Received: 23 August 2016 | Revised: 27 March 2020 | Accepted: 28 April 2020 | Published on: 6 July 2020

DOI: 10.1002/smj.3184

STRATEGIC MANAGEMENT JOURNAL WILEY

RESEARCH ARTICLE

Industry competition and firm conduct: Joint determinants of risk–return relations

Michael Christensen[1] | **Thorbjørn Knudsen**[1] | **Ulrik W. Nash**[1] | **Nils Stieglitz**[2]

[1]Department of Marketing and Management, Strategic Organization Design Unit (SOD), Danish Institute for Advanced Study (DIAS), University of Southern Denmark, Odense M, Denmark

[2]Frankfurt School of Finance and Management, Frankfurt am Main, Germany

Correspondence
Thorbjørn Knudsen, Department of Marketing and Management, Strategic Organization Design Unit (SOD), Danish Institute for Advanced Study (DIAS), University of Southern Denmark, Campusvej 55, DK-5230 Odense M, Denmark.
Email: tok@sam.sdu.dk

Funding information
Danish Council for Independent Research

Abstract

Research Summary: In this study, we offer a novel approach, establishing how firm conduct and competitive interactions among firms jointly provide microfoundations of risk–return relations. Firms influence each other, and the way they reciprocate these influences to cause mutual adjustments is a core feature of industry dynamics. This core feature fundamentally influences risk–return relations. Based on our approach, we develop models that allow us to trace how, at the microlevel, firm conduct in conjunction with competitive interactions generate risk–return relations, and how these are associated with macrolevel measures of industry concentration. That translates into an approach that affords fine-grained predictions of risk–return relations based on the nature of competition in an industry—for example, Cournot or Bertrand—and observations of the industry's competitive intensity and concentration.

Managerial Summary: The risk-return trade-off is of critical importance for strategic management. Yet, the relation between industry conditions and the shape of risk-return relations is often unclear. We develop a framework that allows managers to understand how industry conditions generate risk-return relations, and how these can be inferred from macro-level measures of industry concentration. At the micro-level, we show that low variation in operational implementation of strategies may, under some conditions, be associated with high variation as well as high means in financial returns. Our primary insight is that the risk-return trade-off changes with competitive intensity: Starting from industries with two or three firms, we show that increasing competitive intensity – higher number of firms, higher entry and exit barriers – gives rise to a predictable sequence of risk-return relations that change over the life-cycle of industries. Overall, we contribute a framework that allows managers to infer the risk-return trade-off for the industry that their business units are located in.

KEYWORDS

Bertrand, Bowman paradox, Cournot, firm conduct, industry competition, reflection effect, risk–return

1 | INTRODUCTION

Over more than three decades after Bowman (1980) observed negative correlations between the mean (return) and variance (risk) of accounting-based measures, numerous studies have investigated risk–return relations (e.g., Andersen, Denrell, & Bettis, 2007; Bromiley, 1991; Fiegenbaum and Thomas 1985, 1986; Gooding, Goel, & Wiseman, 1996; Patel, Li, & Park, 2017). The literature is extensive, and the facts are well known. To reiterate the facts, empirical studies continue to find negative risk–return relations, but also continue to report positive and U-shaped risk–return. Some scholars have raised concerns that these empirical findings are artifacts of spurious correlations (e.g., Henkel, 2009). Yet, even when correcting for spurious correlations, recent empirical studies continue to find evidence for negative and U-shaped risk–return relations (see, e.g.. Holder, Petkevich, & Moore, 2016; Patel et al., 2017).

As this summary of the literature highlights, risk–return relations vary across industries. But why are risk–return relations negative in some industries, positive in others, and why do we sometimes see U-shaped risk–return relations? The contingent nature of this phenomenon suggests that risk–return relationships vary with industry settings, but how do we explain this? Facing up to this challenging problem, Andersen et al. (2007) introduced a theoretical explanation and a “strategic fit model” that can explain and predict how risk–return relationships will play out under specific environmental conditions in different settings. They showed that heterogeneity in strategic responsiveness—a firm's ability to obtain strategic fit to environmental conditions—leads to a negative correlation between mean and variance of performance and that the magnitude of this negative correlation varies with the level of dynamism in the environment. The strategic fit model convincingly demonstrated that variation in firms' strategic conduct is a possible driver of risk–return relations. This model also has elements relating to competition, but it does not explicitly capture competitive interactions. Thus, a deeper understanding of competition—and how competition in different industry settings shapes associations between risk and return—remains an important missing piece in the Bowman puzzle.

In this study, we focus on this missing piece. We offer a novel approach, establishing how firm conduct and competitive interactions among firms jointly provide microfoundations of risk–return relations. Firms influence each other, and the way they reciprocate these influences to cause mutual adjustments is a core feature of industry dynamics. This core feature, as we have discovered, fundamentally influences risk–return relations. Based on our approach, we develop models that allow us to trace how, at the microlevel, firm conduct in conjunction with competitive interactions generate risk–return relations, and how these competitive interactions affect macrolevel measures of industry concentration. The result is an approach that affords fine-grained predictions of risk–return relations based on the nature of competition in industry—for example, Cournot or Bertrand competition—and observations of the industry's competitive intensity and concentration.

Our main contribution to theory is to refocus explanations of risk–return relations by bringing industry competition into focus as a natural complement to firm conduct. By firm conduct, we broadly refer to a firm's actions in an industry, including choice of price and output as well as the underlying basis for these choices, which can be summarized in a firm's cost value. By industry competition, we refer to competition among sellers in a market, and we are particularly interested in how the level of strategic interdependence among firms influences risk–return relations. Specifically, we bring into focus what we call the “reflection effect,” capturing how strategic interdependence among competing firms affects the income stream of each firm and creates patterns of risk and return across firms at the industry level.

Competition and strategic interdependence are commonly viewed as fundamental in the study of industries. The relatively little attention they receive when explaining risk–return relations shows an obvious gap in research that we address. Specifically, according to the “reflection effect,” a firm's actions—including the degree of imperfect control in implementation of firm-level strategy—have predictable effects on the mean and variance of a firm's profit *and* the profits earned by its competitors. At the aggregate level, the reflection effect has a decisive influence on an industry's risk–return relations. The intuition is as follows. In a competitive setting, a firm's mistakes not only hurt that firm, they also benefit the competitors. If one firm is unreliably implementing its strategy, the competitors experience an equally unreliable stream of opportunities, which not only increase their average profits but also the variance in their profits. Thus, a firm with low average performance and high variance (negative risk–return) generates high variance and performance for the competitors (positive risk–return). This mechanism explains how the reflection of low performing firms' behavior is a source of positive risk–returns and how it depends on the size of the industry. The smaller the industry, the stronger the reflection effect.

We show that the reflection effect causes a positive risk–return relation which is dampened as the number of incumbents increases. In contrast, imperfect control in the implementation of firm-level strategies and firm-level heterogeneity in cost values jointly causes a negative risk–return effect.[1] These two mechanisms—firm conduct and the reflection effect caused by strategic interactions—jointly predict when risk–return relations will be U-shaped, negative, inverted U-shaped, or positive. In this regard, our approach confirms prior analyses, which is reassuring, but it also implies novel predictions. Thus, we identify conditions where positive risk–return slopes may appear. In addition, our approach offers novel predictions regarding the change of risk–return relations throughout the industry life cycle. The overall implication of our study is that a detailed treatment of industry competition, jointly with firm conduct, must be included in our theories of risk–return relations.

[1]This specification and the underlying mechanism at work can be viewed as a microfounded version of a prior model (Andersen et al., 2007). On the notion of microfoundations, see Felin, Foss, and Ployhart (2015).

The following section details our modeling approach. We then present the main results of our model and unpack the underlying theoretical mechanism. Finally, we discuss how our research informs prior theoretical and empirical work on the Bowman paradox as well as our understanding of risk–return relations, more generally.

2 | MODEL

Our overall aim is to examine risk–return relations in a competitive environment. Proper treatment requires a model structure that both allow us to characterize firm conduct and industry conditions. As in Lenox, Rockart, and Lewin (2006, 2007) and Knudsen, Levinthal, and Winter (2014, 2017), we exclude multiperiod calculations and treat firms as making rational calculations from a one-period, myopic perspective. Thus, firm conduct is captured by the firm's rational adjustment to competitors' actions, given its cost value.

A comparison with the Andersen et al. (2007) model may be useful. In this prior model, firm performance was postulated to be a (parametrized) function of the deviation from an optimal fit, thus eliminating direct competitive interactions because each firm's performance was assumed to be independent of the behavior of other industry incumbents. In contrast, we build our model from first principles using widely accepted models of competitive interactions. This approach allows us to establish microfoundations of risk–return relations encompassing industry competition and firm conduct. Specifically, we add competitive interactions by drawing on the Cournot model. The Cournot model captures industries with capacity constraints where production quantity is the decisive competitive parameter (Acemoglu, Bimpikis, & Ozdaglar, 2009; Kreps & Scheinkman, 1983).

To examine the robustness of our findings, we briefly report results from the Bertrand model. Per Costa, Cool, and Dierickx (2013, p. 447), "[t]hese two models correspond to different market situations. Price competition applies to situations where firms' capacity to produce output is flexible and, as a result, they can meet all the demand that arises at the chosen prices. Quantity competition corresponds to situations where production decisions have to be taken in advance and firms are committed to selling all their output, either because the majority of production costs are sunk or because inventory costs are significant." Thus, the Cournot model captures the situation commonly analyzed in the resource-based view of strategy—where firms operate under significant constraints regarding the expansion of capacity (Knudsen et al., 2014). This is another reason we have chosen this model as the focus of our exposition. That said, our results from the Bertrand model are available in Appendix S1 as they provide useful guidance, for example, for nimble firms competing in fluid market situations.

We use the two models of Lenox et al. (2006) as descriptions of the competitive environment, and the related standard assumptions to characterize firm conduct. In the Cournot model, firms form rational expectations about the amount of output (quantity), given their marginal costs and the distribution of marginal costs of other industry incumbents. A firm's cost value captures its level of efficiency. To keep the presentation tractable and the computations feasible, we use a standard version of the Cournot model with a linear demand function.[2] The competitive environment in the industry comprises several firms (N). Every period, for example,

[2]Results for the Bertrand model from Lenox et al. (2006), which uses a multinomial logit demand function (see Appendix S1), are included for comparison. For further robustness, we also tested a linear Stackelberg model and various bilateral trading models with similar qualitative results. These results are available upon request.

TABLE 1 Summary of the standard Cournot competition model with marginal cost determined by noisy firm conduct

Model component	Symbol/relation
Number of competing firms	N
Total industry supply at time t	$Q_t = \sum q_{i,t}$
Supply of firm i at time t	$q_{i,\ t}$
Demand function	$p_t = p_0 - \alpha Q_t$
Market price at time t	p_t
Maximum price	p_0
Quantity unit	α
Cost function of firm i at time t	$c_{i,\ t} = \beta_{i,t} q_{i,t}$
Unit cost for firm i at time t	$\beta_{i,t} = k_i + d_{i,t}^2$
Efficiency frontier, optimal cost level for firm i	k_i
Deviation from optimal cost, Gaussian noise	$d_{i,t} \in \mathcal{N}\left(0, \sigma_i^2\right)$
Noise scale, or ability to operate efficiently, for firm i	σ_i

Note: Every period, each firm realizes a specific cost as a consequence of firm conduct, after which they rationally set quantities in accordance with the equilibrium point. Mean and variance is measured by observing many such periods (here 10).

quarter or financial year, each firm adapts output to the current industry conditions. Firms are characterized by differential levels of efficiency, which are captured in heterogeneous cost values. Once the distribution of cost is known, the firms are assumed to behave rationally and the Cournot model determines equilibrium quantities.[3]

2.1 | Cournot competition

We use a standard version of the Cournot model to capture period-to-period firm-level competitive interactions (see, e.g., Lenox et al., 2006, 2007; Knudsen et al., 2014, 2017). Table 1 provides an overview of the parameters. In the Cournot model, firms produce undifferentiated goods and rationally commit to equilibrium quantities. The demand function is

$$p_t = p_0 - \alpha Q_t, \tag{1}$$

where p_0 and α are fixed positive constants, p_t is the industry-level equilibrium price at time t, and $Q_t = \sum q_{i,\ t}$ is the aggregate output value at t, summing over all firms' individual output values, $q_{i,t}$. The cost function for firm i in time-period t is linear:

$$c_{i,t} = \beta_{i,t} q_{i,t}, \tag{2}$$

[3]In the Bertrand model, firms rationally set prices given heterogenous fit with consumer preferences. Proof of the equilibrium solutions to the Cournot and Bertrand models can be found in the Appendix S1 to Lenox et al. (2006).

in the output $q_{i,\ t}$ of firm i at time t.[4] The unit cost $\beta_{i,\ t}$ is defined by a quadratic function with minimum k_i:

$$\beta_{i,t} = k_i + d_{i,t}^2. \tag{3}$$

Here k_i is a (firm-specific) baseline describing the optimal unit cost level that this firm can achieve given its existing production technology, that is, the efficiency frontier in terms of the production technology for firm i is k_i. This specification is similar to what Andersen et al. (2007) referred to as a strategic fit, that is, the closer to k_i, the higher the firm's strategic fit.

When all firms have access to identical production technology ($k_i = k$) and can utilize it in the best possible way ($d_{i,t} = 0$), they all operate at the efficiency frontier with minimal, identical unit costs ($\beta_{i,t} = \beta = k$), that is, all firms can be described as having obtained optimal strategic fit. The variable $d_{i,t}$ represents how well managers can control operations and thereby reduce variation in costs (see Knudsen et al., 2014, 2017 for a similar specification). These shocks capture random changes in a firm's production practices that increase deviation from the efficiency frontier. The higher $d_{i,t}$, the larger the distance from the efficiency frontier, and the more variation in a firm's unit costs. Thus, unit costs are heterogeneous if $d_{i,t} > 0$ even when all firms apply identical production technology. The exponent (of $d_{i,t}$) captures the well-documented notion that larger deviations from optimality have more severe consequences for the firm (e.g., Lenox et al., 2006). In each time period, deviations from optimal unit costs are drawn independently from a normal distribution, $d_{i,t} \in \mathcal{N}\left(0, \sigma_i^2\right)$, with zero mean and firm-specific *SD* σ_i. Firms, therefore, vary in their ability (σ_i) to consistently achieve efficient production. For convenience, we interchangeably refer to $d_{i,t}$ as jiggles in cost values beyond the control of the firm, random fluctuations, or simply noise.

While keeping the demand function fixed, we examine industries that differ with respect to the homogeneity (k_i) and reliability (σ_i) of available production technology and capabilities. The two parameters (k_i, σ_i) characterizing each firm's production technology and capabilities are generated at the beginning of the simulation and fixed for all periods. For each firm, σ_i, is drawn independently and uniformly in [0, w] to introduce heterogeneity in strategic fit, where $w = 2$ sets the scale of the reliability of the production technology that is available in the industry. Our baseline model uses homogeneous production technology, that is, $k_i = k$ for all i.

For each period t, a realization of unit costs is observed across firms, and the Cournot equilibrium output of firm i in time-period t is given by

$$q_{i,t}^* = \frac{1}{\alpha}\left(\frac{p_0 + B_t}{N+1} - \beta_{i,t}\right) \tag{4}$$

with $B_t = \sum \beta_{i,t}$ being the sum of all unit costs over all firms i at time t. The equilibrium market price is

$$p^* = \frac{p_0 + B_t}{N+1}, \tag{5}$$

[4]Admitting a non-negative fixed cost, δ_i, the cost function for company i in time period t would be $c_{i,t} = \beta_{i,t} q_{i,t} + \delta_i$, and the fixed cost would likewise have to be subtracted from the profit. Since we abstract from analysis of this case, unit costs are equal to marginal costs.

leading to profitability of firm i:

$$\pi_{i,t} = \frac{1}{\alpha}\left(\frac{p_0 + B_t}{N+1} - \beta_{i,t}\right)^2. \tag{6}$$

2.2 | Market participation and entry/exit dynamics

We offer a cross-sectional analysis that abstracts from entry and exit of firms, that is, we keep the number of firms (N) fixed throughout a simulation. However, even if we do not explicitly consider the arrival of new entrants, a firm may switch between an active and idle state, conditional on expected profits.[5] The firm chooses to remain idle for any period where market participation is expected to be unprofitable and is active otherwise. For firms in a Cournot industry, this occurs when optimal quantities, as given by Equation (4), would otherwise be negative, in which case they mothball their production capacity and produce zero output. When that happens, a new equilibrium solution is calculated for the remaining active firms, and the dropout process is repeated until all remaining firms are viable. During this iterative process, firms drop out according to decreasing unit costs such that high-cost firms drop out first.[6] The distribution of unit costs determines the number of firms that an industry can accommodate (i.e., its carrying capacity). More precisely, the optimal unit cost k relative to market size p_0, and stochastic deviations $d_{i,t}^2$ from the efficiency frontier determines the carrying capacity of the industry. The scenario where k is small, compared to the market size p_0, corresponds to an industry with high carrying capacity where all N firms earn positive profits in every time period. The contrasting scenario has more fluently changing market participation. In each time period, some firms are active while others are idle. Across time, some firms always remain active, some switch between active and idle states, and yet other firms never become active. That is, the parameter N captures the number of firms that potentially may become active in the industry, rather than the actual number of incumbents.

Note that adding an evolutionary component to our model—in terms of selective pressure for firms to exit, and replenishing the population with entrants mimicking successful firms—would drive the distribution of firm-specific costs towards optimal values k_i with a minimal *SD* σ_i. That is, if entry is riskless and barriers to imitation are nonexistent, an industry quickly converges to optimal market outcomes (Alchian, 1950; Lippman & Rumelt, 1982).

3 | SIMULATIONS

We draw the values for firm-specific fluctuations in efficiency—the *SD* σ_i—from a uniform distribution $[0, w]$ with $w = 2$ (lower and higher values of w were tested for robustness). For each instance of an industry that we simulate, we draw these capabilities (σ_i) initially and keep them

[5]This scheme is commonly used in models of entry and exit under uncertainty (e.g. Dixit, 1989). To maintain focus, we abstract from switching costs.

[6]This procedure was used in Knudsen et al. (2014). For robustness, two random dropout processes were tested for the Cournot model for qualitatively similar results (see the robustness section).

fixed for all periods. We focus on the case where firms have access to identical production technology $k_i = k$, and report additional results from a situation where firms are endowed with heterogeneous production technologies, $k_i = k + \epsilon_i$ where the ϵ_i are drawn independently and uniformly from the interval $[0, \epsilon]$.

We set the slope of the linear demand curve in Cournot industries to $\alpha = 1$, and the choke price of the market to $p_0 = 100$ without loss of generality.[7] For the Cournot model, the number of firms (N) and the level of unit cost (k) jointly determine the competitive intensity. We systematically explored the configuration space. In the Cournot model, we varied the unit cost parameter k from 1 to 95 (a value of 100 is equivalent to choke price), and the number of firms, N, from 2 to 100. We use the contribution margin, $p_{i,t} - \beta_{i,t}$ as a basic performance measure for both models.[8] In the absence of fixed costs, the contribution margin is equivalent to a firm's unit profit, a normalized measure that lends itself to comparisons across all instantiations of our models.

For all simulations, we report results from 5,000 data points obtained over a time window of $t_{\max} = 10$ periods. This specification ensures that we obtain a representative time slice of each industry. That is, we generate a set of firm-level profit distributions over the given time frame ($t_{\max} = 10$ periods) by extracting firm-level performance for each firm and each period, t. In each of these periods, the firms make profitability calculations, given the assumption of the Cournot model. Thus, each firm is a data point displaying the mean and variance across the history of $t_{\max} = 10$ periods. We obtain a total of 5,000 data points for each point in the grid defining the configuration space. As the number of firms varies from 2 to 100, the number of industry histories varies. For example, a single history of an industry with 50 incumbents generates 50 data points. Thus, in the case of 50 incumbents, the scatterplot is based on 100 histories. For another example, a single history of an industry with two incumbents (duopoly) generates just two data points, so in this case, the scatterplot is based on 2,500 histories. We use this normalization procedure to provide visually comparable and appealing scatterplots.

4 | RESULTS

Below, we provide detailed results for the Cournot model, followed by additional robustness results, including a summary of results for the Bertrand model. We first highlight selected results that, in detail, show how increasing competition—by increasing the number of incumbents (N = 2, 5, 10, 100)—influences the risk–return relation for a given level of $k = 10$. Recall, that the parameter k is the optimal (lowest) unit cost that any firm can achieve. These four selected results are then related to a heatmap that shows risk–return relations for the entire configuration space, that is, the two-dimensional parameter-grid (N = 2, ..., 100, k = 1, ..., 95).

Figure 1 captures a situation with low competitive intensity—two firms are present in the market ($N = 2$) and the lowest possible (i.e., optimal) unit cost a firm can achieve is set at $k = 10$. Recall that firm-specific stochastic deviations, given by $d_{i,t}$, result in a penalizing random increase in cost levels (by $d_{i\,t^2}$). If $d_{i,\ t} = 0$, we have a standard version of the Cournot model with zero variance in profits. The vertical line, labeled "baseline" in Figure 1, meets the

[7]The two parameters simply set the monetary unit and the quantity unit.

[8]We conducted a series of analyses to examine whether using total profits – instead of contribution margin – makes a difference (available upon request). It turns out that using total profits does not change our results in a qualitative way. It merely changes the scales on the axes of the plots.

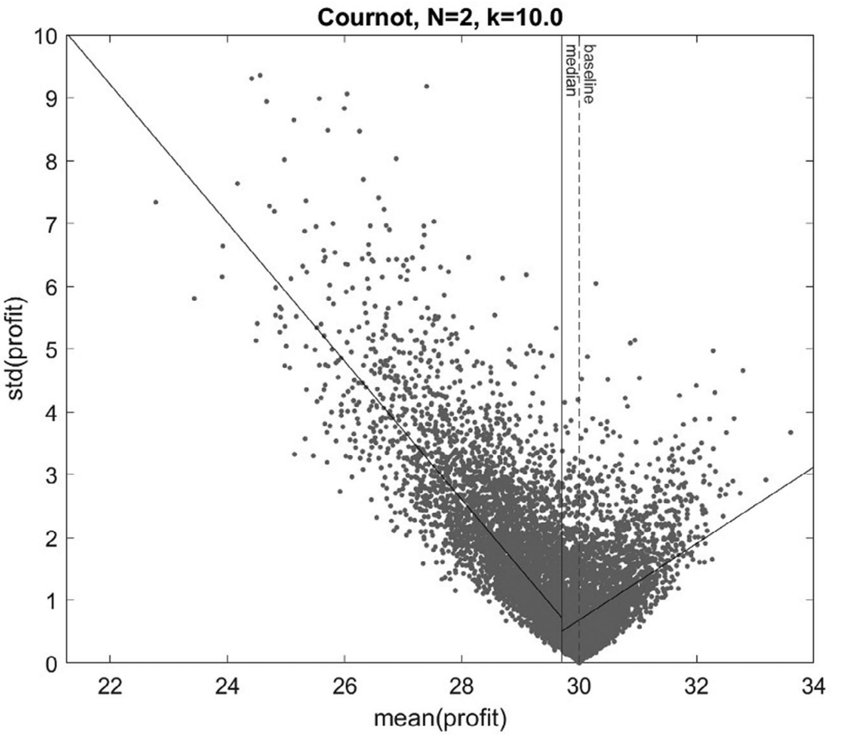


FIGURE 1 Cournot competition, mean and *SD* of unit profits, $N = 2$ firms ($k = 10$, $p_0 = 100$, $w = 2$). Computed from 5,000 data points (firms) each obtained over 10 periods. Baseline is the theoretical Cournot equilibrium point assuming no noise in cost values (dashed line). Median is the median industry performance (profits) computed from realized profits (unbroken line)

x-axis at a point of unit profits $p^* - \beta = \frac{p_0 - k}{3} = 30$, which is the solution to the standard Cournot model of a duopoly.[9]

Figure 1 shows a scatter plot of 5,000 data points obtained from 2,500 histories of a duopoly ($N = 2$), where firms can optimally realize a (minimal) cost value of $k = 10$. Traditionally, the correlation between the average and the variation of profits characterize risk–return relations, which is proportional to the slope of linear regressions to the scatter. Many studies report both total correlation and correlation for worst/best performing firms according to a split point. In Figure 1, we show two vertical lines that both may serve as split points when calculating the risk–return profile (characterized by regression lines fitted below and above the split point). These two vertical lines were computed using different methods, which we refer to as the "baseline split" and the "median split." In Figure 1, the risk–return profile was fitted to the vertical "median." The following explains the difference in estimation methods, why it matters, and why we chose "median split" as our favored method.

The "baseline split" is computed as follows. The data on average unit profits obtained from all realizations of an industry configuration (N, k) are pooled before we compute the split. For each industry configuration, the data are then split into two segments. The point of the split on

[9]Since we have a duopoly with identical cost structures and no time-varying fluctuations in costs, we omit subscripts i and t.

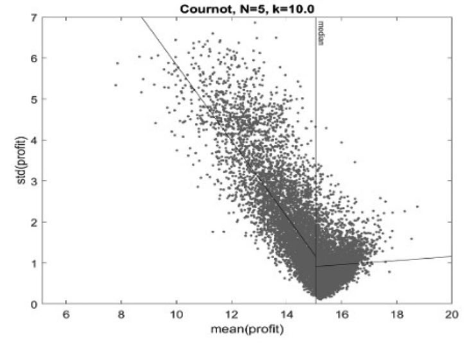
the axis denoted mean (profits) is the Cournot equilibrium, assuming no noise in cost values (i.e., the split point is independent of the simulation results). We then use the least-squares-method to fit lines to the portion of data above and below the point of this split. These two lines jointly provide the risk–return signature. In the case shown in Figure 1, it is U-shaped. Because this method implies that we pool all data across this industry configuration, it may be referred to as a baseline split "across industries." An alternative method would be to compute the baseline split "within industries." This split is done as follows. For each history (10 periods) of an industry configuration (N, k), the data on average unit profits is split into two segments. The point of the split, as before, is the Cournot equilibrium, and regression is performed for each history.[10] Subsequently, averages of the slopes above and below the split point are computed.

To examine whether this way of measuring the effect changes the results, we applied both approaches to the entire parameter space. As it turns out, there is *no* difference between the two ways of measuring risk–return relations. For the "baseline split," it does not matter if we measure the effect "across industries" or "within industries." Why? Because the "baseline split" uses the Cournot equilibrium to identify the point of the data split. This method turns out to be very robust as it identifies the same split point independent of measuring effects "within" or "across" industries. This insight is important because it implies that the Cournot equilibrium is a natural reference point for empirics (more on this, below).

The methods we have outlined, using the "baseline split" extracted from the Cournot equilibrium requires access to cost values. However, empirical research rarely has access to cost values. Rather than using the theoretical "baseline split," empirical research uses a different method that we refer to as "median split." This method commonly applies the "within industries" approach using the median of the industry rather than the theoretical baseline (see, e.g., Gooding et al., 1996). As shown in Figure 1, the two vertical lines extracted using the "median split" and the "baseline split" do not differ much. We conducted a systematic analysis comparing three cases: (a) using the theoretical Cournot equilibrium for the "baseline split," (b) using the median split "across industries," and (c) using the median split "within industries." The results shown in three heatmaps are available in Appendix S1 (Figure S1). They show that there are few and rather marginal differences regarding risk–return signatures across the three methods. This insight is important because it justifies the common empirical practice of using the median split "within industries," while at the same time providing a new, theoretically grounded method, for determining an industry's gain–loss reference point (see Gooding et al., 1996).

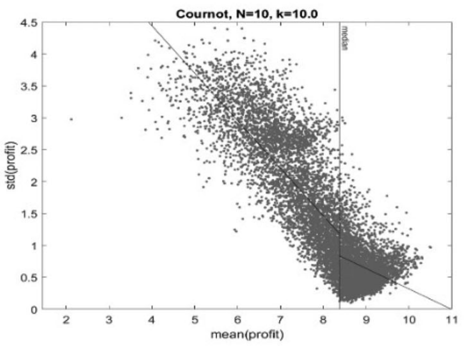
To further assess the effect of using the Cournot equilibrium for the baseline split versus median split "within industries," we computed the absolute difference between the points of the split (used to fit regression lines) that each of the two methods identifies. That is, we compared the noiseless Cournot equilibrium with the average of the median of the industries—as illustrated in Figures 1 and 2. These results, also shown in Appendix S1 (Figures S2 and S3), further support that there are few and rather marginal differences across these methods, predominantly at $N = 2$. These findings are important because they support the idea that the baseline extracted from the Cournot equilibrium can be used as a reference point when assessing empirical research. Thus, a suitably calibrated theoretical baseline split would allow empirical

[10]Technically, for $N < 4$ there is not enough points to perform the two linear regressions and, more generally, for low N the regressions are rather noisy. As a remedy, we observed the same industry (and same firms) across several time windows (10 periods each) until a suitable number of points were available for the regressions. For intermediate and high N, enough points are already available and only one time window was needed for each instance of an industry.



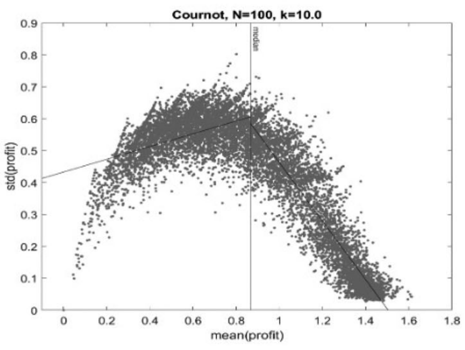
U-shape:

N=5 firms are always active. Noise in own cost values causes 1) negative mean-variance relation in profits, and 2) is reflected in the competitor's profits causing positive slope. When *N*=5, the reflection-effect is dampened, so the positive slope has small inclination.



Negative slope:

N=10 firms are always active. Noise in own cost values causes 1) negative mean-variance relation in profits, and 2) reflection in competitor's profits is muted.



Inverted U-shape:

A fraction of *N* firms is always active, so noise in own cost values causes negative mean-variance relation. More "noisy firms" inactive for longer periods, so have lower mean and variance of profits – causes positive slope.

FIGURE 2 Cournot competition, mean and *SD* of unit profits varying competitive intensity, $N = 5, 10, 100$ (remaining parameters as in Figure 1)

researchers to assess to what extent using a median split "within industries" introduces a bias in the analysis of particular cases. In the following, we favor consistency with empirical research and therefore use only the median split "within industries" in all subsequent figures.

Returning to Figure 1, the scatterplot around the "baseline split" is caused by time-varying stochastic jiggles in firms' unit costs, given by $d_{i,t} \in \mathcal{N}\left(0, \sigma_i^2\right)$, where σ_i is drawn from a uniform distribution $[0,\ w\ =\ 2]$. Some firms experience high penalizing variation in unit costs, while others consistently operate close to the efficiency frontier. Note here that any stochastic variation increases a firm's unit costs, $\beta_{i,t}$ (see Equation (3)). Figure 1 shows realizations of mean and spread of unit profit (recorded over 10 rounds) for 5,000 data points comprising 2,500 histories of duopoly firms. Ignoring the minor difference in the two split points, the vertical line marking the baseline split with zero noise ($d_{i,t} = 0$), divides the scatter in Figure 1 into two portions. First, the portion to the left of the vertical line (average unit profits <30) is scattered around a line with a *negative* slope. This scatter represents firm-specific deviations from optimal unit costs, that is, $\sigma_i > 0$. Firms with low noise levels are close to the baseline (average unit profit = 30), while firms with high noise levels tend to have realizations with lower mean and higher *SD* of profits. This explains the negative relation of mean–variance in profits shown in Figure 1. Thus, noise in own cost values causes a negative mean–variance relation in profits.

Second, the portion of the data to the right of the vertical line (average unit profits >30) is scattered around a line with a *positive* slope, that is, a positive mean–variance relation. This poses two questions: Why do we see a positive mean–variance relation? If noise always increases costs, why do we observe average unit profits greater than the optimal Cournot value of 30? Regarding the latter question, the explanation is as follows. When a duopoly firm experiences a random draw with a high-cost value, this benefits the other firm. As an extreme example, think of a random cost value so high that it precludes profitable production. In the duopoly case, the firm with a relatively lower cost value would enjoy monopoly profits, which, of course, would exceed duopoly profits obtained from the standard Cournot model. Noisy cost values break the symmetry of the Cournot equilibrium, which explains why we observe average unit profits higher than the optimal Cournot value of 30.

Generalizing this intuition also helps understand why we observe that a portion of the data, to the right of the "baseline," has a positive mean–variance relation in Figure 1. The total effect of noisy cost values can be decomposed into two portions. There is a direct effect on the firm's profits, and there is an indirect effect where changes in the firm's cost values affect its competitor's profits. As the noise in the firm's cost values increases, this is reflected in its competitor's profits as they will be higher and more variable. This effect explains the positive relation of mean–variance in profits that can be observed in Figure 1 (for average unit profit >30). In the duopoly case, the effect of stochastic fluctuations, or noise, in cost values encompasses two components: (a) firm conduct, characterized by the level of noise in own cost values, causing a negative mean–variance relation in profits, and (b) strategic interaction, noise in own cost values reflected as a benefit to the competitor, which causes a positive mean–variance relation in the competitor's profits. Jointly, these two components generate a U-shaped relation between mean and variance of profits. According to this observation, the shape of risk–return relations is jointly determined by a firm's conduct, which we capture as variation in a firm's cost values, and competitive interactions, which lead to industry-wide adjustments caused by these cost variations. At least, this finding holds for the duopoly case shown in Figure 1.

It is therefore interesting to examine how going beyond duopoly will affect the relation between mean and variance of profits, and to what extent competitive interactions will influence the mean–variance curve. Figure 2 reports the effect of increasing the number of incumbents (to $N = 5$, 10, 100), holding constant minimal unit costs at $k = 10$. When five firms ($N = 5$) are present in the market, the U-shaped risk–return relation is still present, even if the reflection effect is dampened so that the positive slope has a smaller inclination. However, when the number of incumbents is increased to 10 firms ($N = 10$), the U-shaped risk–return relation that we observed for $N = 2$ and $N = 5$ has disappeared. Instead, we observe a negative risk–return relation.

Why does increasing the number of incumbents generate a negative risk–return relation? The explanation is that the direct effect of noise in own cost values remains, while the indirect effect relating to strategic interdependence is dampened. Increasing the number of incumbents, and thereby increasing competition, does not change the basic effect that noisy cost values penalize the firm. Increasing competition does not change this basic effect because noise in cost values, by construction, moves the firm away from the efficiency frontier, thus increasing its minimal achievable cost k (by d_{i,t^2}). However, from Figure 2, the positive risk–return relation to the "right" of the median split has disappeared.

Why is it that "own noise" no longer promotes a positive mean–variance relation in the competitors' profits? As indicated, the reason is that the reflection of noise in the competitors' cost values is now dampened. From the point of view of any firm, there are now nine competitors. These nine competitors generate a joint distribution of realized cost values such that stochastic variations across these firms tend to be evened out and become a stable background "tapestry" of noise. By contrast, in the duopoly case, a single noisy competitor generates variation in output that is reflected as a bumpy random walk in the competitor's profits.

Even so, in Figure 2 we still observed a scatter with higher profits than in the standard Cournot solution for $N = 10$, which has a mean profit of $90/11 \approx 8.18$. Why? When a few firms operate with low noise, while the remaining firms operate with high noise, the symmetry of the Cournot equilibrium is broken. Some firms benefit from lucky draws, experience low variation in cost values, and consistently operate close to the efficiency frontier. These firms benefit from the penalizing costs that other firms are experiencing. This broken symmetry explains why we observe average unit profits higher than the optimal Cournot value of $\approx$8.18, which is obtained with 10 noiseless, and, therefore, homogenous firms.

In summary, as competition is increased from 2 to 10 incumbents: (a) noise in own cost values causes a negative mean–variance relation in profits, and (b) the reflection of joint stochastic variation in the competitors' cost values is dampened so that the positive risk–return relation vanishes. Jointly, these two effects generate a negative risk–return relation. What, then, happens if we further increase the number of incumbents?

Our final example shown in the lowest panel of Figure 2 investigates a market with $N = 100$ incumbents. Yet again, the increase in competition leads to a new mean and variance signature. Rather than the U-shaped risk-relation observed in the case with 2 or 5 firms, or the negative relation observed in the case with 10 firms, we now see an inverted U-shaped relation. Why? Because although the effect of noise on the firm's own cost values remains, the further increase in the number of competitors generates a joint distribution of realized cost values such that stochastic variations are dampened to the point of becoming a "dull tapestry." This effect explains the straight negative slope in the portion of the scatter with higher mean profits than the median value of $90/101 \approx 0.89$. Again, the symmetry of the Cournot equilibrium is broken so that firms with low noise earn a larger relative share of profits compared to an equilibrium formed by homogenous firms with zero noise. Since noise from the competitors is dampened to a stable background, only noise from the firm's own cost values remain. Thus, the lower the noise in cost values, the higher the mean profit. This relationship explains the negative slope to the right of the median value of $\approx$0.89.

Why, then, does the portion of the scatter with *lower* mean profits than the median value of $\approx$0.89 have a positive slope? A positive slope implies that lower mean profit is associated with lower variance. Since higher variance always increases costs—and thus reduces profits—this result is somewhat puzzling. The explanation is that the competitive intensity is so high that firms with very noisy cost values experience longer spells where it is unprofitable to produce any output. During such spells, they choose to become inactive and mothball production. As a firm earns zero profits during inactive periods, its *SD* (or variance) in profits decreases. The longer the inactive spells, the lower the variance in profits.

In summary, in markets with many incumbents: (a) noise in firms' own cost values cause a negative mean–variance relation in profits when all incumbents remain active (when stochastic cost-components never generate negative profits), (b) noise in firms' own cost values cause a positive mean–variance relation in profits when firms have longer spells where they are inactive (as stochastic cost components would otherwise generate negative profits), (c) the reflection of joint stochastic variation in cost values is completely dampened so that the positive mean–variance relation on any competitor's profits is vanishing. Jointly, these three components generate an inverted U-shaped relation between mean and variance of profits.

4.1 | Analytical results

It is a notable strength of our approach, and a testament to the virtue of our model, that we can provide exact analytically derived risk–return relations for the Cournot model. Per the above exposition, the shape of risk–return curves is jointly determined by a firm's conduct, which we capture as variation in a firm's cost values, and the reflection of these firm-specific variations through competitive interactions. This reflection effect is the underlying cause of the positive risk–return relation above the median split of average profit, as shown in Figures 1 and 2. In a duopoly, the reflection effect is very strong (Figure 1), but as the number of incumbents increases (Figure 2), the reflection effect is dampened. For this reason, the positive risk–return relation above the median shifts from a positive value in industries with few incumbents ($N \leq 5$) to a negative value in industries with many incumbents ($N > 5$). We can analytically demonstrate this contingent relation encompassing firm conduct, competitive interactions, and risk–return relations. We now proceed to show how.

We apply two simplifying assumptions. First, we assume that firms have a homogenous cost structure. Second, we assume that all firms remain active, that is, no firm needs to mothball production. In this case, the contribution margin of firm i is

$$m_i = \frac{\Delta_i - N d_i^2 + \sum_{j \neq i} d_j^2}{N+1}, \tag{11}$$

where the firm-specific constant $\Delta_i \equiv p_0 - Nk_i + \sum_{j \neq i} k_j$ reduces to a common constant $p_0 - k$ when firms have homogenous cost structure, for example, by sharing technology. Calculating the first moments, the average becomes

$$E(m_i) = \frac{\Delta_i - N\sigma_i^2 + \sum_{j \neq i} \sigma_j^2}{N+1} \tag{12}$$

and the variance becomes

$$V(m_i) = (\eta - 1)\left(\frac{N^2}{(N+1)^2}\sigma_i^4 + \frac{1}{(N+1)^2}\sum_{j \neq i} \sigma_j^4\right) \tag{13}$$

where $\eta \equiv E\left(d_j^4\right)/\sigma_j^4 = 3$ for Gaussian noise. From the first relation (Equation (12)), it is evident that increased variation in a firm's cost always reduces average profits, while more variation in competitors' cost values always increases average profits. For the duopoly, the weights on noise in the firm's own cost values versus noise in the competitors' cost values are −2/3 versus 1/3. This relation explains why the left tail (below the baseline split) in Figure 1 is roughly twice as long as the right tail. As competition increases via N, both weights tend to unity. That is, the

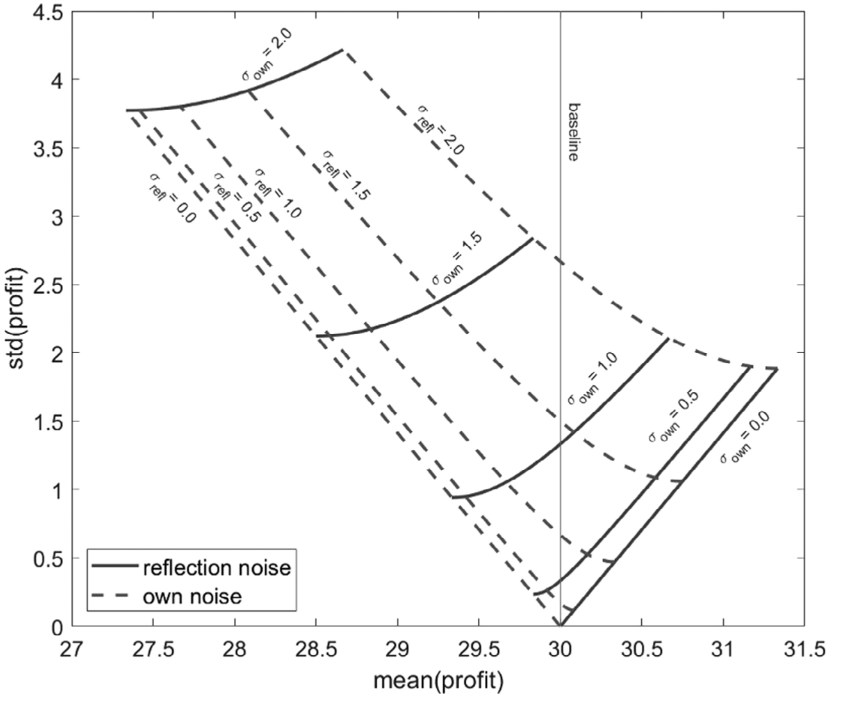


FIGURE 3 Strong reflection noise for $N = 2$. Plot of analytical solution, Cournot model ($N = 2$, $k = 10$, $p_0 = 100$), showing the mean–variance curve if noise on one side (own or others/reflection) is held constant at some mean while the other is varied (σ_{own} is the *SD* on own noise of firm *i*, and σ_{refl} is the *SD* of firm *j*). The case with no noise is the bottom point touching the *x*-axis

firm is always punished by the direct effect of its poor conduct while the reflection of the competitors' noise (the sum) becomes a fixed background. From the second relation (Equation (13)), it is evident that any noise will always contribute positively to variation of profits. The weights of the two types of noise are different, though. As the number of firms increases from $N = 2$ to $N = 100$, the weight on own noise tends from 4/9 to unity, whereas the weight on reflection noise tends from 1/9 to zero. Thus, in highly competitive settings, the variation of profits is determined solely by own conduct as summarized in the level of variation in cost values. These relations are illustrated in Figure 3, which shows the case with strong reflection noise ($N = 2$) and Figure 4, which shows the case with dampened reflection noise ($N = 100$).

The analytical results demonstrate how firm conduct and competitive interactions jointly shape risk–return relations. They also demonstrate the contingent nature of this effect—that is, how competitive interactions mediate firm conduct (variation in a firm's cost values) conditional on competitive intensity (parameter *k*) and the number of incumbents (*N*) present in a market.

5 | RESULTS FROM AN EXHAUSTIVE EXAMINATION OF THE CONFIGURATION SPACE

As the analytical proof was derived under simplifying assumptions, we now turn to an exhaustive examination of the configuration space where the assumption that all firms remain active is lifted (as in Figures 1 and 2). That is, we allow firms to mothball production during periods

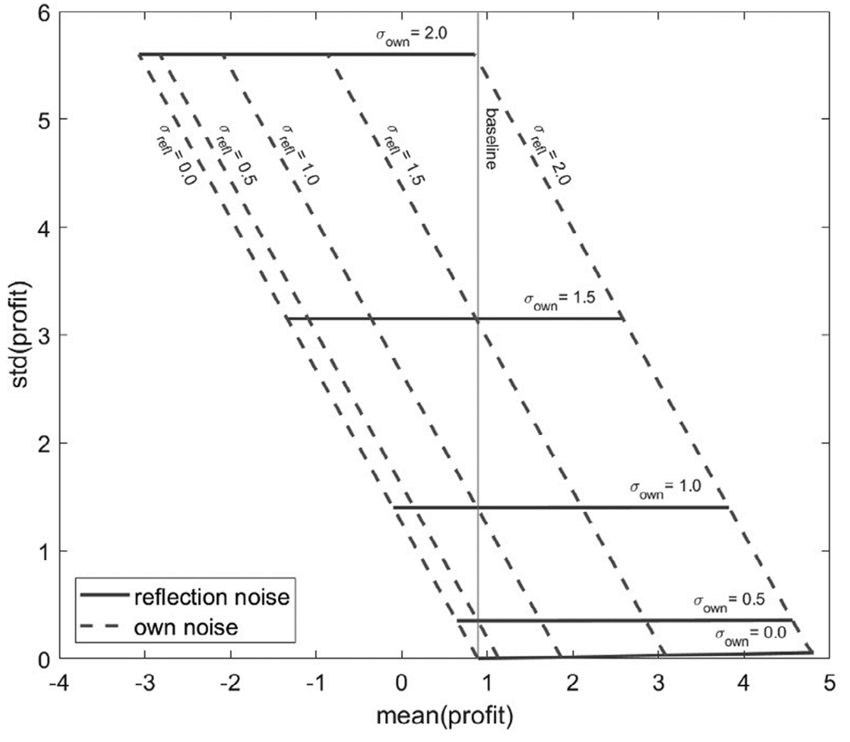


FIGURE 4 Dampening of reflection noise when $N = 50$. Plot of analytical solution, Cournot model ($N = 100$, $k = 10$, $p_0 = 100$), showing the mean–variance curve if noise on one side (own or others/reflection) is held constant at some mean while the other is varied (σ_{own} is the *SD* on own noise of firm i, and σ_{refl} is the *SD* of all other j firms). The case with no noise is the bottom point touching the x-axis

where they would otherwise incur a loss. The robustness section considers the effect of the other limiting assumption, including heterogenous cost structures.

The four examples we have highlighted in Figures 1 and 2 represent points in a larger configuration space spanning a total of 288 combinations of parameters N and k. These four points are shown in the heatmap in Figure 5, which represents the expanded configuration space. This heatmap was produced as follows. For each of the parameter sets in the (N, k) grid shown in Figure 5, we extracted the average slopes of the regression lines below and above the point given by the median split "within industries" (as explained above for Figures 1 and 2). The signs of the slopes of the two regression lines jointly determine the stylized signature, for example, the U-shape is composed of a negative slope below the split point and a positive slope above that point. Our identification of risk–return signatures remains robust across methods (for details, refer to Section 6).

As shown in Figure 5, the three examples from Figures 1 and 2 comprise a set of stylized signatures—U-shape, negative, inverted U-shape—that represent signatures from the entire parameter space. These 3 stylized signatures correspond to three market situations: low competitive intensity (oligopoly with 5 firms, or less), medium competitive intensity (industries with more than 5, but less than 55 firms), and highly competitive intensity (industries with more than 55 firms). As the minimal achievable cost level (efficiency frontier) is pushed toward higher levels, the competitive intensity is further increased, which explains the "shift to the left" in the upper portion of Figure 5 (for values of $k > 20$).

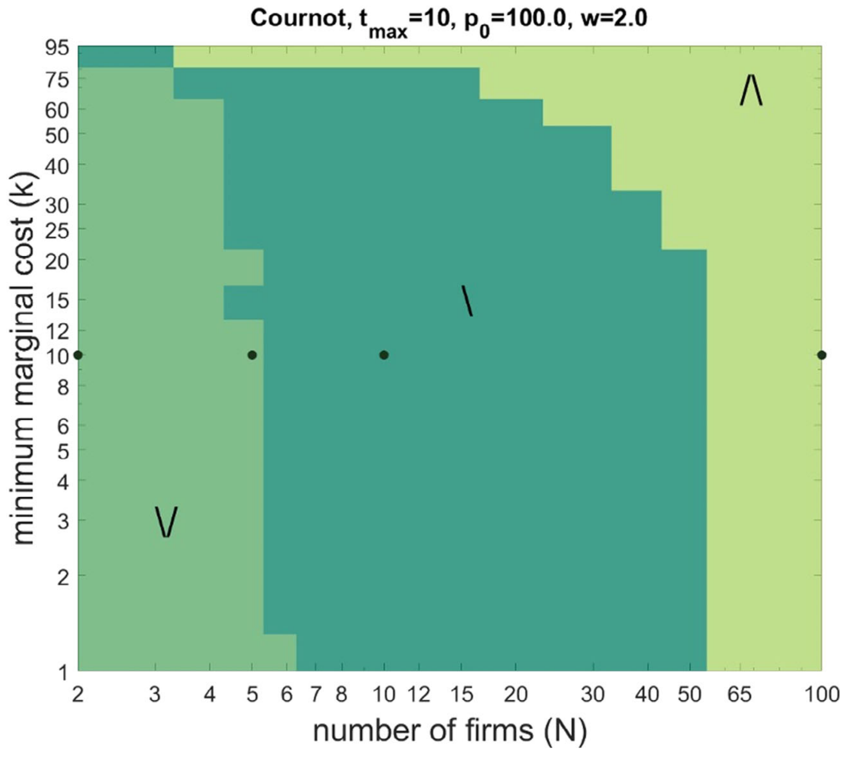


FIGURE 5 Cournot competition, mean and *SD* of unit profits, $N = 2, \ldots, 100$ firms ($k = 1\ldots95$, $p_0 = 100$). Each point in the grid is based on 5,000 data points obtained over 10 periods. Black dots in the grid refer to plots shown in Figures 1 and 2. The signatures denote risk–return relations: \/, U-shape; \, negative; /\, inverted U-shape [Color figure can be viewed at wileyonlinelibrary.com]

Our findings capture markets that are characterized by Cournot competition. While the exact parameter values used in Figure 5 may not characterize a particular market, the qualitative results should hold for empirical cases. Thus, we predict that markets characterized by quantity competition have three distinct risk–return relations that vary with competitive intensity: (a) low competitive intensity—U-shape, (b) medium competitive intensity—negative, and (c) high competitive intensity—inverted U-shape. The parameter k corresponds to the support of the market under fixed demand conditions—as k decreases (e.g., because technical innovation reduces costs), the efficiency frontier moves in the direction of lower minimal costs so that the market supports more firms. This parameter is not always available in empirical studies, so it is reassuring that our predictions are fairly robust to variations in parameter k. Furthermore, our results are consistent with empirical research (e.g., see Andersen et al., 2007; Patel et al., 2017) as the overall correlations were negative for 287 out of the 288 combinations of parameter N and k represented in Figure 5.

In summary, we find an inverse U-shaped risk–return relationship when the competition is intense because of many incumbents (high N) relative to the carrying capacity of the market (high k). This finding is novel. We also find that increasing competitive intensity, starting from a duopoly, gives rise to this sequence of risk–return relations: U-shaped, negative, and inverse U-shaped. This finding implies a novel prediction for the development of risk–return relations over the life cycle of industries. The prediction is U-shape in the early stages of the industry, inverse U-shape around the industry shakeout where the competitive intensity peaks, and then a negative risk–return relation when the industry matures.

6 | ROBUSTNESS

We thoroughly examined the boundary conditions of our model with respect to (a) choice of method to fit regressions and determine risk–return signatures, (b) dropout procedure used to ensure a viable equilibrium, (c) assumptions regarding the heterogeneity of strategic fit, (d) heterogeneous versus homogeneous cost structures, and (e) Bertrand as opposed to Cournot competition.

6.1 | Reference point and fit of regressions

Rather than using the median "within industries" to determine the reference point used to split samples, we examined how using "baseline split" or median split "across industries" would influence our results. We found that the domains of the stylized signatures were largely unaffected by the exact method (see Appendix S1, Figure S1). We also applied alternative methods—spline, quadratic or cubic polynomials—to estimate the regression lines, given this reference point. We found that our results are robust to the way regressions are fitted to determine risk–return signatures

6.2 | Dropout procedure to ensure viable equilibrium

To establish a viable equilibrium in our model, we used a sequential elimination procedure as in Knudsen et al. (2014, 2017). Firms are viable and remain in the market as long as their unit costs are less than the industry price. Because of fluctuations in cost values, firms sometimes become unviable. Such unviable firms are found sequentially by eliminating the firm with the highest cost value and recomputing the industry equilibrium in the absence of this firm. This process is iterated until all remaining firms are viable.

We tested two alternatives to this procedure. In the first alternative approach, all unviable firms are listed, and a random draw from that list determines which firm leaves the market. Yet again, a new list of unviable firms is generated, and a new random draw eliminates the next firm. This process is iterated until all remaining firms are viable. It is reassuring that we observed insignificant changes from exercising this procedure as compared to the method we have used.

In the second alternative procedure, firms in the *entire* population were eliminated at random. It turns out that this random dropout procedure has the interesting effect of generating much more uncertainty. Those firms that remain active (which is a random event) face much more diverse competition compared to the basic model, thus increasing the reflection effect. This leads to a clean upward sloping risk–return curve, an effect similar to applying heterogeneous costs (see below).

Overall, our results are robust to changes in the drop-out procedure that we apply to ensure a viable equilibrium unless this procedure introduces heterogeneity in the population. In the latter case, the dropout procedure will have an effect, which is similar to assuming heterogeneous costs.

6.3 | Heterogenous strategic fit

Strategic fit is an important aspect of firm conduct. As in Andersen et al. (2007), we characterize strategic fit as the level of variation in cost values that a firm can achieve. The lower the variation in cost values, thebetter is the fit. In the above results, the variable characterizing strategic fit σ_i was drawn from a uniform distribution $[0, w]$ with $w = 2$. What happens if we vary this specification?

The effect is contingent on the direction of change in w. An increase in w implies increased heterogeneity in strategic fit. In that case, uncertainty in the industry grows and so does the overall variance and differences in mean return. At higher levels of N and k, increases in w will cause differences in costs that generate systematic sorting on the duration of firms' (in)active spells. In consequence, an increase in w will induce an inverted U-shape.

As w approaches zero, σ_i also approaches zero, and all firms will approach perfect strategic fit. If firms have homogenous cost values below the choke price, all N firms will remain active, and all N firms earn the same profits. In this case, w sets the scale on which differences in mean and variance is measured, but the shape of the mean–variance curve does not change, that is, the portion of Figure 5 with U-shaped and negative risk–returns is robust in this regard. However, as w approaches zero, the inverted U-shaped risk–return signature disappears.

6.4 | Heterogeneous costs

From Equation (12) and (13), it follows that the variation in profits (risk) is unaffected by heterogeneity in costs $k_i = k + \epsilon_i$. However, the introduction of heterogeneous costs will affect average profits (return). Under weak competitive pressure (we tested $\epsilon/k < 1/2$), heterogeneity in costs slightly shifts the risk–return points horizontally as the split points are not identical across simulated industries, making the scatter plots slightly blurred. Under strong competitive pressure (high k) with small N, heterogeneous costs will also concentrate market participation (active states) on low-cost firms. Differences in costs thus generate systematic sorting on the duration of firms' (in)active spells. The higher the costs, the longer the inactive spells suffered by these firms, and the lower the mean and variance of their profits. These conditions produce a strictly upward sloping risk–return curve for the Cournot model in a domain of the configuration space that is comparable to its appearance in the Bertrand model (see below).

Why did we not observe a positive risk–return profile under the assumption of homogenous costs? Because the equilibrium output in the Cournot model is given by a linear equation (Equation (4)), active (and inactive) states are evenly distributed across all firms. On average, no single firm benefits from the absence of others. This explains why a positive risk–return curve is not observed in the Cournot model with homogeneous cost.

The overall conclusion is that introducing differences in firm-specific costs in the Cournot model does not, in a qualitative sense, change our results. However, heterogeneous costs will, under strong competitive pressure in markets with little carrying capacity (small N), give rise to a positive risk–return relation. As explained below, we observe a similar effect in the Bertrand model.

6.5 | Bertrand competition

While industries, such as oil and mining, are capacity constrained, other industries, such as insurance, correspond better to the assumptions of Bertrand competition. It is therefore interesting to see how the form of competition—price versus quantity—possibly influences our predictions for the relation between the mean and variance of profits. The mathematical relations necessary to compute Bertrand results follow the procedure developed by Lenox et al. (2006). They are available in Appendix S1. We used the same procedure as applied to the Cournot model to compute data and extract risk–return relations.

The results from the Bertrand model are comparable to the results from the Cournot model. Overall, we find that markets characterized by Bertrand competition have four distinct risk–return signatures that vary with competitive intensity: (1) low competitive intensity—U-shape, (2) medium competitive intensity—negative, (3) highly competitive markets with a large number of firms—inverted U-shape, and (4) extremely high competition with a small number of firms—positive. Signatures 1–3 were also present in the Cournot model and roughly distributed in the same domains of the configuration space, as showed in the heat-map. The positive risk–return relation (Signature 4), however, appears to be unique to the Bertrand model. In highly competitive markets, we observe a positive risk–return relation, but only when there are few incumbents. Why?

The explanation is that, under high competitive intensity, firms experience spells where their output has a vanishing market share. During such spells, they are practically idle. This is analogous to the Cournot model where firms mothball production if costs exceed industry price, except idleness is a natural feature of the Bertrand model when firms are unable to attract consumers with their inferior products. As a firm earns zero profits during idle periods, its *SD* (or variance) in profits decreases. The longer the inactive spells, the lower the variance in profits. This explains the positive relationship between mean and variance of profits (Signature 4). But why do we not see a negative slope in some portion of the data such that the relationship becomes an inverted U-shape?

This result is explained by the dampening of the "reflection effect." When there are few firms in the Bertrand model, the uncertainty regarding a firm's state, whether it is idle or active, will cause a large variation in profitability. As the number of firms increases, however, some firms with superior conduct are almost always active, and as this number of active participants stabilizes, the variance in profitability is reduced. In the Cournot model, each firm's equilibrium output is given by a linear equation (Equation (4)), which implies that the noise on cost values does not systematically benefit any particular firm. In the Bertrand model, however, equilibrium output is given by an exponential function. This leads to the sorting of firms such that "noisier firms" are inactive for longer periods and therefore have lower variance of profits as well as lower mean profits. While it is difficult to track idle firms in empirical work, data from the U.S. manufacturing industry provides some support for our theoretical finding as a trend toward zero variation for worst performers has been observed (see Figure 2, Patel et al., 2017).

In summary, we have thoroughly examined the boundary conditions of our model and found it remarkably robust to changes in the basic assumptions used in our exposition of results. At the same time, we have identified boundary conditions that will change these results, explained what these changes are, and why they occur. Finally, we have identified a strict positive risk–return profile, which occurs in the Bertrand model, and for the same underlying reason, when firms operate with heterogeneous costs in the Cournot model. The reason we observed positive risk–return profiles is that the duration of spells where the firm remains idle is positively correlated with their cost values. In the following, we further discuss and summarize our findings.

7 | DISCUSSION AND CONCLUSIONS

The Bowman paradox and the literature on risk–return relations it has inspired have advanced our understanding of the contingent nature of the association between firms' mean and variance of profit. The extensive prior literature has emphasized firms' risk-preferences, strategic fit, and other aspects of firm conduct, while it has not linked industry competition to risk–return relations. The present work addresses this research gap as we show that firm conduct and competitive interactions among firms jointly explain risk–return relations.

While prior work has established that firms' actions and the underlying systematic sources of these actions relating to strategic fit and risk-preferences are important drivers of risk–return relations, the literature on industrial organization points to how these actions are reflected in a competitive industry. Firms are interdependent, and the actions one firm takes influence the profitability of the actions that other firms take. This strategic interdependence, which implies that a firm's actions are reflected in other firms' actions and outcomes is captured in classical models of competition, that is, the Cournot, Bertrand, and Stackelberg models. Using these standard models of industry competition, we showed that this "reflection effect" generally causes a positive risk–return relation. Consistent with prior work, we also find that the actions of firms related to imperfect fit cause a negative risk–return effect. As demonstrated analytically, and through extensive computer simulations, we showed that these two mechanisms jointly predict when risk–return relations are U-shaped, negative, inverted U-shaped, or positive. More generally, our perspective emphasizes that variance in profits is generated by a two-sided process, encompassing a firm's choice—as emphasized in behavioral explanations of risk–return relations (Andersen et al., 2007) and prospect theory (Fiegenbaum, 1990; Fiegenbaum & Thomas, 1988)—and the competitive context, which determines how other firms in an industry modify the basis of that choice.

The finding that the reflection of a firm's actions influences the actions and outcomes of other firms in an industry is not novel. Our contribution is rather to relate an industry's competitive interactions to the shape of risk–return relations. This is a novel finding. As we have shown, the "reflection effect" generates a positive risk–return relation in the portion of the data above the median industry profitability. The underlying reason is that high variance in profitability, which low-performers experience, is reflected in the high-performers' adjustments. Our results thus suggest a strong general theorem (or conjecture) that "minimum variance in profits will occur for a firm that does not have maximum expected performance." The (normative) implication is that firms should aim at benefiting from competitors' mistakes rather than minimizing variance in (own) profits. We have shown that the reflection effect, which generates this result depends on the number of incumbents in an industry.

The reflection effect, which is obviously strongest in a duopoly, is dampened when more incumbents are present in an industry. This dampening effect occurs because the joint effect of performance variance will tend to cancel out as the incumbents increase in number. Thus, the positive portion of risk–return relations becomes weaker as the number of industry incumbents increases, which in turn will show up as a stronger overall negative risk–return relation. This finding implies that risk–return relations are U-shaped in early stages of an industry, where markets expand and competition is lenient, while they become negative after the shake-out where the industry has reached a steady state. This prediction of shifts in risk–return relations over the life cycle of an industry adds new insight to the literature on firm strategy and industry dynamics (Klepper, 1997; Klepper & Graddy, 1990; Knudsen et al., 2014). A further related insight is implied. In industries where competitive responses are delayed, or muted, it follows that the reflection effect is dampened so that, overall, risk–return relations become more negative.

While it is beyond the scope of the present article to test this prediction, note that our approach directly relates to measures on concentration commonly used in the literature on industrial organization. In this regard, it should be noted that a test of our prediction requires that industries are well defined, in the sense that otherwise unrelated firms are not lumped together because of coarse-grained market definitions. Thus, we wish to call for future work that applies a disaggregated empirical approach. While our model, consistent with empirical findings (Patel et al., 2017) and in agreement with prior models (Andersen et al., 2007), predict that the overall correlation between risk and returns for most industries at any given time is negative, a more finely decomposed analysis reveals hidden underlying patterns. A first step in that direction would be to analyze panels of risk–return relations for single industries. That is, avoid grouping of data from different Standard Industrial Classification (SIC) codes. If possible, add a dummy to indicate whether the industry is postshakeout. While this approach in and of itself is interesting, it would potentially serve the further purpose of revealing an underlying reason for unexplained variations in the strength of the negative correlations observed at higher levels of aggregation across multiple four-digit SIC codes (e.g., as in Andersen et al., 2007 or Patel et al., 2017).

Another note relates to the study of duopolies or oligopolies with three or four firms. Such industries are less common than larger industries but obviously important for the literature on firm strategy. For example, some commonly mentioned examples of duopolies that have existed for some duration of time include Apple and Android (Smartphones), MasterCard and Visa (Electronic payments), and Coca-Cola and Pepsi (Soft drinks). While data may currently be sparse on such industries, we hold that a good theory should also be able to explain exceptional cases.

Our examination of the dynamics of competition also uncovered situations where industries are characterized by a positive risk–return relation (no U-shape). This positive relationship occurred when firms experience idle spells because it is unprofitable to produce any output. As a firm earns zero profits during idle periods, its variance in profits decreases. The longer the inactive spells, the lower the variance in profits. As the number of firms increases, some firms with superior conduct are almost always active, and as this number of active participants stabilizes, the variance in their profitability is reduced. If cost values are heterogeneous or competitive pressures strong (Bertrand model), the result is a sorting of firms that generates a positive risk–return profile as predicted in standard theories of corporate finance. While this situation may not be widespread, there is both a theoretical and an empirical basis for developing this conjecture. A well-known empirical example is the Aluminum industry, where small firms comprising a price-taking fringe are temporarily idle (see Yang, 2001). For a theoretical example of such situations, Dixit's (1989) analysis of entry and exit decisions under uncertainty employs the assumption that firms may become idle to avoid negative profits. More to this point, the U.S. Census Bureau has, for decades, used a category of "temporarily or seasonally inactive" firms. Perhaps the future will see more examples of businesses that produce based on temporary spot contracts while enduring idle periods? In such cases, our model predicts a positive risk–return relation.

While our study is theoretical, we thoroughly examined the effect of using the industry median as a reference point for splitting the data in a high- and low-performing portion. Compared to using the theoretical Cournot point to perform this split, we found that the domains of the stylized signatures were largely unaffected by the exact method. Thus, our results are robust to the way regressions are fitted to determine risk–return signatures. The implication is that performing a median split in theoretical research is generally robust. The situations with notable deviations are industries with few incumbents. For these industries, a suitably calibrated theoretical split would allow empirical researchers to assess to what extent using the median split introduces a bias in the analysis of particular cases.

Per the above mentioned substantive results, the overall implication of our study is that a detailed treatment of industry competition, jointly with firm conduct, must be included in our theories of risk–return relations. Thus, we hope that we have provided an argument for bringing the question of industry dynamics and competitive interactions among firms to the foreground of our understanding of risk–return relations.

ACKNOWLEDGEMENTS

The authors benefited from comments on previous drafts from Kevin Du, Tobias Kretschmer, and two anonymous reviewers, as well as from participants at the TOM Scoiety's yearly conference and the Vienna Conference on Strategy, Organizational Design and Innovation. This work was supported by the Danish Council for Independent Research.

ORCID

Michael Christensen https://orcid.org/0000-0002-0299-1237
Thorbjørn Knudsen https://orcid.org/0000-0003-2798-7485
Ulrik W. Nash https://orcid.org/0000-0002-1819-7281
Nils Stieglitz https://orcid.org/0000-0002-2158-6691

REFERENCES

Acemoglu, D., Bimpikis, K., & Ozdaglar, A. (2009). Price and capacity competition. *Games and Economic Behavior*, *66*(1), 1–26.

Alchian, A. (1950). Uncertainty, evolution and economic theory. *Journal of Political Economy*, *58*(3), 211–221.

Andersen, T. J., Denrell, J., & Bettis, R. A. (2007). Strategic responsiveness and Bowman's risk-return paradox. *Strategic Management Journal*, *28*(4), 407–429.

Bowman, E. H. (1980). A risk/return paradox for strategic management. *Sloan Management Review*, *21*, 17–33.

Bromiley, P. (1991). Testing a causal model of corporate risk taking and performance. *Academy of Management*, *34*(1), 37–59.

Costa, L. A., Cool, K., & Dierickx I. (2013). The competitive implications of the deployment of unique resources. *Strategic Management Journal*, *34*(4), 445–463.

Dixit, A. K. (1989). Entry and exit decisions under uncertainty. *Journal of Political Economy*, *97*(3), 620–638.

Felin, T., Foss, N. J., & Ployhart, R. E. (2015). The microfoundations movement in strategy and organization theory. *The Academy of Management Annals*, *9*(1), 575–632.

Fiegenbaum, A., & Thomas H. (1985). An Examination of the Structural Stability of Boman's Risk-Return Paradox. *Academy of Management Proceedings*, *1985*(1), 7–10.

Fiegenbaum, A., & Thomas H. (1986). Dynamic and risk measurement perspectives on bowman's risk-return paradox for strategic management: An empirical study. *Strategic Management Journal*, *7*(5), 395–407.

Fiegenbaum, A. (1990). Prospect theory and the risk-return association. An empirical examination in 85 industries. *Journal of Economic Behavior and Organization*, *14*(2), 187–203.

Fiegenbaum, A., & Thomas, H. (1988). Attitudes toward risk and the risk-return paradox: Prospect theory explanations. *Academy of Management Journal*, *31*(1), 85–106.

Gooding, R. Z., Goel, S., & Wiseman, R. M. (1996). Fixed versus variable reference points in the risk-return relationship. *Journal of Economic Behavior and Organization*, *29*(2), 331–350.

Henkel, J. (2009). The risk-return paradox for strategic management: Disentangling true and spurious effects. *Strategic Management Journal*, *30*(3), 287–303.

Holder, A. D., Petkevich, A., & Moore, G. (2016). Does managerial myopia explain Bowman's paradox. *American Journal of Business*, *31*(3), 102–122.

Klepper, S. (1997). Industry life cycles. *Industrial and Corporate Change*, *6*(1), 144–181.

Klepper, S., & Graddy, E. (1990). The evolution of new industries and the determinants of market structure. *RAND Journal of Economics*, *21*(1), 27–44.

Knudsen, T., Levinthal, D. A., & Winter, S. G. (2014). Hidden but in plain sight: The role of scale adjustment in industry dynamics. *Strategic Management Journal*, *35*(11), 1569–1584.

Knudsen, T., Levinthal, D. A., & Winter, S. G. (2017). Systematic differences and random rates: Reconciling Gibrat's law with firm differences. *Strategy Science*, *2*(2), 111–120.

Kreps, D. M., & Scheinkman, J. A. (1983). Quantity precommitment and Bertrand competition yield Cournot outcomes. *The Bell Journal of Economics*, *14*(2), 326–337.

Lenox, M., Rockart, S., & Lewin, A. (2007). Interdependency, competition, and industry dynamics. *Management Science*, *53*(4), 599–615.

Lenox, M. J., Rockart, S. F., & Lewin, A. Y. (2006). Interdependency, competition, and the distribution of firm and industry profits. *Management Science*, *52*(5), 757–772.

Lippman, S. A., & Rumelt, R. P. (1982). Uncertain imitability: An analysis of interfirm differences in efficiency under competition. *The Bell Journal of Economics*, *13*(2), 418–438.

Patel, P. C., Li, M., & Park, H. D. (2017). A replication of Bowman's paradox phenomenon across 28 countries. *Journal of Innovation & Knowledge*, *3*(3), 128–142.

Yang Sheng-Peng. (2001). *Review of Industrial Organization*, *19*(3), 365–380.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of this article.

How to cite this article: Christensen M, Knudsen T, Nash UW, Stieglitz N. Industry competition and firm conduct: Joint determinants of risk–return relations. *Strat Mgmt J*. 2020;41:2315–2338. https://doi.org/10.1002/smj.3184