i = c$ for all $i = 1, ..., n$. In this case, we can solve equation (7) to get

$$(14) \qquad t^* = \frac{\sqrt{(\beta + c)/\lambda} - 2(\beta + c)h}{2\beta + 2c + \gamma(n - 1)}.$$

It is evident from Eq. (14) that $dt^*/dn < 0$ and $dt^*/d\gamma < 0$. Thus, an increase in competition as measured by either an increase in the number of firms or a decrease in the degree of product differentiation reduces the level of information precision acquired by all firms. We next turn to Monte Carlo simulation to see the effect of increasing competition on investment inefficiency.

Suppose first that the n firms are symmetric and have the same cost parameter set equal to unity. As in the duopoly case, we employ the following parameter values: $\beta = 1$, $\lambda = 0.05$, $\mu = 50$, and $h = 0.1$. The simulation proceeds in two steps. In the first step, we solve for the perfect Bayesian equilibrium precision level, t^*. In the second step, we obtain a random draw of α from $N(50, 10)$, and n independent random draws of the noise, $\varepsilon_1, ..., \varepsilon_n$, from $N(0, 1/t^*)$. We then compute the ex-post difference between the investment cost under incomplete information and that under full information for each firm. We repeat the second step 1 million times. The sample mean of the 1 million ex-post differences is the estimated measure of investment inefficiency, D, for each firm.

We use two measures of competition intensity: (i) the number of firms, n, and (ii) the Herfindahl-Hirschman Index (HHI), which is defined as

$$(15) \qquad \text{HHI} = \sum_{i=1}^{n} s_i^2 = \sum_{i=1}^{n} \left(\frac{p_i q_i}{\sum_{j=1}^{n} p_j q_j} \right)^2,$$

where s_i is the market share of firm i in the industry. We use equation (15) to calculate HHI in each

of the 1 million simulations. The sample mean gives us the estimated HHI. In Figure 3, we illustrate the simulation results for the cases where the degree of product differentiation is set to $\gamma = 1$ and $\gamma = 0.5$ in the symmetric case.

Graph A and B of Figure 3 feature the case when $\gamma = 1$ while Graph C and Graph D correspond to the case where $\gamma = 0.5$. We see that increases in the number of firms increase inefficiency while increases in HHI decrease inefficiency. The effect is stronger when γ is larger. When we increase the number of firms from 2 to 30, inefficiency increases from about 8 basis points to about 25 basis points for the case where $\gamma = 1$, but it only increases from 8 to about 16 basis points if $\gamma = 0.5$. The effect, however, would be zero if $\gamma = 0$.

Finally, we divide firms into two groups. One group of firms is large, with $c_i = 0.5$ and another group of firms small with $c_j = 1$. We again vary the number of firms. Likewise, Graph A and Graph B of Figure 4 correspond to the case where $\gamma = 1$ while Graph C and D correspond to the case where $\gamma = 0.5$. We see both types of firms' investment efficiency are very sensitive to changing the number of firms. Larger firms perform better in that given the same number of firms, they out perform in terms of less investment inefficiency. The higher the value of γ, the stronger the sensitivity.

Based on the simulations we see that the results from Proposition 2 carry over to the oligopolistic case of n firms. We therefore conclude that more competition, measured either by the number of firms or the HHI, decreases investment efficiency. The intuition for this result is as before: the volatility of production levels in response to imperfectly observed signals interacts with Jensen's inequality to yield investment inefficiency.

Of course in general the level of competition is exogenous. In Internet Appendix D, we show that our results are are preserved in an extension to the model with endogenous (costly) entry.

FIGURE 3
n Symmetric Firms Case

The four sub figures plot the relationship between investment inefficiency and competition for the n symmetric firms case under two scenarios. In the first two sub figures, γ is set to 1. In the last two subfigures γ equals 0.5. Graph A and C use number of firms in the x-axis. Graph B and D use HHI as x-axis. Parameters used: $\mu = 50, h = 0.1, \beta = 1, \lambda = 0.05, c_i = 1$.
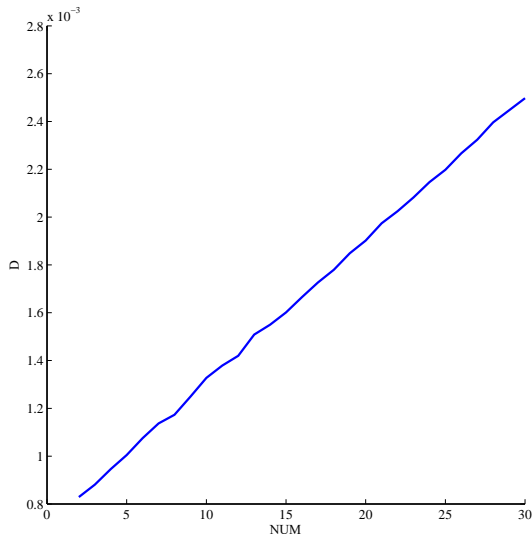
Graph A. Relation between NUM and D
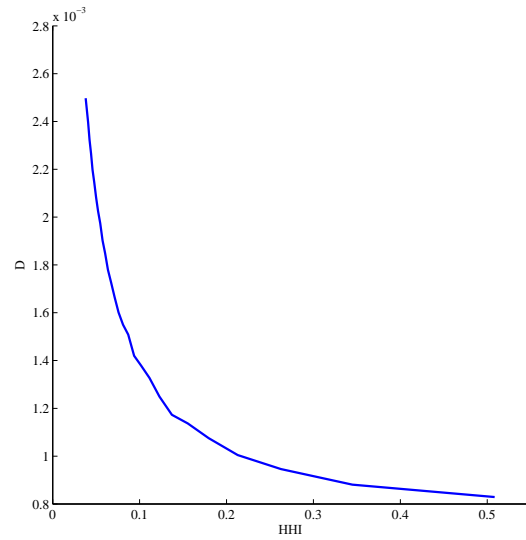
Graph B. Relation between HHI and D

Graph C. Relation between NUM and D

Graph D. Relation between HHI and D

23
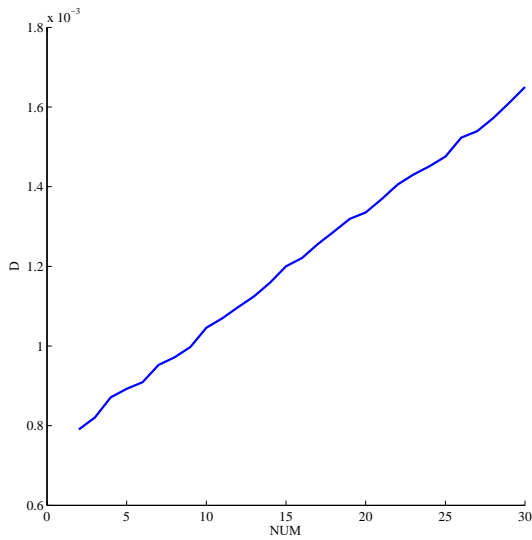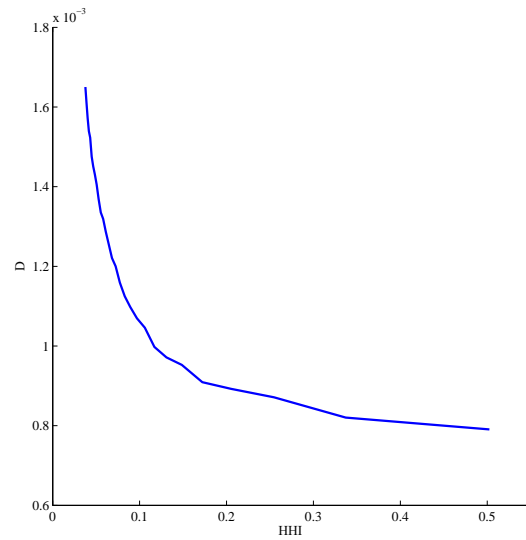
FIGURE 4
n **Asymmetric Firms Case**

The four sub figures plot the relationship between investment inefficiency and competition for the n asymmetric firms case under two scenarios, where half of the firms are large with $c_i = 0.5$ and another half small with $c_j = 1$. In the first two figures, γ is set to 1. While in the last two γ equals 0.5. Graph A and C use number of firms in the x-axis. Graph B and D use HHI as x-axis. The parameters used are: $\beta = 1, \lambda = 0.05, h = 0.1$.
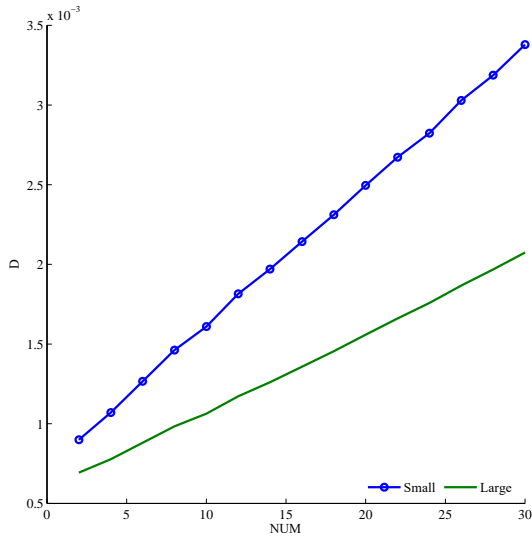
Graph A. Relation between NUM and D
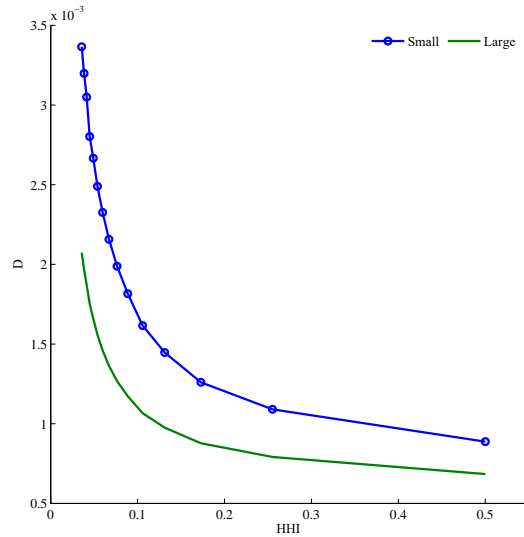
Graph B. Relation between HHI and D
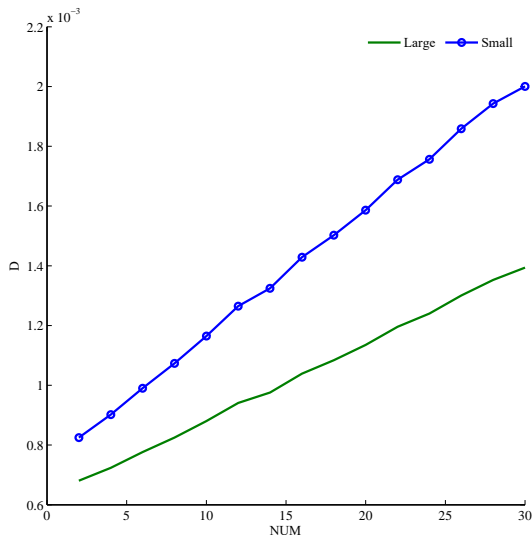
Graph C. Relation between NUM and D
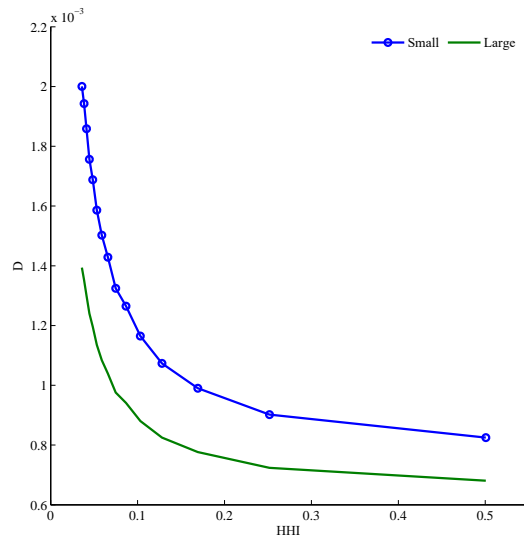
Graph D. Relation between HHI and D

F. Summary of the Theoretical Model

We now summarize and provide the intuition for our theoretical results and derive two hypotheses to be tested. Since firms in our model cannot collude and coordinate their levels of information production, information is only gathered when it is rewarded in the form of higher ex ante profits. We have shown that based on a number of competition measures, firms have greater incentives to invest in costly information production in more concentrated industries, especially when they are larger. The intuition is that equilibrium quantities and prices are a function of all firms' signals. The impact of a single firm's signal is less when outcomes are influenced by a lot of other firms and therefore there is less incentive to invest in information precision. As a result even if there are more firms, and therefore more independent signals, overall the precisions are lower and the entire market winds up with less information embedded in prices and quantities than with industry concentration. We therefore find that:

H1: An increase in product market competitiveness, as measured by lower HHI or higher number of firms, decreases information acquisition by firms in the industry. As a result, the measure of investment inefficiency increases.

H2: Within an industry, large firms are less sensitive to competition. Small firms are more sensitive to this effect and their information acquisition is monotonically decreasing with competition.

III. Empirical Estimation

We present the empirical results in this section where we will describe the data, empirical proxies, regression results and robustness issues.

A. Data Sources and Variable Definitions

We use an unbalanced panel of firms from Compustat. The time period stretches from 1980 to 2012. Regulated utilities (SIC 4900-4999) and financial firms (SIC 6000-6999) are excluded from our analysis[7]. Unbounded variables are winsorized at the 1% level. We also use data from CRSP, trade data from Peter Schott's website and competition measures from Hoberg Phillips data library. Finally for our inefficiency approach based on downstream firms, we use the Input-Output Accounts data from the Bureau of Economic Analysis (BEA) and analysts forecasts of earnings from I/B/E/S.

1. Measures of Competition

To explore the relationship between competition and investment efficiency, we need to measure the level of competitiveness in an industry. The most frequently utilized measure in the literature is the usual Herfindahl-Hirschman Index (HHI) that measures the product market concentration of the industry. Therefore competition in industry j is calculated as $\text{HHI}_j = \sum_{i=1}^{I} s_{ij}^2$, where s_{ij} is the sales share of firm i in industry j.

Because of recent developments in the literature we employ three different HHI measures. First, we use Compustat data from 1980-2012 to compute the HHI for all firms in the database with the exception of financial and utilities. Here we define industries by three-digit SIC codes. Second, following Ali, Klasa, and Yeung (2009), we obtain the census HHI figures for manufacturing firms from 1982, 1987, 1992, 1997, 2002 and 2007. These data are more robust in the sense that they include private as well as publicly traded firms. In these data, four-digit SICs define industries for the first three years and the NAICS definitions are used for the subsequent three years. Finally we use

---

[7]We exclude them because of the different characteristics of investment for these firms. Nevertheless results including these are available from the authors.

the 10K text-based competition measure following the procedures outlined in Hoberg and Phillips (2010a), (2010c). These HHI figures also use Compustat data. This is referred to as the HP_HHI data. In all cases lower HHI signifies more intense competition in the industry.[8] We also employ the simpler measure of the number of firms divided by 1,000 in an industry, obtained also from Compustat data. Here the number of firms is a proxy for increasing competition.

2. Measure of Inefficiency

The principal dependent variable in our analysis is investment inefficiency. Inefficiency is the deviation of the firm's investment expenditures from its full information outcome. In order to measure this we employ a popular model of efficient investment from the accounting literature due to Richardson (2006). Richardson shows that the optimal investment level of a firm is related to its growth opportunities, leverage, cash, size, stock return and lagged investment. Recall that our theoretical efficient investment benchmark represents all available information including signals of firms other than the one in question. It also depends on competition. Therefore we slightly generalize the Richardson model to allow for competition neasures as an additional explanatory variable. We estimate the optimal investment relation over all firms and years. Equation (16) gives the specification of the regression

$$
\begin{aligned}
(16) \quad I_{i,t} \;=\; & \beta_0 + \beta_1 V/P_{i,t-1} + \beta_2 \text{LEVERAGE}_{i,t-1} + \beta_3 \text{CASH}_{i,t-1} \\
& + \beta_4 \text{SIZE}_{i,t-1} + \beta_5 \text{STOCK RETURN}_{i,t-1} + \beta_6 \text{AGE}_{i,t-1} \\
& + \beta_7 I_{i,t-1} + \beta_8 \text{COMPETITION}_{j,t-1} + \mu_i + \eta_t + \epsilon_{i,t}
\end{aligned}
$$

---

[8]Other competition measures will be discussed in the robustness section.

where i stands for firm, j for industry and t for time. V/P measures the growth opportunities of the firm, and involves a residual income model of the assets in place, V, divided by the market value of equity, P.[9] We expect that higher V/P reduces investment. $I_{i,t}$ is the measure of new investment that consists of capital expenditure plus research and development expenditure plus acquisition minus sale of property, plant and equipment and amortization and depreciation scaled by total assets.[10] LEVERAGE is the sum of the book value of short term and long-term debt divided by the sum of the book value of total debt and equity. CASH is the balance of cash and short term investment scaled by total assets of the firm. SIZE is the logarithm of total assets. STOCK RETURN is the change in the market capitalization of the firm. AGE is calculated by the number of years the firm has been listed on CRSP. COMPETITION consists of various measures such as the number of firms and HHI.

Since equation (16) is estimated using panel data involving all firms and all years, the residual, $\epsilon_{i,t}$, from this regression reflects inefficiency in the model. By construction the average value of all residuals is zero; therefore to measure inefficiency on a firm by firm basis, we utilize the absolute value of the residual as our measure of investment inefficiency, which we label INF hereafter, that is for each firm, i, at time t, $\text{INF}_{i,t} = |\epsilon_{i,t}|$.

The investment regression results are reported in Table 1. Since we have four measures of competition, we have a total of four regressions. The first column in the table uses three-digit SIC code HHI as the measure of competition, the second uses the number of firms scaled by 1000, the third

---

[9] See Internet Appendix E for variable definitions. Using book to market as the measure of investment opportunities does not change our results.

[10] When the R&D variable is missing in Compustat, we set it equal to zero as is standard practice. We have also run tests without R&D expenditure and by omitting firms with missing R&D entries.

utilizes the Census HHI while the last uses the HP_HHI. We include firm and year fixed effects in all the regressions except for column 3 where we only use year fixed effects. This is because Census HHI is only available every five years, and the fixed effect tends to pick up the variation in competition with the more limited data set. Higher V/P, larger Size, higher Leverage and older Age reduce investment as expected while higher Cash, Return, and lagged investment increases investment.

Firms in concentrated industries invest more than those in competitive industries as the coefficient on HHI and HP_HHI is negative and significant. The loadings on NUM and Census HHI have the same relation but the coefficients are insignificant.

(Insert Table 1 Here)

3.   Summary Statistics

We present summary statistics in Table 2. Average investment is 10% of lagged assets. As for the inefficiency measure, the average INF is 10% of lagged assets, which is about the same magnitude as average investment. On average, competition measured by Compustat HHI is 0.22 while the number of firms is 40 in a three-digit SIC industry. We have 60,202 firm-year entries in total for the investment variables. We have far fewer observations for the census HHI due to the restriction to manufacturing firms and the five year interval period for the data.

(Insert Table 2 Here)

B. Summary Evidence

1. Inefficiency across Industries

How does inefficiency vary across industries with different industry concentration? To see this, we sort firms by their level of competition in each fiscal year. Four groups of firms are formed based on HHI each year. We calculate the mean of inefficiency for each group each year. Figure 5 presents the average inefficiency of firms within the group with highest HHI and that of firms in the group with lowest HHI across years. The inefficiency measure is consistently higher for firms in the lowest HHI (competitive) group compared with that of firms in the highest HHI (concentrated) group. We then average the time-series means of inefficiency of the four groups of firms by HHI quartile. Figure 6 plots the means of the four groups. Firms in the first group are in the most competitive industries while the fourth group firms face the least competition. A declining trend in terms of mean inefficiency is clearly observed, with the mean of inefficiency in the first group being 0.10 while the last group is equal to 0.07.

C. Hypothesis Tests

1. Competition and Efficiency

We now examine the relationship between competition and efficiency using the following specification:

$$(17) \qquad \text{INF}_{i,t} = \beta_0 + \beta_1 \text{COMPETITION}_{j,t-1} + \beta' \mathbf{X_{i,t-1}} + \epsilon_{i,t}$$

FIGURE 5
Mean of Inefficiency Across Time

In each year, firms are divided into 4 groups by their HHI level in ascending order. In each year, we calculate the mean of the inefficiency within the lowest and the highest HHI groups and plot the mean across time. The sample covers firm-year data from 1980 to 2012.
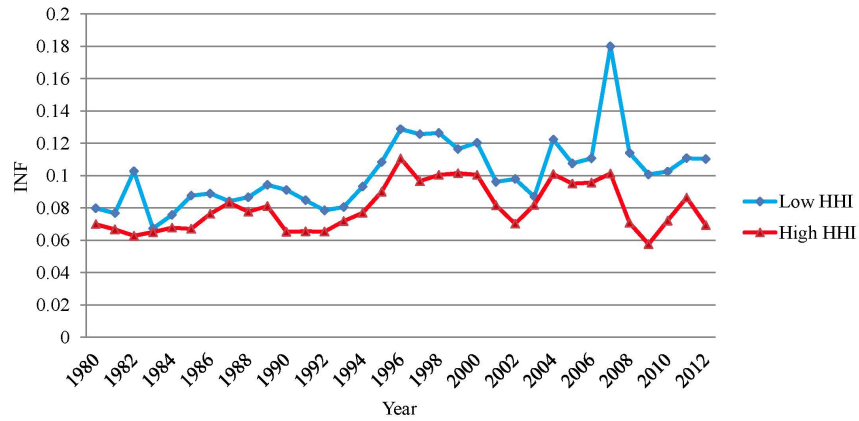
FIGURE 6
Mean Deviation Across Groups

Firms are divided into 4 groups by their HHI level in ascending order. We calculate the mean of each group across time as well as across firms. Group 1 are firms with lowest HHI. The sample covers firm-year data from 1980 to 2012.
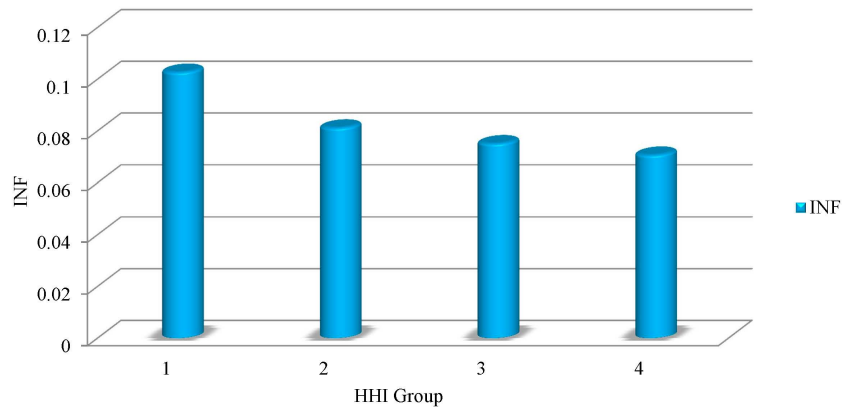
where $INF_{i,t}$ is our empirical measure of inefficiency. Competition is either measured by our three HHI measures or the number of firms divided by 1000. **X** is a set of firm level control variables, including market to book, leverage of the firm, cash holdings, size of the firm, tangibility of assets and the age of the firm. Increasing firm size is predicted by the model to be related to greater efficiency of investment. Firms that are older have more experience, thus should invest more efficiently. Firms with more cash tend to invest less efficiently which could be due to agency problems. Tangibility is a measure of financial constraints, i.e. the ability to borrow. Other variables control for the fundamental difference of firms with different growth opportunities (MTB) and different risk levels (Leverage). Note that we have four sets of inefficiency measures depending on the competition measures used in regression (16). The same competition measures will be used in both stages. Namely if we take INF from column 1 of Table 1 where HHI is used, the same HHI will be used as the competition measure here.

Table 3 presents the regression results. We obtain strong results confirming our hypothesis regarding the link between inefficiency and industry concentration. In the first two columns, we find that the coefficient on Compustat HHI is either -0.044 or -0.022 both significant at the 1% level, without and with controls respectively. Columns 5 and 6 give corresponding results for the census HHI and columns 7 and 8 provide the result for the HP_HHI. All of these coefficients are of the correct sign and similar in terms of magnitude. Standard errors are clustered by firm and year. We find support as well by using the number of firms as an indication of competition. Here the coefficient is positive, as expected.

In terms of economic significance, a one standard deviation increase in Compustat HHI (0.18) decreases the INF of an average firm by 5% to 10%. We find similar figures for census HHI, where the economic significance of a one standard deviation (0.06) move leads to roughly 8% to 9% in

terms of inefficiency. A one standard deviation increase in the number of firms (0.06) decreases inefficiency by 9% to 17%. A one standard deviation increase in HP_HHI decreases inefficiency by about 6%. These results are reassuring and our findings confirm our hypothesis H1 where competition increases inefficiency.

In terms of control variables, firms that are less financially constrained, i.e. with more cash or tangibility of assets invest less efficiently. Firms that are older in age or larger in size invest more efficiently as predicted by the model where information accumulated across the years or more resources could be used to acquire information for large firms. More leverage increases inefficiency while higher market to book ratios increases inefficiency.

<center>(Insert Table 3 Here)</center>

2. Competition and Efficiency: Non-Linearities

As seen from the simulation, there is a possible non-linearity between competition and efficiency. We test this by including a quadratic term of the competition measure in the regression:

$$(18) \qquad \text{INF}_{i,t} = \beta_0 + \beta_1 \text{COMPETITION}_{j,t-1} + \beta_2 \text{COMPETITION}^2_{j,t-1} + \beta' \mathbf{X_{i,t-1}} + \epsilon_{i,t}$$

Table 4 reports the regression results. We discover a strong convex non-linear effect with respect to Compustat HHI, which is consistent with our simulation results. The coefficients on HHI remain negative while the coefficients on HHI^2 are positive. The effect is similar for the HP_HHI measure. Unfortunately we could not confirm this non-linear effect using the Census HHI data. Here the nonlinearity seems to decrease even the significance of the linear term. With respect to the number of firms, while the sign is correct with respect to the linear term (positive), we find a

concave relationship, as opposed to the linear relationship we found in the simulation studies.

(Insert Table 4 Here)

3. Competition and Efficiency: Size Effect

Hypothesis H2 of our model is that firms that are larger in size are more efficient and are less sensitive to the impact of competition. To test this hypothesis, we introduce an interaction term between size and competition into the regression:

$$
(19) \quad \text{INF}_{i,t} = \beta_0 + \beta_1 \text{COMPETITION}_{j,t-1} + \beta_2 \text{SIZE}_{i,t-1}
$$

$$
+ \beta_3 \text{SIZE}_{i,t-1} \times \text{COMPETITION}_{j,t-1} + \beta' \mathbf{X_{i,t-1}} + \epsilon_{i,t}
$$

β_1 measures the impact of competition and β_2 measures the size effect, which is expected to be negative as larger firms are more efficient. β_3 is the coefficient of interest here. As larger firms are less influenced by competition, it should be of the opposite sign as compared with β_1.

Table 5 reports the results. When competition is measured as Compustat HHI in columns 1 and 2, β_1 is negative, confirming again that less competition reduces inefficiency. The impact of size as reflected in β_2 is also negative so that size reduces inefficiency. Importantly β_3 is of the opposite sign, although it loses significance when controls are added. Nevertheless with or without controls it stands up in the census HHI and HP_HHI. The size effect on efficiency is stronger when there is more concentration. Firms that are larger in size are less affected by competition in terms of investment efficiency. The results are maintained when competition is measured using NUM as in columns 3 and 4. The coefficient on size, β_2, is consistently negatively and significant in all settings. Other control variables have the same sign as in previous regressions. Thus, our hypotheses H2 is

confirmed here using three out of four measures. Large firms are less sensitive to competition and invest more efficiently.

(Insert Table 5 Here)

D. A Downstream Based Inefficiency Measure

Our previous measure of investment inefficiency has relied on the Richardson (2006) model, which utilizes firm-level data to explain future investment. One possible criticism of this approach is that if the firm is behaving inefficiently presumably some of the explanatory variables in the Richardson model such as lagged investment, stock price and cash levels will not be efficient.

Therefore, we construct an alternative measure of investment efficiency based on on downstream industry-level data that the firm presumably reacts to in setting its investment level.[11] This procedure gives us an additional measure of inefficiency that we label as INF(DOWN).

We use the Input-Output Accounts data from the Bureau of Economic Analysis (BEA) to identify an unique downstream industry for each upstream industry to which a firm belongs.[12] We utilize the Use Table that identifies the amount of a commodity purchased by each industry as an intermediate input into the industry's production. The Input-Output table has been widely used to identify the links between industries (for example by Fan and Goyal (2006); Menzly and Ozbas (2010); Ahern and Harford (2014)). For each firm, i, in industry j, we identify a single downstream industry, k, the industry that uses the most dollar amount of goods from industry j. For example, the petroleum refineries industry is identified as the downstream industry of oil and gas extraction firms. We use

---

[11] We thank an anonymous referee for suggesting this method of measuring inefficiency.

[12] We use the 2007 version of the data that includes 389 industries. Since the output data for detailed industries were first made available in 1997, our sample period starts in 1998.

four downstream *industry level* variables to which the investment of upstream firms should respond to.[13]

We include the total output of our identified downstream industry (DOWN_OUTPUT) to reflect the demand from the downstream industry. We also include the average growth in sales (DOWN_SALEG) of downstream industry firms as a proxy for investment opportunities of the upstream firm. In addition, we include two variables that are market-based and forward-looking: (i) the average market-to-book ratio of firms in the downstream industry (DOWN_MTB), and (ii) the average of long-term growth forecasts of firms by analysts in the downstream industry (DOWN_ANALYST). Finally, as our model predicts competition levels influence efficient investment, we include the competition measure in the first-stage regression.

Specifically, we employ the following model that relates efficient investment in firm i in upstream industry j to downstream industry k:

$$
(20) \qquad
\begin{aligned}
I_{i,t} \ = \ & \alpha + \beta_1 \text{DOWN\_OUTPUT}_{k,t-1} + \beta_2 \text{DOWN\_SALEG}_{k,t-1} \\
& + \beta_3 \text{DOWN\_MTB}_{k,t-1} + \beta_4 \text{DOWN\_ANALYST}_{k,t-1} \\
& + \beta_5 \text{COMPETITION}_{j,t-1} + \eta_t + \epsilon_{i,t}
\end{aligned}
$$

Table 6 reports the first-stage regression results using different competition measures. All four downstream industry level variables are positively correlated with firm-level investment, as might be expected. Note that the downstream analysis confirms the firm level analysis since book-to-market was negatively related to firm investment in the first specification and market-to-book is positive in this specification. Further we find corresponding results with respect to our competition mea-

---

[13]Internet Appendix E provides detailed definitions of variables.

sures since HHI and Hoberg Philips HHI are significantly negatively related to investment in both specifications.

(Insert Table 6 Here)

Taking the absolute value of the residual gives rise to the new measure of efficiency, INF(DOWN)[14]. The correlation between our original measure, INF, and the new measure, INF(DOWN), is given in panel A of Table A.1. We find that these two measures are significantly positively correlated and equal to 0.75 when HHI is used as the competition measure in the first stage.[15].

(Table A.1 Available in the Internet Appendix)

As with our earlier approach, we run equation (17) using the new inefficiency measures we calculate using the downstream industry approach. Table 7 reports the regression results. Our main results are confirmed again. All coefficients of the competition variables are of the right (negative) sign and are highly significant. In fact our results using this downstream analysis are even stronger than with the Richardson approach as indicated by the higher R^2 in the multivariate regressions and also the fact that the competition coefficients are almost always greater in magnitude as well as statistical significance.

(Insert Table 7 Here)

E. Robustness

In this section, we vary the measure of competition for robustness purposes. We also separate samples by either free cash flow or the idiosyncratic volatility of the firm to verify that our results

---

[14]The values of INF(DOWN) differ slightly depending on the competition measures used in the first stage.

[15]The correlation matrices of other sets of competition measures are similar and are therefore not reported.

are not driven by these two factors that might also be related to investment efficiency.

1. Different Measures of Competition

All the previous competition measures are more or less endogenous. Thus, we follow the literature on endogeneity by adopting trade cost as another measure of competition, which is exogenous to the influence of the industry. The higher the trade cost, the lower the competition. We also utilize the price-cost margin as additional measures of competition. The higher the price-cost margin, the lower the competition level. Panel B of Table A.1 shows the correlation matrix of the eight competition measures. All of the measures are positively correlated with Compustat HHI, of course, with the exception of the number of firms which is negatively correlated. HP_HHI behaves similarly. The census HHI variable has a correlation that is negatively related to HP_HHI and is positively related to the number of firms. This could be due to the restriction to manufacturing firms which is not representative of the broader sample of firms applied in the remaining competition measures.

Using the robust measures of competition, we rerun equation (16) to calculate inefficiency. Table A.2 presents the first stage result. Inefficiency measures are then used in equation (17). Table A.3 presents the result using trade cost (tariff plus freight cost), or tariff and freight cost separately and EPCM. EPCM is significantly negative, confirming our previous results using HHI and number of firms. Competition as measured by trade data shows that higher competition (lower trade cost) increases the inefficiency of investment of the firm. The impact of tariffs is larger than that of freight cost. A one standard deviation increase of tariff (0.04) reduces inefficiency by 0.00868. Given the mean of inefficiency being 0.08, it corresponds to an increase of a 11% increase. The magnitude is similar with our previous results.

(Table A.2 and A.3 Available in the Internet Appendix)

2. Sub-group Results

Giroud and Mueller (2010) find that good corporate governance benefits firms, but the results are only significant for firms in concentrated industries. Because agency problems are also a possibility to influence firm investment efficiency and industries with different concentration might also differ in the nature of the agency problem they face, our results that industry concentration helps efficiency might be the result of an agency problem that correlates with industry concentration. If that were the case, when we divide firms into high free cash flow and low free cash flow firms, we would expect the industry concentration measures would no longer be significant.

We calculate free cash flow (FCF) as the cash flow from operating activities minus depreciation and amortization plus research and development minus expected investment in year t as in the optimal investment regression (16). In year t, if a firm is below the median in terms of FCF then it is in the low FCF group. Otherwise it is in the high FCF sample.

Table A.4 reports the results. In columns 3 and 4, firms are divided into subsamples and the coefficient on competition is considered in the case of the number of firms. We find no difference with respect to the FCF subsamples. Similarly our results are preserved when looking at Census HHI measures in columns 5 and 6. There is also confirmation of our previous results when considering the HP_HHI. Unfortunately it appears that the reduction of sample size results in less significance with respect to our HHI variable in column 1. Nevertheless the coefficient remains with the correct negative sign. Our main conclusion is that higher concentration reduces investment inefficiency is not driven by the agency problem of free cash flows.

(Table A.4 Available in the Internet Appendix)

Another potential concern is that firms can learn from the stock market (Dow and Gorton

(1997), Subrahmanyam and Titman (1999), and Polk and Sapienza (2009)). Perhaps it is information communication through financial market prices that assists with efficiency. Hence we check if private information stemming from the stock market helps firms in concentrated industries invest more efficiently. Peress (2010) showed that firms with more market power have more informative stock prices. This, however, was from the view of investors. From the view of managers, insider trading information is something already known by the manager as argued by Chen, Goldstein, and Jiang (2007). It is private information in the stock market that is otherwise unknown to the manager that potentially might help determine investment. The more private information the manager can get from the stock market for free, the better is investment efficiency. Following Morck, Yeung, and Yu (2000), we use idiosyncratic volatility as a proxy for private information from the stock market. For each firm, idiosyncratic volatility is calculated from $1 - R^2$ of the regression $r_{i,j,t} = \beta_{i,0} + \beta_{i,m}r_{m,t} + \beta_{i,j}r_{j,t} + \epsilon_{i,t}$, where i stands for firm, j stands for industry and t stands for time. $r_{i,j,t}$ is the return of firm i in industry j at time t. $r_{m,t}$ is the market return at time t. $r_{j,t}$ is the average return of industry j at time t. Therefore idiosyncratic volatility actually measures the returns variation of a firm that cannot be accounted for by the market and the industry to which it belongs. A firm is in the Low IDVOL group in year t if the IDVOL of it is below its mean in year t.

Table A.5 presents the results. Our results remain strong in each of the sub groups for all of our measures. This indicates our results are not driven by stock market informativeness.

(Table A.5 Available in the Internet Appendix)

IV. Conclusion

This paper has sought to test a ``Schumpeterian" hypothesis that monopoly power is a necessary ingredient to promote valuable information production and achieve investment efficiencies. As such we have developed a model of competition with differentiated products along with an innovation stage in which firms invest in acquiring more precise information signals about product demand. We find that larger firms act more aggressively than do smaller firms for given levels of information precision. This effect is reinforced when we consider endogenous information acquisition. Now we demonstrate that in a duopoly between two firms of different sizes, the large firm also acquires more precise information. Our results in the duopoly case can also be extended to derive predictions about the level of investment efficiency. Here we describe inefficiency as the *ex ante* deviation between the costs of investment given the specific signal and the limiting case of an infinitely precise signal. Intuitively this results from the fact that the larger firm acquires more precise information than does the small firm.

Our duopoly results are then generalized to more general industry structures. Using Monte Carlo simulation techniques we have discovered that information acquisition diminishes when the products are more substitutable (i.e., less differentiated). In general our result that large firms respond more aggressively is preserved. For the n firm case, we consider two types of scenarios. One is where firm structure is symmetric in the sense that size of all firms is the same; the second is where firms are either of a large type or a smaller type (with corresponding cost implications). We measure competition in the n firm case using one of two possibilities. First, we can consider the number of firms as representative of competition. This is an exogenous measure. Second, we compute the Hirfindahl-Hirschman Index (HHI), which is the sum of market shares using the equilibrium price

and quantities. This measure represents an endogenous measure of competition. For the symmetric firm case we find that investment inefficiencies increase with respect to the number of firms, and with a reduction in the realized HHI. For the asymmetric case, small firms retain the same relationship with respect to competition - they are adversely affected. Large firms are as well, although the results are less significant.

Our empirical test is based on unbalanced panel regressions using data from 1980-2012. In order to do so, we must take a stand on how to model optimal investment activities and therefore to measure the deviation from optimal practices. We begin with the Richardson (2006) model from the accounting literature, as it has garnered wide acceptance. The model predicts investment using the present value of growth opportunities, leverage, cash, size, momentum, age, and past investment. Since our theory is about competition we generalize the Richardson model to allow for this as well. The residuals from the Richardson model are then utilized to compute the investment inefficiency, which is then related to competition. We find very strong results from a univariate regression test relating either competition measure (the number of firms or the HHI) to investment inefficiency. In addition some support for a nonlinear effect is also included. Investment inefficiency is very severe with either a weak industry concentration or a large number of firms in the industry. Our size effects are found by utilizing an interactive term in the panel regression. Now, the competition effects are lower for larger firms. We also utilize an alternative measure for investment inefficiency based on input-output tables from the BEA. Our results regarding competition and efficiency are even stronger with respect to this alternative specification.

Our theoretical and empirical results have been found to be robust to endogenous competition. Specifically we use tariff and freight cost as a source of exogenous variation that would affect all firms in the industry but be outside the control of any of them. We also employ the Lerner index

(price-cost margin). We find that all of our results remain strong and economically significant. Two remaining robustness checks are also carried out: we find that private information does not impact our results nor does a measure of the agency costs of free cash flows.

In closing it is worthwhile mentioning that we have taken the viewpoint that a concentrated industry is not a competitive one. Other authors, e.g., Sutton (1991) have pointed out that in such cases, entry would be potentially more attractive, and the influence from the potential entrant makes this a more competitive product market. Our model--although it can be generalized to allow for an entry decision that endogenizes the number of firms--does not consider investment as a motive for preemption and therefore analyzes these effects for a given market size. This has the empirical advantage that one can utilize existing data, rather than having to develop proxies for entry costs and characteristics of firms that do not exist. Nevertheless, a better understanding of entry models in the area of information production would be a highly desirable future research area in our view.

References

Ahern, K. R., and J. Harford. "The Importance of Industry Links in Merger Waves." *The Journal of Finance*, 69 (2014), 527–576.

Akdoğu, E., and P. MacKay. "Investment and Competition." *Journal of Financial and Quantitative Analysis*, 43 (2008), 299–330.

Ali, A.; S. Klasa; and E. Yeung. "The Limitations of Industry Concentration Measures Constructed with Compustat Data: Implications for Finance Research." *Review of Financial Studies*, 22 (2009), 3839–3871.

Ali, A.; S. Klasa; and E. Yeung. "Industry Concentration and Corporate Disclosure Policy." *Journal of Accounting and Economics*, 58 (2014), 240–264.

Chen, Q.; I. Goldstein; and W. Jiang. "Price Informativeness and Investment Sensitivity to Stock Price." *Review of Financial Studies*, 20 (2007), 619–650.

Darrough, M. N. "Disclosure Policy and Competition: Cournot vs. Bertrand." *Accounting Review*, 68 (1993), 534–561.

Darrough, M. N., and N. M. Stoughton. "Financial Disclosure Policy in an Entry Game." *Journal of Accounting and Economics*, 12 (1990), 219–243.

Dow, J., and G. Gorton. "Stock Market Efficiency and Economic Efficiency: Is There a Connection?" *The Journal of Finance*, 52 (1997), 1087–1129.

Fan, J. P., and V. K. Goyal. "On the Patterns and Wealth Effects of Vertical Mergers." *The Journal of Business*, 79 (2006), 877–902.

Foucault, T., and L. Frésard. "Corporate Strategy, Conformism, and the Stock Market." HEC Paris Research Paper No. FIN-2015-1099, (2016).

Gal-Or, E. "Information Transmission-Cournot and Bertrand equilibria." *The Review of Economic Studies*, 53 (1986), 85–92.

Giroud, X., and H. M. Mueller. "Does Corporate Governance Matter in Competitive Industries?" *Journal of Financial Economics*, 95 (2010), 312–331.

Giroud, X., and H. M. Mueller. "Corporate Governance, Product Market Competition, and Equity Prices." *The Journal of Finance*, 66 (2011), 563–600.

Hoberg, G., and G. M. Phillips. "Product Market Synergies and Competition in Mergers and Acquisitions: A Text-based Analysis." *Review of Financial Studies*, 23 (2010a), 3773–3811.

Hoberg, G., and G. M. Phillips. "Real and Financial Industry Booms and Busts." *The Journal of Finance*, 65 (2010b), 45–86.

Hoberg, G., and G. M. Phillips. "Text-Based Network Industries and Endogenous Product Differentiation." Working Paper, National Bureau of Economic Research, (2010c).

Hou, K., and D. T. Robinson. "Industry Concentration and Average Stock Returns." *The Journal of Finance*, 61 (2006), 1927–1956.

Hwang, H.-S. "Optimal Information Acquisition for Heterogenous Duopoly Firms." *Journal of Economic Theory*, 59 (1993), 385–402.

MacKay, P., and G. M. Phillips. "How does Industry Affect Firm Financial Structure?" *Review of Financial Studies*, 18 (2005), 1433–1466.

Menzly, L., and O. Ozbas. "Market Segmentation and Cross-predictability of Returns." *The Journal of Finance*, 65 (2010), 1555–1580.

Morck, R.; B. Yeung; and W. Yu. "The Information Content of Stock Markets: Why do Emerging Markets have Synchronous Stock Price Movements?" *Journal of Financial Economics*, 58 (2000), 215–260.

Morellec, E.; B. Nikolov; and F. Zucchi. "Competition, Cash Holdings, and Financing Decisions." Swiss Finance Institute Research Paper, (2013).

Nickell, S. J. "Competition and Corporate Performance." *Journal of Political Economy*, (1996), 724–746.

Peress, J. "Product Market Competition, Insider Trading, and Stock Market Efficiency." *The Journal of Finance*, 65 (2010), 1–43.

Polk, C., and P. Sapienza. "The Stock Market and Corporate Investment: A Test of Catering Theory." *Review of Financial Studies*, 22 (2009), 187–217.

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

Schumpeter, J. A. *Capitalism, Socialism and Democracy*. Harper & Row New York (1962).

Singh, N., and X. Vives. "Price and Quantity Competition in a Differentiated Duopoly." *The RAND Journal of Economics*, (1984), 546–554.

Subrahmanyam, A., and S. Titman. "The Going-Public Decision and the Development of Financial Markets." *The Journal of Finance*, 54 (1999), 1045–1082.

Sundaram, A. K.; T. A. John; and K. John. "An Empirical Analysis of Strategic Competition and Firm Values the Case of R&D Competition." *Journal of Financial Economics*, 40 (1996), 459–486.

Sutton, J. *Sunk Costs and Market Structure: Price Competition, Advertising, and the Evolution of Concentration.* MIT press (1991).

Verrecchia, R. E. "Discretionary Disclosure." *Journal of Accounting and Economics*, 5 (1983), 179–194.

Vives, X. "Aggregation of Information in Large Cournot Markets." *Econometrica*, (1988), 851–876.

Vives, X. "Innovation and Competitive Pressure." *The Journal of Industrial Economics*, 56 (2008), 419–469.

Internet Appendix Available at

`https://drive.google.com/open?id=0B-mpAWChDJTidWxSQkNRUGVXWnc`

TABLE 1
Optimal Investment Regression

The regression is as follows:

$$I_{i,t} = \beta_0 + \beta_1 VP_{i,t-1} + \beta_2 \text{LEVERAGE}_{i,t-1} + \beta_3 \text{CASH}_{i,t-1} + \beta_4 \text{SIZE}_{i,t-1}$$

$$+ \beta_5 \text{STOCK RETURN}_{i,t-1} + \beta_6 \text{AGE}_{i,t-1} + \beta_7 I_{i,t-1} + \beta_8 \text{COMPETITION}_{j,t-1} + \mu_i + \eta_t + \epsilon_{i,t}$$

Table 1 presents the result of the OLS investment regression. 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 except for no firm fixed effects for column 3 where Census data is available only every five years. 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	-0.031***	-0.031***	-0.052***	-0.032***
	(-10.22)	(-9.97)	(-4.25)	(-5.73)
LEVERAGE	-0.159***	-0.158***	-0.081***	-0.168***
	(-17.05)	(-16.99)	(-4.68)	(-10.74)
CASH	0.059***	0.059***	0.159**	0.082***
	(4.94)	(4.94)	(2.44)	(3.79)
SIZE	-0.063***	-0.063***	-0.010*	-0.110***
	(-15.67)	(-15.49)	(-1.86)	(-13.89)
RETURN	0.019***	0.019***	0.014	0.014***
	(10.18)	(10.11)	(1.39)	(5.17)
AGE	-0.010*	-0.010**	-0.014**	0.015
	(-1.86)	(-1.99)	(-1.98)	(1.22)
$I_{i,t-1}$	0.004	0.004	0.165**	0.008
	(0.39)	(0.40)	(2.20)	(0.39)
HHI	-0.030***			
	(-3.05)			
NUM		0.150		
		(1.42)		
CENSUS HHI			-0.090	
			(-1.47)	
HP_HHI				-0.068***
				(-4.90)
CONSTANT	0.432***	0.422***	0.211***	0.660***
	(23.01)	(22.71)	(7.96)	(17.08)
OBSERVATIONS	60,170	60,202	5,289	31,134
Firm Fixed Effects	Y	Y	N	Y
Year Fixed Effects	Y	Y	Y	Y
Adjusted R^2	0.279	0.279	0.049	0.285

TABLE 2
Summary Statistics

The sample includes all Compustat annual data from 1980 to 2012. RETURN and AGE are constructed from CRSP. We exclude regulated utilities (SIC 4900-4999) and financial firms (SIC 6000-6999). The upper part of the table summarizes variables that are used in the optimal investment regression. Investment is deflated by total assets. VP measures growth opportunity. LEVERAGE is the sum of the book value of short term and long-term debt divided by the sum of the book value of total debt and equity. CASH is the balance of cash and short term investment scaled by total assets of the firm. SIZE is the logarithm of total assets. STOCK RETURN is the change of the market value of the firm. AGE is calculated by the number of years the firm has been listed on CRSP. Next we present measures of inefficiency grouped by the competition measure used in the first-stage estimation, followed by various measures for competition. In the remaining part of the table, we report other variables that have been used.

Variable	N	Mean	S.D.	Min	P25	P50	P75	Max
Investment Regression								
$I_{i,t}$	60202	0.10	0.27	-4.87	0.00	0.05	0.13	29.38
VP	60202	0.64	0.55	-2.54	0.30	0.55	0.89	3.40
LEVERAGE	60202	0.28	0.26	0.00	0.03	0.24	0.45	1.27
CASH	60202	0.21	0.25	0.00	0.03	0.10	0.29	1.53
SIZE	60202	4.75	1.93	0.50	3.34	4.63	6.03	10.66
RETURN	60202	0.21	0.72	-0.81	-0.24	0.05	0.43	4.59
AGE	60202	2.44	0.79	0.00	1.79	2.48	3.00	4.48
Efficiency								
INF(HHI)	60170	0.08	0.19	0.00	0.02	0.04	0.09	25.60
INF(NUM)	60202	0.08	0.19	0.00	0.02	0.04	0.09	25.59
INF(CENSUS HHI)	5289	0.11	0.45	0.00	0.03	0.07	0.12	29.05
INF(HP_HHI)	31134	0.09	0.22	0.00	0.02	0.05	0.09	25.14
INF(EPCM)	59684	0.08	0.15	0.00	0.02	0.04	0.09	10.36
INF(TARIFF)	14911	0.08	0.16	0.00	0.02	0.04	0.09	5.12
INF(FREIGHT COST)	14911	0.08	0.16	0.00	0.02	0.04	0.09	5.12
INF(TRADE COST)	14911	0.08	0.16	0.00	0.02	0.04	0.09	5.12
Competition								
HHI	60170	0.22	0.18	0.03	0.09	0.16	0.28	1.00
NUM	60202	0.04	0.06	0.00	0.01	0.02	0.06	0.29
CENSUS HHI	5289	0.07	0.06	0.00	0.03	0.05	0.08	0.30
HP_HHI	31134	0.21	0.22	0.01	0.07	0.12	0.27	1.00
EPCM	59684	-0.22	1.40	-12.73	-0.05	0.00	0.05	6.53
TARIFF	22707	0.04	0.04	0.00	0.01	0.03	0.05	0.68
FREIGHT COST	22707	0.05	0.04	0.00	0.02	0.03	0.06	1.40
TRADE COST	22707	0.09	0.07	0.00	0.05	0.07	0.11	1.60
Other Variables								
TANGIBLE	60160	0.28	0.22	0.00	0.10	0.22	0.39	0.90
MTB	59442	1.70	1.79	0.11	0.79	1.15	1.90	28.45
FCF	48837	-0.05	0.33	-47.18	-0.12	-0.02	0.05	26.87
INFO	57886	0.82	0.19	0.00	0.75	0.89	0.95	1.00

TABLE 3
Inefficiency and Competition

The regression is as follows:

$$INF_{i,t} = \beta_0 + \beta_1 COMPETITION_{j,t-1} + \beta'\mathbf{X}_{i,t-1} + \epsilon_{i,t}$$

Table 3 reports OLS regression results of INF on Competition (measured by Compustat HHI, number of firms in the industry in each year over 1,000, U.S. Census HHI or Hoberg Philips HHI). **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
HHI	-0.044***	-0.022***						
	(-6.56)	(-3.89)						
NUM			0.232***	0.120***				
			(6.36)	(4.93)				
CENSUS HHI					-0.142***	-0.172***		
					(-3.60)	(-3.90)		
HP_HHI							-0.024***	-0.027***
							(-3.62)	(-3.86)
MTB		0.012***		0.012***		0.010**		0.011***
		(10.48)		(10.36)		(2.44)		(7.76)
LEVERAGE		0.024***		0.026***		-0.006		0.035***
		(4.94)		(5.21)		(-0.83)		(5.00)
CASH		0.040***		0.035***		0.116**		0.031***
		(5.14)		(4.87)		(2.18)		(3.59)
SIZE		-0.006***		-0.006***		-0.011***		-0.010***
		(-6.68)		(-6.80)		(-3.49)		(-6.38)
TANGIBLE		0.035***		0.039***		-0.027		0.017**
		(6.05)		(6.99)		(-0.85)		(2.05)
AGE		-0.003		-0.002		-0.000		0.000
		(-1.25)		(-1.06)		(-0.03)		(0.02)
CONSTANT	0.092***	0.076***	0.072***	0.065***	0.122***	0.137***	0.092***	0.100***
	(26.81)	(12.51)	(33.40)	(9.95)	(7.70)	(5.84)	(23.56)	(16.78)
Observations	60,170	59,372	60,202	59,401	5,289	5,260	31,134	30,798
Adjusted R^2	0.002	0.025	0.005	0.026	0.000	0.015	0.001	0.020

TABLE 4

Inefficiency and Competition with a Quadratic Term

The regression is as follows:

$$INF_{i,t} = \beta_0 + \beta_1 COMPETITION_{j,t-1} + \beta_2 COMPETITION^2_{j,t-1} + \beta' \mathbf{X}_{i,t-1} + \epsilon_{i,t}$$

Table 4 reports OLS regression results of INF on COMPETITION and COMPETITION squared to capture the possible non-linear effects between INF and Competition. 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	-0.139***	-0.067***						
	(-8.21)	(-4.69)						
HHI²	0.123***	0.057***						
	(6.76)	(3.67)						
NUM			0.575***	0.267***				
			(7.68)	(4.62)				
Num²			-1.674***	-0.693***				
			(-5.89)	(-2.74)				
CENSUS HHI					0.218	-0.083		
					(1.16)	(-0.55)		
CENSUS HHI²					-1.610*	-0.398		
					(-1.80)	(-0.52)		
HP_HHI							-0.065***	-0.058***
							(-3.20)	(-2.90)
HP_HHI²							0.049**	0.036*
							(2.44)	(1.90)
MTB		0.012***		0.012***		0.010**		0.011***
		(10.38)		(10.22)		(2.42)		(7.74)
LEVERAGE		0.025***		0.026***		-0.006		0.036***
		(5.04)		(5.31)		(-0.80)		(4.99)
CASH		0.039***		0.034***		0.116**		0.030***
		(4.93)		(4.61)		(2.20)		(3.53)
SIZE		-0.006***		-0.006***		-0.011***		-0.010***
		(-6.80)		(-6.74)		(-3.46)		(-6.38)
TANGIBLE		0.036***		0.039***		-0.027		0.016*
		(6.19)		(6.95)		(-0.85)		(1.90)
AGE		-0.002		-0.002		-0.000		0.000
		(-1.19)		(-1.10)		(-0.02)		(0.11)
CONSTANT	0.103***	0.081***	0.066***	0.062***	0.110***	0.134***	0.096***	0.104***
	(25.07)	(13.73)	(28.80)	(9.58)	(10.35)	(5.33)	(21.29)	(17.37)
Observations	60,170	59,372	60,202	59,401	5,289	5,260	31,134	30,798
Adjusted R²	0.003	0.026	0.007	0.026	0.000	0.014	0.001	0.020

TABLE 5
Inefficiency and Firm Size

The regression is as follows:

$$INF_{i,t} = \beta_0 + \beta_1 COMPETITION_{j,t-1} + \beta_2 SIZE_{i,t-1} + \beta_3 SIZE_{i,t-1} \times COMPETITION_{j,t-1} + \beta' \mathbf{X_{i,t-1}} + \epsilon_{i,t}$$

Table 5 reports OLS regression results of INF (MAIN) on COMPETITION and an interaction term with size to capture the possible differentiated effects between INF and COMPETITION for firms of different size. SIZE is measured by the log of total assets. 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	-0.073***	-0.041***						
	(-3.81)	(-2.64)						
HHIxSIZE	0.007**	0.004						
	(2.31)	(1.57)						
NUM			0.467***	0.310***				
			(3.84)	(3.04)				
NUMxSIZE			-0.056***	-0.041**				
			(-2.77)	(-2.18)				
CENSUS HHI					-0.250***	-0.344***		
					(-2.85)	(-4.46)		
CENSUS HHIxSIZE					0.035**	0.034***		
					(2.18)	(2.86)		
HP_HHI							-0.089***	-0.069***
							(-3.99)	(-3.33)
HP_HHIxSIZE							0.010***	0.009***
							(2.73)	(2.60)
MTB		0.012***		0.012***		0.010**		0.011***
		(10.55)		(10.59)		(2.43)		(7.77)
LEVERAGE		0.024***		0.025***		-0.007		0.035***
		(4.94)		(5.24)		(-0.90)		(4.97)
CASH		0.040***		0.035***		0.116**		0.031***
		(5.14)		(4.85)		(2.18)		(3.57)
SIZE	-0.009***	-0.007***	-0.005***	-0.004***	-0.019***	-0.013***	-0.013***	-0.011***
	(-6.32)	(-5.32)	(-5.71)	(-5.03)	(-5.62)	(-4.13)	(-6.20)	(-5.97)
TANGIBLE		0.035***		0.038***		-0.026		0.018**
		(6.08)		(6.69)		(-0.82)		(2.20)
AGE		-0.003		-0.003		-0.001		0.000
		(-1.29)		(-1.23)		(-0.05)		(0.09)
CONSTANT	0.135***	0.080***	0.096***	0.057***	0.206***	0.149***	0.165***	0.108***
	(14.17)	(11.09)	(18.40)	(6.94)	(6.80)	(6.23)	(11.57)	(15.38)
Observations	60,170	59,372	60,202	59,401	5,289	5,260	31,134	30,798
Adjusted R^2	0.008	0.026	0.011	0.027	0.005	0.015	0.009	0.020

TABLE 6
Investment based on downstream analysis

The regression is as follows:

$$I_{i,t} = \alpha + \beta_1 \text{DOWN\_OUTPUT}_{k,t-1} + \beta_2 \text{DOWN\_SALEG}_{k,t-1} + \beta_3 \text{DOWN\_MTB}_{k,t-1}$$

$$+ \beta_4 \text{DOWN\_ANALYST}_{k,t-1} + \beta_5 \text{COMPETITION}_{j,t-1} + \eta_t + \epsilon_{i,t}$$

Table 6 presents the result of the OLS investment regression. Column 1 uses HHI as the measure of competition. Column 2 uses NUM as the measure of competition. Column 3 uses Census HHI as the measure of competition. Column 4 uses Hoberg Philips HHI as the measure of competition. We use the fitted value of this regression as the expected investment and the absolute value of residuals as the proxy for inefficiency. The sample covers firm-year data from 1998 to 2012. 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 $I_{i,t}$
DOWN_OUTPUT	0.019***	0.006***	0.028***	0.021***
	(10.97)	(3.31)	(4.53)	(11.42)
DOWN_SALEG	0.040***	0.054***	0.088***	0.042***
	(3.40)	(4.72)	(3.13)	(3.44)
DOWN_MTB	0.026***	0.007**	0.014	0.026***
	(7.11)	(2.20)	(1.60)	(7.01)
DOWN_ANALYST	0.168***	0.017	0.398***	0.241***
	(3.15)	(0.31)	(2.93)	(4.23)
HHI	-0.175***			
	(-20.95)			
NUM		1.312***		
		(27.39)		
CENSUS HHI			0.088	
			(1.37)	
HP_HHI				-0.109***
				(-15.24)
CONSTANT	-0.136***	-0.020	-0.332***	-0.180***
	(-6.18)	(-0.93)	(-5.35)	(-8.01)
Observations	15,555	15,558	1,683	14,655
Year Fixed Effects	Y	Y	Y	Y
Adjusted R^2	0.051	0.102	0.077	0.046

TABLE 7
Inefficiency based on downstream analysis

The regression is as follows:

$$INF_{i,t} = \beta_0 + \beta_1\,COMPETITION_{j,t-1} + \beta'\mathbf{X}_{i,t-1} + \epsilon_{i,t}$$

Table 7 reports regression results of INF (DOWN) on Competition (measured by Compustat HHI, number of firms in the industry in each year over 1,000, U.S. Census HHI or Hoberg Philips HHI). **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(DOWN)	2 INF(DOWN)	3 INF(DOWN)	4 INF(DOWN)	5 INF(DOWN)	6 INF(DOWN)	7 INF(DOWN)	8 INF(DOWN)
HHI	-0.095***	-0.064***						
	(-7.34)	(-6.19)						
NUM			0.755***	0.677***				
			(10.72)	(10.65)				
CENSUS HHI					-0.119***	-0.140***		
					(-4.56)	(-7.50)		
HP_HHI							-0.075***	-0.081***
							(-9.48)	(-8.39)
MTB		0.017***		0.014***		0.010***		0.018***
		(5.44)		(4.70)		(5.40)		(5.15)
LEVERAGE		0.047***		0.031***		0.029*		0.049***
		(7.43)		(5.25)		(1.96)		(7.33)
CASH		0.020		-0.025*		0.031***		0.011
		(1.49)		(-1.93)		(3.47)		(0.78)
SIZE		-0.011***		-0.010***		-0.002		-0.013***
		(-7.42)		(-6.96)		(-0.57)		(-7.44)
TANGIBLE		0.048***		0.062***		-0.006		0.045***
		(3.70)		(4.92)		(-1.05)		(3.27)
AGE		-0.010***		-0.008**		-0.002		-0.009***
		(-3.28)		(-2.55)		(-0.61)		(-2.68)
CONSTANT	0.146***	0.161***	0.078***	0.110***	0.131***	0.114***	0.141***	0.172***
	(26.68)	(13.71)	(13.06)	(8.30)	(4.92)	(22.74)	(29.95)	(12.70)
Observations	15,555	15,546	15,558	15,549	1,683	1,683	14,655	14,648
Adjusted R^2	0.007	0.051	0.041	0.069	0.001	0.024	0.007	0.052

54