Data and Markups: A Macro-Finance Perspective[*]

Jan Eeckhout — UPF Barcelona[†]
Laura Veldkamp — Columbia[‡]

March 2, 2026

Abstract

What does market power look like in a data economy? Economists typically use markups to measure market power. We use a simple model to show how firms' growing stocks of data can change markups, by affecting the firm's ability to reduce uncertainty. Data's effects depend on how markups are aggregated. Growing data can produce differences in markup measures that match empirical facts. Markup aggregation wedges can measure data stocks and offer a way to purge markups of data's effects to reveal market power.

Keywords: Information frictions, data, macroeconomy, learning, capital allocation, endogenous markups.
JEL. C6. D4. D5. L1.

[*]We thank Vasco Carvalho, Joel David, Maryam Farboodi, Basile Grassi, Hugo Hopenhayn, Ariel Pakes, Bruno Pellegrino and Karthik Sastry as well as seminar audiences for insightful discussions. We benefited from superb research assistance by Renjie Bao, Ankit Bhutani, Alessandro Fornassari, Yihao Li, Zihao Li, Huasheng Nie, Xiaobo Yu, and Judy Yue. Eeckhout gratefully acknowledges support from the ERC, Advanced grant 882499, and from PID2022-138443NB-I00. This paper originally circulated under the title, "Data and Market Power."
[†]jan.eeckhout@upf.edu – ICREA-BSE-CREI
[‡]lv2405@columbia.edu

Changes in firms' market power and the sources of those changes have become the focus of intense debate. Economists point to rising markups, economies of scale in information and the dominance of large, data-intensive firms as evidence that the unequal accumulation of data is responsible for a decline in competition (Jarsulic 2019). There are two main barriers to investigating this phenomenon. One is that we cannot observe firms' stocks of data. Second, even if we could observe data, we would not know how to establish whether or not any covariances one might find are indicative of data. This paper uses theory to first illustrate the difficulty: Data may raise some markups but lower others. Then, it uncovers new ways of measuring data and new predictions that distinguish data from other forces that generate markups. The paper does not provide a definitive answer about the cause of markups. But it takes a critical first step in that agenda.

Modern big-data algorithms are prediction technologies. They might predict the next word, phrase, image or financial outcome. But, at their essence, these algorithms use data to predict uncertain outcomes. In reality, firms use data to predict demand, costs, procurement delays, hiring needs, financial outcome and many more uncertain outcomes. To keep the analysis tractable, we focus on one source of uncertainty. For this purpose, we choose demand to be the uncertain outcome firms predict. We represent data as a tool to reduce prediction errors and enable more profitable production decisions.

This simple representation of data as a prediction input can generate a rich array of markup effects. Besides the power for firms to affect their market prices, there are three additional forces that show up in product markups: cost advantages, compensation for risk and data barter. For firm and industry markups, additional aggregation effects arise. Data interacts with all of these. We first use a model with exogenous data to explain how data affects markups through cost, risk and aggregation. Later, we endogenize data to show how data barter depresses markups as well. Our work does not prove that each assumption is necessary, although the model is consistent with many empirical findings. Rather, the model shines a light on a number of potential issues and proposes a new data measurement technique. The goal is to explain relevant data-economy forces and point to unexplored aspects of markups, whose measurement might change their interpretation.

The essence of data is that it reduces uncertainty. Just like a larger data set reduces the econometrician's standard error, more data for firms reduces their prediction errors. Making future events more predictable improves firms' decisions and expected profits and reduces their uncertainty and risk. While abstracting from risk is appropriate to study many questions, abstracting from risk and the price of risk when studying data masks its essential character and misses measurement opportunities. Section I introduces a model where data is used to guide firms' production decisions, in an uncertain environment.

Section II shows that if firms use data to improve their forecasts and make their revenues more predictable and less risky, they produce more, which lowers price. So, holding costs fixed, the "risk channel" implies that more data reduces markups. However, data also makes it more profitable to grow large. In the model, firms choose an up-front investment, which lowers their future marginal cost of production. When data lowers the risk of investment by predicting future demand, firms invest more. This is *investment-data complementarity*. More investment means the firm grows larger, produces at lower marginal cost, and earns higher markups. What we call "investment channel" might be described in the language of Sutton (1991, 2001) as: our firms strategically use data to differentiate themselves and create a dominant position. While this logic is not novel or unique to data, it is prevalent in the data debate and worth including for consideration, alongside the other data-related effects.

A key theme of the paper is that, to disentangle data from markups, we should look to covariances. What distinguishes data from other assets and forces is that data facilitates prediction. Better predictions alter the composition of products and firms. Data-rich firms produce more of goods that their data predicts are likely to be profitable (high-markup goods). Thus, the firm with more data uses data to skew its product mix toward higher-markup goods. The covariance that data allows firms to achieve between production and markups is exactly the covariance that creates aggregation effects. This is useful because it means that aggregation wedges, in this case the difference between the average product markup and the firm's markup, can be used to infer firms' data. Not only can data explain composition effects in markups, but some sort of information is necessary to explain the change in composition. Every firm would like to produce more of the more profitable goods and less of the less profitable ones. This is only a feasible (measurable) strategy if the firm can predict what will be profitable and what will not. Good prediction requires good data.

Data also causes the markup of an average firm and its industry to diverge. Data-investment complementarity ensures that high-data firms invest more and sell more. But if these high-data firms also have high firm-level markups because they skew the composition of their goods toward high-markup goods, then high-markup firms are also larger firms. This effect is stronger in recessions where demand is most uncertain.

These findings rationalize a curious feature of markup measurement that has been at the heart of a debate. From one perspective, markets are just as competitive today as in the past because good-level markups are stable (see Anderson, Rebelo, and Wong [2018]). Instead, growing firm-level and industry markups and increased industry concentration are evidence of declining competition (see Philippon [2019]; Furman and Orszag [2015]; Grullon, Larkin, and Michaely [2016]; De Loecker, Eeckhout, and Unger [2020]; Hall [2018]). Section IV shows that the level of aggregation also changes the measured cyclicality of markups, in the way the model predicts. The model lends an economic interpretation to these facts, beyond the explanation that composition effects are at work. The differences no longer represent an aggregation problem, but rather a useful measure of firms' data.

The static model takes data to be exogenous. That assumption matters. When data becomes a by-product of economic activity in Section VI, a new effect on markups emerges. To do more transactions, a firm must lower its price and thus its markup. The solution reveals that price and data enter as substitutes. Firms are effectively paid either with money or with their customers' data. The idea that price does not fully capture the value of a transaction to a firm provides one more reason that markups fail to capture market power in a data economy.

Section V calibrates the model using data on firms' markups and their prediction of future sales. Through the lens of our model, we interpret firms that predict future sales better as firms with more data. We use the quantification to ask questions, such as: What is the difference across firms in their access to data? How large might data's effect on product markups be?

Section VII presents a linear rotation of our model that allows for realistic cross-product elasticities. In this setting, we re-interpret the model to be about firms that use data for product innovation. This version of the model connects with the literature on rising niche consumption.

This is the beginning of an agenda, not its conclusion. These findings do not prove that data is responsible for markup trends. Instead, they suggestively match a variety of facts and point us in a new measurement direction: to explore the wedge between product markups and firm markups, in order to obtain firm-level measures of data.

RELATED EMPIRICAL EVIDENCE Our work obviously speaks to the large literature on markup measurement and complements it by providing new interpretations of results about trends and fluctuations in markups. With growing stocks of data, our model speaks to the following facts. 1) Industry measures of markups are larger than firm-level markup measures. 2) Sales-weighted markups exceed cost-weighted markups, on average De Loecker, Eeckhout, and Unger (2020). 3)

The gap between cost-weighted and sales-weighted markups is growing over time De Loecker, Eeckhout, and Unger (2020). 4) Firm markups are more counter-cyclical than product markups Burstein, Carvalho, and Grassi (2020). (See contrasting results in Nekarda and Ramey (2020); Bils (1985, 1987) for more evidence of counter-cyclical markup wedges.) 5) Firms with high risk premia have higher markups, on average (Corhay, Kung, and Schmid 2020).

While it is possible to craft an alternative explanation for each one of these facts, a model simply based on the premise that firms use data to preduct uncertain outcomes and shift resources accordingly predicts all five of these outcomes.

Additional non-markup facts point to the specific role of data. Galdon-Sanchez, Gil, and Uriz-Uharte (2023) report on an experiment by a bank that provides small and medium size enterprises with information about sales in their industry. They show that the firms that access the information increase their revenue between 4.5% and 9%, and show that data allows the firms to identify the more profitable sales opportunities from among the existing ones and exploit them. Jaimovich, Rebelo, and Wong (2019) finds similar re-allocation between goods, in response to the business cycle, which constitutes a realization of important macro data. In a randomized control trial, Kumar, Gorodnichenko, and Coibion (2023) provide some firm managers with predictive data and show that these firms invest and produce more, as in the model. Finally, Kwon, Ma, and Zimmermann (2022) argue that the timing and degree of rising concentration in an industry correlate closely with the industry's investment in information technology.

Although the assumption that firms respond to risk is unusual in the firm competition literature, it is a bedrock principle of the field of corporate finance. Brealey, Myers, and Allen (2003); Eckbo (2008) support the idea that this effect is worth considering as one of many possible links between firm data and markups.

RELATED THEORIES The technical contribution of the model is to extend the tools in Lambert, Ostrovsky, and Panov (2018) for strategic play in informationally complex environments to the case of endogenous payoffs. The payoff to firm production is the price, which depends on what others firms do. Relative to Cournot-based frameworks like Pellegrino (2023), we add uncertainty and data. These additions are not trivial and are essential to understand the role of data in firm competition.

Because we model data as digitized information, our approach is most similar to those in the information frictions literature in macroeconomics. Work by Lorenzoni (2009), Angeletos and La'O

(2013), Asriyan, Laeven, and Martin (2022), David and Venkateswaran (2019), Nimark (2014) and Maćkowiak and Wiederholt (2009) feature similar information frictions, used to explain features of business cycles. Work by Rostek and Weretka (2012) explore a reverse question: the effect of market size and market power on price informativeness. Similar tools are used in models of banking competition as well (Vives and Ye 2021), where banks use information for forecasting and pricing risk. However, banks differ from firms: while goods-producing firms choose freely how many units of a good to produce, lenders typically cannot lend twice the requested amount to a promising borrower. The ability to scale production is central to market power. Finally, firms use data to forecast price, which depends on others' actions. Capturing this strategic use of data requires new solution tools to layer a forecasting-the-forecasts-of-others problem, on top of a imperfect information portfolio problem.

In the data economy literature, Jones and Tonetti (2020) explore what data ownership rules facilitates economic growth. In Kirpalani and Philippon (2020), data enables directed two-sided search. Acemoglu et al. (2022) and Bergemann and Bonatti (2019) model data as information and explore whether data markets are efficient. Ichihashi (2020) shows how firms can use consumer data to price discriminate, while Liang and Madsen (2021) explore the use of data in labor markets. In De Ridder (2021) information technology raises fixed costs and reduces marginal costs. We do not dispute that data can be used for all of these purposes. However, we introduce uncertainty, aggregation effects and data barter and show how these affect markup levels, trends and cyclicality.

The recent work by Burstein, Carvalho, and Grassi (2020) combines markup theory and data. In addition to documenting how the sign of markup cyclicality varies with aggregation, their model shows how cyclical firm weights can explain this pattern. Our results complement these insights by proposing a specific mechanism that causes markups and weights to covary.

Some new papers model the mechanisms that give rise to trending markups (see for example De Loecker, Eeckhout, and Mongey [2021]). Those models and Edmond, Midrigan, and Xu (2019) evaluate the welfare consequences of markups. Our approach differs because we focus on firms' use of data.

Related work by Kohlhas and Asriyan (2025) analyzes the impact of data on market power through the reduction of uncertainty. Theirs is a model of monopolistic competition as in Dixit and Stiglitz (1977) without strategic interaction. The firm faces a concave revenue function because of financial frictions: firms choose quantities before the realization of shocks, and due to Jensen's inequality, they produce less than under full information. Firms thus also price risk as in our setting, but they do it because of financial frictions. The paper's focus is on price discrimination by monopolistic firms as a source of inefficiency, and how the inefficiency is affected by the use of data. In our setting, the key feature is the interaction between data and strategic interaction between oligopolistic firms rather than monopoly with price discrimination. Their work inspired us to use forecast data to measure the impact of data on market power.

I Model

To explore the how data interacts with measures of market power, we build a model with multiple possible interactions. This model is not the simplest explanation for any given set of facts. Rather it is like a checklist of possible interactions to look for, meant to guide a measurement agenda to determine which are more relevant. The key assumptions are as follows. First, firms face uncertainty about consumer demand. It is not essential that uncertainty is about demand, rather than advertising, hiring, product placement or costs. We simply need a variable that is profit-relevant and uncertain. Second, data is used to resolve this uncertainty. Data is used to predict the profitability of various actions. Third, firms face a cost of bearing risk. This price of risk is what governs the magnitude of the link between data, uncertainty, and investment. Fourth, to explore the relationship between data and the composition of the goods a firm produces, we model firms that choose quantities of multiple goods. Finally, since the data competition hypothesis is about high-data firms growing large, we allow firms to choose an initial investment, which reduces their marginal cost of production. While this final effect on markups is not surprising, we would be remiss to discuss the relationship between data and markups and not include it.

We first explore these features in a static model. Since our question is about what effects data has on competition measures, we take data to be exogenous and move it around in the model to observe its effect. Later, Section VI introduces dynamics and endogenizes data as a by-product of economic activity and something that can be purchased or sold. The forces we describe here will survive in that dynamic setting.

I.A Setup

FIRMS There are n_F firms, indexed by i: $i \in \{1, 2, \ldots, n_F\}$. The product space has N goods, indexed by j. Firm production profit π_i depends on quantities of each good, which are entries in the $N \times 1$ vector $\mathbf{q}_i$, the market price of each good, $\mathbf{p}$ of dimension ($N \times 1$), and the marginal cost of production, $\mathbf{c}_i$ (also $N \times 1$):

$$\pi_i = \mathbf{q}_i' \left(\mathbf{p} - \mathbf{c}_i\right). \tag{1}$$

Each firm chooses the number of units of each good they want to produce, an $N \times 1$ vector $\mathbf{q}_i$, to maximize risk-adjusted profit, where the price of risk is ρ_i.

$$U_i = \mathbf{E}\left[\pi_i|\mathcal{I}_i\right] - \frac{\rho_i}{2}\mathbf{Var}\left[\pi_i|\mathcal{I}_i\right] - g(\chi_c, \mathbf{c}_i). \tag{2}$$

The last term in (2) is each firm's up-front investment. Let $\mathbf{c}_i$ be the ($N \times 1$) vector of marginal production costs for a unit of each good. The up-front investment choice is modeled as a choice of $\mathbf{c}_i$ at an investment cost $g(\chi_c, \mathbf{c}_i)$ to maximize $E[U_i]$. Assume that $g(\chi_c, \mathbf{c}_i)$ is additively separable and strictly decreasing in each entry of the vector $\mathbf{c}_i$. Since lower choices of $\mathbf{c}_i$ require a greater up-front investment, we interpret this as choosing a larger firm. Since we want to interpret χ_c as a parameter that governs the marginal cost of investment, we impose $\partial^2 g/\partial\chi_c\partial c_i(j) < 0$, for each entry j of the vector $\mathbf{c}_i$. To guarantee non-negative interior marginal cost choices, $g\left(\chi_c, \mathbf{c}_i\right)$ is convex over $\mathbf{c}_i$, with $g\left(\chi_c, \overline{p}\right) = 0$, where $\bar{p}$ is the highest possible price, and $\lim_{\mathbf{c}\to 0} g\left(\chi_c, \mathbf{c}\right) = +\infty$.

PRICE Our demand system is Cournot with exogenous shocks. However, we later show how this structure can be mapped into a linear hedonic demand structure, which embodies the idea that goods with similar attributes are partial substitutes for each other.

For now, each good j has an average market price that depends on an good-specific constant and on the total quantity of that good that all firms produce:

$$p_j^M = \bar{p}_j - \frac{1}{\phi}\sum_{i=1}^{n_F} q_{ij}. \tag{3}$$

This demand system has the advantage that it holds the power to affect prices dp^M/dq_i fixed. The assumption is surely not true. But it ensures that when markups change, it is from effects other than the power to affect price.

Each firm does not receive the market price, but rather faces an uncertain price that depends on a demand shock $\mathbf{b}_i$. The demand shock $\mathbf{b}_i$ is a vector with jth element b_{ij}. This vector is random and unknown to the firm: $\mathbf{b}_i \sim N(0, I)$.[1] Demand shocks can covary across firms: $\xi =$

[1] The fact that the variance matrix is diagonal is without loss. We show later how to embed this model in a linear hedonic demand structure, where goods are combination of attributes. Those attributes are then defined to be an

$\mathbf{Cov}(b_{ij}, b_{lj})\ \forall j\ \forall i \neq l$. The price a firm receives for a unit of good j is thus $p_j + b_{ij}$. We can express firm i's price in vector form as

$$\mathbf{p}_i = \left[p_1^M, p_2^M, \ldots, p_N^M\right]' + \mathbf{b}_i. \tag{4}$$

In the case where the shocks b are identical across firms, then this is a uniform price across firms, still with price differences across goods.

INFORMATION Each firm generates n_{di} data points. Each data point is a signal about the demands for each good: $\mathbf{s}_{i,z} = \mathbf{b}_i + \boldsymbol{\varepsilon}_{i,z}$, where $\boldsymbol{\varepsilon}_{i,z} \sim N(\mathbf{0}, \mathbf{I})$ is an $N \times 1$ vector. Signal noises are uncorrelated across goods and across firms.[2]

Because we are interested in how data affects competition, we take data n_{di} as given and exogenously change the amount of firms' data, n_{di}. Section VI explores what aspects of the results change when data is endogenously generated as a by-product of economic transactions.

We consider two possible information structures. In one, firms observe only their own data: $\mathcal{I}_i = \{\mathbf{s}_{i,z}\}_{z=1}^{ndi}$. In the other, data is public: all firms can observe every piece of data about every firm ($\mathcal{I}_i = \{\{\mathbf{s}_{i',z}\}_{z=1}^{ndi}\}_{i'=1}^{nF}$). This structure simplifies results and clarifies that the mechanism does not rely on asymmetric information. When more relevant data exists about a firm i, similar results arise, regardless of who else observes that data.

EQUILIBRIUM

1. Each firm chooses a vector of marginal costs $\mathbf{c}_i$, taking as given other firms' cost choices. Since the data realizations are unknown in this ex ante investment stage, the objective is the unconditional expectation of the utility in (2).

2. After observing the realized data, each firm updates beliefs with Bayes' law and then chooses the vector $\mathbf{q}_i$ of quantities to maximize conditional expected utility in (2), taking as given other firms' best responses.

3. Prices are given by (3) and (4).

orthogonal decomposition of the demand variance-covariance space. Not only can we create covariance. The variances can also change. We investigate the effect of changing $Var(b_i)$ in Section IV.

[2]For now, we assume each firm has n_{di} data points about every gods. Section VI relaxes this and allows data to differ by goods and their attributes. The assumption of signal variance $\mathbf{I}$ is then without loss. By Bayes' law, two independent data points function in the same way as one data point with twice the precision. By assuming $\mathbf{Var}(\boldsymbol{\varepsilon}_{i,z}) = \mathbf{I}$, we are simply letting n_{di} govern data precision.

I.B Discussion of Assumptions

FIRMS THAT PRICE RISK. Risk pricing means that firms do less of activities that are perceived as risky. A price of risk could represent a convex cost, firm bankruptcy risk, the risk-aversion of a manager with equity compensation or a firm's cost of capital. Risk pricing is a bedrock principle of corporate finance and is well supported by numerous empirical studies.[3] Under the capital cost interpretation, our ρ_i term in (2) captures both the price of risk and the covariance of the firm shock with market risk (the firm's beta).

Our firms may also price firm-specific risk because we are exploring a market with large players where firm-specific risk is not diversifiable. There is growing evidence that even idiosyncratic risk is priced, especially when firms face financial constraints (Hennesy and Whited 2007).

Pricing risk is not essential for the main aggregation results. The results cover the case where risk is not priced or even incentivizes more production ($\rho \leq 0$ as in the Oi-Hartman-Abel effect). However, when firms face uncertain profits, evidence suggests we should take the effect of data on risk into account. Boar, Gorea, and Midrigan (2022) find that most of the variation in private business markups is compensation for risk. In a randomized control trial where firms are treated with predictive data, Kumar, Gorodnichenko, and Coibion (2023) show that these firms invest more and produce more.

DATA ABOUT CONSUMER DEMAND. One might question whether data is used to forecast demand or marginal cost. Conceptually, it shouldn't matter. If data helps firms reduce profit risk, whether from the cost or the revenue side, it should embolden them to invest more and produce more at a lower market price. The same forces operate. Why then choose to model demand uncertainty? Markups are price divided by marginal cost. Having the random variable in the denominator makes it nearly impossible to characterize the average value of markups. If one wanted to study inverse markups, then it would be more practical to model cost uncertainty.

[3]The Handbook of Empirical Corporate Finance fills chapter 18 with the evidence that firms price risk (Eckbo 2008). Most of the other chapters contain evidence in support of theories that are premised on risk pricing. For a textbook treatment of the topic, see Welch (2009), chapter 9. More recent work on this topic explores whether male and female CFOs are equally risk averse (Doan and Iskandar-Datta 2020). Management and psychology scholars (Lovallo et al. 2020) find that firms place too high a price on risk. In international economics, David, Ranciere, and Zeke (2022) document that multinational firms facing more risk hire less and compensate their capital owners with a greater share of income. Specifically, Brealey, Myers, and Allen (2003) argues for a price of risk ρ that matches the risk premium on the S&P 500. If a firm gets less return per unit of risk than this, the firm would be better off not investing in production and instead returning the cash to investors to invest in a market portfolio of equity.

COMPETITION IN QUANTITIES. We model conduct by firms as competition in quantities because it holds market power ($\partial p/\partial q_i$) fixed. That assumption is surely wrong. But it holds fixed the non-data sources of markups, in order to see the effect of data more clearly. Section V shows how this simple Cournot-like setup maps into Pellegrino (2023)'s Generalized Hedonic-Linear demand system, used to study market power in a network economy with realistic cross-product elasticities.

Of course, there are other conduct assumptions and frictions, including Bertrand competition, information frictions, transaction cost and barriers to entry. Appendix D.2. explores Bertrand competition. As is well-known, markups are lower under Bertrand than Cournot. Furthermore, while the risk effects may differ, our main results are robust.

NO PRICE DISCRIMINATION. We assume a uniform price for all consumers, at least for a given firm. The consumer demand uncertainty could represent uncertainty about how to best price-discriminate. The solution would be different. However, the idea that data reduces uncertainty and encourages firm growth, as well as the later results about covariances that information makes possible, would all still make sense. Since this agenda is still in its beginning stages, it makes sense to first understand uniform pricing, which corresponds to the behavior of the vast majority of firms. We leave the price discrimination version of the model for future work.

NO VARIABLE CAPITAL COST. We made the investment in technology an up-front fixed cost. That means that the cost of capital is not part of the marginal cost that enters the markup calculation. One might object to that assumption on the grounds that the cost of capital is what captures the price of risk. Including a capital cost with a risk premium in marginal cost arguably absorbs the effect of risk on markups. This objection is tenuous. First, the capital cost is typically a borrowing cost. The risk premium on debt is not the same as the risk premium on equity. The firm cares about the variance of its cash flows, which is an equity claim. Second, the long-horizon risk that lenders care about is not the same as the short-term demand or cost fluctuations that data helps firms to forecast. These are substantially different risks. While including a variable capital cost with a risk premium in markup calculations probably improves their accuracy, this risk compensation has very little interaction with the way in which data helps to reduce operational uncertainties.

EXOGENOUS DATA. Section VI endogenizes data. The static forces are still present in that model and one new force emerges.

No entry or exit. Adding entry would undoubtedly bring new insights. But that would also require a dynamic framework and a different paper. Since the static problem is not well understood, we start there. However, recent work by Baqaee and Farhi (2021) suggests that the aggregate distortions from market power are even larger once there is entry.

I.C General Solution

We solve the model by backwards induction, starting with the quantity choices and then working backwards to determine optimal firm investments in lowering marginal costs $\mathbf{c}_i$.

Optimal Production The first-order condition with respect to goods production $\mathbf{q}_i$ is $\partial U_i / \partial \mathbf{q}_i$: $\mathbf{E}[\mathbf{p}_i|\mathcal{I}_i] - \mathbf{c}_i + \frac{\partial \mathbf{E}[\mathbf{p}_i|\mathcal{I}_i]}{\partial \mathbf{q}_i}\mathbf{q}_i - \rho_i \mathbf{Var}[\mathbf{p}_i|\mathcal{I}_i]\,\mathbf{q}_i = 0$. Rearranging delivers optimal production:

$$\mathbf{q}_i = \left(\rho_i \mathbf{Var}[\mathbf{p}_i|\mathcal{I}_i] - \frac{\partial \mathbf{E}[\mathbf{p}_i|\mathcal{I}_i]}{\partial \mathbf{q}_i}\right)^{-1} (\mathbf{E}[\mathbf{p}_i|\mathcal{I}_i] - \mathbf{c}_i). \tag{5}$$

The second term tells us that firms produce more of goods that have high expected prices, relative to their marginal costs. The first term tells us that uncertainty (conditional variance) or market power cause the firm to scale back their production response to changes in expected profit. By improving forecasts, data reduces uncertainty.

A key reason one should think about priced risk in this context is that risk mimics market power. Because market power enters only through the denominator in (5), more market power is mathematically equivalent to increasing the conditional variance $\mathbf{Var}[\mathbf{p}_i|\mathcal{I}_i]$. Both risk and market power restrain production. Both make firms less sensitive to expected changes in price or cost. In one case, it is because a risk-averse firm makes more conservative production decisions to manage its risk. In the other case, the firm makes more conservative decisions to minimize its price impact. This is one reason that data and market power are difficult to disentangle.

From differentiating the pricing function (3), we find that the price impact of one additional unit of output is[4]

$$\frac{\partial \mathbf{E}[\mathbf{p}_i|\mathcal{I}_i]}{\partial \mathbf{q}_i} = -\frac{1}{\phi}\mathbf{I}_N. \tag{6}$$

[4] In a typical setting with supply-function competition, such as Rostek and Weretka (2012), the price impact of a change in quantity depends on the amount of data. Here, price impact depends only on the parameter ϕ. The difference arises because in the standard setting, agents learn from prices, before choosing quantities. Price information gives rise to the relationship between data and price impact. In this setting, firms choose quantities, before they know what the price will be.

I.D Solution with Public Data and Firm-Specific Shocks

While not as realistic, the special case of firm-specific shocks ($\xi = 0$) and public data allows us to more clearly illustrate the model's mechanics and the main economic forces of data.

Bayesian Updating According to Bayes' law for normal variables, observing n_{di} signals, each with signal noise variance Σ_e, is the same as observing the average signal $\mathbf{s}_i = (1/n_{di}) \sum_{z=1}^{n_{di}} \mathbf{s}_{iz} = \mathbf{b}_i + \boldsymbol{\varepsilon}_i$, where the variance of $\boldsymbol{\varepsilon}_i$ is $\Sigma_{\epsilon_i} = 1/n_{di} \mathbf{I}_N$.

Define $\mathbf{K}_i$ to be the sensitivity of price beliefs to the signal s_i:[5] $\mathbf{K}_i := \mathbf{I}_N \cdot n_{di} / (1 + n_{di})$. Then, firm i's expected value of the shock $\mathbf{b}_i$ can be expressed simply as $\mathbf{E}[\mathbf{b}_i|\mathcal{I}_i] = \mathbf{K}_i \mathbf{s}_i$. The expectation and variance of the pricing function (3) are

$$\begin{aligned} \mathbf{E}\left[\mathbf{p}_i|\mathcal{I}_i\right] &= \bar{\mathbf{p}} + \mathbf{K}_i \mathbf{s}_i - \frac{1}{\phi} \sum_{i'=1}^{n_F} \mathbf{q}_{i'}, \\ \mathbf{Var}\left[\mathbf{p}_i|\mathcal{I}_i\right] &= \mathbf{Var}\left[\mathbf{b}_i|\mathcal{I}_i\right] = \frac{1}{1+n_{di}} \mathbf{I}_N \end{aligned} \tag{7}$$

Optimal Production Define the sensitivity of production to a change in expected profit as

$$\hat{\mathbf{H}}_i \equiv \left(\rho_i \mathbf{Var}\left[\mathbf{p}_i|\mathcal{I}_i\right] + \frac{\mathbf{I}_N}{\phi} \right)^{-1}. \tag{8}$$

If we rewrite (5), replacing the price impact (6) and conditional expectation (7) we get firm i's output, in terms of other firms' output, i's costs and i's data $\mathbf{s}_i$: $\mathbf{q}_i = \hat{\mathbf{H}}_i \left(\overline{\mathbf{p}} + \mathbf{K}_i \mathbf{s}_i - \frac{1}{\phi} \sum_{i'} \mathbf{q}_{i'} - \mathbf{c}_i \right)$.

Next, sum production $\mathbf{q}_i$ over all firms i to get total production of each good $\sum_{i'} \mathbf{q}_{i'}$. This sum has a $\sum_{i'} \mathbf{q}_{i'}$ on both the left- and right-hand sides. Collect these terms and rearrange to get $\sum_{i'} \mathbf{q}_{i'} = \left(\mathbf{I} + \frac{1}{\phi} \sum_i \hat{\mathbf{H}}_i \right)^{-1} \left[\sum_i \hat{\mathbf{H}}_i \left(\overline{\mathbf{p}} + \mathbf{K}_i \mathbf{s}_i - \mathbf{c}_i \right) \right]$. Substituting this total production expression for $\sum_{i'=1}^{n_F} \mathbf{q}_{i'}$ in firm i's optimal production ($\mathbf{q}_i{}^*$) yields the optimal production of each good by each firm i.[6]

[5] In a dynamic model, K_i would be called the Kalman gain.

[6] Since all signals are normally distributed, this formula does tell us that production can potentially be negative. We could bound choices to be non-negative, but this would make analytical solutions for covariances impossible. If parameters are such that all firms want negative production of a good, then the solution is simply to redefine the product as its opposite. In the numerical results, we simply choose parameters that make negative production extremely unlikely.

Equilibrium Price Substituting this aggregate quantity in the pricing function (3) yields an equilibrium average price of each good:

$$\mathbf{p}^M = \overline{\mathbf{p}} - \left(\phi \mathbf{I}_N + \sum_i \hat{\mathbf{H}}_i\right)^{-1} \left[\sum_i \hat{\mathbf{H}}_i \left(\overline{\mathbf{p}} + \mathbf{K}_i \mathbf{s}_i - \mathbf{c}_i\right)\right]. \tag{9}$$

Optimal Investment Choices Firm i chooses cost $\mathbf{c}_i$ to maximize its unconditional expected utility $\mathbf{E}[U_i]$, taking all other firms' investment choices as given.

The optimal cost $\mathbf{c}_i$ for an interior solution satisfies (see Appendix A. for derivation):

$$\frac{\partial \mathbf{E}[U_i]}{\partial \mathbf{c}_i} = \left(\frac{\phi \hat{\mathbf{H}}_i + 1}{2\phi}\right) \frac{\partial \mathbf{E}[\mathbf{q}_i]' \mathbf{E}[\mathbf{q}_i]}{\partial \mathbf{c}_i} - \frac{\partial g(\chi_c, \mathbf{c}_i)}{\partial \mathbf{c}_i} = 0, \tag{10}$$

The first term is the marginal benefit. Lower production costs enable production at a greater scale and higher profit per unit. The second term is the marginal cost of the up-front investment.

I.E Solution with Private Data and Common Shocks

The optimal production takes the same form as before. The difference is in the expectation $\mathbf{E}\left[\mathbf{p}_i | \mathcal{I}_i\right]$ and the conditional variance.

The solution to this model is complicated by firms' need to forecast what other firms know, as in Angeletos and Pavan (2007). Because firms do not know other firms' data, they face strategic uncertainty. They use their own data to forecast what other firms will do. Data thus reduces risk in two ways—by predicting demand for the firm's products and by predicting the production decisions of other firms. This strengthens the risk channel because data reduces both demand uncertainty and strategic uncertainty.

Lemma 1. *With private data and common shocks, the equilibrium price takes the form*

$$\mathbf{p} = \bar{\mathbf{p}}^M + \mathbf{F}\mathbf{b} - \frac{1}{\phi} \sum_{i=1}^{n_F} \hat{\mathbf{H}}_i \mathbf{K}_i \boldsymbol{\varepsilon}_i \qquad \text{where } \bar{\mathbf{p}}^M,\ \mathbf{F},\ \mathbf{K}_i,\ \text{and } \hat{\mathbf{H}}_i \ \text{are reported in Appendix A.} \tag{11}$$

Solving this problem requires a technical innovation. Bayesian updating weights always depend on the covariance between the observed data s_i and the price the firm needs to forecast p_i. But, in this model, that covariance is endogenous. It depends on firms' production choices. To solve the fixed point of beliefs and output requires a state-space updating approach for informationally complex environments, as in Lambert, Ostrovsky, and Panov (2018). A state-space approach defines all the relevant model objects in terms of exogenous, orthogonal shocks and weights on those shocks. We extend this approach to the case of endogenous weights on each shock. To do that, choose Bayesian weights that maximize firms' objectives, given the beliefs that result.

When data is firm-specific, more data always reduces uncertainty. But when data is aggregate, it is possible for data to increase uncertainty about the price.[7] Since a firm would never choose data that raises uncertainty, going forward, we assume that the risk price of all firms is sufficiently low to ensure that $\partial \hat{\mathbf{H}}_i(j,j)/n_{di}(j) \geq 0 \ \forall j$. The parameter bound also ensures that when firm i acquires more data about good j, it does not make other firms less uncertain about the price of good j: $\partial \hat{\mathbf{H}}_{i'}(j,j)/n_{di}(j) \ \forall j$.

II Data and Product Markups

This section does not contain the main results. Higher or lower markups can be explained by many factors. However, these are a stepping stone to the aggregation results, which are more specific to data. We begin by exploring just two of the ways in which data affects markups, through cost and risk. The goal in this section is simply to establish the foundation of ideas upon which we build later.

By reducing the uncertainty a firm faces about consumer demand, data encourages the firm to produce more for a given level of investment. Reducing uncertainty also emboldens the firm to invest more in infrastructure that enables them to produce at a lower marginal cost. These two forces have opposite effects on markups. More production lowers prices, which in turn lowers markups. More initial investment lowers marginal cost, which raises markups. This section explores that tension.

Definition 1 (Product markup)**.** *The product-level markup for product j produced by firm i is $M^p_{ij} := \mathbf{E}[\mathbf{p}_i(j)]/\mathbf{c}_i(j)$. Firm i's average product-level markup is $\overline{M}^p_i := \sum_j E[\mathbf{q}_i(j)]M^p_{ij}/(\sum_j E[\mathbf{q}_i(j)])$.*

To derive an expression for the product markup in the model, we take the expectation of each

[7] The reason is that if one firm is better informed, their beliefs might be less noise and more predictable to a competitor. Because the competitor can predict his rival's action, the competitor's uncertainty declines, causing the competitor to produce more aggressively, in response to his data. This aggressive response of the competitor to data unknown to the original firm could, in theory, overcome the original decline in uncertainty about the demand shock and cause a rise in uncertainty.

expected product price, using (9) and $\mathbf{E}[\mathbf{s}_i] = \mathbf{0}$, and divide by the marginal cost $\mathbf{c}_i$ of that product:

$$M_i^p = \left(\overline{\mathbf{p}} - \left(\phi \mathbf{I}_N + \sum_{i=1}^{n_F} \hat{\mathbf{H}}_i \right)^{-1} \left(\sum_i \hat{\mathbf{H}}_i \left(\overline{\mathbf{p}} - \mathbf{c}_i \right) \right) \right) ./ \mathbf{c}_i, \tag{12}$$

where ./ denotes element-by-element division of two vectors. The average markup weights the markup on each product by the average amount the firm produces of that product.

Some of these causes of high markups in equation (12) are not surprising. For example, producing goods with high expected value $\bar{p}$, fewer firms (low n_F) or low price elasticity (low ϕ) raise markups. Two forces show up in the markup formula that are affected by how much data a firm has. Those forces are risk and cost. We state and explain each in turn.

DATA, INVESTMENT, OUTPUT, AND MARKUPS The first two results encapsulate the standard logic about data and competition: Data enables firms to grow larger (invest more). These larger firms charge higher markups.

Lemma 2. ***Data-investment complementarity.*** *A firm with more data chooses a lower marginal cost $\mathbf{c}_i$, which entails a higher cost investment.*

The proofs of this and all further results are in Appendix B.. The role of investment in data is to reduce the conditional variance of the firm's stochastic demand, which encourages the firm to produce more. Data increases the expected revenue of a firm by allowing it to produce more in states in which the price will be high. It also reduces the uncertainty around that investment and lowers the risk of the firm. Both of these effects increase the marginal benefit of production and the marginal benefit of investment. What this means is that high-data firms invest more and grow larger. As the next result shows, higher investment is also a channel through which data increases product markups.

Lemma 3. ***Higher investment raises product markups.*** *More investment (lower $\mathbf{c}_i$ choice) in any good j raises firm i's markup on good j.*

A firm that invests in producing a good can produce that good at a lower cost. Since markups are price divided by marginal cost, a lower cost raises the markup. Of course, a lower cost also lowers the equilibrium price of the good. However, the proof shows that price does not fall as much as cost. Therefore, the markup rises.

The model teaches us that there is a second channel through which data affects markups: data reduces the risk of production, induces more production, and thereby lowers prices and markups.

Lemma 4. ***Data reduces product markups (risk premium channel).*** *Holding all firms' investments fixed (χ_c sufficiently high), an increase in any firm's data about good j reduces their markup on good j.*

Data reduces markups because it reduces the risk in production. This induces firms to produce more. This effect can be seen in the firm's first-order condition (5) where the conditional variance in the denominator represents risk. When this variance declines, optimal production rises. More production lowers price and lowers markups. This force shows up in the markup (12) as high ρ_i makes $\hat{\mathbf{H}}_i$ low. When we reduce risk with data, firms do not need as much markup compensation to be willing to produce.

The restriction on χ_c is there to shut down then investment channel, to isolate the risk effect. When χ_c or ρ is low, this risk premium channel is still present. But it may be overpowered by the investment channel working in the opposite direction.

Proposition 1. ***Data in(de)creases product markups when risk price or marginal cost of investment is sufficiently low (high).*** *If the price of risk ρ is sufficiently low or the investment cost χ_c is sufficiently low, then an increase in any firm's data about good j increases the average markup on good j. Otherwise, an increase in any firm's data about good j reduces the markup of good j.*

When firm investments greatly decrease marginal cost (low χ_c), then the cost channel is dominant and more data primarily increases investment, lowers costs, and raises markups. When the cost-reduction investment is inefficient (high χ_c), then data still prompts more investment, but this has little effect on marginal cost. Instead, the dominant force is risk reduction. Similarly, if the price of risk is high, risk reduction is also the dominant force. A data-rich firm faces less cost from taking on more risk with a large production plan. By producing more, data-rich firms drive prices down and lower markups. Which scenario prevails depends on the strength of each force in a particular industry.

The main point is not whether data increases or decreases the markup. It is that, despite holding market power fixed, markups are contaminated by firms' use of data. To solve this problem, we need to know how much data firms have.

III Measuring Markups and Measuring Data

In empirical work, markups are often measured at the firm or industry level. Measuring markups at these more aggregated levels often yields different answers about how competition is evolving. The next set of results show why aggregate markups differ from product-level markups in ways that vary systematically with the amount of data firms have. The difference between a firm's product- and firm-level markups turns out to be a good bound for the amount or quality of data that a firm must have.

These composition effects are quantitatively important. De Loecker, Eeckhout, and Unger (2020) find that two-thirds of the increase in measured industry markups comes from such composition effects. Crouzet and Eberly (2018) link the trend increase in markups to intangible assets, a broader category that includes data assets. They find that intangible-abundant firms have higher markups and that intangible-abundant industries have even higher markups. The results that follow contribute to this discussion by explaining why firms' use of predictive data can generate such statistical patterns and providing tools to measure firms' data.

III.A Firm Markups

Economists have long known that difference in markup measurement at different levels of aggregation represent composition effects. What is less well understood is why such composition effects might change. We show how firms' data accumulation naturally gives rise to changes in the composition of firms' products. Data is what makes it possible for the firm to skew the composition of their products in the direction of high-markup goods. So, data strengthens the composition effect and makes firm markups larger and larger relative to that firm's average product markup.

Definition 2 (Firm Markup)**.** *The firm markup for firm i is the firm's revenue divided by the firm's total variable costs:*

$$M_i^f := \frac{\mathbf{E}[\mathbf{q}_i'\mathbf{p}_i]}{\mathbf{E}[\mathbf{q}_i'\mathbf{c}_i]}. \tag{13}$$

We can rewrite the expectation of the product of price and quantity as the product of expectations, plus a covariance term (trace of matrix):

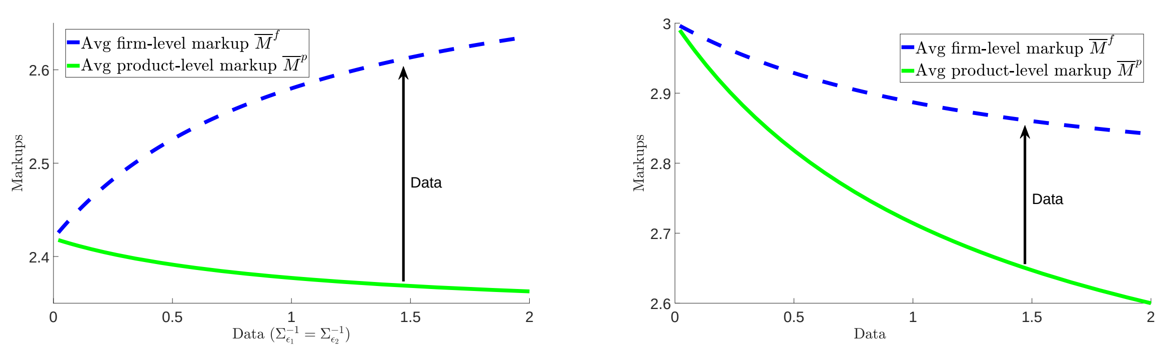
$$M_i^f = \frac{\mathbf{E}[\mathbf{q}_i]'\mathbf{E}[\mathbf{p}_i] + \mathbf{tr}\left[\mathbf{Cov}(\mathbf{p}_i, \mathbf{q}_i)\right]}{\mathbf{E}[\mathbf{q}_i'\mathbf{c}_i]} = \underbrace{\frac{\sum_{j=1}^N M_{ij}^p \mathbf{c}_i(j)\mathbf{E}[\mathbf{q}_i(j)]}{\sum_{j=1}^N \mathbf{c}_i(j)\mathbf{E}[\mathbf{q}_i(j)]}}_{\text{Cost-weighted product markups}} + \frac{\mathbf{tr}[\mathbf{Cov}\left(\mathbf{p}_i, \mathbf{q}_i\right)]}{\sum_{j=1}^N \mathbf{c}_i(j)\mathbf{E}[\mathbf{q}_i(j)]}. \tag{14}$$

The second equality just comes from using the definition of the product markup to substitute: $\mathbf{E}[\mathbf{p}_i] = M_i^{p\prime}\mathbf{c}_i$ and then rewriting the vector products as sums. We learn that the firm markup is a cost-weighted sum of product markups, plus a term that depends on the covariance of prices and quantities. Firm data acts on this covariance term. It allows firms to produce more of goods that turn out to have high demand and thus high price.

Define a firm's average product price to be $\overline{M}_i^p := \sum_j \mathbf{E}[\mathbf{q}_i(j)]M_{ij}^p/(\sum_j \mathbf{E}[\mathbf{q}_i(j)])$. This weights the markup on each product by the average amount the firm produces of that product.

Proposition 2. ***Data accumulation widens the wedge between product and firm markups.*** *Holding all firms' investments fixed ($(\mathbf{c}_1, ..., \mathbf{c}_{n_F})$ given), an increase in firm i's data about any good weakly increases* $\mathbf{E}[M_i^f - \overline{M}_i^p]$.

Firm markups rise when data increases the covariance of firm's production decision $\mathbf{q}_i$ with the price $\mathbf{p}_i$ in (14). Without any data to predict demand, this covariance is low because firms cannot know which markups would be high and which goods to produce more of. The positive effect of data on the price-quantity covariance shows up in the production first-order condition (5), where a reduction in the conditional variance of demand makes production decisions $\mathbf{q}_i$ more sensitive to expected changes in price $\mathbf{p}_i$. That higher sensitivity is a higher covariance.



(a) Low price of risk.

(b) High price of risk.

Figure 1: More data may raise or lower markups but always causes product and firm markups to diverge. Parameters used: $\bar{p} = 5$, $\phi = 0.1$. Firm marginal costs are not chosen here. They are fixed as $c_1 = c_2 = 1$. On the left, $\rho_1 = \rho_2 = 1$. On the right, $\rho_1 = \rho_2 = 10$.

AN ILLUSTRATIVE EXAMPLE OF THE PRODUCT-FIRM MARKUP WEDGE To illustrate the mechanisms at work, Figure 1 plots the competing effects data has on product and firm/industry markups in a specific example. When the price of risk is high, the product-level markup falls as both firms' data rises. The reason the product markup is falling is that data is resolving risk.

It is allowing the firms to be less uncertain because data allows them to forecast demand more precisely. Firms that are less uncertain require a lower markup to compensate them for the lower risk. When the price of risk is low, more data may result in higher firm markups, as high-data firms invest, grow, and lower their marginal costs.

What the model teaches us so far is that increases or decreases in markups, at either the product level or the firm level, are not indicative of a firm that has a larger stock of data. As a firm accumulates more data, both product and firm markups may increase, both may decrease, or they may move in opposite directions. Instead, data changes the composition of products and firms and makes firm and product markups diverge. This composition effect is a theme that will recur as we proceed to explore markups at the industry level.

III.B Action-Payoff Covariance: A More General Measure of Data

The gap between average product markups and firm markups arises because of a covariance between a firm's action (the quantity they choose) and a payoff (the price they earn per unit). Action-payoff covariances are indicative of data, across a wide variety of settings. The reason is that an agent cannot systematically take actions that covary with a payoff, unless they can predict that payoff sufficiently well. Since agents choose actions, they have to have at least as much information as what the action choice itself contains. The result below formalizes this logic for normal variables. Then we describe how to extend the logic beyond the class of normals.

Consider data that conveys information about X, where $X \sim \mathcal{N}(\mu_X, \sigma_X^2)$ is a choice-relevant variable, not directly observable by a firm. Without loss of generality, we can call one data point however much information conveys one unit of precision about X. If the noise in data is normal, we can express one data point as $s = X + \epsilon$, where $\epsilon \sim N(0, 1)$ is independent across data points. Let the firm's chosen action be given by Y, where Y can depend on their belief about the state $\hat{X}$. If the action is increasing in the belief about the state, then define $g(\hat{X}) = Y$. If the action is decreasing in the state, then define $g(\hat{X}) = -Y$. The g function encapsulates the model: the agent's preferences, payoffs and any constraints the agent may face.

Proposition 3. ***Action-payoff covariance requires data.*** *Suppose g is a differentiable, weakly increasing function (almost everywhere) and strictly increasing on a subset with positive measure. If the payoff-action covariance is $Cov(X, g(\hat{X})) = z$, then the firm must have $\psi(z)$ data points, where ψ is a strictly increasing function of the covariance z.*

The magnitude of the mapping between covariance and data points depends on the prior variance σ^2 and the economic model, captured by the g function. But what is general is that there is no model g and no prior σ^2 that can generate covariance, in the absence of data (if $n_d = 0$).

For non-normal variables, covariance and number of data points do not correctly measure information content. Instead, mutual information is a measure of information that generalizes to any distribution. Given data on actions and payoffs, the mutual information $MI(Y, X)$ is measurable. It is the minimum amount of data, including prior information, that the firm must have about X. The proof is that the firm must know action Y, in order to choose it. Knowing only the action they chose, the firm already has $MI(Y, X)$ amount of information about X. It is possible they know more. But this establishes a lower bound on the amount of data they must have.

III.C Industry Markup Definitions

Typically, researchers are interested in the markup for an industry because the regulatory question of interest is whether that industry is a competitive one or not. However, there are multiple ways to aggregate markups into a single industry measure. We examine four of the most common measures inside the model. The model lends an interpretation to the differing trends arising from the different ways empirical researchers measure industry markups.

Definition 3. **a.** *The equally-weighted average firm markup in an industry is*

$$\bar{M}^f := (1/N) \sum_{i=1}^{n_F} M_i^f. \tag{15}$$

b. *The cost-weighted markup for an industry is*

$$M^c := \sum_{i=1}^{n_F} w_i^c M_i^f \quad \textit{where cost weights are} \quad w_i^c = \frac{\mathbf{E}\left[\mathbf{q}_i' \mathbf{c}_i\right]}{\sum_{i=1}^{n_F} \mathbf{E}\left[\mathbf{q}_i' \mathbf{c}_i\right]}. \tag{16}$$

c. *The sales-weighted markup is*

$$M^s := \sum_{i=1}^{n_F} w_i^s M_i^f \quad \textit{where sales weights are} \quad w_i^s = \frac{\mathbf{E}\left[\mathbf{q}_i' \mathbf{p}_i\right]}{\sum_{i=1}^{n_F} \mathbf{E}\left[\mathbf{q}_i' \mathbf{p}_i\right]}. \tag{17}$$

d. *The industry- aggregates markup is total industry sales over the total industry variable cost:*

$$M^{ind} := \frac{\mathbf{E}\left[\sum_{i=1}^{n_F} \mathbf{q}_i' \mathbf{p}_i\right]}{\mathbf{E}\left[\sum_{i=1}^{n_F} \mathbf{q}_i' \mathbf{c}_i\right]}. \tag{18}$$

III.D Data's Effect on Industry Markup Measures

Industry aggregation effects from the covariance of data and firm size. Data-investment complementarity means that more data makes larger up-front investment (larger firms) optimal. So, high-data, large firms are weighted more, relative to the unweighted firm average. High-data firms use data to skew their production toward high-markup goods (Proposition 2). Thus, the measures that weight large, high-data firms more will also weight high-markup firms more, generating a higher predicted industry markup.

Proposition 4. ***Growing data increases the wedges between industry markup measures.*** *An increase in firm i's data about any good they produce*

a. increases the difference between cost-weighted and unweighted firm markups $\mathbf{E}[M^c - \overline{M}^f]$*, if the price of risk is not too high* ($\rho_i < \bar{\rho}$)*;*

If, in addition, all firms are initially symmetric, then an increase in firm i's data about any good

b. increases the difference between sales-weighted and cost-weighted markups $\mathbf{E}[M^s - M^c]$*, and*

c. increases the difference between the sales-weighted and industry-aggregates markup $\mathbf{E}[M^s - M^{ind}]$.

Mathematically, the key to each of these results is a covariance. In the first case (a), the covariance is between the firm markup and the total production of a firm. If the risk channel is not so strong that it overpowers both the cost channel and the firm markup aggregation effect, then high-data firms are high-markup firms and these firms get weighted more than small firms by the cost weights.

Economically, this effect arises because data has economies of scale. Firms get the most value from their data if they grow large. The way they get value from data is to use the data to forecast which goods are high-margin and produce more of them. Thus, more data increases the covariance between size and markups and makes the aggregate markup larger than the average firm markup.

In cases (b) and (c), the key covariance is between a firm's markup and the firm's revenue, relative to its costs. High-data firms are firms that are able to produce more of the products that have high price relative to their cost of production. Therefore, these high-data firms have higher sales-weighted markups relative to their cost-weighted markups.

Part (c) follows from (b) because industry-aggregates markups are identical to cost-weighted markups:

$$M^c := \sum_{i=1}^{N} \frac{\mathbf{E}\left[\mathbf{q}_i'\mathbf{c}_i\right]}{\sum_{i=1}^{N}\mathbf{E}\left[\mathbf{q}_i'\mathbf{c}_i\right]} M_I^f = \sum_{i=1}^{N} \frac{\mathbf{E}\left[\mathbf{q}_i'\mathbf{c}_i\right]}{\sum_{i=1}^{N}\mathbf{E}\left[\mathbf{q}_i'\mathbf{c}_i\right]} \frac{\mathbf{E}\left[\mathbf{q}_i'\mathbf{p}_i\right]}{\mathbf{E}\left[\mathbf{q}_i'\mathbf{c}_i\right]} = \frac{\mathbf{E}\left[\sum_{i=1}^{N}\mathbf{q}_i'\mathbf{p}_i\right]}{\mathbf{E}\left[\sum_{i=1}^{N}\mathbf{q}_i'\mathbf{c}_i\right]} := M^{ind}. \tag{19}$$

Identical holds in theory. In practice, with different sources of measurement error at the firm and aggregate level, they differ somewhat, but have similar trends.

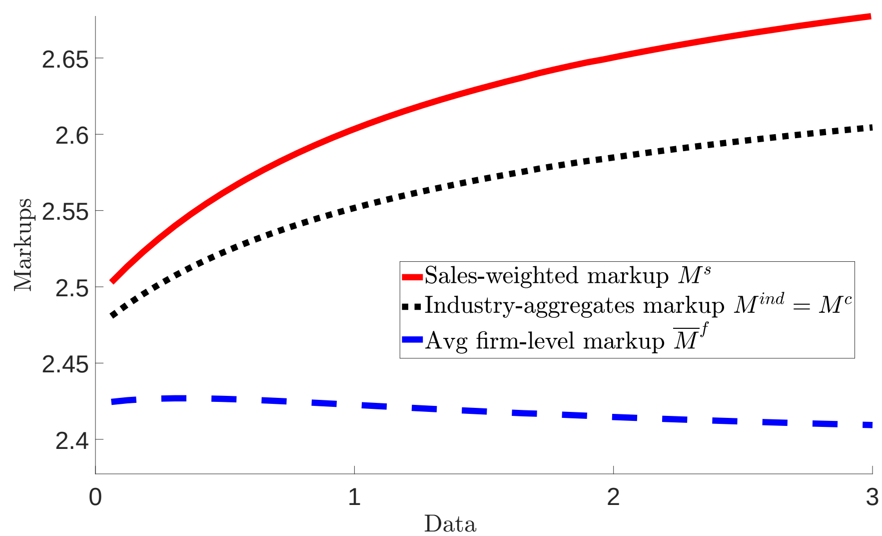


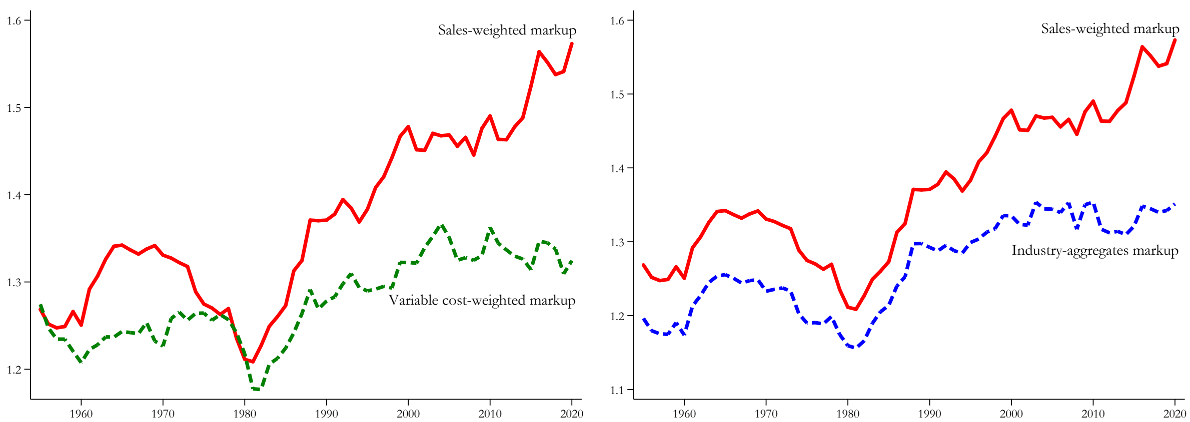
Figure 2: Data Accumulation Makes Industry Markup Measures Diverge. Investment cost function is $g(\chi_c, c_i) = \chi_c / c_i^2$, with $\chi_c = 1$. Parameters are $\bar{p} = 5$, $\rho_1 = 1$, $\rho_2 = 5$, $\phi = 0.8$, and $A = I$. Firm 1's data is measured on the x-axis. Firm 2's data is fixed at $\Sigma_{\epsilon_2}^{-1} = 1$.

An Illustrative Example of Industry Markup Divergence Figure 2 illustrates the divergence. The firm-level markup (dashed line at the bottom) rises less than the industry-aggregates markup (middle dotted line), which rises less than that sales-weighted markup (solid, top line). That result suggests that as firms process more data over time, the differences between markups measured at various levels of aggregation will continue to grow.

With aggregate markups, there are now four ways in which data affects markups, aside from true changes in market power. Data increases markups because of reduced cost, cross-product aggregation, and cross-firm aggregation. Data decreases markups because it induces firms to produce more (risk premium channel).

III.E Empirical Evidence from Industry Markups

The empirical literature finds that there is a wedge between the sales-weighted markup and the cost-weighted markup, and that this wedge is growing over time since the early 1980s (see Figure 3a, from De Loecker, Eeckhout, and Unger (2020)). Firms that have market power sell at higher prices and therefore have higher revenue and relatively lower costs. This difference between sales and costs therefore drives a wedge between sales- and cost-weighted markup measures. This is consistent with what we find as firms that have market power boost their sales with fewer inputs since they have higher markups. In our model, firms who invest heavily in data do exactly that, and the more important the role of data, the bigger the wedge between the input- and output-weighted aggregate markup. Our contribution is to propose a theory based on the role of data in creating these wedges, and how they grow as the data becomes more abundant.



(a) Sales-weighted markups, M^s, (solid line) vs. cost-weighted markups, M^c, (dashed line)

(b) Sales-weighted markups, M^s, (solid line) vs. industry-aggregates markups, M^{ind}, (dashed line)

Figure 3: Markups Measured and Aggregated in Different Ways Diverged Over Time.
Note. From De Loecker, Eeckhout, and Unger (2020) (updated), Figure XVI.A (left panel) and Figure V (right panel).

Our theory predicts differences between product, firm, and industry markups. To date, there is still limited evidence comparing product versus firm markups using the same data source. However, there is consistent evidence comparing firm markups to industry markups. In fact, the seminal paper on markup measurement by means of the production approach by Hall (1988) uses industry, not firm-level, data to construct aggregate markup measures (see also Hall [2018] for recent industry estimates using KLEMS data). With firm-level data and industry classification codes, we can mimic the industry aggregates using exactly the same set of firms underlying the industry aggregates. Based on De Loecker, Eeckhout, and Unger (2020) using data on publicly traded firms, Figure 3b shows that industry markups (blue dashed line) have increased by half as much as sales-weighted firm markups (red line). In other words, they find that there is a wedge between the industry markup and the sales-weighted firm markup, and that wedge is increasing as investment in data increases. Note that industry markups (in Figure 3b) look remarkably similar to cost-weighted firm markups (in Figure 3a). This is due to the systematic relation between input-weighting and industry aggregates in equation (19).[8]

IV Cyclicality of Markups

A key question for mainstream New Keynesian models of the type often used by central banks is whether markups are countercyclical. This question has created stark disagreement. Researchers who measure markups at the firm or industry level find clear evidence of countercyclical markups (Bils 1985, 1987). In contrast, researchers who measure markups at the product level do not find evidence of countercyclicality (Nekarda and Ramey 2020). Our model offers a way to reconcile these facts.

Our explanation builds on the progress in Burstein, Carvalho, and Grassi (2020). They show analytically how composition changes can turn procyclical markups into countercyclical ones, depending on how markups are aggregated. Our model provides a specific economic mechanism for these composition changes. The cyclical markup evidence, in turn, supports the realism of the model's assumptions.

To use the model to explore the cyclicality of markups, we first need to define what is a boom or recession in the context of this model. There are two relevant features of a boom: 1) demand rises and 2) the variances of demand and of output fall. In contrast, recessions are volatile, uncertain times. To formalize this new assumption, we introduce a variable $Boom \in \{0, 1\}$ that makes the level of demand procyclical and the demand variance countercyclical:

$$\bar{p} = d_0 + d_1 * Boom \qquad \text{where } \; d_0, d_1 \geq 0, \tag{20}$$

[8] Rossi-Hansberg, Sarte, and Trachter (2018) discovered a different divergence in measures of market power, one between local and national markets. That difference in market power is not expressed in markups but in concentration indices such as HHI (Herfindahl-Hirschman Index). Expressed in markups, there is no documented local-national divergence. Benkard, Yurukoglu, and Zhang (2021) argue that HHI is defined over the market where consumers are located, whereas data used to measure HHI is based on the location of production, which leads to misleading and inconsistent findings when aggregating. Eeckhout (2020) shows that the discrepancy stems from a mechanical relation between population size and the market definition.

$$\Sigma_b = d_2 - d_3 * Boom \qquad \text{where } \ d_2, d_3 \geq 0. \tag{21}$$

High demand in a boom (20) regulates how countercyclical or acyclical product markups are. Falling variance in a boom (21) is what makes the cyclical behavior of aggregate markups differ relative to product markups. The second statement is formalized in the next proposition.

Proposition 5. ***Product markups diverge from firm and industry markups when volatility rises.*** *Suppose the investment cost structure is such that firms choose identical investments ($\mathbf{c}_i = \mathbf{c}_j \ \ \forall\, i, j$).*

a. The product-level markup is strictly increasing in demand variance, $\partial \mathbf{E}[M^p_{ij}]/\partial \Sigma_{b,j} > 0$, and converges to a constant as $\Sigma_{b,j} \to \infty$.

b. If demand variance is large enough, firm and industry markups are strictly increasing, $\partial \mathbf{E}[M^f_{ij}]/\partial \Sigma_{b,j} > 0$ and $\partial \mathbf{E}[M^m_{ij}]/\partial \Sigma_{b,j} > 0$, and asymptotic to a function increasing in variance, $\lim_{\Sigma_{b,k} \to \infty} \partial \mathbf{E}[M^f_{ij}]/\partial \Sigma_{b,j}, \partial \mathbf{E}[M^m_{ij}]/\partial \Sigma_{b,j} > 0$.

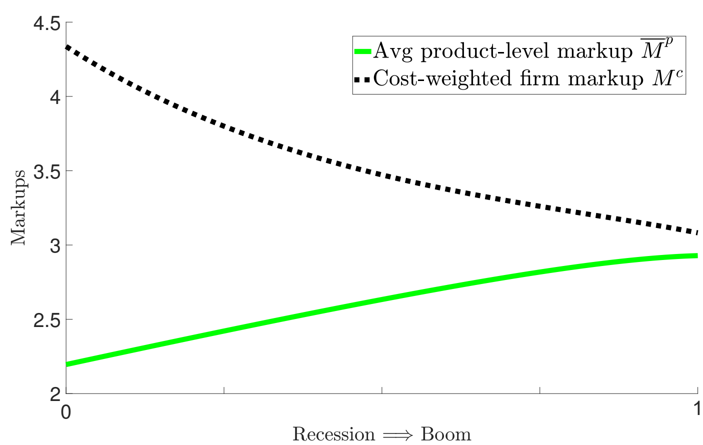


Figure 4: Procyclical product markups can coexist with countercyclical firm / industry markups. Left (right) on the x-axis represents recessions (booms), as described in (20) and (21), where $d_0 = 7/2$, $d_1 = 2$, $d_2 = 5$, and $d_3 = 4$. A decreasing line represents a countercyclical markup. Remaining parameters are $\rho_1 = \rho_2 = 1$, $\mathbf{c}_1 = \mathbf{c}_2 = 1$, $\phi = 1$, and $\Sigma_{\epsilon_1} = \Sigma_{\epsilon_2} = 1$.

The coexistence of a procyclical product markup and a countercyclical firm or industry markup is illustrated in Figure 4. The covariance of demand and a firm's output is what makes firm markups different from product markups. Firms have higher markups in more volatile environments because that volatility allows them to produce more of products that have extremely high markups. In other words, when variance of demand rises, covariance rises as well. This strengthens the composition effects that push firm markups up higher than product markups.

V Quantifying Markup Aggregation Wedges

We have shown that the model can replicate trend and cyclical behavior of markups qualitatively. Next, we explore the model's quantitative predictions. Specifically, we calibrate to match recent aggregate markup measures and then use the model to decompose aggregate markups into their various components: the risk premia effect, the investment cost effect, and the firm and industry aggregation wedges.

V.A Model calibration

We calibrate the model to aggregate moments of the U.S. economy. We target measured industry markups, so that when we decompose these markups into their constituent pieces, we are decomposing a realistically-sized markup. We use measures of firm-level uncertainty to calibrate the quantity and response to data. We calibrate a two-good, two-firm model in the simpler, public signal setting. Neither firm has data about the first good. One firm has data on the second good. For a two-firm model, there are 7 parameters. One parameter $\bar{p}$ governs the price level and sets the units of prices. That is a normalization. One is set to zero to represent firms with little data. That leaves 5 independent parameters. One is externally calibrated. The remaining four parameters are jointly estimated. We describe the details of each of these moments and the estimation procedure below. The parameter values are summarized in Table 1.

Price level $\bar{p}$. Since the model is a real one, scaling all prices leaves the real results unchanged. That implies we can choose one parameter to determine the price level. Therefore, the price upper bound $\bar{p}$ is set to 10.

Price impact of output ϕ. We match the parameter ϕ directly to estimates of firms' ability to affect prices. The firm's ability to affect prices is guided by the demand elasticity. In the Atkeson and Burstein (2008) model, there is a distribution of elasticities, with estimates ranging between 1.2 and 5.75. Therefore, we follow De Loecker, Eeckhout, and Mongey (2021) and set ϕ, which is the inverse of the elasticity, equal to 0.3, so the resulting elasticity of 3.33. This is approximately the midpoint of the measured elasticity distribution.

Data quantities $n_{d1}\sigma_\epsilon^2$ and $n_{d2}\sigma_\epsilon^2$.

Firms are identical ex ante, except in the number of signals they obtain. We use data on firms' ability to predict their future revenue to infer the amount of data firms use to improve their predictions. What we target in calibration is the ratio of the forecast errors of low-data and high-data firms. We use the ratio as a target because it is scale neutral. It allows us to calibrate the relative amout of data that various firms have. This relative data amount is key for the aggregation effects because it determines the degree of difference between the otherwise identical firms.

Low-data firms: The majority of firms, even public firms, issue no revenue forecasts. We assume such firm have no data because most firms fail to use data to enhance their forecasting or decision-making capabilities.[9] Thus, the low-data firm is assumed to have no data ($n_{d1} = 0$). This does not imply that such a firm knows nothing. It knows the true variance of demand shocks and their prior realizations.

Defining forecast errors. The forecast error is the percentage difference between the manager's revenue forecast, $E[\mathbf{q}'_{i,t}\mathbf{p}_{i,t}|\mathcal{I}_{i,t}]$, at the start of year t and $\mathbf{q}'_i\mathbf{p}_i$, the realized revenue at the end of the year:

$$FE_{i,t} = \frac{E[\mathbf{q}'_{i,t}\mathbf{p}_{i,t}|\mathcal{I}_{i,t}] - \mathbf{q}'_{i,t}\mathbf{p}_{i,t}}{E[\mathbf{q}'_{i,t}\mathbf{p}_{i,t}|\mathcal{I}_{i,t}]} \tag{22}$$

By the definition of variance, the expected squared forecast error is a conditional variance. The expected squared error is approximately the sample average squared error, which we normalize by the average squared mean:

$$\bar{FE}_i^2 = \frac{1}{T}\sum_{t=1}^{T} FE_{i,t}^2. \tag{23}$$

The moment that we compute in the data and match in the model simulations is the forecast error variance ratio of low data to high data firms: $\bar{FE}^2_{L,t}/\bar{FE}^2_{H,t}$, where L refers to the low-data firm forecast and H refers to the high-data firms.

Measuring forecast errors. To measure forecast errors, we merge data from IBES and Compustat. IBES has information on the forecast of managers about the firm revenue one year ahead. Compustat has information on the income statement and balance sheet of the firm, which allows us to calculate markups. In Appendix E. we elaborate in further detail on our data sources.

For calibration, we use the average from 2003-2019, because the large, unexpected Covid profit shocks in 2020 are not something the model is designed to explain.

[9] Bryniolfsson and Hitt (2000) find that this is because they lack the complementary organizational changes required to generate value from information technologies. Contemporary industry evidence reinforces this pattern, with executives themselves reporting that the vast majority of organizations still fail to become data-driven in practice Bean and Davenport (2022).

For each year, we split the sample into two groups. Firms with squared forecast errors ($FE^2_{i,t}$) below the median are the "high-data firms." Firms with ($FE^2_{i,t}$) above the median are the "low-data firms." In each year, , within each group, we take the cross-sectional variance of $FE_{i,t}$, which is the same as averaging the $FE^2_{i,t}$ values if they are mean-zero.[10] Then, we average that value over the years 2003-2019. The resulting variance of the median high-data firm forecast errors is 4.25%. The variance of the low-data firm forecast errors is 6.41%. The ratio of those is 1.51. This ratio is the calibration target.

We choose to target the ratio, instead of the magnitude of forecast errors, because it is important that the amount of data we give the high-data firm is realistic. Data reduces the conditional variance, relative to the unconditional one. If we targeted the conditional variance outright, then estimation error of the unconditional variance would contaminate the estimate of the amount of data. Using the ratio of firms with data to those without isolates the effect of data more cleanly.

In the model, we have one high data firm and one low-data firm. We compute the same variance ratio in each simulation and average across the 10,000 simulations. The model can match the empirical forecast error ratio (1.51) to within 2 decimal places.

Price of risk ρ. The price of risk determines how responsive a firm's actions and revenues are to a change in its conditional variance. We determine this parameter jointly with those below. However, one moment is particularly informative about ρ.

A key finding in Kohlhas and Asriyan (2025) is that a 1% point decrease in squared relative revenue forecast error (e.g., 0.03 to 0.02) is associated with a 1.9% increase in profits.

The corresponds to a semi-elasticity of -1.9%. The squared forecast error in the data corresponds to the conditional variance of demand shocks in the model. Thus the model's equivalent semi-elasticity is

$$\frac{\Delta \mathbf{q}_i' \mathbf{p}_i}{\Delta \mathrm{Var}(b_i|\mathcal{I}_i)} \tag{24}$$

To compute this elasticity, we subtract one percent of the conditional variance by giving the high data firm more data and then re-simulate. We then compare the new average revenue to the old average revenue and compute the percentage change: $(\mathbf{q}_i'\mathbf{p}_i^{NEW} - \mathbf{q}_i'\mathbf{p}_i)/\mathbf{q}_i'\mathbf{p}_i$. Our calibration produces 0.01896, which corresponds to a semi-elasticity of revenue to changes in conditional variance of -1.896%.

[10]We have nearly identical results with the sales-weighted average $Var(group) = \frac{\sum_{group,year} w_{i,t} * FE^2_{i,t}}{\sum_{group,year} w_{i,t}}$, where $w_{i,t}$ is the firm's sales share in year t.

Demand variance σ_b^2 and the cost of production $\bar{c}$.

To simulate the model, we need to assume a parametric form for the cost of investment. It needs to be a convex cost that is decreasing in the firm's choice of $\tilde{c}_{i,j}$, $\forall i, j$. For simplicity, we use a quadratic form. The cost function for up-front firm investment is: [11]

$$g_j(\chi_c, \tilde{c}_{i,j}) = \frac{1}{2}(\bar{c}_j - \tilde{c}_{i,j})^2.$$

The base marginal cost ($\bar{c}$) and the variance of the demand shock (σ_b^2) are jointly calibrated to match the empirical cost-weighted and sales-weighted industry markups. Both $\bar{c}$ and σ_b^2 are common to both firms and both goods.

For the purpose of deriving theoretical results, we defined markups and markup weights as ratios of averages. We did this because taking expectations of ratios of normal variables is not analytically tractable. However, when calibrating the model to data, we compute markups in the same way in the model, as in the data.

De Loecker, Eeckhout, and Unger (2020) measure cost-weighted markups as (26) and find the 2019 average industry markup to be 1.29. They measure the sales-weighted markup as in (25) and find the 2019 average industry markup to be 1.54.

$$SW = \frac{1}{n_F} \sum_{i=1}^{nF} \left(\frac{\mathbf{q}_i' \mathbf{p}_i}{\sum_{i=1}^{n_F} \mathbf{q}_i' \mathbf{p}_i} \right) M_i^f \tag{25}$$

$$CW = \frac{1}{n_F} \sum_{i=1}^{nF} \left(\frac{\mathbf{q}_i' \mathbf{c}_i}{\sum_{i=1}^{n_F} \mathbf{q}_i' \mathbf{c}_i} \right) M_i^f \tag{26}$$

For the model, with each parameter configuration, we compute (25) and(26), for our two-good, two-firm ($n_F = 2$) model, and average the SW and CW realizations across the simulations. Our parameters σ_b^2 and $\bar{c}$ are chosen to match SW and CW exactly, up to 3 significant digits.

Simulation details We used the Method of Simulated Moments (MSM) to choose $\bar{c}$, σ_b^2, n_{d2} and ρ to minimize the distance between empirically observed aggregate markups and those generated by the model. We use a minimum of 10,000 simulations, with more if the objective is not stable.

For each parameter configuration, we simulated the model at least 10,000 times, drawing four demand shocks (2 firms times 2 goods each) and four signal noise realizations. For each random

[11] The Appendix shows that the results are robust to a more general cost function that has an arbitrary multiplicative parameter in front, instead of the 1/2. The reason is that when re-calibrated, $\bar{c}_j$ is chosen to ensure the same marginal cost.

Table 1: Calibrated Parameters for 2019

Parameter	Description	Value	Calibration Target
$\bar{p}$	Price level	10	Normalization of price units
ϕ	Price impact of output	10/3	De Loecker, Eeckhout, and Mongey (2021)
σ_b^2	Demand variance	69.3	Sales- and cost-weighted markups
ρ	Price of risk	0.91	Semi-elasticity of revenue to forecast error
n_{d1}	Data quantity	0	Firm 1 is a no-data firm
n_{d2}	Data quantity	0.47	Forecast error variance ratio
$\bar{c}$	Cost of production	9.99	Calibrated to sales- and cost-weighted markups

draw, we compute SW and CW, the forecast error (23) and the semi-elasticity (24). Then, we average these outputs across each of the simulations. The algorithm searches for the parameter values that minimize the distance between the average realizations and their empirical counterparts.

For the optimal cost choice in a two-firm Nash equilibrium, we use the analytical solution (eq. 197 in Appendix E.).

Quantity choices for the second good are truncated at 0, and the opposing firm's quantities are recomputed to ensure a best-response. Truncating quantities ensures that total production is always positive. Because both firms have no data about the first good, it ensures that total production costs are never zero. This places a natural upper bound on firm markups and prevents near-infinite markup realizations from driving the results.

The minimization was performed using a differential evolution algorithm, a global stochastic optimization method robust to a non-smooth objective surface. This was necessary because the truncation at zero produced kinks and flat regions of the objective. This algorithm was able to look for minimia on the other side of the flat, zero-quantity region of the objective. To ensure numerical stability, the exogenous shocks and signal noise are drawn once and frozen prior to the optimization routine.

The finding that the maximum market market price is close to the marginal cost makes it look as though there are no markups here. However, what this really implies is that firms often do not find it profitable to produce good 2. When a firm does produce, it often has large and variable demand shocks that can make its markup sizable. Appendix E. reports additional calibration details.

V.B Decomposition Results: How Much of Industry Markups Come from Each Source?

The calibration targets the cost-weighted and sales-weighted markups because these are the objects directly observed in the data. While these moments discipline the model along empirically relevant dimensions, it is not obvious ex ante that a single parameterization can jointly match the observed markup wedges while also delivering a realistic price of risk and forecast error variance. The calibration therefore poses a nontrivial question: What do the observed markup wedges imply about the quantitative importance of the other mechanisms embedded in the model?

The central finding is that aggregation effects account for the bulk of the observed markup differences. Matching the wedge between cost-weighted and sales-weighted markups requires substantial volatility in firm-level revenue shocks. At the same time, the calibration implies only a modest price of risk. Taken together, these features have strong implications for the model's internal decomposition of markups.

High volatility in demand, combined with a moderate price of risk, amplifies firms' incentives to reallocate production toward products with higher expected markups. As a result, product-to-firm aggregation effects become quantitatively large (Figure 5). These effects propagate further through firm-to-cost-weighted aggregation and, finally, from cost-weighted to sales-weighted markups. In the calibrated economy, this chain of aggregation effects explains most of the gap between observed markup measures (Figure 6).

Because aggregation effects absorb so much of the observed variation in markups, there is limited scope left for the cost channel to play a dominant role. Although firms do invest in cost reduction, and lower marginal costs do contribute to higher markups at the product level, this mechanism is quantitatively secondary once the model is disciplined to match the empirical markup wedges. Instead, cost-weighted markups are driven primarily by composition effects: firms produce relatively more of high-markup products, high-markup firms receive greater weight in cost aggregation, and firms with higher revenues receive disproportionate weight in sales-weighted measures.

Overall, the results suggest that observed differences between cost-weighted and sales-weighted markups should be interpreted largely as evidence of aggregation and composition effects rather than as direct measures of changes in marginal costs or market power. However, this does not imply that market power is not part of the story of markups. In fact, market power may behind firm-specific demand shocks and the reason some firms amass so much data that allows them to select the high-markup products to produce. But this points to a mechanism that market power, or other forces, operate through. We further explore this idea of endogenous data in the next section.

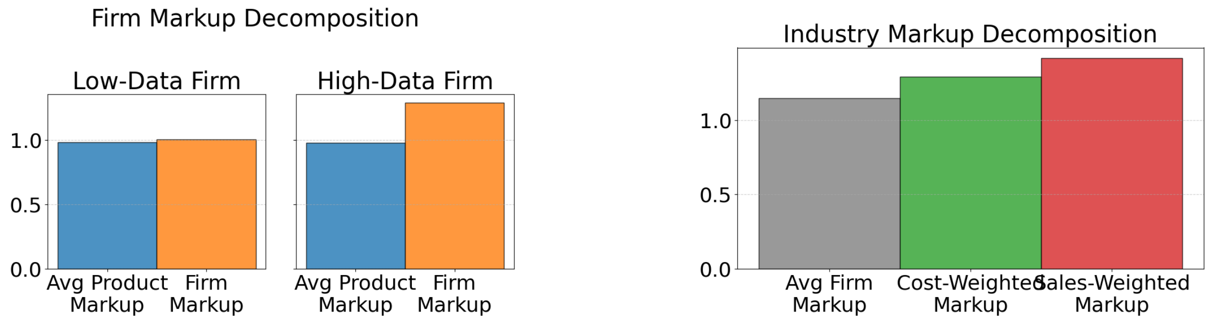


Figure 5: Markups levels for product, firm and industry.

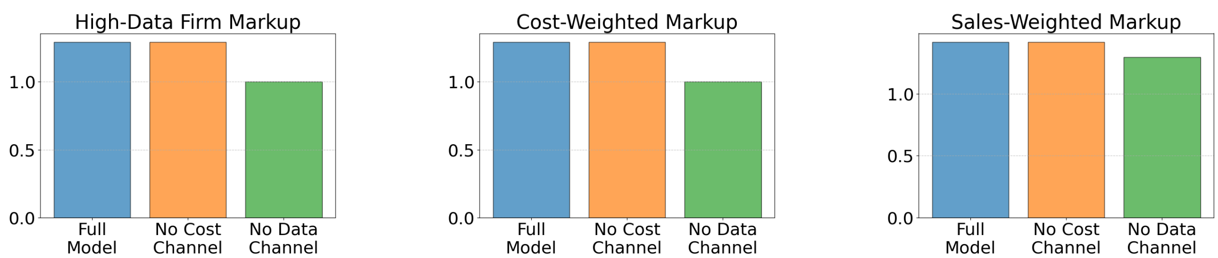


Figure 6: Decomposition of firm and industry markups by data-difference and cost-difference channels.

VI Data Barter and Markups: A Dynamic Model

In a dynamic economy, there is one more effect of data on markups: a role for data as a means of payment. We call this data barter. The dynamic model teaches us that, when assessing competition, customers can pay with money or data. This is not just true for free digital apps. The price of every good should be affected. Despite this additional force, the original insights derived from the static model survive, even when data comes from firms' transactions with their customers.

Dynamic model setup. As before, firms choose a quantity $\mathbf{q}_{it}$ to produce to maximize expected profit, with a price of risk adjustment as in equation (2). There are n goods and demand shocks for each of those goods.

What makes data an asset that retains value over multiple periods is that those demand shocks are persistent. If they were not persistent, if demand were independent each period, then data about yesterday's demand would have no value in predicting today's demand. Data would have one-period value. It would not be a long-lived asset. Therefore, we assume a persistent demand process that is an AR(1):

$$\mathbf{b}_t = \rho \mathbf{b}_{t-1} + \boldsymbol{\eta}_{bt}, \qquad \boldsymbol{\eta}_{bt} \sim iid\ N(0, \sigma_\eta \mathbf{I}). \tag{27}$$

At the same time, there needs to be some transitory noise in prices. If there were not, the price of a good would be a sufficient statistic for all past data. If prices revealed all the information in past data, then data would confer only a one-period advantage. It would also not be a long-lived asset. Therefore the demand shock for each good $\tilde{\mathbf{b}}_t$ is the persistent process (27), plus some transitory noise:

$$\tilde{\mathbf{b}}_t = \mathbf{b}_t + \boldsymbol{\epsilon}_{bt} \qquad \boldsymbol{\epsilon}_{bt} \sim iid\ N(0, \sigma_\epsilon \mathbf{I}) \tag{28}$$

Data is produced as a by-product of economic activity. In other words, the more a firm produces and sells, the more it learns about its customers, its suppliers and its optimal choices. We can model this as a number of data points about each product that depends on the amount produced of each product: $\mathbf{n}_{it} = \mathbf{q}_{i,t-1}$. Notice that this makes production a form of active experimentation. Firms are like gamblers in a classic bandit problem, learning about the profitability of each action by observing its result.

The amount of data that a firm has depends on data production, as well as data purchases or sales: $D_{it} = \text{diag}(\mathbf{n}_{it} + \mathbf{m}_{it})$ where $\mathbf{m}_{it}$ is the amount of data purchased by firm i at date t and diag turns the vector into a diagonal matrix. Firms also choose an amount of data to sell $\mathbf{l}_{it}$. Since data is non-rival, data that is sold is not lost. However, selling data may not be optimal if better-informed rivals reduce a firm's own production and profits.

Each data point is a signal about the demand shock vector $\mathbf{b}_t$, with precision σ_e^{-1} per signal. Firms update demand forecasts using Bayes law. Thus, when a firm obtains D_{it} units of data about each good's demand, Bayes law tells the firm to average the signals[12] to arrive at a composite signal that has precision $D_{it}\sigma_e^{-1}$.

[12]Notice that Bayes' law allows us to incorporate non-integer numbers of signals. So we can proceed considering $\mathbf{d}_{it}$ and D_{it} to be vectors of any real, non-negative numbers.

Dynamic model solution Let Ω_t be the set of all firm's data precisions $\{\omega_{it}\}_{i=1}^N$. The firms' optimal production $\{\mathbf{q}_{i,t}, a_{i,t}\}$ and data purchases / sales $\{\mathbf{m}_{i,t}, \mathbf{l}_{i,t}\}$ solve:

$$V(\Omega_t) = \max_{\mathbf{q}_{i,t}, \mathbf{a}_{i,t}, \mathbf{m}_{i,t}, \mathbf{l}_{i,t}} E_t \left[(\mathbf{p}_t - \mathbf{c}_i) \mathbf{q}_{i,t} a_{i,t} + \mathcal{P}_t (\mathbf{l}_{i,t} - \mathbf{m}_{i,t}) + \left(\frac{1}{1+r} \right) V(\Omega_{t+1}) \right], \tag{29}$$

where $\mathcal{P}_t$ is the time-t market price per unit of data and the law of motion for firm i's data stock $\omega_{i,t}$ is

$$\omega_{i,t+1} = \left[\rho^2 \omega_{i,t}^{-1} + \sigma_\epsilon I \right]^{-1} + D_{i,t} \sigma_e^{-1} \tag{30}$$

and the firm's number of new data points, $\mathbf{q}_{i,t} + \mathbf{m}_{i,t}$, is the vector on the diagonal of matrix $D_{i,t}$.

The first-order condition for the quantity of production looks similar to (5) in the static model. Optimal production depends on risk and price impact, in the denominator, and expected profit $(\mathbf{p} - \mathbf{c}_i)$, in the numerator.

$$\mathbf{q}_i = \left(\rho_i \mathbf{Var} \left[\mathbf{p} | \mathcal{I}_i \right] + \frac{\partial \mathbf{E} \left[\mathbf{p} | \mathcal{I}_i \right]}{\partial \mathbf{q}_i} \right)^{-1} \left(\mathbf{E} \left[\mathbf{p} | \mathcal{I}_i \right] + \frac{\partial V(\Omega_{t+1})}{\partial \mathbf{q}_i} - \mathbf{c}_i \right) \tag{31}$$

However, there is one new term in dynamic model: $\partial V / \partial \mathbf{q}_i$ is the increase in the future value of the firm, from producing data.

Notice that the future value of data enters additively with the price. This means that monetary payments and data payments are substitutes for the firm. In other words, customers pay for goods, in part with data. This is a partial barter trade where goods are partly paid for with data, as when you receive a loyalty card discount at a supermarket or pharmacy. These discounts are similar to those in customer acquisition models (Nakamura and Steinsson 2011).

Data barter changes the interpretation of markups. The solution (31) reveals that the price of a good is not the complete payment for the good. The relevant measure of income from selling a unit of a good is $\mathbf{p} + \partial V / \partial \mathbf{q}_i$. So markups underestimate market power because they fail to account for the data payment that accompanies the monetary payment from customers. Firms in areas of the product space where data is valuable should keep their measured markups low, in order to generate more transactions, to generate more valuable data.

While this dynamic extension introduced new ideas about the interaction of data and markups, it did not change the main conclusions of the static model. Data still complicates the interpretation of markups as measures of market power. In this model, there are three main forces at work in dynamic product markups: (i) the classic effect of market power, (ii) a risk premium, and (iii) data barter. In a data-intensive sector, markups reflect the value of data and its effect on risk as well. Data still shows up as a force that changes how markups are aggregated. Firms use data to predict which goods will have high demand and produce more of those goods. Firms that do this prediction well will have higher firm markups and will grow bigger and get higher weights in their industry markup. But this model suggests that simply correcting markups for a risk premium will not be enough to solve the problem of measuring competition in data-intensive industries.

VII A Linear Hedonic Demand System and Product Innovation

Recent work shows that a modification of our linear setup fits the cross-product elasticity facts well. A rotation of our model builds on Pellegrino (2023)'s Generalized Hedonic-Linear demand system, used to study market power in a network economy (see also Galeotti et al. [2022]). A feature of this model is the declining demand elasticity in firm size. This generates realistic higher markups for larger firms. Using the nonparametric estimates of Baqaee, Farhi, and Sangani (2021) for the demand, Ederer and Pellegrino (2022) show that the linear demand system fits the data better than the iso-elastic demand system.

If one wanted to change the relationship between elasticity and firm size, introducing a function $\phi(c_i)$ would only change the magnitudes of our results. The same trade-offs arise.

To map our model into Pellegrino (2023)'s structure, we rename the model object that was celled "goods" and instead call that an "attribute." Then goods are bundles (linear combinations) of attributes.

The product space has N attributes, indexed by $j \in \{1, 2, \ldots, N\}$ and N goods, indexed by k, that are combinations of attributes. We use tildes to denote quantities and prices of attributes. Each good $k \in \{1, 2, \ldots, N\}$ can be represented as an $N \times 1$ vector $\mathbf{a}_k$ of weights that a good places on each attribute. The jth entry of vector $\mathbf{a}_k$ describes how much of attribute j the kth good requires. This collection of weights describes a good's location in the product space. Let the collection of $\mathbf{a}_k$'s be an $N \times N$, full-rank matrix $\mathbf{A}$, such that

$$\mathbf{q}_i = \mathbf{A}^{-1}\tilde{\mathbf{q}}_i. \tag{32}$$

Conversely, the quantity of attributes that a firm i produces is a vector $\tilde{\mathbf{q}}_i$, with jth element $\tilde{q}_{ij}$. Similarly, we represent the price and the marginal cost of production of goods as the linear combinations of the vector of prices and costs of the attributes: $\mathbf{p}_i = \mathbf{A}'\tilde{\mathbf{p}}_i$ and $\mathbf{c}_i = \mathbf{A}'\tilde{\mathbf{c}}_i$.

FIRMS There are n_F firms, indexed by i: $i \in \{1, 2, \ldots, n_F\}$. Firm production profit π_i depends on quantities of each good or each attribute, as follows:

$$\pi_i = \mathbf{q}_i' (\mathbf{p} - \mathbf{c}_i) = \tilde{\mathbf{q}}_i' \mathbf{A}^{-1\prime} \left(\mathbf{A}'\mathbf{p} - \mathbf{A}'\mathbf{c}_i \right) = \tilde{\mathbf{q}}_i' (\tilde{\mathbf{p}}_i - \tilde{\mathbf{c}}_i) \tag{33}$$

Therefore, the firm's objective (2) is unchanged and we can write the choice problem of the firm equivalently in terms of its goods or its attributes.

The last term in (2) is each firm's up-front investment. The up-front investment choice is modeled as a choice of attribute costs $\widetilde{\mathbf{c}}_i$.

PRICE Our demand system embodies the idea that goods with similar attributes are partial substitutes for each other. Therefore, the price of good k can depend on the amount every firm produces of every good.

Equation 3 now describes the market price of each attribute j. As before, each firm faces an uncertain price that depends on a demand shock $\mathbf{b}_i$. This vector is random and unknown to the firm: $\mathbf{b}_i \sim N(\mathbf{0}, \mathbf{I})$. The fact that the variance matrix is diagonal is without loss. We simply define attributes to be an orthogonal decomposition of the demand variance-covariance space.

The price a firm receives for a unit of attribute j is thus $\tilde{p}_j^M + b_{ij}$. We can express firm i's price in vector form as $\tilde{\mathbf{p}}_i = \left[\tilde{p}_1^M, \tilde{p}_2^M, \ldots, \tilde{p}_N^M\right]' + \mathbf{b}_i$.

The firm's price of each good depends on its attributes. The price of good k is the units of each attribute a_{jk} times the price of each attribute $\tilde{p}_j$, summed over all the attributes:

$$p_{ik} = \sum_{j=1}^{N} a_{jk} \tilde{p}_{ij}. \tag{34}$$

Model results for the linear hedonic demand system All the results in the appendix are proven for this linear hedonic demand system. They hold for the previous model as well because that is the special case of this model where each good loads only on one attribute: $\mathbf{A} = \mathbf{I}$. The solution works by first solving the model for attributes – or independent goods – and then projecting the solution back on to the space of goods that are linear combinations of attributes.

Using data for product innovation One use of data is for product innovation. A simple redefinition of the choice variable allows our model to capture this use of data. Instead of products having fixed weights on attributes, suppose that $\mathbf{a}_i$ is now a vector of weights that firms choose.

Whereas q_i was a vector of quantities of each product a firm produces, we now restrict that to be a scalar – the quantity of the one good the firm chooses to produce.

Equilibrium: After observing the realized data, each firm updates beliefs with Bayes' law . Each firm then chooses the attribute weight vector a_i and the scalar number of units of their chosen product q_i to maximize conditional expected utility in (2), taking as given other firms' best responses. Prices for each attribute are given by (3) and (4). The price of firm i's good is the linear combination of its attributes, as in (34).

The solution to this model is just a linear rotation of the original model solution. What was a portfolio of products becomes a portfolio of attributes in the firm's chosen product design. What was the optimal $\mathbf{q}_i$, now becomes the optimal $q_i\mathbf{a}_i$, where the entries of $\mathbf{a}_i$ must sum to one. Before, firms used data to tilt their product mix toward high-markup products. In this model, firms use data to tilt their product design toward valuable, high-markup attributes. The prediction that firms use growing amounts of data to identify new product niches is supported by Neiman and Vavra (2023).

VIII Conclusion

The hypothesis that data encourages large firms to grow larger and gain market power is both plausible and incomplete. Because data improves both prediction and firms' profitability, we need to consider competitive effects using a framework where firms compete and face uncertain outcomes that require prediction. In other words, wrestling with the competitive effects of data requires incorporating risk.

We used the a model to guide new measurement to disentangle data's effects from market power. One way in which data and market competition differ is in the covariance of actions and payoff-relevant outcomes. In this model, data boosts the covariance between price and quantity by allowing firms to have better forecasts of demand and thereby price. Market competition also changes this covariance by making production decisions more sensitive to expected price changes. But data enhances that sensitivity and also makes expected price and actual price more informative about each other. A second approach to measuring data is to recognize that data enables more accurate forecasting, while market competition does not. We show how to use the accuracy of firms' sales forecasts to infer data and quantify its effects on markups.

We find that high-data firms do invest more, grow larger, and exert more impact on prices.

However, if uncertain firms scale back production, then more data that resolves their uncertainty also pushes markups down. The effect of data may not be seen in the level of markups.

Instead, the effects of data should show up in markup aggregation. Firms react to data about demand by shifting their production to high-demand goods. These are high-markup goods. So data changes the composition of production. This composition effect leads firms to shift production toward high-markup goods, which raises markups. The tug-of-war between risk reduction and the composition effects induced by data plays out differently for product, firm, and industry markups. A model designed to explore the logic of data and large firms turned out to explain why econometricians got different answers about what was happening to markups over time when they measured at different levels of aggregation. Our model suggests a new interpretation of existing facts. Constant product markups and rising firm and industry markups are not competing facts. They are consistent with an economy where firms are getting better and better at forecasting future demand. Both are helpful in the attempt to understand and measure firms' use of data.

References

Abis, Simona, and Laura Veldkamp, "The Changing Economics of Knowledge Production," *The Review of Financial Studies*, 37 (2024), 89–118, Published online August 2, 2023; first posted SSRN April 6, 2020.

Acemoglu, Daron, Ali Makhdoumi, Azarakhsh Malekian, and Asu Ozdaglar, "Too Much Data: Prices and Inefficiencies in Data Markets," *American Economic Journal: Microeconomics*, forthcoming (2022).

Anderson, Eric, Sergio Rebelo, and Arlene Wong, "Markups across Space and Time," NBER Working Paper 24434, 2018.

Angeletos, George-Marios, and Jennifer La'O, "Sentiments," *Econometrica*, 81 (2013), 739–779.

Angeletos, George-Marios, and Alessandro Pavan, "Efficient Use of Information and Social Value of Information," *Econometrica*, 75 (2007), 1103–1142.

Asriyan, Vladimir, Luc Laeven, and Alberto Martin, "Collateral Booms and Information Depletion," *Review of Economic Studies*, 89 (2022), 517–555.

Atkeson, Andrew, and Ariel Burstein, "Pricing-to-Market, Trade Costs, and International Relative Prices," *American Economic Review*, 98 (2008), 1998–2031.

Baqaee, David, and Emmanuel Farhi, "Entry vs. Rents: Aggregation with Economies of Scale," NBER Working Paper 27140, 2021.

Baqaee, David Rezza, Emmanuel Farhi, and Kunal Sangani, "The Darwinian Returns to Scale," NBER Working Paper 27139, 2021.

Bean, Randy, and Thomas H. Davenport, "2022 Data and AI Executive Survey: The Journey to Becoming Data-Driven," NewVantage PartnersTechnical report, 2022.

Benkard, Lanier, Ali Yurukoglu, and Anthony Lee Zhang, "Concentration in Product Markets," NBER Working Paper 28745, 2021.

Bergemann, Dirk, and Alessandro Bonatti, "The Economics of Social Data: An Introduction," Cowles Foundation Discussion Paper 2171R, 2019.

Bils, Mark, "The Cyclical Behavior of Marginal Cost and Price," *American Economic Review*, 77 (1987), 838–855.

Bils, Mark J, "Real Wages over the Business Cycle: Evidence from Panel Data," *Journal of Political Economy*, 93 (1985), 666–689.

Boar, Corina, Denis Gorea, and Virgiliu Midrigan, "Why Are Returns to Private Business Wealth So Dispersed?" *NBER Working Paper No. 29705*.

Brealey, RA, SC Myers, and F. Allen, *Principles of Corporate Finance*, 7th ed. (New York, NY: McGraw-Hill, 2003).

Brynjolfsson, Erik, and Lorin M. Hitt, "Information Technology and Organizational Transformation: Firm-Level Evidence," *Management Science*, 46 (2000), 548–562.

Burstein, Ariel, Vasco M Carvalho, and Basile Grassi, "Bottom-Up Markup Fluctuations," NBER Working Paper 27958, 2020.

Corhay, Alexandre, Howard Kung, and Lukas Schmid, "Competition, markups, and predictable returns," *The Review of Financial Studies*, 33 (2020), 5906–5939.

Crouzet, Nicolas, and Janice Eberly, "Understanding Weak Capital Investment: The Role of Market Concentration and Intangibles," in *Jackson Hole Economic Policy Symposium*, 87–149, 2018.

David, Joel M, Romain Ranciere, and David Zeke, "International Diversification, Reallocation, and the Labor Share," Chicago Fed Mimeo, 2022.

David, Joel, and Venky Venkateswaran, "The Sources of Capital Misallocation," *American Economic Review*, 109 (2019), 2531–2567.

De Loecker, Jan, Jan Eeckhout, and Simon Mongey, "Quantifying Market Power and Business Dynamism in the Macroeconomy," NBER Working Paper 28761, 2021.

De Loecker, Jan, Jan Eeckhout, and Gabriel Unger, "The Rise of Market Power and the Macroeconomic Implications," *Quarterly Journal of Economics*, 135 (2020), 561–644.

De Ridder, Maarten, "Market Power and Innovation in the Intangible Economy," Cambridge Working Paper 1931, 2021.

Dixit, Avinash K, and Joseph E Stiglitz, "Monopolistic competition and optimum product diversity," *The American economic review*, 67 (1977), 297–308.

Doan, Trang, and Mai Iskandar-Datta, "Are Female Top Executives More Risk-Averse or More Ethical? Evidence from Corporate Cash Holdings Policy," *Journal of Empirical Finance*, 55 (2020), 161–176.

Eckbo, B. Espen, ed. *Handbook of Empirical Corporate Finance Set* (Amsterdam: Elsevier, 2008).

Ederer, Florian, and Bruno Pellegrino, "A Tale of Two Networks: Common Ownership and Product Market Rivalry," NBER Working Paper 30004, 2022.

Edmond, Chris, Virgiliu Midrigan, and Daniel Yi Xu, "How Costly Are Markups?", NBER Working Paper 24800, 2019.

Eeckhout, Jan, "Comment on: Diverging Trends in National and Local Concentration," in *NBER Macroeconomics Annual 2020*, 35, 151–166: NBER), 2020.

Furman, Jason, and Peter Orszag, "A Firm-Level Perspective on the Role of Rents in the Rise in Inequality," Speech, 2015.

Galdon-Sanchez, Jose Enrique, Ricard Gil, and Guillermo Uriz-Uharte, "The Value of Information in Competitive Markets: Evidence from Small and Medium Enterprises," 2023, Queens University mimeo.

Galeotti, Andrea, Benjamin Golub, Sanjeev Goyal, Eduard Talamàs, and Omer Tamuz, "Taxes and Market Power: A Principal Components Approach," Northwestern University Mimeo, 2022.

Grullon, Gustavo, Yelena Larkin, and Roni Michaely, "Are US Industries Becoming More Concentrated?", mimeoTechnical report, 2016.

Hall, Robert E., "The Relation between Price and Marginal Cost in U.S. Industry," *Journal of Political Economy*, 96 (1988), 921–947.

Hall, Robert E, "New Evidence on the Markup of Prices over Marginal Costs and the Role of Mega-Firms in the US Economy," NBER Working Paper 24574, 2018.

Hennesy, Christopher, and Toni Whited, "How Costly Is External Financing? Evidence from a Structural Estimation," *Journal of Finance*, 62 (2007), 1705–1745.

Ichihashi, Shota, "Online Privacy and Information Disclosure by Consumers," *American Economic Review*, 110 (2020), 569–595.

Jaimovich, Nir, Sergio Rebelo, and Arlene Wong, "Trading down and the business cycle," *Journal of Monetary Economics*, 102 (2019), 96–121, Carnegie-Rochester-NYU ConferenceSeries on Public Policy"A Conference Honoring the Contributions ofCharles Plosser to Economics"Held at the University of RochesterSimon Business SchoolApril 20-21, 2018.

Jarsulic, Marc, "Antitrust Enforcement for the 21st Century," *The Antitrust Bulletin*, 64 (2019), 514–530.

Jones, Charles, and Chris Tonetti, "Nonrivalry and the Economics of Data," *American Economic Review*, 110 (2020), 2819–58.

Kirpalani, Rishabh, and Thomas Philippon, "Data Sharing and Market Power with Two-Sided Platforms," NBER Working Paper 28023, 2020.

Kohlhas, Alexandre, and Vladimir Asriyan, "The Macroeconomics of Data: Scale, Product Choice, and Pricing in the Information Age,"Technical report, 2025.

Kumar, Saten, Yuriy Gorodnichenko, and Olivier Coibion, "The Effect of Macroeconomic Uncertainty on Firm Decisions," *Econometrica*, 91(4) (2023).

Kwon, Spencer Yongwook, Yueran Ma, and Kaspar Zimmermann, "100 Years of Rising Corporate Concentration," SAFE Working Paper 359, 2022.

Lambert, Nicolas S., Michael Ostrovsky, and Mikhail Panov, "Strategic Trading in Informationally Complex Environments," *Econometrica*, 86 (2018), 1119–1157.

Liang, Annie, and Erik Madsen, "Data and Incentives," mimeoTechnical report, 2021.

Lorenzoni, Guido, "A Theory of Demand Shocks," *American Economic Review*, 99 (2009), 2050–2084.

Lovallo, Dan, Tim Koller, Robert Uhlaner, and Daniel Kahneman, "Your Company Is Too Risk-Averse," *Harvard Business Review*.

Maćkowiak, Bartosz, and Mirko Wiederholt, "Optimal Sticky Prices Under Rational Inattention," *American Economic Review*, 99 (3) (2009), 769–803.

Nakamura, Emi, and Jón Steinsson, "Price setting in forward-looking customer markets," *Journal of Monetary Economics*, 58 (2011), 220–233.

Neiman, Brent, and Joseph Vavra, "The rise of niche consumption," *American Economic Journal: Macroeconomics*, 15 (2023), 224–264.

Nekarda, Christopher, and Valerie Ramey, "The Cyclical Behavior of the Price-Cost Markup," *Journal of Money, Credit, and Banking*, 52 (2020), 319–353.

Nimark, Kristoffer P., "Man-Bites-Dog Business Cycles," *American Economic Review*, 104 (2014), 2320–2367.

Pellegrino, Bruno, "Product Differentiation and Oligopoly: a Network Approach," *American Economic Review*.

Philippon, Thomas, *The great reversal: How America gave up on free markets*: Harvard University Press, 2019).

Rossi-Hansberg, Esteban, Pierre-Daniel Sarte, and Nicholas Trachter, "Diverging Trends in National and Local Concentration," NBER Working Paper 25066, 2018.

Rostek, Marzena, and Marek Weretka, "Price Inference in Small Markets," *Econometrica*, 80 (2012), 687–711.

Sutton, John, *Sunk Costs and Market Structure: Price Competition, Advertising, and the Evolution of Concentration* (Cambridge, MA: MIT Press, 1991).

——— *Technology and Market Structure: Theory and History* (Cambridge, MA: MIT Press, 2001).

Vives, Xavier, and Zhiqiang Ye, "Information Technology and Bank Competition," IESE Working Paper, 2021.

Welch, Ivo, *Corporate Finance An Introduction* (Boston, MA: Pearson, 2009).

Online Appendix

A. Appendix: Solution Details

We present all solutions and results for the model with product attributes. Allowing those goods to have correlated attributes, as in Pellegrino (2023), makes data relevant to multiple goods. Since the case without attributes is simply the case where $A = I$, this general formulation allows us to present the results just once.

We start by solving the model with firm-specific shocks and public information. Then we solve the model where shocks are aggregate, and derive analogous expressions for two key objects in the model that govern the sensitivity of beliefs to signals and the sensitivity of production to changes in expected price. Then, in Appendix B., we use these solutions to prove the propositions and show that the same properties hold for both models. Appendix C has the most technical lemmas that are inputs into the proposition proofs.

A.1. Preliminaries

PRODUCTS AND ATTRIBUTES We consider n_F firms, indexed by i. Each firm potentially produces N goods indexed by k. The product space has N independent attributes indexed by $j \in \{1, 2, ..., N\}$. Therefore, each good $k \in 1, 2, ..., N$ can be represented by an $N \times 1$ vector $\mathbf{a}_k$ of weights that good k places each attribute. The collection of weights describes a good's location in the product space. Let the collection of weights ($\mathbf{a}_k{}'s$) be an $N \times N$ full-rank matrix $\boldsymbol{A}$, such that

$$\boldsymbol{q}_i = \boldsymbol{A}\tilde{\mathbf{q}}_i$$

The linear mapping $\boldsymbol{A}$ between good and attribute spaces allows us to transform the original model into attribute-competition model in which n_F firms choose upfront investments and attributes to maximize their mean-variance utility. As the attributes are assumed to be orthogonal, the model can be solved by considering one attribute at a time.

PRICES AND COSTS Similarly, we can represent the price and marginal cost of production of goods as the linear combinations of the vector of prices and costs of the attributes:

$$\boldsymbol{p}_i = \boldsymbol{A}\tilde{\mathbf{p}}_i \tag{35}$$

$$\boldsymbol{c}_i = \boldsymbol{A}\tilde{\mathbf{c}}_i. \tag{36}$$

Each attribute j has an an average market price $\tilde{p}_j^M$ that depends on an attribute-specific constant $(\bar{p}_j)$ and on the total quantity of that attribute that all firms produce. It is given by the inverse demand function, which holds for each attribute j:

$$\tilde{p}_j^M = \bar{p}_j - \frac{1}{\phi}\sum_{i=1}^{n_F} \tilde{q}_{ij} \tag{37}$$

INFORMATION Each firm i has n_{di} data points, each of which is a signal of the attribute demand shock $\boldsymbol{s}_{i,z} = \boldsymbol{b_i} + \boldsymbol{\varepsilon}_{i,z}$ where $z = 1, \ldots, n_{di}$. We assume signal noises are uncorrelated and normally distributed with zero mean and precision 1. This is equivalent to a compound signal $\boldsymbol{s}_i$ with total data precision $n_{di} = \sum_{z=1}^{n_{di}} 1$, the sum of the precisions of all of data points about firm i's demand. Therefore, we define a composite signal for each firm-attribute pair, which is the average of the n_{di} data points, each of which is the jth entry of the vector $\boldsymbol{s}_{i,z}$:

$$s_{i,j} = \frac{1}{n_{di}}\sum_{z=1}^{n_{di}} \boldsymbol{s}_{i,z}(j) \tag{38}$$

This signal is a sufficient statistic for all data observed about firm i, attribute j. The posterior variance of the demand shock, conditional on the composite signal $s_{i,j}$, is $\mathbb{V}\left(b_{i,j}|s_{i,j}\right) = \frac{1}{1+n_{di}}$. The posterior mean (the prediction) of demand is $\mathbb{E}[b_{i,j}|s_{i,j}] = (n_{di}/(1+n_{di}))s_{i,j}$, because the prior mean of $b_{i,j}$ is zero, with precision 1.

For results that refer to more data, we mean a derivative with respect to the number of data points about a firm, n_{di}.

A.2. Solving the model with firm-specific shocks and public information

As mentioned earlier, we show our results for two specifications of the model. In the first specification, there are firm-specific shocks and all data is public. Therefore, instead of facing the average market price $\tilde{p}_j^M$ for attribute j, each firm

i faces a different price $\tilde{p}_{i,j} = \bar{p}_j^M + b_{i,j}$ where $b_{i,j}$ is mean-zero shock distributed randomly with variance 1. The vector of prices faced by firm i is

$$\tilde{\mathbf{p}}_i = [\tilde{p}_1^M, \tilde{p}_2^M, ..., \tilde{p}_N^M] + \boldsymbol{b_i} \tag{39}$$

When other firms' outputs are known, the only uncertainty is about the b shock. Therefore conditional mean and variance of prices is a linear function of the mean and variance of b. For firm i and attribute j,

$$\begin{aligned} \mathbb{E}\left[\tilde{p}_{i,j}|\mathcal{I}_i\right] &= \bar{p}_j + \mathbb{E}\left[b_{i,j}|\mathcal{I}_i\right] - \frac{1}{\phi}\sum_{i'=1}^{n_F}\tilde{q}_{i',j} = \bar{p}_j + K_{i,j}s_{i,j} - \frac{1}{\phi}\sum_{i'=1}^{n_F}\tilde{q}_{i',j} \\ \mathbb{V}\left[\tilde{p}_{i,j}|\mathcal{I}_i\right] &= \mathbb{V}\left[b_{i,j}|\mathcal{I}_i\right] = \frac{1}{1+n_{di}} \end{aligned} \tag{40}$$

where $K_{i,j}$ is defined to be the weight that firm i puts on its signal about attribute j, when forming its expectation about the price. In this model, $K_{i,j} = \frac{n_{di}}{n_{di}+1}$. When others' actions are not observed, K takes another form.

MAXIMIZING RISK-ADJUSTED PROFIT Taking first-order condition of firm's utility function, we get an expression for optimal attribute choices.

$$\tilde{q}_{i,j} = \left(\rho_i \mathbb{V}\left[\tilde{p}_{i,j}|\mathcal{I}_i\right] - \frac{\partial \mathbb{E}\left[\tilde{p}_{i,j}|\mathcal{I}_i\right]}{\partial \tilde{q}_{i,j}}\right)^{-1} \left(\mathbb{E}\left[\tilde{p}_{i,j}|\mathcal{I}_i\right] - \tilde{c}_{i,j}\right) \tag{41}$$

Differentiating the inverse demand curve $\tilde{p}_{i,j} = \bar{p}_j + b_{i,j} - \frac{1}{\phi}\sum_{i'=1}^{n_F}\tilde{q}_{i',j}$ reveals that market power is constant:

$$\frac{\partial \mathbb{E}\left[\tilde{p}_{i,j}|\mathcal{I}_i\right]}{\partial \tilde{q}_{i,j}} = \frac{\partial \mathbb{E}\left[p_{i,j}|\mathcal{I}_i\right]}{\partial q_{i,j}} = -\frac{1}{\phi} \tag{42}$$

Define the sensitivity of supply to a change in the expected profit as:

$$\hat{H}_{i,j} := \left(\frac{1}{\phi} + \rho_i \mathbb{V}\left[\tilde{p}_{i,j}|\mathcal{I}_i\right]\right)^{-1}. \tag{43}$$

Substituting this constant market power into the first order condition for optimal output yields the next expression for optimal attribute production. But this expression has $\tilde{q}_{i,j}$ on both the left and the right sides of the equality. It arises on the right side because firm i's production choice $\tilde{q}_{i,j}$ affects the expected price $\mathbb{E}\left[\tilde{p}_{i,j}|\mathcal{I}_i\right]$. Therefore, we substitute in the price and re-arrange to collect all $\tilde{q}_{i,j}$ terms and reveal the optimal production choice. Use (42) to substitute out $\partial\mathbb{E}\left[\tilde{p}_{i,j}|\mathcal{I}_i\right]/\partial\tilde{q}_{i,j}$. Then use Bayes law to replace the expectation $\mathbb{E}\left[b_{i,j}|\mathcal{I}_i\right]$ with the weighted sum of signals $K_{i,j}s_{i,j}$, with the Bayesian updating weight $K_{i,j} = \frac{n_{di}}{n_{di}+1}$. Using these three substitutions, we can rewrite (41) as:

$$\tilde{q}_{i,j} = \hat{H}_{i,j}\left(\bar{p}_j + K_{i,j}s_{i,j} - \frac{1}{\phi}\sum_{i'=1}^{n_F}\tilde{q}_{i',j} - \tilde{c}_{i,j}\right) \tag{44}$$

The solution above generates the best-response function, given the aggregate output. When all data is public, this aggregate output is known. But firm i output choice still shows up on the left and right sides.

Before solving for aggregate output and prices, we first stop to correctly express firm i's best response, as a function of other firms' output choices. Gathering all the terms involving $\tilde{q}_{i,j}$ on the left:

$$\begin{aligned} \tilde{q}_{i,j}\left(1 + \frac{\hat{H}_{i,j}}{\phi}\right) &= \hat{H}_{i,j}\left(\bar{p}_j + K_{i,j}s_{i,j} - \frac{1}{\phi}\sum_{i'\neq i}\tilde{q}_{i',j} - \tilde{c}_{i,j}\right) \\ \tilde{q}_{i,j} &= \frac{\hat{H}_{i,j}}{1+\frac{\hat{H}_{i,j}}{\phi}}\left(\bar{p}_j + K_{i,j}s_{i,j} - \frac{1}{\phi}\sum_{i'\neq i}\tilde{q}_{i',j} - \tilde{c}_{i,j}\right) \\ &= H_{i,j}\left(\bar{p}_j + K_{i,j}s_{i,j} - \frac{1}{\phi}\sum_{i'\neq i}\tilde{q}_{i',j} - \tilde{c}_{i,j}\right) \end{aligned} \tag{45}$$

where $H_{i,j}$ is the sensitivity of firm i's output choice to the expected profit induced by other firms' choices. In contrast to $\hat{H}_{i,j}$, $H_{i,j}$ excludes the influence of i's own choices on profits.

$$H_{i,j} \equiv \frac{\hat{H}_{i,j}}{1+\frac{\hat{H}_{i,j}}{\phi}} = \frac{1}{\hat{H}_{i,j}^{-1}+\frac{1}{\phi}} = \frac{1}{\rho_i \mathbb{V}\left[\tilde{p}_{i,j}|\mathcal{I}_i\right]+\frac{1}{\phi}+\frac{1}{\phi}} = \left(\rho_i \mathbb{V}\left[\tilde{p}_{i,j}|\mathcal{I}_i\right]+\frac{2}{\phi}\right)^{-1}. \tag{46}$$

In other words, $H^{-1} = \hat{H}^{-1} + \frac{1}{\phi}$.

SUB-GAME EQUILIBRIUM We solve the sub-game Nash equilibrium by summing both sides of (44) over all firms to express the aggregate output:

$$\begin{aligned}
\sum_i \tilde{q}_{i,j} &= \sum_i \hat{H}_{i,j}\left(\bar{p}_j + K_{i,j}s_{i,j} - \frac{1}{\phi}\sum_{i=1}^{n_F}\tilde{q}_{i,j} - \tilde{c}_{i,j}\right) \\
\left(1+\frac{1}{\phi}\sum_i \hat{H}_{i,j}\right)\sum_i \tilde{q}_{i,j} &= \sum_i \hat{H}_{i,j}\left(\bar{p}_j + K_{i,j}s_{i,j} - \tilde{c}_{i,j}\right) \\
\sum_i \tilde{q}_{i,j} &= \left(1+\frac{1}{\phi}\sum_i \hat{H}_{i,j}\right)^{-1}\left(\bar{p}_j\sum_i \hat{H}_{i,j} + \sum_i \hat{H}_{i,j}K_{i,j}s_{i,j} - \sum_i \hat{H}_{i,j}\tilde{c}_{i,j}\right)
\end{aligned} \tag{47}$$

Denote,

$$\begin{aligned}
C_j &= \sum_i \hat{H}_{i,j}\tilde{c}_{i,j} \\
\kappa_j &= \sum_i \hat{H}_{i,j}K_{i,j}s_{i,j} \\
\bar{H}_j &= \sum_i \hat{H}_{i,j}
\end{aligned} \tag{48}$$

As $\hat{H}_{i,j}$ denotes the sensitivity of supply to a change in the expected profit, C_j represents the sensitivity weighted average cost of attribute j, hereafter referred to as the average cost of attribute j. Similarly, κ_j represents the average signal about the demand shock for attribute j weighted by the sensitivity term $\hat{H}_{i,j}$ and the weight $K_{i,j}$ that the firm i puts on its signal about attribute j , and $\bar{H}_j$ represents the average sensitivity of supply to expected profit.

Then, aggregate output can be expressed as

$$\sum_i \tilde{q}_{i,j} = \frac{\bar{p}_j\bar{H}_j + \kappa_j - C_j}{1+\frac{\bar{H}_j}{\phi}} \tag{49}$$

From aggregate output, we can express the market price of each attribute as $p_j^M = \bar{p}_j - \frac{1}{\phi}\sum_i \tilde{q}_{i,j}$. Using the expressions derived above, this can be written as

$$\begin{aligned}
p_j^M &= \bar{p}_j - \frac{1}{\phi}\left(\frac{\bar{p}_j\bar{H}_j + \kappa_j - C_j}{1+\frac{\bar{H}_j}{\phi}}\right) \\
&= \frac{\bar{p}_j + \frac{C_j}{\phi}}{1+\frac{\bar{H}_j}{\phi}} - \frac{\frac{\kappa_j}{\phi}}{1+\frac{\bar{H}_j}{\phi}}
\end{aligned} \tag{50}$$

Define $\bar{p}_j^M$ to be the expected market price for attribute j while p_j^M denotes the realized market price. Given that the signals $s_{i,j}$ have mean zero, κ_j equals zero in expectation. Therefore, the expected market price is

$$\bar{p}_j^M = \frac{\bar{p}_j + \frac{C_j}{\phi}}{1+\frac{\bar{H}_j}{\phi}} = \left(1+\frac{1}{\phi}\sum_i \hat{H}_{i,j}\right)^{-1}\left(\bar{p}_j + \frac{1}{\phi}\sum_i \hat{H}_{i,j}\tilde{c_{i,j}}\right) \tag{51}$$

Note that this also shows the how the realized market price p_j^M differs from its expectation $\bar{p}_j^M$ because of a weighted sum of the firms' random data realizations $s_{i,j}$. More specifically,

$$
\begin{aligned}
p_j^M &= \bar{p}_j^M - \frac{\kappa_j}{\phi(1+\frac{\bar{H}_j}{\phi})} \\
&= \bar{p}_j^M - \left(\phi + \sum_i \hat{H}_{i,j}\right)^{-1} \sum_i \hat{H}_{i,j} K_{i,j} s_{i,j}
\end{aligned}
\tag{52}
$$

Finally, we can express the equilibrium output as functions of marginal costs, parameters and firms' data $s_{i,j}$. Starting from (44), we substitute the definition of p^M to obtain:

$$
\begin{aligned}
\tilde{q}_{i,j} &= \hat{H}_{i,j}\left(p_j^M + K_{i,j}s_{i,j} - \tilde{c}_{i,j}\right) \\
&= \hat{H}_{i,j}\left(\bar{p}_j^M - \frac{\kappa_j}{\phi(1+\frac{\bar{H}_j}{\phi})} + K_{i,j}s_{i,j} - \tilde{c}_{i,j}\right) \\
&= \hat{H}_{i,j}(\bar{p}_j^M - \tilde{c}_{i,j}) + \hat{H}_{i,j}K_{i,j}s_{i,j} - \frac{\hat{H}_{i,j}\kappa_j}{\phi(1+\frac{\bar{H}_j}{\phi})} \\
&= \hat{H}_{i,j}(\bar{p}_j^M - \tilde{c}_{i,j}) + \hat{H}_{i,j}K_{i,j}s_{i,j} - \frac{\hat{H}_{i,j}}{\phi}\left(1 + \frac{1}{\phi}\sum_{i'=1}^{n_F}\hat{H}_{i',j}\right)^{-1}\sum_{i'=1}^{n_F}\hat{H}_{i',j}K_{i',j}s_{i',j}
\end{aligned}
\tag{53}
$$

where the second equality uses the relationship between market price and aggregate quantity, and the third equality uses (51). Similarly, equilibrium price available to firm i for attribute j is given by:

$$
\tilde{p}_{i,j} \equiv p_j^M + b_{i,j} = \bar{p}_j^M + b_{i,j} - \frac{1}{\phi}\left(1 + \frac{1}{\phi}\sum_{i'=1}^{n_F}\hat{H}_{i',j}\right)^{-1}\sum_{i'=1}^{n_F}\hat{H}_{i',j}K_{i',j}s_{i',j}
\tag{54}
$$

A.3. Solving the model with aggregate shocks and private information

So far, we have established the equilibrium solution for the model with private shocks and public data. However, the results generalize easily to a model with aggregate shocks and private data as well. Compared to the model discussed above, there are two main changes as discussed below.

Demand: The first change is that in this model, shocks affect the demand for attributes. These shocks affect all firms whose product load on these attributes. The market price $\tilde{p}_j$ for attribute j is given by the following inverse demand function

$$
\tilde{p}_j = \bar{p} + b_j - \frac{1}{\phi}\sum_{i=1}^{n_F}\tilde{q}_{i,j}
\tag{55}
$$

where b_j is the aggregate shock to the market price of attribute j. Recall that $\tilde{q}_{i,j}$ is the quantity produced by firm i and $\bar{p}_j$ is a constant. We assume that the shocks $[b_1, \ldots, b_N]$ are normally distributed with mean 0 and variance 1.

Information: The second change is that instead of the data being public, now each firm sees a private signal $s_{i,j} = b_j + \varepsilon_{i,j}$ where the variance of $\varepsilon_{i,j}$ depends on t he number of data points (n_{di}) available to firm i and equals $1/n_{di}$. We assume that $\varepsilon_{i,j}$ is i.i.d across firms. However, the signals $s_{i,j}$ are correlated across firms because of the common aggregate component b_j.

The new complication in this model is that the solution is not explicit. This is because, from (55), the price $\tilde{p}_j$ of an attribute j depends on the aggregate demand shock b_j as well as the quantity choices made by all firms. The quantity choice $\tilde{q}_{i,j}$ made by firm i for attribute j, depends on the noisy signal $s_{i,j}$, which is private information of the firm i, as well as it's beliefs about the signals (and therefore, the quantity choices) of other firms. As such, Bayesian updating weights always depend on the covariance between the observed signal $s_{i,j}$ and the price $\tilde{p}_{i,j}$ the firm needs to forecast. The additional complexity in this model is that this covariance is endogenous and depends on the production choices of all other firms. Therefore, we use a state-space approach and solve for the fixed point between beliefs and output. A state-space approach defines all the relevant model variables in terms of exogenous, orthogonal shocks and weights on those shocks. The innovation in this paper is that we extend the approach to the case of endogenous weights on each shock. Firms choose the Bayesian weights in a way that maximizes their objective.

We now prove Lemma 1 which provides the equilibrium solution for this model.

Proof of Lemma 1: Pricing with Aggregate Shocks

Proof. Following the state space updating approach, we guess and verify a linear price function and then solve for the coefficients at the end. A linear ansatz takes the following form with coefficients D_j, F_j, and $\{h_{i,j}\}_{i\in[1,\ldots,n_F]}$

$$\tilde{p}_j = D_j + F_j b_j + \sum_{i=1}^{n_F} h_{i,j}\varepsilon_{i,j} \tag{56}$$

Since firm i could only observe s_i, its expectation of the price is

$$\mathbb{E}[p_j|\mathcal{I}_i] = D_j + K_{i,j}s_{i,j} \tag{57}$$

where

$$K_{i,j} = \frac{\text{Cov}(\tilde{p}_{i,j}, s_{i,j})}{\mathbb{V}(s_{i,j})} = \frac{F_j + \frac{h_{i,j}}{n_{d_i}}}{1+\frac{1}{n_{d_i}}} \tag{58}$$

where the second equality follows from using the guess (56) and the fact that $s_{i,j} = b_j + \varepsilon_{i,j}$ where b_j has variance 1 and $\varepsilon_{i,j}$ has variance $\frac{1}{n_{d_i}}$. Similarly, the variance of price forecast error is given as follows

$$\begin{aligned}\mathbb{V}[\tilde{p}_{i,j}|s_{i,j}] &= \mathbb{V}(\tilde{p}_j) - \frac{\text{Cov}(\tilde{p}_j, s_{i,j})^2}{\mathbb{V}(s_i)} \\ &= F_j^2 + \frac{1}{n_{d_i}}\sum_{i'=1}^{n_F} h_{i',j}^2 - \frac{\left(F_j + \frac{h_{i,j}}{n_{d_i}}\right)^2}{1+\frac{1}{n_{d_i}}}\end{aligned} \tag{59}$$

Note that each firm has the same mean-variance objective. It's first-order condition with respect to $\tilde{q}_{i,j}$ is

$$\tilde{q}_{i,j} = \left(\rho_i \mathbb{V}[\tilde{p}_j|\mathcal{I}_i] - \frac{\partial \mathbb{E}[\tilde{p}_j|\mathcal{I}_i]}{\partial \tilde{q}_{i,j}}\right)^{-1} (\mathbb{E}[p_j|\mathcal{I}_i] - \tilde{c}_{i,j}) \tag{60}$$

From the pricing function (55), we obtain the price impact of one additional unit of attribute output as:

$$\frac{\partial \mathbb{E}[\tilde{p}_j|\mathcal{I}_i]}{\partial \tilde{q}_{i,j}} = -\frac{1}{\phi} \tag{61}$$

Defining $\hat{H}_{i,j} \equiv \left(\rho_i \mathbb{V}[\tilde{p}_j|\mathcal{I}_i] + \frac{1}{\phi}\right)^{-1}$, we obtain the expression for optimal production as

$$\tilde{q}_{i,j} = \hat{H}_{i,j}(\mathbb{E}[\tilde{p}_j|\mathcal{I}_i] - \tilde{c}_{i,j}) = \hat{H}_{i,j}(D_j + K_{i,j}s_{i,j} - \tilde{c}_{i,j}) \tag{62}$$

Plugging in the expression for $\tilde{q}_{i,j}$ from (62) and the guess for price from (56) into the pricing function (55), we obtain

$$D_j + F_j b_j + \sum_{i=1}^{n_F} h_{i,j}\varepsilon_{i,j} = \bar{p} + b_j - \frac{1}{\phi}\sum_{i=1}^{n_F} \hat{H}_{i,j}(D_j + K_{i,j}(b_j + \varepsilon_{i,j}) - \tilde{c}_{i,j}) \tag{63}$$

Rearranging to collect coefficients for each variable, we get

$$\left(F_j - 1 + \frac{1}{\phi}\sum_{i=1}^{n_F} \hat{H}_{i,j}K_{i,j}\right) b_j + \sum_{i=1}^{n_F}\left(h_{i,j} + \frac{1}{\phi}\hat{H}_{i,j}K_{i,j}\right)\varepsilon_{i,j} + D_j\left(1 + \frac{1}{\phi}\sum_{i=1}^{n_F}\hat{H}_{i,j}\right) - \bar{p}_j - \frac{1}{\phi}\sum_{i=1}^{n_F}\hat{H}_{i,j}\tilde{c}_{i,j} = 0 \tag{64}$$

As the above equation must hold for all values of b_j and $[\varepsilon_{1,j}, \ldots, \varepsilon_{n_F,j}]$, and these are all independent variables, it must be that their coefficients in the above equation are zero. This gives

$$F_j = 1 - \frac{1}{\phi}\sum_{i=1}^{n_F}\hat{H}_{i,j}K_{i,j} \tag{65}$$

$$h_{i,j} = -\frac{1}{\phi}\hat{H}_{i,j}K_{i,j} \tag{66}$$

$$D_j = \left(1 + \frac{1}{\phi}\sum_{i=1}^{n_F}\hat{H}_{i,j}\right)^{-1}\left(\bar{p}_j + \frac{1}{\phi}\sum_{i=1}^{n_F}\hat{H}_{i,j}\tilde{c}_{i,j}\right) \tag{67}$$

where the last equation follows from by evaluating the expression (64) at $b_j = \varepsilon_{1,j} = \ldots = \varepsilon_{n_F,j} = 0$ where $K_{i,j}$ is given by (58) and $\hat{H}_{i,j}$ solves the following implicit equation obtained by using (59) and the definition of $\hat{H}_{i,j}$

$$\begin{aligned}
\hat{H}_{i,j} &= \left(\rho_i \mathbb{V}[\tilde{p}_j|\mathcal{I}_i] + \frac{1}{\phi}\right)^{-1} \\
&= \left(\rho_i \left(F_j^2 + \frac{1}{n_{d_i}} \sum_{i'=1}^{n_F} h_{i',j}^2 - \frac{\left(F_j + \frac{h_{i,j}}{n_{d_i}}\right)^2}{1 + \frac{1}{n_{d_i}}}\right) + \frac{1}{\phi}\right)^{-1}
\end{aligned} \tag{68}$$

As these calculations can be done independently for each attribute and yield identical expressions, we can write the final result in matrix form as:

$$\begin{aligned}
\hat{\boldsymbol{H}}_i &= \left[\rho_i \left(\boldsymbol{F}\boldsymbol{F}' + \frac{1}{n_{d_i}} \sum_{i'=1}^{n} \boldsymbol{h}_{i'}^2 - \frac{1}{1 + \frac{1}{n_{d_i}}} \left(\boldsymbol{F} + \frac{\boldsymbol{h}_i}{n_{d_i}}\right)\left(\boldsymbol{F} + \frac{\boldsymbol{h}_i}{n_{d_i}}\right)'\right) + \frac{\boldsymbol{I}_N}{\phi}\right]^{-1} \\
\boldsymbol{F} &= \boldsymbol{I}_N - \frac{1}{\phi} \sum_{i'=1}^{n_F} \hat{\boldsymbol{H}}_{i'} \boldsymbol{K}_{i'} \\
\boldsymbol{h}_i &= -\frac{1}{\phi} \hat{\boldsymbol{H}}_{\boldsymbol{i}} \boldsymbol{K}_{\boldsymbol{i}} \\
\boldsymbol{K}_i &= \frac{\boldsymbol{F} + \frac{\boldsymbol{h}_i}{n_{d_i}}}{1 + \frac{1}{n_{d_i}}} \\
\boldsymbol{D} &= \left(\boldsymbol{I}_N + \frac{1}{\phi} \sum_{i=1}^{n_F} \hat{\boldsymbol{H}}_i\right)^{-1} \left(\bar{p} + \frac{1}{\phi} \sum_{i=1}^{n_F} \hat{\boldsymbol{H}}_i \tilde{\boldsymbol{c}}_i\right)
\end{aligned}$$

where $\hat{\boldsymbol{H}}_i, \boldsymbol{F}, \boldsymbol{h}_i, \boldsymbol{K}, \boldsymbol{D}$ are all diagonal matrices with entry (j, j) corresponding to the j^{th} attribute.

□

As we can see, apart from the differences in the expression for $\hat{\boldsymbol{H}}_i$ and $\hat{\boldsymbol{K}}_i$, the expressions for expected price and expected quantity take similar form as for the model with private shocks and public data. For example, output and expected output are

$$\begin{aligned}
\tilde{\boldsymbol{q}}_i &= \hat{\boldsymbol{H}}_i(\bar{\boldsymbol{p}}^M - \boldsymbol{c}_i + \boldsymbol{K}_i \boldsymbol{s}_i) \\
\mathbb{E}[\tilde{\boldsymbol{q}}_i] &= \hat{\boldsymbol{H}}_i(\bar{\boldsymbol{p}}^M - \boldsymbol{c}_i).
\end{aligned} \tag{69}$$

Of course, these expressions still contain endogenous variables. Appendix C. digs deeper to express price in terms of underlying parameters. It also shows conditions under which H, F and K have derivatives with respect to more data that have the same sign as in the public information model. Because H, F and K react similarly to more data, many of the the same results will hold for both the models.

A.4. Markups

In this section, we define the markups based on different levels of aggregation and characterize how various components of the markups are affected by data holding the marginal cost term $\tilde{c}_{i,j}$ as fixed. The next section deals with the optimal choice of marginal cost. Finally, propositions 1 - 4 put these two together and discuss the impact of more data while accounting for its effect on the choice of marginal cost. For the ease of exposition, wherever we delve into the details related to the number of data points in the next section, we show the calculations using the model with private shocks and public data. For the reader interested in more details about these calculations for the model with aggregate shocks and private data, we refer them to Appendix C..

PRODUCT-LEVEL MARKUP The product-level markup of product j for firm i is defined as $M_{i,j}^{\tilde{p}} := E[\tilde{p}_{i,j}]/\tilde{c}_{i,j}$. The expected price of product j by firm i is the expectation of (54). Note that $b_{i,j}$ and $s_{i,j}$ are mean-zero random variables. Thus from (54), the mean of $\tilde{p}_{i,j}$ is $\bar{p}_j^M$. Thus, the product-level markup, averaged over products (attributes) and over firms is

$$\overline{M}^{\tilde{p}} = \frac{1}{N} \frac{1}{n_F} \sum_{i=1}^{n_F} \sum_{j=1}^{N} M_{i,j}^{\tilde{p}} = \frac{1}{n_F N} \sum_{i=1}^{n_F} \sum_{j=1}^{N} \frac{E[\tilde{p}_{i,j}]}{\tilde{c}_{i,j}} = \frac{1}{n_F N} \sum_{i=1}^{n_F} \sum_{j=1}^{N} \frac{\bar{p}_j^M}{\tilde{c}_{i,j}} \tag{70}$$

To see the effect of obtaining more data points on the average product level mark-up, we need to calculate $\partial \overline{M}^{\tilde{p}} / \partial n_{di}$. Using the expression above, this becomes

$$\frac{\partial \overline{M}^{\tilde{p}}}{\partial n_{di}} = \frac{1}{n_F N} \sum_{i=1}^{n_F} \sum_{j=1}^{N} \frac{1}{\tilde{c}_{i,j}} \frac{\partial \bar{p}_j^M}{\partial n_{di}} \tag{71}$$

Using (51) and defining $\bar{H}_j = \sum_{i'} \hat{H}_{i',j}$, we get

$$\begin{aligned} \frac{\partial \bar{p}_j^M}{\partial n_{di}} &= \frac{(1+\frac{\bar{H}_j}{\phi})\frac{1}{\phi}\frac{\partial C_j}{\partial n_{di}} - (\bar{p}_j + \frac{C_j}{\phi})\frac{1}{\phi}\frac{\partial \bar{H}_j}{\partial n_{di}}}{(1+\frac{\bar{H}_j}{\phi})^2} \\ &= \frac{(1+\frac{\bar{H}_j}{\phi})\frac{1}{\phi}\frac{\partial C_j}{\partial n_{di}} - (1+\frac{\bar{H}_j}{\phi})\frac{\bar{p}_j^M}{\phi}\frac{\partial \bar{H}_j}{\partial n_{di}}}{(1+\frac{\bar{H}_j}{\phi})^2} \\ &= \frac{\frac{1}{\phi}\frac{\partial C_j}{\partial n_{di}} - \frac{\bar{p}_j^M}{\phi}\frac{\partial \bar{H}_j}{\partial n_{di}}}{(1+\frac{\bar{H}_j}{\phi})} \end{aligned} \tag{72}$$

Using the expressions for $\bar{H}_j$ and C_j from (48), we get

$$\begin{aligned} \frac{\partial \bar{H}_j}{\partial n_{di}} &= \frac{\partial \hat{H}_{i,j}}{\partial n_{di}} + \sum_{i' \neq i} \frac{\partial \hat{H}_{i',j}}{\partial n_{di}} \\ \frac{\partial C_j}{\partial n_{di}} &= \tilde{c}_{i,j}\frac{\partial \hat{H}_{i,j}}{\partial n_{di}} + \sum_{i' \neq i} \tilde{c}_{i',j}\frac{\partial \hat{H}_{i',j}}{\partial n_{di}} \end{aligned} \tag{73}$$

As a firm's sensitivity does not depend on any other firm's data, the cross-derivative terms are all zero i.e. $\frac{\partial \hat{H}_{i',j}}{\partial n_{di}} = 0 \forall i' \neq i$. Now, using (40) to replace the variance term in (43) in terms of n_{di} and taking the partial derivative yields

$$\frac{\partial \hat{H}_{i,j}}{\partial n_{di}} = \rho_i \left(\frac{\hat{H}_{i,j}}{1+n_{di}} \right)^2 > 0 \tag{74}$$

Using the two expressions above, we can now write:

$$\begin{aligned} \frac{\partial \bar{p}_j^M}{\partial n_{di}} &= \frac{1}{\phi}\frac{\partial \hat{H}_{i,j}}{\partial n_{di}} \left(1+\frac{\bar{H}_j}{\phi}\right)^{-1} (\tilde{c}_{i,j} - \bar{p}_j^M) \\ &= \frac{\rho_i}{\phi} \left(\frac{\hat{H}_{i,j}}{1+n_{di}} \right)^2 \left(1+\frac{\bar{H}_j}{\phi}\right)^{-1} (\tilde{c}_{i,j} - \bar{p}_j^M) < 0 \end{aligned} \tag{75}$$

where the last inequality follows from the fact that for a firm to have non-negative expected profits, the cost $\tilde{c}_{i,j}$ must be less than the average market price $\bar{p}_j^M$. If it were not so, the firm i would not choose to produce positive quantity (see (41)).

Using this result in (71), we obtain that $\partial \overline{M}^{\tilde{p}} / \partial n_{di} < 0$

FIRM-LEVEL MARKUP The firm-level markup for firm i is the quantity-weighted prices divided by quantity-weighted costs:

$$M_i^f = \frac{\mathbb{E}[\tilde{\boldsymbol{q}}_i' \tilde{\boldsymbol{p}}_i]}{\mathbb{E}[\tilde{\boldsymbol{q}}_i' \tilde{\boldsymbol{c}}_i]} = \frac{\mathbb{E}[\tilde{\boldsymbol{q}}_i]' \mathbb{E}[\tilde{\boldsymbol{p}}_i] + \textbf{Trace}[\textbf{Cov}(\tilde{\boldsymbol{p}}_i, \tilde{\boldsymbol{q}}_i)]}{\mathbb{E}[\tilde{\boldsymbol{q}}_i' \tilde{\boldsymbol{c}}_i]} \tag{76}$$

As for the denominator, note that $\mathbb{E}\tilde{q}_{i,j} = \hat{H}_{i,j}(\bar{p}_j^M - \tilde{c}_{i,j})$. So, the equilibrium output increases with more data since

$$\begin{aligned}
\frac{\partial \mathbb{E}\tilde{q}_{i,j}}{\partial n_{di}} &= (\bar{p}_j^M - \tilde{c}_{i,j})\frac{\partial \hat{H}_{i,j}}{\partial n_{di}} + \hat{H}_{i,j}\frac{\partial \bar{p}_j^M}{\partial n_{di}} \\
&= (\bar{p}_j^M - \tilde{c}_{i,j})\frac{\partial \hat{H}_{i,j}}{\partial n_{di}} + \frac{\hat{H}_{i,j}}{\phi}\left(\frac{\tilde{c}_{i,j} - \bar{p}_j^M}{1 + \frac{\bar{H}_j}{\phi}}\right)\frac{\partial \hat{H}_{i,j}}{\partial n_{di}} \\
&= (\bar{p}_j^M - \tilde{c}_{i,j})\frac{\partial \hat{H}_{i,j}}{\partial n_{di}}\left(1 - \frac{\hat{H}_{i,j}}{\phi + \bar{H}_j}\right) > 0
\end{aligned} \tag{77}$$

where the second equality follows from (72) and (73), and the last inequality uses $\bar{H}_j = \sum_{i'} \hat{H}_{i',j} > \hat{H}_{i,j}$ and $\bar{p}_j^M > \tilde{c}_{i,j}$ as argued earlier.

Although price decreases with more data, the revenue rises.

$$\begin{aligned}
\frac{\partial \mathbb{E}\tilde{q}_{i,j}\mathbb{E}\tilde{p}_{i,j}}{\partial n_{di}} &= \mathbb{E}\tilde{q}_{i,j}\frac{\partial \mathbb{E}\tilde{p}_{i,j}}{\partial n_{di}} + \mathbb{E}\tilde{p}_{i,j}\frac{\partial \mathbb{E}\tilde{q}_{i,j}}{\partial n_{di}} \\
&= \hat{H}_{i,j}(\bar{p}_j^M - \tilde{c}_{i,j})\frac{\partial \bar{p}_j^M}{\partial n_{di}} + \bar{p}_j^M(\bar{p}_j^M - \tilde{c}_{i,j})\frac{\partial \hat{H}_i, j}{\partial n_{di}}\left(1 - \frac{\hat{H}_{i,j}}{\phi + \bar{H}_j}\right) \\
&= (\bar{p}_j^M - \tilde{c}_{i,j})\left(\frac{\hat{H}_{i,j}}{\phi}\frac{\partial \hat{H}_{i,j}}{\partial n_{di}}\left(1 + \frac{\bar{H}_j}{\phi}\right)^{-1}(\tilde{c}_{i,j} - \bar{p}_j^M) + \bar{p}_j^M\frac{\partial \hat{H}_i, j}{\partial n_{di}}\left(1 - \frac{\hat{H}_{i,j}}{\phi + \bar{H}_j}\right)\right) \\
&= \frac{\partial \hat{H}_{i,j}}{\partial n_{di}}\left(\frac{\bar{p}_j^M - \tilde{c}_{i,j}}{\phi + \bar{H}}\right)\left(\hat{H}_{i,j}(\tilde{c}_{i,j} - \bar{p}_j^M) + \bar{p}_j^M\left(\phi + \bar{H}_j - \hat{H}_{i,j}\right)\right) \\
&= \frac{\partial \hat{H}_{i,j}}{\partial n_{di}}\left(\frac{\bar{p}_j^M - \tilde{c}_{i,j}}{\phi + \bar{H}}\right)\left(\hat{H}_{i,j}\tilde{c}_{i,j} + \bar{p}_j^M\left(\phi - \hat{H}_{i,j} + \sum_{k \neq i}\hat{H}_{k,j}\right)\right) > 0
\end{aligned} \tag{78}$$

The second equality above uses (77), the third equality uses (75), and the final inequality follows from the fact that $\phi > \hat{H}_{i,j}$ by definition (see (43)).

COST-WEIGHTED INDUSTRY MARKUP The industry markup weighted by cost is

$$M^c := \frac{\mathbb{E}\left[\sum_{i=1}^{n_F} \tilde{\boldsymbol{q}}_i' \tilde{\boldsymbol{p}}_i\right]}{\mathbb{E}\left[\sum_{i=1}^{n_F} \tilde{\boldsymbol{q}}_i' \tilde{\boldsymbol{c}}_i\right]} = \frac{\sum_{i=1}^{n_F} \mathbb{E}\left[\tilde{\boldsymbol{q}}_i' \tilde{\boldsymbol{p}}_i\right]}{\sum_{i=1}^{n_F} \mathbb{E}\left[\tilde{\boldsymbol{q}}_i' \tilde{\boldsymbol{c}}_i\right]} = \sum_{i=1}^{n_F} w_i^{cost} M_i^f \quad \text{where} \quad w_i^{cost} = \frac{\mathbb{E}\left[\tilde{\boldsymbol{q}}_i' \tilde{\boldsymbol{c}}_i\right]}{\sum_{i=1}^{n_F} \mathbb{E}\left[\tilde{\boldsymbol{q}}_i' \tilde{\boldsymbol{c}}_i\right]}. \tag{79}$$

Denote $w_{i,j}^{cost} = \frac{\mathbb{E}[\tilde{q}_{i,j}\tilde{c}_{i,j}]}{\sum_{i=1}^{n_F} \mathbb{E}[\tilde{\boldsymbol{q}}_i' \tilde{\boldsymbol{c}}_i]}$. Then, $w_i^{cost} = \sum_{j=1}^{N} w_{i,j}^{cost}$. The weight $w_{i,j}^{cost}$ increases with more data n_{di} as

$$\frac{\partial w_{i,j}^{cost}}{\partial n_{di}} = \frac{\tilde{\boldsymbol{c}}_{i,j}}{\left(\sum_{i=1}^{n_F} \mathbb{E}\left[\tilde{\boldsymbol{q}}_i' \tilde{\boldsymbol{c}}_i\right]\right)^2}\left[\frac{\partial \mathbb{E}\tilde{\boldsymbol{q}}_{i,j}}{\partial n_{di}}\left(\sum_{k=1,k\neq i}^{n_F} \mathbb{E}\left[\tilde{\boldsymbol{q}}_k' \tilde{\boldsymbol{c}}_k\right]\right) - \mathbb{E}\tilde{\boldsymbol{q}}_{i,j}\left(\sum_{k=1,k\neq i}^{n_F} \frac{\partial \mathbb{E}(\tilde{\boldsymbol{q}}_{k,j})}{\partial n_{di}}\tilde{\boldsymbol{c}}_{k,j}\right)\right] > 0 \tag{80}$$

The last inequality is due to the existing results $\frac{\partial \mathbb{E}\tilde{\boldsymbol{q}}_{i,j}}{\partial n_{di}} > 0$ and $\frac{\partial \mathbb{E}(\tilde{\boldsymbol{q}}_{k,j})}{\partial n_{di}} = \hat{H}_{k,j}\frac{\partial \bar{\boldsymbol{p}}_j^M}{\partial n_{di}} < 0$.

SALES-WEIGHTED INDUSTRY MARKUP The industry markup weighted by sales is

$$M^s := \sum_{i=1}^{n_F} w_i^s M_i^f = \frac{\sum_{i=1}^{n_F} \frac{\mathbb{E}^2[\tilde{\boldsymbol{q}}_i' \tilde{\boldsymbol{p}}_i]}{\mathbb{E}[\tilde{\boldsymbol{q}}_i' \tilde{\boldsymbol{c}}_i]}}{\sum_{i=1}^{n_F} \mathbb{E}\left[\tilde{\boldsymbol{q}}_i' \tilde{\boldsymbol{p}}_i\right]} \quad \text{where} \quad w_i^s = \frac{\mathbb{E}\left[\tilde{\boldsymbol{q}}_i' \tilde{\boldsymbol{p}}_i\right]}{\sum_{i=1}^{n_F} \mathbb{E}\left[\tilde{\boldsymbol{q}}_i' \tilde{\boldsymbol{p}}_i\right]}. \tag{81}$$

A.5. Firm's choice of marginal cost

In the previous section, we treated the marginal cost $\tilde{c}_{i,j}$ as fixed. In this section, we discuss how a firm makes the choice of marginal cost.

EX-ANTE EXPECTED PAYOFF To solve for the firms' cost choices, we need to solve for expected utility of each firm, before the realization of data signals are observed. This before-data expectation is what we call ex-ante. Substituting in the definition of the risk-adjusted firm profit into the objective function (2), one can re-arrange to write the ex-ante payoff as additively separate in each attribute (or each good, if goods are independent):

$$\begin{aligned}
\mathbb{E}[U_i] &= \mathbb{E}\left[\tilde{\boldsymbol{q}}_i'\left(\mathbb{E}[\tilde{\boldsymbol{p}}_i|\mathcal{I}_i]-\tilde{\boldsymbol{c}}_i\right)\right]-\frac{\rho_i}{2}\mathbb{E}\left[\tilde{\boldsymbol{q}}_i'\mathbb{V}\left[\tilde{\boldsymbol{p}}_i|\mathcal{I}_i\right]\tilde{\boldsymbol{q}}_i\right]-g(\chi_c,\tilde{\boldsymbol{c}}_i)\\
&= \mathbb{E}\left[\sum_{j=1}^{N}\tilde{q}_{i,j}\left(\mathbb{E}[\tilde{p}_{i,j}|\mathcal{I}_i]-\tilde{c}_{i,j}\right)\right]-\frac{\rho_i}{2}\mathbb{E}\left[\sum_{j=1}^{N}\tilde{q}_{i,j}\mathbb{V}\left[\tilde{p}_{i,j}|\mathcal{I}_i\right]\tilde{q}_{i,j}\right]-\sum_{j=1}^{N}g_j(\chi_c,\tilde{c}_{i,j})\\
&= \sum_{j=1}^{N}\left(\mathbb{E}\left[\tilde{q}_{i,j}\left(\mathbb{E}[\tilde{p}_{i,j}|\mathcal{I}_i]-\tilde{c}_{i,j}\right)\right]-\frac{\rho_i}{2}\mathbb{E}\left[\tilde{q}_{i,j}^2\mathbb{V}\left[\tilde{p}_{i,j}|\mathcal{I}_i\right]\right]-g_j(\chi_c,\tilde{c}_{i,j})\right) \quad = \sum_{j=1}^{N}\mathbb{E}[U_{i,j}]
\end{aligned} \tag{82}$$

where $U_{i,j} = \left[\tilde{q}_{i,j}\left(\mathbb{E}[\tilde{p}_{i,j}|\mathcal{I}_i]-\tilde{c}_{i,j}\right)\right]-\frac{\rho_i}{2}\mathbb{E}\left[\tilde{q}_{i,j}^2\mathbb{V}\left[\tilde{p}_{i,j}|\mathcal{I}_i\right]\right]-g_j(\chi_c,\tilde{c}_{i,j})$ represents the utility of firm i from attribute j based on firm's data. In deriving the above expression, we have used the assumption that the signal noises are uncorrelated among attributes which makes $\mathbb{V}\left[\tilde{\boldsymbol{p}}_i|\mathcal{I}_i\right]$ a diagonal matrix. We have also used the assumption that $g(\chi_c,\tilde{\boldsymbol{c}}_i)$ is additively separable w.r.t to the attributes. Because of the independence of attributes, we can now simplify each $\mathbb{E}[U_{i,j}]$ term separately and add them up to get the total expected utility.

Note that the first term in the expected profits expression for good j is $\mathbb{E}\left[\tilde{q}_{i,j}\left(\mathbb{E}\left[\tilde{p}_{i,j}|\mathcal{I}_i\right]-\tilde{c}_{i,j}\right)\right]$. Using the first order condition (41), we can substitute $\left(\mathbb{E}\left[\tilde{p}_{i,j}|\mathcal{I}_i\right]-\tilde{c}_{i,j}\right)$ out with $\hat{H}_{i,j}^{-1}\tilde{q}_{i,j}$. Thus, the first term can be expressed as $\tilde{q}_{i,j}^2\hat{H}_{i,j}^{-1}$. That substitution allows us to factor out the $\tilde{q}_{i,j}^2$ term from the mean and variance and write the firm ex-ante objective for attribute j as

$$\begin{aligned}
\mathbb{E}U_{i,j} &= \mathbb{E}[\tilde{q}_{i,j}^2]\left(\hat{H}_{i,j}^{-1}-\frac{\rho_i}{2}\mathbb{V}\left[\tilde{p}_{i,j}|\mathcal{I}_i\right]\right)-g_j(\chi_c,\tilde{c}_{i,j})\\
&= \mathbb{E}[\tilde{q}_{i,j}^2]\left(\frac{1}{\phi}+\frac{\rho_i}{2}\mathbb{V}\left[\tilde{p}_{i,j}|\mathcal{I}_i\right]\right)-g_j(\chi_c,\tilde{c}_{i,j})\\
&= \frac{1}{2}\mathbb{E}[\tilde{q}_{i,j}^2]H_{i,j}^{-1}-g_j(\chi_c,\tilde{c}_{i,j})
\end{aligned} \tag{83}$$

where the last line follows from the relationship that $H^{-1}=\hat{H}^{-1}+\frac{1}{\phi}$ (eq. 46).

The formula for the ex-ante variance of output is $\mathbb{V}[\tilde{q}_{i,j}] = \hat{H}_{i,j}^2\mathbb{V}\left(\tilde{p}_{i,j}\right)$. Notice that this part of expected utility is independent of the firm's cost choices. Thus, $\partial\mathbb{E}[\tilde{q}_{i,j}^2]/\partial\tilde{c}_{i,j} = \partial\mathbb{E}[\tilde{q}_{i,j}]^2/\partial\tilde{c}_{i,j}$. This also implies that the sensitivity of output to expected price changes $H_{i,j}$ is also not dependent on cost c.

OPTIMAL CHOICE OF MARGINAL COST The first and second order condition for the optimal marginal cost choice $\tilde{c}_{i,j}$ is

$$\begin{aligned}
\frac{\partial\mathbb{E}[U_{i,j}]}{\partial\tilde{c}_{i,j}} &= \frac{1}{2}\frac{\partial\mathbb{E}[\tilde{q}_{i,j}]^2H_{i,j}^{-1}}{\partial\tilde{c}_{i,j}}-\frac{\partial g_j(\chi_c,\tilde{c}_{i,j})}{\partial\tilde{c}_{i,j}}=0\\
\frac{\partial^2\mathbb{E}[U_{i,j}]}{\partial\tilde{c}_{i,j}^2} &= \frac{H_{i,j}^{-1}}{2}\frac{\partial^2\mathbb{E}[\tilde{q}_{i,j}]^2}{\partial\tilde{c}_{i,j}^2}-\frac{\partial^2 g_j(\chi_c,\tilde{c}_{i,j})}{\partial\tilde{c}_{i,j}^2}\leq 0
\end{aligned} \tag{84}$$

Substitute for $\mathbb{E}\tilde{q}_{i,j}$ using the fact that $\mathbb{E}\tilde{q}_{i,j} = \hat{H}_{i,j}(\bar{p}_j^M-\tilde{c}_{i,j})$. Since signal noise is diagonal, we have $H_{i,j}^{-1} = \frac{2}{\phi}+\rho_i\mathbb{V}\left[b_i|\mathcal{I}_i\right]$ and $\mathbb{V}\left[b_i|\mathcal{I}_i\right]=(1+n_{di})^{-1}$. Thus, the FOC and SOC could be written as

$$\begin{aligned}
\frac{\partial\mathbb{E}[U_{i,j}]}{\partial\tilde{c_{i,j}}} &= H_{i,j}^{-1}\mathbb{E}[\tilde{q}_{i,j}]\frac{\partial\mathbb{E}[\tilde{q}_{i,j}]}{\partial\tilde{c}_{i,j}}-\frac{\partial g_j(\chi_c,\tilde{c}_{i,j})}{\partial\tilde{c}_{i,j}}\\
&= \left(\bar{p}_j^M-\tilde{c}_{i,j}\right)H_{i,j}^{-1}\hat{H}_{i,j}^2\left(\frac{\frac{1}{\phi}\hat{H}_{i,j}}{1+\frac{1}{\phi}\bar{H}}-1\right)-\frac{\partial g_j(\chi_c,\tilde{c}_{i,j})}{\partial\tilde{c}_{i,j}}=0\\
\frac{\partial^2\mathbb{E}[U_i]}{\partial\tilde{c}_{i,j}^2} &= \hat{H}_{i,j}^2H_{i,j}^{-1}\left(\frac{\frac{1}{\phi}\hat{H}_{i,j}}{1+\frac{1}{\phi}\bar{H}}-1\right)^2-\frac{\partial^2 g_j(\chi_c,\tilde{c}_{i,j})}{\partial\tilde{c}_{i,j}^2}
\end{aligned} \tag{85}$$

since the average market price p_j^M for attribute j is

$$\bar{p}_j^M = \frac{\bar{p}_j + \frac{1}{\phi}\sum_{s=1}^{n_F} \hat{H}_{s,j}\tilde{c}_{s,j}}{1 + \frac{1}{\phi}\bar{H}} \quad \text{and} \quad \frac{\partial \bar{p}_j^M}{\partial \tilde{c}_{i,j}} = \frac{\frac{1}{\phi}\hat{H}_{i,j}}{1 + \frac{1}{\phi}\bar{H}} \tag{86}$$

B. Proofs

In Appendix A., we established the equilibrium solution. In this appendix, we prove the main results discussed in the paper.

Proof of Lemma 2: Data-Investment Complementarity

Proof. To show this complementarity between information and costs, we first differentiate $\mathbb{E}\left[U_{i,j}\right]$ in (83) with respect to marginal cost. Here, $\tilde{c}_{ij}$ denotes firm i's marginal cost of producing attribute j. $\widehat{H}_{ij}$ denotes the jj-th entry of the diagonal matrix $\widehat{\boldsymbol{H}}_i$, which captures the sensitivity of i's production of attribute j to a marginal change in the expected profit of producing attribute j. We can simplify the expression derived in (85) a bit further as follows:

$$\begin{aligned}
\frac{\partial \mathbb{E}[U_{i,j}]}{\partial \tilde{c}_{i,j}} &= \left(\bar{p}_j^M - \tilde{c}_{i,j}\right) H_{i,j}^{-1}\hat{H}_{i,j}^2 \left(\frac{\frac{1}{\phi}\hat{H}_{i,j}}{1 + \frac{1}{\phi}\bar{H}_j} - 1\right) - \frac{\partial g_j(\chi_c, \tilde{c}_{i,j})}{\partial \tilde{c}_{i,j}} \\
\frac{\partial^2 \mathbb{E}[U_{i,j}]}{\partial \tilde{c}_{i,j}\partial \hat{H}_{i,j}} &= \left(\bar{p}_j^M - \tilde{c}_{i,j}\right) \frac{\partial}{\partial \hat{H}_{i,j}} \left(H_{i,j}^{-1}\hat{H}_{i,j}^2 \left(\frac{\frac{1}{\phi}\hat{H}_{i,j}}{1 + \frac{1}{\phi}\bar{H}_j} - 1\right)\right) \\
&= \left(\bar{p}_j^M - \tilde{c}_{i,j}\right) \frac{\partial}{\partial \hat{H}_{i,j}} \left(\left(1 + \frac{\hat{H}_{i,j}}{\phi}\right)\hat{H}_{i,j} \left(\frac{\frac{1}{\phi}\hat{H}_{i,j}}{1 + \frac{1}{\phi}\bar{H}_j} - 1\right)\right) \\
&= \frac{\bar{p}_j^M - \tilde{c}_{i,j}}{\phi} \frac{\partial}{\partial \hat{H}_{i,j}} \left(\left(\phi + \hat{H}_{i,j}\right)\hat{H}_{i,j} \left(\frac{\hat{H}_{i,j}}{\phi + \bar{H}_j} - 1\right)\right)
\end{aligned} \tag{87}$$

where the third equality uses the fact that $H_{i,j}^{-1} = \hat{H}_{i,j}^{-1} + \frac{1}{\phi}$.

Let $T = \hat{H}_{i,j}(\phi + \hat{H}_{i,j})\left(\frac{\hat{H}_{i,j}}{\phi + \bar{H}_j} - 1\right)$. Taking log and differentiating w.r.t $\hat{H}_{i,j}$, we obtain

$$\begin{aligned}
\log T &= \log \hat{H}_{i,j} + \log(\phi + \hat{H}_{i,j}) + \log\left(\frac{\hat{H}_{i,j}}{\phi + \bar{H}_j} - 1\right) \\
\frac{1}{T}\frac{\partial T}{\hat{H}_{i,j}} &= \frac{1}{\hat{H}_{i,j}} + \frac{1}{\phi + \hat{H}_{i,j}} + \left(\frac{\hat{H}_{i,j}}{\phi + \bar{H}_j} - 1\right)^{-1} \left(\frac{(\phi + \bar{H}_j) - \hat{H}_{i,j}\sum_{k=1}^{n_F}\frac{\partial \hat{H}_{k,j}}{\partial \hat{H}_{i,j}}}{(\phi + \bar{H}_j)^2}\right) \\
&= \frac{1}{\hat{H}_{i,j}} + \frac{1}{\phi + \hat{H}_{i,j}} + \left(\frac{\phi + \bar{H}_j}{\hat{H}_{i,j} - \phi + \bar{H}_j}\right) \left(\frac{(\phi + \bar{H}_j) - \hat{H}_{i,j}\sum_{k=1}^{n_F}\frac{\partial \hat{H}_{k,j}}{\partial \hat{H}_{i,j}}}{(\phi + \bar{H}_j)^2}\right)
\end{aligned} \tag{88}$$

For the model with private shocks and public data, for $i \neq k$, $\frac{\partial \hat{H}_{i,j}}{\partial \hat{H}_{k,j}} = 0$. So, the above equality becomes

$$\begin{aligned}
\frac{1}{T}\frac{\partial T}{\partial \hat{H}_{i,j}} &= \frac{1}{\hat{H}_{i,j}} + \frac{1}{\phi + \hat{H}_{i,j}} + \frac{1}{\hat{H}_{i,j} - \phi - \bar{H}_j}\left(\frac{\phi + \bar{H}_j - \hat{H}_{i,j}}{\phi + \bar{H}_j}\right) \\
&= \frac{1}{\hat{H}_{i,j}} + \frac{1}{\phi + \hat{H}_{i,j}} - \frac{1}{\phi + \bar{H}_j} \\
\implies \frac{\partial T}{\partial \hat{H}_{i,j}} &= T\left(\frac{1}{\hat{H}_{i,j}} + \frac{1}{\phi + \hat{H}_{i,j}} - \frac{1}{\phi + \bar{H}_j}\right)
\end{aligned} \tag{89}$$

Note that the as $\bar{H}_j = \sum_{k=1}^{n_F} \hat{H}_{k,j} > \hat{H}_{i,j}$, we have that $\frac{1}{\phi+\hat{H}_{i,j}} > \frac{1}{\phi+\bar{H}_j}$. So, the second term on the RHS above is positive. However, $T = \hat{H}_{i,j}(\phi + \hat{H}_{i,j})\left(\frac{\hat{H}_{i,j}}{\phi+\bar{H}_j} - 1\right) < 0$ as $\phi + \sum_{k=1}^{n_F} \hat{H}_{k,j} > \hat{H}_{i,j}$. Hence, we get that $\frac{\partial T}{\partial \hat{H}_{i,j}} < 0$ and therefore, $\frac{\partial^2 \mathbb{E}[U_i]}{\partial \tilde{c}_{ij} \partial \hat{H}_{i,j}} < 0$, which means the marginal benefit from reducing costs is higher (more negative) when firms have better information (higher sensitivity $\hat{H}_{i,j}$).

Another way to look at this result is to note that from FOC (85), and the fact that $H_{i,j}^{-1} = \hat{H}_{i,j}^{-1} + \frac{1}{\phi}$, we have

$$\frac{\partial \mathbb{E}[U_{i,j}]}{\partial \tilde{c}_{i,j}} = \left(\bar{p}_j^M - \tilde{c}_{i,j}\right)\left(1 + \frac{\hat{H}_{i,j}}{\phi}\right)\hat{H}_{i,j}\left(\frac{\frac{1}{\phi}\hat{H}_{i,j}}{1 + \frac{1}{\phi}\bar{H}_j} - 1\right) - \frac{\partial g_j(\chi_c, \tilde{c}_{i,j})}{\partial \tilde{c}_{i,j}} = 0 \tag{90}$$

Define $F(\tilde{c}_{i,j}, \hat{H}_{i,j}) = \left(\bar{p}_j^M - \tilde{c}_{i,j}\right) H_{i,j}^{-1} \hat{H}_{i,j}^2 \left(\frac{\frac{1}{\phi}\hat{H}_{i,j}}{1+\frac{1}{\phi}\bar{H}_j} - 1\right) - \frac{\partial g_j(\chi_c, \tilde{c}_{i,j})}{\partial \tilde{c}_{i,j}}$ Close to the optimal choice of $(\tilde{c}_{i,j}, \hat{H}_{i,j})$, we have that $F(\tilde{c}_{i,j}, \hat{H}_{i,j}) = 0$. Using implicit function theorem, we can obtain

$$\frac{d\tilde{c}_{i,j}}{d\hat{H}_{i,j}} = -\frac{\frac{\partial F}{\partial \hat{H}_{i,j}}}{\frac{\partial F}{\partial \tilde{c}_{i,j}}} \tag{91}$$

Using the notation above, $\frac{\partial F}{\partial \hat{H}_{i,j}} = (\bar{p}_j^M - \tilde{c}_{i,j})\frac{\partial T}{\partial \hat{H}_{i,j}} < 0$. Also, $\frac{\partial F}{\partial \tilde{c}_{i,j}} = \frac{\partial}{\partial \tilde{c}_{i,j}}\left(\frac{\partial \mathbb{E}[U_{i,j}]}{\partial \tilde{c}_{i,j}}\right) < 0$ by the second order condition (85). Combining the above two results, we get the required result

$$\frac{d\tilde{c}_{i,j}}{d\hat{H}_{i,j}} < 0 \tag{92}$$

□

Proof of Lemma 3: Greater investment raises a firm's product markup.

Proof. More investment would lower marginal cost $\tilde{\mathbf{c}}_{i,j}$. The markup effect is

$$\frac{\partial M_{i,j}^{\tilde{p}}}{\partial \tilde{c}_{i,j}} = \frac{\frac{\partial \bar{p}_j^M}{\partial \tilde{c}_{i,j}}\tilde{c}_{i,j} - \bar{p}_j^M}{\tilde{c}_{i,j}^2} = \frac{\frac{1}{\phi}\hat{H}_{i,j}\tilde{c}_{i,j} - \bar{p}_j - \frac{1}{\phi}\sum_{s=1}^{n_F}\hat{H}_{s,j}\tilde{c}_{s,j}}{\tilde{c}_{i,j}^2\left(1 + \bar{H}_j\right)} = -\frac{\bar{p}_j + \frac{1}{\phi}\sum_{s=1, s\neq i}^{n_F}\hat{H}_{s,j}\tilde{c}_{s,j}}{\tilde{c}_{i,j}^2\left(1 + \bar{H}_j\right)} \leq 0 \tag{93}$$

The negative derivative confirms that more investment leads to higher attribute-level markup. Similarly, for the other attributes j' we have $\frac{\partial M_{i,j'}^{\tilde{p}}}{\partial \tilde{\mathbf{c}}_{i,j}} = 0$.

Next, differentiate this product markup with respect to the marginal cost of attribute j. Consider the markup on product k that used attribute j ($A_{kj} > 0$). This markup is $\sum_j A_{k,j}\mathbb{E}[\tilde{p}_{i,j}]/(\sum_j A_{kj}\tilde{\mathbf{c}}_{i,j})$. Its derivative is

$$\frac{dM_{i,k}}{d\tilde{c}_{i,j}} = \frac{[\sum_j A_{kj}\tilde{c}_{i,j}]A_{kj}\frac{\partial}{\partial \tilde{c}_{i,j}}\mathbb{E}[\tilde{p}_{i,j}] - [\sum_j A_{k,j}\mathbb{E}[\tilde{p}_{i,j}]]A_{kj}}{[\sum_j A_{kj}\tilde{c}_{i,j}]^2} \tag{94}$$

$$= \frac{A_{kj}}{\sum_j A_{kj}\tilde{c}_{i,j}}\left[\frac{\partial}{\partial \tilde{c}_{i,j}}\mathbb{E}[\tilde{p}_{i,j}] - M_{i,k}\right]. \tag{95}$$

We know that $\frac{\partial}{\partial \tilde{c}_{i,j}}\mathbb{E}[p_{i,j}] < M_{i,j}^{\tilde{p}}$ because earlier in the proof, we established that

$$\frac{dM_{i,j}^{\tilde{p}}}{d\tilde{c}_{i,j}} = \frac{1}{\tilde{c}_{i,j}}\left[\frac{\partial}{\partial \tilde{c}_{i,j}}\mathbb{E}[\tilde{p}_{i,j}] - M_{i,j}^{\tilde{p}}\right] \leq 0. \tag{96}$$

Therefore, (95) is negative if the markup on product k is greater than the markup on attribute j: $M_{i,k} \geq M_{i,j}^{\tilde{p}}$. □

Proof of Lemma 4: (Risk premium channel) Product-level markup decreases in data. When investment is sufficiently inflexible (high χ_c), and product i loads positively on all attributes ($a_{ij} \geq 0$), then the product markup $\mathbb{E}\left(p_i/c_i\right) = \mathbb{E}\left(p_i\right)/c_i$ is decreasing in data.

Proof. Assume each firm is endowed with a fixed investment ($\boldsymbol{c}_i$). By continuity, the result will extend to cases where the investment is close to fixed, which is when χ_c is sufficiently high. The markup on the attribute j, produced by firm i is $M^{\tilde{p}}_{i,j} := \mathbb{E}[\tilde{p}_{i,j}]/\tilde{c}_{i,j}$. The average markup on the attributes is

$$\overline{M}^{\tilde{p}} = \frac{1}{N}\frac{1}{n_F}\sum_{i=1}^{n_F}\sum_{j=1}^{N} M^{\tilde{p}}_{i,j} = \frac{1}{n_F N}\sum_{i=1}^{n_F}\sum_{j=1}^{N}\frac{\mathbb{E}[\tilde{p}_{i,j}]}{\tilde{c}_{i,j}} = \frac{1}{n_F N}\sum_{i=1}^{n_F}\sum_{j=1}^{N}\frac{\bar{p}^M_j}{\tilde{c}_{i,j}} \tag{97}$$

The j^{th} term of equilibrium price $\bar{\boldsymbol{p}}^M$ is

$$\bar{p}^M_j = \frac{\phi\bar{p}_j + \sum_{i=1}^{n_F}\hat{H}_{i,j}\tilde{c}_{i,j}}{\phi + \bar{H}_j} \quad \text{where} \quad \hat{H}_{i,j} = \left(\rho_i \mathbf{Var}[\tilde{p}_{i,j}|\mathcal{I}_i] + \frac{1}{\phi}\right)^{-1} \quad \text{and} \quad \bar{H}_j = \sum_{i=1}^{n_F}\hat{H}_{i,j} \tag{98}$$

The positive output means $\bar{p}^M_j \geq \tilde{c}_{i,j}$, thus

$$\begin{aligned}
\frac{\partial \bar{p}^M_j}{\partial \hat{H}_{i,j}} &= (\phi + \bar{H}_j)^{-2}\left((\phi + \bar{H}_j)\sum_{k=1}^{n_F}\tilde{c}_{k,j}\frac{\partial \hat{H}_{k,j}}{\partial \hat{H}_{i,j}} - \left(\phi\bar{p}_j + \sum_{i=1}^{n_F}\hat{H}_{i,j}\tilde{c}_{i,j}\right)\sum_{k=1}^{n_F}\frac{\partial \hat{H}_{k,j}}{\partial \hat{H}_{i,j}}\right)\\
&= (\phi + \bar{H}_j)^{-2}\left((\phi + \bar{H}_j)\sum_{k=1}^{n_F}\tilde{c}_{k,j}\frac{\partial \hat{H}_{k,j}}{\partial \hat{H}_{i,j}} - \bar{p}^M_j(\phi + \bar{H}_j)\sum_{k=1}^{n_F}\frac{\partial \hat{H}_{k,j}}{\partial \hat{H}_{i,j}}\right)\\
&= (\phi + \bar{H}_j)^{-1}\sum_{k=1}^{n_F}(\tilde{c}_{k,j} - \bar{p}^M_j)\frac{\partial \hat{H}_{k,j}}{\partial \hat{H}_{i,j}}\\
&= (\phi + \bar{H}_j)^{-1}\left(\tilde{c}_{k,j} - \bar{p}^M_j + \sum_{k\neq i}(\tilde{c}_{k,j} - \bar{p}^M_j)\frac{\partial \hat{H}_{k,j}}{\partial \hat{H}_{i,j}}\right)
\end{aligned} \tag{99}$$

We know from earlier results that for the private shocks model, $\frac{\partial \hat{H}_{k,j}}{\partial \hat{H}_{i,j}} = 0$ for $i \neq k$. For the aggregate shocks model also, $\frac{\partial \hat{H}_{k,j}}{\partial \hat{H}_{i,j}}$ is negligible and can be ignored in comparison to the effect of increase in $\hat{H}_{i,j}$. Therefore,

$$\frac{\partial \bar{p}^M_j}{\partial \hat{H}_{i,j}} = (\phi + \bar{H}_j)^{-1}\left(\tilde{c}_{k,j} - \bar{p}^M_j\right) \leq 0 \tag{100}$$

Since the price of a good is a_i times the vector of attribute prices, and all the attribute prices are decreasing in data, the good price and thus the product-level markup is decreasing in data as well.

We prove the negative first order derivative for fixed choices of cost $\tilde{\boldsymbol{c}}_i$, which corresponds to infinitely high marginal cost $\chi_c \to \infty$. This result is strictly negative and continuous in $\tilde{\boldsymbol{c}}_i$. If we assume χ_c is sufficiently high, this is arbitrarily close to fixed c. By continuity, the inequality will still hold. □

□

Proof of Proposition 1: Product markups increase or decrease in data (net change).

Proof. The product-level markup is $M^{\tilde{p}}_{i,j} = \mathbb{E}[\tilde{p}_{i,j}]/\tilde{c}_{i,j} = \bar{p}^M_j/\tilde{c}_{i,j}$. Its partial derivative to the sensitivity $\hat{H}_{i,j}$ is

$$\frac{\partial M^{\tilde{p}}_{i,j}}{\partial n_{di}} = \frac{1}{\tilde{c}^2_{i,j}}\left(\frac{\partial \bar{p}^M_j}{\partial \hat{H}_{i,j}}\tilde{c}_{i,j} - \bar{p}^M_j\frac{\partial \tilde{c}_{i,j}}{\partial \hat{H}_{i,j}}\right)\frac{\partial \hat{H}_{i,j}}{\partial n_{di}} \tag{101}$$

We have already established that under the conditions considered for this proposition, $\frac{\partial \hat{H}_{i,j}}{\partial n_{di}} > 0$. From (51), we have

$\bar{p}_j^M = \left(\phi + \bar{H}_j\right)^{-1}\left(\phi\bar{p}_j + \sum_i \hat{H}_{i,j}\tilde{c_{i,j}}\right)$. For the private-shocks-public-data model, we have

$$\begin{aligned}\frac{\partial \bar{p}_j^M}{\partial \hat{H}_{i,j}} &= \frac{(\phi+\bar{H}_j)\tilde{c}_{i,j} - (\phi\bar{p}_j + \sum_k \hat{H}_{k,j}\tilde{c}_{k,j})}{(\phi+\bar{H}_j)^2} \\ &= \frac{(\phi+\bar{H}_j)\tilde{c}_{i,j} - \bar{p}_j^M(\phi+\bar{H}_j)}{(\phi+\bar{H}_j)^2} \\ &= \frac{\tilde{c}_{i,j} - \bar{p}_j^M}{\phi+\bar{H}_j} < 0\end{aligned} \tag{102}$$

Now, from Lemma 2, we have

$$\frac{\partial \tilde{c}_{i,j}}{\partial \hat{H}_{i,j}} < 0 \tag{103}$$

If marginal cost $\tilde{c}_{i,j}$ or price of risk ρ_i is sufficiently low, the second term in the numerator $-\bar{p}_j^M \frac{\partial \tilde{c}_{i,j}}{\partial n_{di}} > 0$ dominates the marginal effect, thus increasing product markups.

□

Proof of Proposition 2: The firm-level markup wedge increases in data.

Proof. The firm-level markup wedge is given by

$$M_i^f - \bar{M}_i^p = \frac{\mathbb{E}[\tilde{\boldsymbol{q}}_i'\tilde{\boldsymbol{p}}_i]}{\mathbb{E}[\tilde{\boldsymbol{q}}_i'\tilde{\boldsymbol{c}}_i]} - \frac{1}{N}\sum_{j=1}^{n_F}\mathbf{E}\left[\frac{\tilde{p}_{i,j}}{\tilde{c}_{i,j}}\right] \tag{104}$$

Data has free disposal. So, the expected utility $\mathbf{E}[U_i]$ of firm i must be (weakly) increasing in its data (holding the data of all other firms fixed). By Lemma 2 (data-investment complementarity), the upfront investment $(g(\chi_c, \tilde{\boldsymbol{c}}_i))$ is increasing in data. Therefore, it follows from (2) that when the price of risk is low, the expected profits of firm i must be increasing in data i.e. the the following must hold

$$\frac{\partial}{\partial n_{di}}\left(\mathbf{E}\left[\tilde{\mathbf{q}}_i'\tilde{\mathbf{p}}_i\right] - \mathbf{E}\left[\tilde{\mathbf{q}}_i'\tilde{\mathbf{c}}_i\right]\right) > 0 \tag{105}$$

(105) indicates that starting from any level of $(\tilde{\mathbf{p}}_i, \tilde{\mathbf{q}}_i, \tilde{\mathbf{c}}_i)$, when a firm i expects to get more data (n_{di}), either $\mathbf{E}\left[\tilde{\mathbf{q}}_i'\tilde{\mathbf{p}}_i\right]$ goes up or $\mathbf{E}\left[\tilde{\mathbf{q}}_i'\tilde{\mathbf{c}}_i\right]$ goes down, or both. Any of these three changes would result in an increase in firm markup M_i^f. Therefore, M_i^f goes up unambiguously. By Lemma 4, the average product mark-up $\bar{M}_i^p$ goes down with data (n_{di}). Therefore, the firm-level markup wedge $M_i^f - \bar{M}_i^p$ increases in data.

□

Proof of Proposition 4a: Wedge between cost-weighted firm markup and average firm markup.

This proof shows that high-data firms produce more on average. Therefore, they have larger impacts on cost-weighted industry markup, increasing the industry-level markup wedge.

Proof. The cost weight for firm i is

$$w_i^{cost} = \frac{\mathbb{E}\left[\tilde{\boldsymbol{q}}_i'\tilde{\boldsymbol{c}}_i\right]}{\sum_{k=1}^{n_F}\mathbb{E}\left[\tilde{\boldsymbol{q}}_k'\tilde{\boldsymbol{c}}_k\right]} = \frac{\sum_{l=1}^{N}\mathbb{E}[\tilde{q}_{i,l}]\tilde{c}_{i,l}}{\sum_{k=1}^{n_F}\sum_{l=1}^{N}\mathbb{E}[\tilde{q}_{k,l}]\tilde{c}_{k,l}} \tag{106}$$

We show below that this weight is increasing in data for the firm i. Taking the partial derivative of the weight with respect to the number of data points n_{di}, we get

$$\frac{\partial w_i^{cost}}{\partial n_{di}} = \frac{\left(\sum_{k=1}^{n_F}\mathbb{E}\left[\tilde{\boldsymbol{q}}_k'\tilde{\boldsymbol{c}}_k\right]\right)\left(\sum_{j=1}^{N}\tilde{c}_{i,j}\frac{\partial\mathbb{E}[\tilde{q}_{i,j}]}{\partial n_{di}}\right) - \mathbb{E}[\tilde{\boldsymbol{q}}_i'\tilde{\boldsymbol{c}}_i]\left(\sum_{j=1}^{N}\tilde{c}_{i,j}\frac{\partial\mathbb{E}[\tilde{q}_{i,j}]}{\partial n_{di}} + \sum_{k=1,k\neq i}^{n_F}\sum_{j=1}^{N}\tilde{c}_{k,j}\frac{\partial\mathbb{E}[\tilde{q}_{k,j}]}{\partial n_{di}}\right)}{\left(\sum_{k=1}^{n_F}\mathbb{E}\left[\tilde{\boldsymbol{q}}_k'\tilde{\boldsymbol{c}}_k\right]\right)^2} \tag{107}$$

Now, we use that fact that $\mathbb{E}[\tilde{q}_{k,j}] = \hat{H}_{k,j}(\bar{p}_j^M - \tilde{c}_{k,j})$. Differentiating this w.r.t n_{di}, we obtain

$$\frac{\partial\mathbb{E}[\tilde{q}_{k,j}]}{\partial n_{di}} = (\bar{p}_j^M - \tilde{c}_{k,j})\frac{\partial \hat{H}_{k,j}}{\partial n_{di}} + \hat{H}_{k,j}\frac{\partial \bar{p}_j^M}{\partial n_{di}} \tag{108}$$

(75) shows that $\frac{\partial \bar{p}_j^M}{\partial n_{di}} < 0$. For the model with private shocks and public data, $\hat{H}_{k,j}$ does not depend on the amount of data for any other firm. Therefore, for $k \neq i$, $\frac{\partial \hat{H}_{k,j}}{\partial n_{di}} = 0$. For the model with aggregate shocks and private data, $k \neq i$, $\frac{\partial \hat{H}_{k,j}}{\partial n_{di}}$ is negligible in comparison to the change in market price $\frac{\partial \bar{p}_j^M}{\partial n_{di}}$. Therefore, in either case, for $k \neq i$, $\frac{\partial \mathbb{E}[\tilde{q}_{k,j}]}{\partial n_{di}} < 0$. Using this fact in the (107), we obtain

$$\begin{aligned}\frac{\partial w_i^{cost}}{\partial n_{di}} &\geq \frac{\left(\sum_{k=1}^{n_F} \mathbb{E}\left[\tilde{\boldsymbol{q}}_k' \tilde{\boldsymbol{c}}_k\right]\right)\left(\sum_{j=1}^N \tilde{c}_{i,j} \frac{\partial \mathbb{E}[\tilde{q}_{i,j}]}{\partial n_{di}}\right) - \mathbb{E}[\tilde{\boldsymbol{q}}_i' \tilde{\boldsymbol{c}}_i]\left(\sum_{j=1}^N \tilde{c}_{i,j} \frac{\partial \mathbb{E}[\tilde{q}_{i,j}]}{\partial n_{di}}\right)}{\left(\sum_{k=1}^{n_F} \mathbb{E}\left[\tilde{\boldsymbol{q}}_k' \tilde{\boldsymbol{c}}_k\right]\right)^2} \\ &\geq \frac{\left(\sum_{k=1,k\neq i}^{n_F} \mathbb{E}\left[\tilde{\boldsymbol{q}}_k' \tilde{\boldsymbol{c}}_k\right]\right)\left(\sum_{j=1}^N \tilde{c}_{i,j} \frac{\partial \mathbb{E}[\tilde{q}_{i,j}]}{\partial n_{di}}\right)}{\left(\sum_{k=1}^{n_F} \mathbb{E}\left[\tilde{\boldsymbol{q}}_k' \tilde{\boldsymbol{c}}_k\right]\right)^2} > 0\end{aligned} \tag{109}$$

where the last inequality follows because we know from (77) that $\forall j \in [1, \ldots, N]$, $\frac{\partial \mathbb{E}[\tilde{q}_{i,j}]}{\partial n_{di}} > 0$. We know from (77) that high-data firms produce more on average and the result above indicates that these firms have larger impacts on cost-weighted industry markup than their low-data counterparts. We know from Proposition 2 that the firm-level markup increases in data if the price of risk ρ is sufficiently small. Since more data increases both w_i^{cost} and M_i^f for a firm, it makes the expected product $E[w_i^{cost} M_i^f]$ greater than the unweighted sum $\bar{M}^f$. This logic holds for fixed costs $\tilde{c}_i$. However, Lemma 3 shows that data reduces the costs firms choose, which increases markups. Thus the cost channel increases markups even more, for the highly-weighted firms.

□

Proof of Proposition 4b: Sales weighted vs cost-weighted markup

Proof. The result assumes that firms are ex ante identical. Under this assumption, when one firm gets more data, the wedge between the sales weighted markup and cost weighted markup goes up. Using the expressions for sales weighted markup (17) and cost weighted markup (16), we can express the wedge as

$$\begin{aligned} M^s - M^c &= \sum_{k=1}^{n_F}\left(\underbrace{\frac{\mathbf{E}\left[\tilde{\mathbf{q}}_k'\tilde{\mathbf{p}}_k\right]}{\sum_{k=1}^{n_F}\mathbf{E}\left[\tilde{\mathbf{q}}_k'\tilde{\mathbf{p}}_k\right]}}_{w_k^s} - \underbrace{\frac{\mathbf{E}\left[\tilde{\mathbf{q}}_k'\tilde{\mathbf{c}}_k\right]}{\sum_{k=1}^{n_F}\mathbf{E}\left[\tilde{\mathbf{q}}_k'\tilde{\mathbf{c}}_k\right]}}_{w_k^c}\right)\underbrace{\frac{\mathbf{E}\left[\tilde{\mathbf{q}}_i'\tilde{\mathbf{p}}_k\right]}{\mathbf{E}\left[\tilde{\mathbf{q}}_k'\tilde{\mathbf{c}}_k\right]}}_{M_k^f} \\ &= \sum_{k=1}^{n_F}\left(w_k^s - w_k^c\right)M_k^f = \sum_{k=1}^{n_F}\hat{w}_k M_k^f \end{aligned} \tag{110}$$

where $\hat{w}_k \equiv w_k^s - w_k^c$ denotes the difference in weights for firm k. It is easy to see that $\sum_{k=1}^{n_F} \hat{w}_k = 0$ as $\sum_{k=1}^{n_F} w_k^s = \sum_{k=1}^{n_F} w_k^c = 1$. Next, we define the average firm markup as $\bar{M}^f = \frac{1}{n_F}\sum_{k=1}^{n_F} M_k^f$. We can now rewrite the wedge as

$$\begin{aligned} M^s - M^c &= \sum_{k=1}^{n_F}\hat{w}_k(M_k^f - \bar{M}^f + \bar{M}^f) \\ &= \sum_{k=1}^{n_F}\hat{w}_k(M_k^f - \bar{M}^f) + \bar{M}^f\underbrace{\sum_{k=1}^{n_F}\hat{w}_k}_{=0} \\ &= \sum_{k=1}^{n_F}\hat{w}_k(M_k^f - \bar{M}^f) = \sum_{k=1}^{n_F}\hat{w}_k\hat{M}_k^f \end{aligned} \tag{111}$$

where $\hat{M}_k^f \equiv M_k^f - \bar{M}^f$. When firms are identical, this wedge is zero as by symmetry, $M_k^f = \bar{M}^f \implies \hat{M}_k^f = 0$. Therefore, it is sufficient to show that if some firm i gets more data, the wedge becomes positive. We prove this by showing that an increase in n_{di} makes each term of the summation in (111) positive.

First note that by definition $\sum_{k=1}^{n_F} \hat{M}_k^f = 0$. Differentiating this expression with respect to n_{di}, we obtain

$$\frac{\partial \hat{M}_i^f}{\partial n_{di}} + \sum_{k\neq i}\frac{\partial \hat{M}_k^f}{\partial n_{di}} = 0 \tag{112}$$

From Proposition 2, we know that the firm-markup increases in firm i's data n_{di}. Therefore, $\frac{\partial \hat{M}_i^f}{\partial n_{di}} > 0$. Now, note that starting from all firms being identical, the $n_F - 1$ firms which do not get additional data are still identical to each other. Using this in the expression above, we obtain for any firm $k \neq i$,

$$\begin{aligned}\frac{\partial \hat{M}_i^f}{\partial n_{di}} + (n_F - 1)\frac{\partial \hat{M}_k^f}{\partial n_{di}} &= 0 \\ \frac{\partial \hat{M}_k^f}{\partial n_{di}} &= -\frac{1}{n_F - 1}\frac{\partial \hat{M}_i^f}{\partial n_{di}} < 0 \end{aligned} \tag{113}$$

By similar calculations for $\sum_{k=1}^{n_F} \hat{w}_k$, we obtain that

$$\frac{\partial \hat{w}_k}{\partial n_{di}} = -\frac{1}{n_F - 1}\frac{\partial \hat{w}_i}{\partial n_{di}} \tag{114}$$

As we start the analysis from identical firms where $\hat{M}_k^f = \hat{w}_k = 0, \forall k \in \{1, \dots, n_F\}$, the above inequalities imply that if a firm i gets more data, $\hat{M}_i^f$ becomes positive and $\hat{M}_k^f$ becomes negative for $k \neq i$. For the weight differences, we have that the change in $\hat{w}_k$ for all the other firms $k \neq i$ goes in the opposite direction to the change for firm i. To obtain the required result, it is now sufficient to show that an increase in n_{di} makes $\hat{w}_i$ positive. To this end, note that

$$\frac{\partial \hat{w}_i}{\partial n_{di}} = \frac{\partial w_i^s}{\partial n_{di}} - \frac{\partial w_i^c}{\partial n_{di}} \tag{115}$$

We can calculate the expressions for each of these terms separately and then combine them. For the first term, we obtain from the definition of sales weight

$$\begin{aligned}\frac{\partial w_i^s}{\partial n_{di}} &= \frac{\partial}{\partial n_{di}} \frac{\mathbf{E}\left[\tilde{\mathbf{q}}_i'\tilde{\mathbf{p}}_i\right]}{\sum_{k'=1}^{n_F}\mathbf{E}\left[\tilde{\mathbf{q}}_{k'}'\tilde{\mathbf{p}}_{k'}\right]} \\ &= w_i^s\left(\frac{1}{\mathbf{E}\left[\tilde{\mathbf{q}}_i'\tilde{\mathbf{p}}_i\right]}\frac{\partial \mathbf{E}\left[\tilde{\mathbf{q}}_i'\tilde{\mathbf{p}}_i\right]}{\partial n_{di}} - \frac{1}{\sum_{k'}\mathbf{E}\left[\tilde{\mathbf{q}}_{k'}'\tilde{\mathbf{p}}_{k'}\right]}\sum_{k'}\frac{\partial \mathbf{E}\left[\tilde{\mathbf{q}}_{k'}'\tilde{\mathbf{p}}_{k'}\right]}{\partial n_{di}}\right) \\ &= w_i^s\left(\frac{1}{\mathbf{E}\left[\tilde{\mathbf{q}}_i'\tilde{\mathbf{p}}_i\right]}\frac{\partial \mathbf{E}\left[\tilde{\mathbf{q}}_i'\tilde{\mathbf{p}}_i\right]}{\partial n_{di}} - \frac{1}{\sum_{k'}\mathbf{E}\left[\tilde{\mathbf{q}}_{k'}'\tilde{\mathbf{p}}_{k'}\right]}\frac{\partial \mathbf{E}\left[\tilde{\mathbf{q}}_i'\tilde{\mathbf{p}}_i\right]}{\partial n_{di}} - \frac{1}{\sum_{k'}\mathbf{E}\left[\tilde{\mathbf{q}}_{k'}'\tilde{\mathbf{p}}_{k'}\right]}\sum_{k'\neq i}\frac{\partial \mathbf{E}\left[\tilde{\mathbf{q}}_{k'}'\tilde{\mathbf{p}}_{k'}\right]}{\partial n_{di}}\right)\end{aligned}$$

Using the assumption that all firms are ex-ante identical and all the $n_F - 1$ firms which do not get additional data are ex-post identical to each other, we obtain

$$\frac{\partial w_i^s}{\partial n_{di}} = \frac{n_F - 1}{n_F^2\mathbf{E}\left[\tilde{\mathbf{q}}_k'\tilde{\mathbf{p}}_k\right]}\left(\frac{\partial \mathbf{E}\left[\tilde{\mathbf{q}}_i'\tilde{\mathbf{p}}_i\right]}{\partial n_{di}} - \frac{\partial \mathbf{E}\left[\tilde{\mathbf{q}}_k'\tilde{\mathbf{p}}_k\right]}{\partial n_{di}}\right) \tag{116}$$

Similarly for the cost weight, we get

$$\frac{\partial w_i^c}{\partial n_{di}} = \frac{n_F - 1}{n_F^2\mathbf{E}\left[\tilde{\mathbf{q}}_k'\tilde{\mathbf{c}}_k\right]}\left(\frac{\partial \mathbf{E}\left[\tilde{\mathbf{q}}_i'\tilde{\mathbf{c}}_i\right]}{\partial n_{di}} - \frac{\partial \mathbf{E}\left[\tilde{\mathbf{q}}_k'\tilde{\mathbf{c}}_k\right]}{\partial n_{di}}\right) \tag{117}$$

Combining the results for firm i, we get

$$\begin{aligned}\frac{\partial}{\partial n_{di}}(w_i^s - w_i^c) &= \frac{n_F - 1}{n_F^2\mathbf{E}\left[\tilde{\mathbf{q}}_k'\tilde{\mathbf{p}}_k\right]}\left(\frac{\partial \mathbf{E}\left[\tilde{\mathbf{q}}_i'\tilde{\mathbf{p}}_i\right]}{\partial n_{di}} - \frac{\partial \mathbf{E}\left[\tilde{\mathbf{q}}_k'\tilde{\mathbf{p}}_k\right]}{\partial n_{di}}\right) \\ &\quad - \frac{n_F - 1}{n_F^2\mathbf{E}\left[\tilde{\mathbf{q}}_k'\tilde{\mathbf{c}}_k\right]}\left(\frac{\partial \mathbf{E}\left[\tilde{\mathbf{q}}_i'\tilde{\mathbf{c}}_i\right]}{\partial n_{di}} - \frac{\partial \mathbf{E}\left[\tilde{\mathbf{q}}_k'\tilde{\mathbf{c}}_k\right]}{\partial n_{di}}\right) \\ &= \frac{n_F - 1}{n_F^2}\left(\frac{\frac{\partial \mathbf{E}[\tilde{\mathbf{q}}_i'\tilde{\mathbf{p}}_i]}{\partial n_{di}} - \frac{\partial \mathbf{E}[\tilde{\mathbf{q}}_k'\tilde{\mathbf{p}}_k]}{\partial n_{di}}}{\mathbf{E}\left[\tilde{\mathbf{q}}_k'\tilde{\mathbf{p}}_k\right]} - \frac{\frac{\partial \mathbf{E}[\tilde{\mathbf{q}}_i'\tilde{\mathbf{c}}_i]}{\partial n_{di}} - \frac{\partial \mathbf{E}[\tilde{\mathbf{q}}_k'\tilde{\mathbf{c}}_k]}{\partial n_{di}}}{\mathbf{E}\left[\tilde{\mathbf{q}}_k'\tilde{\mathbf{c}}_k\right]}\right)\end{aligned} \tag{118}$$

where k is any firm different from i. Next, we use the results derived for markups earlier in the proof to show that the RHS of the above expression (118) is positive. From Proposition 2 and (113), we know that $\frac{\partial \hat{M}_i^f}{\partial n_{di}} > 0$ and $\frac{\partial \hat{M}_k^f}{\partial n_{di}} < 0$.

Therefore,

$$\begin{aligned}
& \frac{\partial \hat{M}_i^f}{\partial n_{di}} > \frac{\partial \hat{M}_k^f}{\partial n_{di}} \\
\iff \quad & \frac{\partial M_i^f}{\partial n_{di}} - \frac{\partial \bar{M}^f}{\partial n_{di}} > \frac{\partial M_k^f}{\partial n_{di}} - \frac{\partial \bar{M}^f}{\partial n_{di}} \\
\iff \quad & \frac{\partial}{\partial n_{di}} \frac{\mathbf{E}\left[\tilde{\mathbf{q}}_i'\tilde{\mathbf{p}}_i\right]}{\mathbf{E}\left[\tilde{\mathbf{q}}_i'\tilde{\mathbf{c}}_i\right]} > \frac{\partial}{\partial n_{di}} \frac{\mathbf{E}\left[\tilde{\mathbf{q}}_k'\tilde{\mathbf{p}}_k\right]}{\mathbf{E}\left[\tilde{\mathbf{q}}_k'\tilde{\mathbf{c}}_k\right]} \\
\iff \quad & \frac{1}{\mathbf{E}\left[\tilde{\mathbf{q}}_i'\tilde{\mathbf{c}}_i\right]} \frac{\partial \mathbf{E}\left[\tilde{\mathbf{q}}_i'\tilde{\mathbf{p}}_i\right]}{\partial n_{di}} - \frac{\mathbf{E}\left[\tilde{\mathbf{q}}_i'\tilde{\mathbf{p}}_i\right]}{(\mathbf{E}\left[\tilde{\mathbf{q}}_i'\tilde{\mathbf{c}}_i\right])^2} \frac{\partial \mathbf{E}\left[\tilde{\mathbf{q}}_i'\tilde{\mathbf{c}}_i\right]}{\partial n_{di}} > \frac{1}{\mathbf{E}\left[\tilde{\mathbf{q}}_k'\tilde{\mathbf{c}}_k\right]} \frac{\partial \mathbf{E}\left[\tilde{\mathbf{q}}_k'\tilde{\mathbf{p}}_k\right]}{\partial n_{di}} - \frac{\mathbf{E}\left[\tilde{\mathbf{q}}_k'\tilde{\mathbf{p}}_k\right]}{(\mathbf{E}\left[\tilde{\mathbf{q}}_k'\tilde{\mathbf{c}}_k\right])^2} \frac{\partial \mathbf{E}\left[\tilde{\mathbf{q}}_k'\tilde{\mathbf{c}}_k\right]}{\partial n_{di}} \\
\iff \quad & \frac{1}{\mathbf{E}\left[\tilde{\mathbf{q}}_i'\tilde{\mathbf{c}}_i\right]} \frac{\partial \mathbf{E}\left[\tilde{\mathbf{q}}_i'\tilde{\mathbf{p}}_i\right]}{\partial n_{di}} - \frac{1}{\mathbf{E}\left[\tilde{\mathbf{q}}_k'\tilde{\mathbf{c}}_k\right]} \frac{\partial \mathbf{E}\left[\tilde{\mathbf{q}}_k'\tilde{\mathbf{p}}_k\right]}{\partial n_{di}} > \frac{\mathbf{E}\left[\tilde{\mathbf{q}}_i'\tilde{\mathbf{p}}_i\right]}{(\mathbf{E}\left[\tilde{\mathbf{q}}_i'\tilde{\mathbf{c}}_i\right])^2} \frac{\partial \mathbf{E}\left[\tilde{\mathbf{q}}_i'\tilde{\mathbf{c}}_i\right]}{\partial n_{di}} - \frac{\mathbf{E}\left[\tilde{\mathbf{q}}_k'\tilde{\mathbf{p}}_k\right]}{(\mathbf{E}\left[\tilde{\mathbf{q}}_k'\tilde{\mathbf{c}}_k\right])^2} \frac{\partial \mathbf{E}\left[\tilde{\mathbf{q}}_k'\tilde{\mathbf{c}}_k\right]}{\partial n_{di}} \\
\iff \quad & \frac{1}{\mathbf{E}\left[\tilde{\mathbf{q}}_k'\tilde{\mathbf{p}}_k\right]} \left(\frac{\partial \mathbf{E}\left[\tilde{\mathbf{q}}_i'\tilde{\mathbf{p}}_i\right]}{\partial n_{di}} - \frac{\partial \mathbf{E}\left[\tilde{\mathbf{q}}_k'\tilde{\mathbf{p}}_k\right]}{\partial n_{di}} \right) > \frac{1}{\mathbf{E}\left[\tilde{\mathbf{q}}_k'\tilde{\mathbf{c}}_k\right]} \left(\frac{\partial \mathbf{E}\left[\tilde{\mathbf{q}}_i'\tilde{\mathbf{c}}_i\right]}{\partial n_{di}} - \frac{\partial \mathbf{E}\left[\tilde{\mathbf{q}}_k'\tilde{\mathbf{c}}_k\right]}{\partial n_{di}} \right) \qquad (119)
\end{aligned}$$

where the last inequality follows because we have assumed that the firms are ex-ante identical. In such a setting, the expected costs and revenues are the same for all the firms a priori (i.e. $\mathbf{E}\left[\tilde{\mathbf{q}}_i'\tilde{\mathbf{c}}_i\right] = \mathbf{E}\left[\tilde{\mathbf{q}}_k'\tilde{\mathbf{c}}_k\right]$ and $\mathbf{E}\left[\tilde{\mathbf{q}}_i'\tilde{\mathbf{p}}_i\right] = \mathbf{E}\left[\tilde{\mathbf{q}}_k'\tilde{\mathbf{p}}_k\right]$)

Using the above result (119) in (118) calculated earlier, we conclude that $\hat{w}_i$, the difference in sales-weight and cost-weight for firm i, increases with data n_{di} and for $k \neq i$, $\hat{w}_k$ decreases with n_{di} (using 114). Therefore, we have shown that when firm i gets more data, $\hat{M}_i^f$ and $\hat{w}_i$ become positive and for all the remaining firms $k \neq i$, $\hat{M}_k^f$ and $\hat{w}_k$ become negative. Rewriting (111) below

$$M^s - M^c = \sum_{k=1}^{n_F} \hat{w}_k \hat{M}_k^f \qquad (120)$$

we note that each term in the summation is positive which means that $M^s - M^c$ is positive. When all firms are identical, the wedge $M^s - M^c = 0$ and when one firm gets more data, the wedge becomes positive. Therefore, an increase in one firm's data increases the wedge. □

Proof of Proposition 4c: Sales-weighted vs. industry aggregates markup The reason this corollary follows directly from Proposition 4b, that the cost-weighted industry markup and the aggregate markup are the same, in our setting. This is a version of the aggregation results of Edmond, Midrigan, and Xu (2019), extended to our linear demand system. The proof is just algebraic manipulation:

$$M^{ag} := \frac{\mathbb{E}\left[\sum_{i=1}^N \boldsymbol{q}_i'\boldsymbol{p}_i\right]}{\mathbb{E}\left[\sum_{i=1}^N \boldsymbol{q}_i'\boldsymbol{c}_i\right]} = \frac{\sum_{i=1}^N \mathbb{E}\left[\boldsymbol{q}_i'\boldsymbol{p}_i\right]}{\sum_{i=1}^N \mathbb{E}\left[\boldsymbol{q}_i'\boldsymbol{c}_i\right]} = \sum_{i=1}^N w_i^m M_i^f = M^m \quad \text{where} \quad w_i^c = \frac{\mathbb{E}\left[\boldsymbol{q}_i'\boldsymbol{c}_i\right]}{\sum_{i=1}^N \mathbb{E}\left[\boldsymbol{q}_i'\boldsymbol{c}_i\right]}. \qquad (121)$$

Proof of proposition 5: Cyclical Markups Part a: product markups are increasing in demand variance and converge to a constant.

Proof. Let $\sigma_b I_N$ denote the variance of demand shocks b. According to the definition of $\hat{\boldsymbol{H}}_i$, we have

$$\begin{aligned}
\hat{\boldsymbol{H}}_i = \left(\frac{\boldsymbol{I}_N}{\phi} + \rho_i \mathbf{Var}(\tilde{\boldsymbol{p}}_i|\mathcal{I}_i)\right)^{-1} \quad \text{and} \quad \mathbf{Var}(\tilde{\boldsymbol{p}}_i|\mathcal{I}_i) = \left(\sigma_b^{-1} + n_{di}\right)^{-1} \\
\Rightarrow \lim_{\sigma_b \to \infty} \mathbf{Var}(\tilde{\boldsymbol{p}}_i|\mathcal{I}_i) = 1/n_{di}, \quad \tilde{\hat{\boldsymbol{H}}}_i := \lim_{\sigma_b \to \infty} \hat{\boldsymbol{H}}_i = \left(\frac{\boldsymbol{I}_N}{\phi} + \rho_i/n_{di}\right)^{-1}
\end{aligned} \qquad (122)$$

The equilibrium price is given by

$$\mathbb{E}\left[\tilde{\boldsymbol{p}}_i\right] = \bar{\boldsymbol{p}}^M = \left(\boldsymbol{I}_N + \frac{1}{\phi}\sum_{i=1}^{n_F} \hat{\boldsymbol{H}}_i\right)^{-1} \left(\bar{\boldsymbol{p}} + \frac{1}{\phi}\sum_{i=1}^{n_F} \hat{\boldsymbol{H}}_i \boldsymbol{c}_i\right) \qquad (123)$$

It clearly converges due to convergent $\hat{\boldsymbol{H}}_i$, so we have

$$\begin{aligned}\tilde{\boldsymbol{p}} := \lim_{\sigma_b\to\infty} \mathbb{E}\left[\tilde{\boldsymbol{p}}_i\right] &= \left(\boldsymbol{I}_N + \frac{1}{\phi}\sum_{i=1}^{n_F}\lim_{\sigma_b\to\infty}\hat{\boldsymbol{H}}_i\right)^{-1}\left(\bar{\boldsymbol{p}} + \frac{1}{\phi}\sum_{i=1}^{n_F}\lim_{\sigma_b\to\infty}\hat{\boldsymbol{H}}_i\boldsymbol{c}_i\right) \\ &= \left[\boldsymbol{I}_N + \sum_{i=1}^{n_F}\left(\boldsymbol{I}_N + \phi\rho_i/n_{di}\right)^{-1}\right]^{-1}\left[\bar{\boldsymbol{p}} + \sum_{i=1}^{n_F}\boldsymbol{c}_i\left(\boldsymbol{I}_N + \phi\rho_i/n_{di}\right)^{-1}\right]\end{aligned} \tag{124}$$

This result implies convergent product-level markup on the attributes as $\lim_{\sigma_b\to\infty}\overline{M}^p$ exists. Since equilibrium price on the goods is a linear combination of weight matrix $\boldsymbol{A}$ and $\tilde{\boldsymbol{p}}_i$, the product-level markup on the goods converges.

$$\boldsymbol{q}_i = \boldsymbol{A}\tilde{\boldsymbol{q}}_i \quad \text{and} \quad \boldsymbol{p}_i = \boldsymbol{A}\tilde{\boldsymbol{p}}_i \Rightarrow \overline{M}^p = \frac{1}{N}\frac{1}{n_F}\sum_{i=1}^{n_F}\sum_{j=1}^{N}\frac{\left(\boldsymbol{A}\mathbb{E}\left[\tilde{\boldsymbol{p}}_i\right]\right)_j}{\left(\boldsymbol{A}\boldsymbol{c}_i\right)_j} \quad \text{converges.} \tag{125}$$

If all the firms have identical sizes ($\boldsymbol{c}_i = \bar{\boldsymbol{c}}$), the derivative of equilibrium price for specific attribute j is

$$\frac{\partial\mathbb{E}[\tilde{\boldsymbol{p}}_{i,j}]}{\partial\Sigma_{b,j}} = \frac{(\bar{\boldsymbol{c}}_j - \bar{\boldsymbol{p}}_j)\frac{1}{\phi}\sum_{i=1}^{n_F}\frac{\partial\hat{\boldsymbol{H}}_{i,j}}{\partial\Sigma_{b,j}}}{\left(1 + \frac{1}{\phi}\sum_{i=1}^{n_F}\hat{\boldsymbol{H}}_{i,j}\right)^2} \quad \text{and} \quad \frac{\partial\hat{\boldsymbol{H}}_{i,j}}{\partial\Sigma_{b,j}} = -\frac{\hat{\boldsymbol{H}}_{i,j}^2\rho_i/n_{di}^2}{\left(\Sigma_{b.j} + 1/n_{di}\right)^2} \leq 0 \tag{126}$$

Since positive production implies lower marginal cost ($\bar{\boldsymbol{c}}_j \leq \bar{\boldsymbol{p}}_j$), the numerator of the derivative is positive. □

Part b: Firm and industry level markups are increasing in demand variance. They asymptote to a linearly increasing function of demand variance.

Proof. Recall that firm, and thus industry markups are increasing in the covariance of price and quantity for each firm i. Thus is suffices to show that the trace of the covariance $\mathbf{tr}[\mathbf{Cov}\left(\tilde{\boldsymbol{p}}_i, \tilde{\boldsymbol{q}}_i\right)]$ is increasing in demand variance σ_b^2.

From the first order condition (5), we know that the quantity choice is a linear function of the expected price: $\tilde{\boldsymbol{q}}_i = \hat{\boldsymbol{H}}_i(E[\tilde{\boldsymbol{p}}_i|\mathcal{I}_i] - \boldsymbol{c}_i)$. Because expectations are optimal linear projections, the realized price must be the expected price, plus a mean-zero expectation error that is orthogonal to the expectation. Thus, $\mathbf{Cov}\left(\tilde{\boldsymbol{p}}_i, \tilde{\boldsymbol{q}}_i\right) = \hat{\boldsymbol{H}}_i\mathbf{Var}(E[\tilde{\boldsymbol{p}}_i|\mathcal{I}_i])$. The covariance is proportional to the unconditional variance of price forecasts, which vary because of the random realization of data signals.

Using the equilibrium price equation (3), we see that firm i's expected price depends on their expectation of their demand shock b_i and their expectation of the aggregate quantity produced:

$$E[\tilde{\boldsymbol{p}}_i|\mathcal{I}_i] = \bar{p} + E[\tilde{\boldsymbol{b}}_i|\mathcal{I}_i] - \frac{1}{\phi}\sum_{i'}\hat{\boldsymbol{H}}_{i'}(E[\tilde{\boldsymbol{p}}_{i'}|\mathcal{I}_{i'}] - \boldsymbol{c}_{i'})$$

Since all price shocks are orthogonal, the variance of expected price for all other firms $i' \neq i$ is simply additive to firm i's shock. Thus, the variance of expected price can be expressed as

$$\mathbf{Var}(E[\tilde{\boldsymbol{p}}_i|\mathcal{I}_i]) = (1 - (1/\phi)\hat{\boldsymbol{H}}_i)'\mathbf{Var}(E[\tilde{\boldsymbol{b}}_i|\mathcal{I}_i])(1 - (1/\phi)\hat{\boldsymbol{H}}_i) + \frac{1}{\phi}\sum_{i'\neq i}\zeta_{i'}\mathbf{Var}(E[\tilde{\boldsymbol{b}}_j|\mathcal{I}_j])\zeta_{i'}$$

for the matrix ζ that maps firm i's expected demand shock into their output. By symmetry, using the solution for firm i above, it must be that $\zeta_{i'} = \hat{\boldsymbol{H}}_{i'}(1 - (1/\phi)\hat{\boldsymbol{H}}_{i'})$.

For each firm j and each good i, an increase in the variance of demand σ_b^2 increases the unconditional variance of their expectation of the shock. Too see this, note that by Bayes law for normal variables, the expectation is the precision-weighted average of the prior (0) and the signal: $E[\tilde{\boldsymbol{b}}_{ij}|\mathcal{I}_i] = n_{di}s_{ij}/(\sigma_b^{-2} + n_{di})$. Since the unconditional variance of the signal is the variance of demand σ_b^2 plus the variance of signal noise n_{di}^{-1},

$$\mathbf{Var}(E[\tilde{\boldsymbol{b}}_i|\mathcal{I}_i]) = \left(\frac{n_{di}}{\sigma_b^{-2} + n_{di}}\right)^2(\sigma_b^2 + n_{di}^{-1}).$$

This is increasing in demand variance σ_b^2 and has no finite upper bound. Since the variance of each element of the diagonal variance-covariance matrix is increasing, the trace of the matrix is also increasing: $\lim_{\sigma_b^2\to\infty} Tr[\mathbf{Var}(E[\tilde{\boldsymbol{b}}_i|\mathcal{I}_i])] = \infty$.

An increase in demand variance also affects $\hat{\boldsymbol{H}}_i$. However, it converges to a constant: $\lim_{\sigma_b^2\to\infty}\hat{\boldsymbol{H}}_i = (1/\phi + \rho_i/n_{di})^{-1}$. Thus, the products are increasing in the limit: $\lim_{\sigma_b^2\to\infty}\mathbf{tr}[\mathbf{Cov}\left(\tilde{\boldsymbol{p}}_i, \tilde{\boldsymbol{q}}_i\right)] = \infty$.

This proof held marginal costs $\tilde{c}$ fixed. However, changes in marginal cost do no affect $\hat{\boldsymbol{H}}_i$ or the variance of prices. (See the derivation of ex-ante payoff.) So this is without loss. □

C. Aggregate Shock Model Results

This section contains proofs of lemmas that are used in the proofs of propositions and additional welfare analysis. First, we derive the price in terms of parameters. The solution is still an implicit solution to a system of equations. Then, we derive conditions under which the H terms, which represent the sensitivity of quantity to a chance in expected price, increase/decrease in data. Finally, we consumer surplus welfare when firms are asymmetric.

Because the production problem is additively separable in attributes, without loss of generality, we focus on the single-attribute problem. The results the generalize to the the multi-attribute case because they hold for each one of the attributes.

Lemma 5. *In the equilibrium, each firm i's quantity is linear in their signal, that is to say,* $\boldsymbol{q}_i = k_i \mathbf{s}_i + p_i^M$. *And* k_i *and* p_i^M *are solutions of the following set of equations.*

$$k_i = \frac{\frac{n_{di}}{n_{di}+1} - \frac{1}{\phi}(\sum_{j=1}^{n_F} k_i)}{\rho_i[(1-\frac{\sum_{j\neq i}^{n_F} k_j}{\phi})^2 \frac{1}{n_{di}+1} + \frac{1}{\phi^2}\sum_{j\neq i}^{n_F} k_j^2 \frac{1}{n_{dj}+1}] + 1/\phi}$$

and

$$p_i^M = \frac{\bar{p} - \frac{1}{\phi}\sum_{j=1}^{n_F} p_i^M - \boldsymbol{c}_i}{\rho_i[(1-\frac{\sum_{j\neq i}^{n_F} k_j}{\phi})^2 \frac{1}{n_{di}+1} + \frac{1}{\phi^2}\sum_{l\neq i}^{n_F} k_l^2 \frac{1}{n_{dl}+1}] + 1/\phi}.$$

Consider the denominator. The term $\frac{1}{\phi^2}\sum_{j\neq i}^{n_F} k_j^2 \frac{1}{n_{dj}+1}$ represents strategic uncertainty. This is the uncertainty about the price that comes from not knowing what competitors will do. The first term in brackets represents the way in which other firms' choices reduce uncertainty: $(1-\frac{\sum_{j\neq i}^{n_F} k_j}{\phi})^2 \frac{1}{n_{di}+1}$. Other firms' responses to their data (if $k_l \geq 0\ \forall l$) reduce the squared term, which is a component of price variance. The idea is that if every firm's quantity choice is positively correlated with their signal, when you receive a positive shock signal, you know the prices will increase, but you also know that other firms are likely to receive positive shock signals so that they will produce more. Others producing more reduces the price and offsets some of the price increase, which reduces price variance.

Proof. Taking the variance of the pricing rule $\boldsymbol{p} = \bar{p} - \frac{1}{\phi}\sum_{i=1}^{n_F} \boldsymbol{q}_i + \mathbf{b}$ yields

$$\begin{aligned}
\mathbf{Var}[\boldsymbol{p}|\mathcal{I}_i] &= \mathbf{Var}\left[\frac{1}{\phi}\sum_{j\neq i}^{n_F} \boldsymbol{q}_j - \mathbf{b}|\mathbf{s}_i\right] \\
&= \mathbf{Var}\left[\mathbf{E}\left[\frac{1}{\phi}\sum_{j\neq i}^{n_F} \boldsymbol{q}_j|\mathbf{b}\right] - \mathbf{b}|\mathbf{s}_i\right] + \mathbf{E}\left[\mathbf{Var}\left[\frac{1}{\phi}\sum_{j\neq i}^{n_F} \boldsymbol{q}_j|\mathbf{b}\right]|\mathbf{s}_i\right].
\end{aligned}$$

Recall the first order condition $\boldsymbol{q}_i = (\rho_i \mathbf{Var}[\boldsymbol{p}|\mathcal{I}_i] + 1/\phi)^{-1}(\mathbf{E}[\boldsymbol{p}|\mathcal{I}_i] - \boldsymbol{c}_i)$. Because our random variables are normal and functions are linear in those random variables, $\mathbf{E}[\boldsymbol{p}|\mathbf{s}_i]$ should be linear in $\mathbf{s}_i$, and $\mathbf{Var}[\boldsymbol{p}|\mathbf{s}_i]$ should not change with respect to the realization of $\mathbf{s}_i$. Thus, we can represent equilibrium production as $\boldsymbol{q}_i = k_i \mathbf{s}_i + p_i^M$.

Note conditional on $\mathbf{s}_i$, $\mathbf{b}$ is distributed as $N(\frac{n_{di}}{n_{di}+1}\mathbf{s}_i, \frac{1}{n_{di}+1})$. Thus, we have

$$\begin{aligned}
\mathbf{E}[\boldsymbol{p}|\mathbf{s}_i] &= \mathbf{E}\left[\bar{p} - \frac{1}{\phi}\sum_{i=1}^{n_F} \boldsymbol{q}_i + \mathbf{b}|\mathbf{s}_i\right] \\
&= \mathbf{E}\left[\bar{p} - \frac{1}{\phi}\sum_{i=1}^{n_F}(k_i \mathbf{s}_i + p_i^M) + \mathbf{b}|\mathbf{s}_i\right] \\
&= \bar{p} + \frac{n_{di}}{n_{di}+1}\mathbf{s}_i - \frac{1}{\phi}(\sum_{j=1}^{n_F} k_i)\mathbf{s}_i - \frac{1}{\phi}\sum_{j=1}^{n_F} p_i^M
\end{aligned}$$

$$
\begin{aligned}
\mathbf{Var}\left[\mathbf{E}\left[\frac{1}{\phi}\sum_{j\neq i}^{n_F}\boldsymbol{q}_j|\mathbf{b}\right]-\mathbf{b}|\mathbf{s}_i\right] &= \mathbf{Var}\left[\mathbf{E}\left[\frac{1}{\phi}\sum_{j\neq i}^{n_F}(k_j\mathbf{s}_j+a_j)|\mathbf{b}\right]-\mathbf{b}|\mathbf{s}_i\right]\\
&= \mathbf{Var}\left[\frac{1}{\phi}\sum_{j\neq i}^{n_F}(k_j\mathbf{b}+a_j)-\mathbf{b}|\mathbf{s}_i\right]\\
&= (1-\frac{\sum_{j\neq i}^{n_F}k_j}{\phi})^2\frac{1}{n_{di}+1}
\end{aligned}
$$

$$
\begin{aligned}
\mathbf{E}\left[\mathbf{Var}\left[\frac{1}{\phi}\sum_{j\neq i}^{n_F}\boldsymbol{q}_j|\mathbf{b}\right]|\mathbf{s}_i\right] &= \mathbf{E}\left[\mathbf{Var}\left[\frac{1}{\phi}\sum_{j\neq i}^{n_F}\boldsymbol{q}_j|\mathbf{b}\right]|\mathbf{s}_i\right]\\
&= \mathbf{E}\left[\frac{1}{\phi^2}\sum_{j\neq i}^{n_F}\mathbf{Var}\left[\boldsymbol{q}_j|\mathbf{b}\right]|\mathbf{s}_i\right]\\
&= \mathbf{E}\left[\frac{1}{\phi^2}\sum_{j\neq i}^{n_F}k_j^2\frac{1}{n_{dj}+1}|\mathbf{s}_i\right]\\
&= \frac{1}{\phi^2}\sum_{j\neq i}^{n_F}k_j^2\frac{1}{n_{dj}+1}
\end{aligned}
$$

Thus, we have for any i

$$
k_i\mathbf{s}_i+p_i^M=\frac{\bar{p}+\frac{n_{di}}{n_{di}+1}\mathbf{s}_i-\frac{1}{\phi}(\sum_{j=1}^{n_F}k_i)\mathbf{s}_i-\frac{1}{\phi}\sum_{j=1}^{n_F}p_i^M-\boldsymbol{c}_i}{\rho_i[(1-\frac{\sum_{j\neq i}^{n_F}k_j}{\phi})^2\frac{1}{n_{di}+1}+\frac{1}{\phi^2}\sum_{j\neq i}^{n_F}k_j^2\frac{1}{n_{dj}+1}]+1/\phi}.
$$

If we match coefficients, we have

$$
k_i=\frac{\frac{n_{di}}{n_{di}+1}-\frac{1}{\phi}(\sum_{j=1}^{n_F}k_i)}{\rho_i[(1-\frac{\sum_{j\neq i}^{n_F}k_j}{\phi})^2\frac{1}{n_{di}+1}+\frac{1}{\phi^2}\sum_{j\neq i}^{n_F}k_j^2\frac{1}{n_{dj}+1}]+1/\phi}
$$

and

$$
p_i^M=\frac{\bar{p}-\frac{1}{\phi}\sum_{j=1}^{n_F}p_i^M-\boldsymbol{c}_i}{\rho_i[(1-\frac{\sum_{j\neq i}^{n_F}k_j}{\phi})^2\frac{1}{n_{di}+1}+\frac{1}{\phi^2}\sum_{j\neq i}^{n_F}k_j^2\frac{1}{n_{dj}+1}]+1/\phi}.
$$

□

Lemma 6. ***Data reduces price uncertainty:*** *Suppose ρ_1,ρ_2 are sufficiently small, and n_{d1} is sufficiently large, then $\mathbf{Var}\left[\boldsymbol{p}|\mathcal{I}_1\right]$ decreases as n_{d1} increases. Furthermore, if $\rho_1\gg\rho_2$, then $\frac{\partial\mathbf{Var}[\boldsymbol{p}|\mathcal{I}_1]}{n_{d1}}\gg\frac{\partial\mathbf{Var}[\boldsymbol{p}|\mathcal{I}_2]}{\partial n_{d1}}$.*

Proof. Let

$$
f_1(k_1,k_2,n_{d1})=\rho_1[(1-\frac{1}{\phi}(n_F-1)k_2)^2\frac{1}{n_{d1}+1}+\frac{1}{\phi^2}(n_F-1)k_2^2\frac{1}{n_{d2}+1}]+1/\phi,
$$

then we have

$$
\begin{aligned}
\frac{\partial f_1}{\partial k_1}&=0\\
\frac{\partial f_1}{\partial k_2}&=\rho_1[\frac{1}{n_{d1}+1}(\frac{2(n_F-1)^2k_2}{\phi^2}-\frac{2(n_F-1)}{\phi})+\frac{1}{n_{d2}+1}\frac{1}{\phi^2}2(n_F-1)k_2]\\
\frac{\partial f_1}{\partial n_{d1}}&=-\rho_1(1-\frac{1}{\phi}(n_F-1)k_2)^2\frac{1}{(n_{d1}+1)^2}
\end{aligned}
\tag{127}
$$

Let

$$
f_2(k_1,k_2,n_{d1})=\rho_2[(1-\frac{1}{\phi}(n_F-2)k_2-\frac{1}{\phi}k_1)^2\frac{1}{n_{d2}+1}+\frac{1}{\phi^2}[(n_F-2)k_2^2\frac{1}{n_{d2}+1}+k_1^2\frac{1}{n_{d1}+1}]]+1/\phi,
$$

then we have

$$\frac{\partial f_2}{\partial k_1} = \rho_2[\frac{1}{n_{d2}+1}\frac{2}{\phi}[\frac{1}{\phi}k_1 + \frac{1}{\phi}(n_F-1)k_2 - 1] + \frac{1}{n_{d1}+1}\frac{1}{\phi^2}2k_1]$$

$$\frac{\partial f_2}{\partial k_2} = \rho_2[\frac{1}{n_{d2}+1}\frac{2(n_F-2)}{\phi}[\frac{1}{\phi}(n_F-2)k_2 - 1 + \frac{1}{\phi}k_1)] + \frac{1}{n_{d2}+1}[\frac{1}{\phi^2}(n_F-2)2k_2]]$$

$$\frac{\partial f_2}{\partial n_{d1}} = -\rho_2\frac{1}{\phi^2}k_1^2\frac{1}{(n_{d1}+1)^2}$$

Define

$$g_1(k_1,k_2,d_1) = \frac{\frac{n_{d1}}{n_{d1}+1} - \frac{1}{\phi}k_1 - \frac{1}{\phi}(n_F-1)k_2}{\rho_1[(1-\frac{1}{\phi}(n_F-1)k_2)^2\frac{1}{n_{d1}+1} + \frac{1}{\phi^2}(n_F-1)k_2^2\frac{1}{n_{d2}+1}] + 1/\phi} - k_1$$

$$g_2(k_1,k_2,d_1) = \frac{\frac{n_{d2}}{n_{d2}+1} - \frac{1}{\phi}k_1 - \frac{1}{\phi}(n_F-1)k_2}{\rho_2[(1-\frac{1}{\phi}(n_F-2)k_2 - \frac{1}{\phi}k_1)^2\frac{1}{n_{d2}+1} + \frac{1}{\phi^2}[(n_F-2)k_2^2\frac{1}{n_{d2}+1} + k_1^2\frac{1}{n_{d1}+1}]] + 1/\phi} - k_2$$

Then

$$\frac{\partial g_1}{\partial k_1} = \frac{-\frac{1}{\phi}f_1 - (\frac{n_{d1}}{n_{d1}+1} - \frac{1}{\phi}k_1 - \frac{1}{\phi}(n_F-1)k_2)\frac{\partial f_1}{\partial k_1}}{f_1^2} - 1 = -\frac{\frac{1}{\phi} + k_1\frac{\partial f_1}{\partial k_1}}{f_1} - 1$$

$$\frac{\partial g_1}{\partial k_2} = \frac{-\frac{1}{\phi}(n_F-1)f_1 - (\frac{n_{d1}}{n_{d1}+1} - \frac{1}{\phi}k_1 - \frac{1}{\phi}(n_F-1)k_2)\frac{\partial f_1}{\partial k_2}}{f_1^2} = -\frac{\frac{1}{\phi}(n_F-1) + k_1\frac{\partial f_1}{\partial k_2}}{f_1}$$

$$\frac{\partial g_2}{\partial k_1} = \frac{-\frac{1}{\phi}f_2 - (\frac{n_{d2}}{n_{d2}+1} - \frac{1}{\phi}k_1 - \frac{1}{\phi}(n_F-1)k_2)\frac{\partial f_2}{\partial k_1}}{f_2^2} = -\frac{\frac{1}{\phi} + k_2\frac{\partial f_2}{\partial k_1}}{f_2}$$

$$\frac{\partial g_2}{\partial k_2} = \frac{-\frac{1}{\phi}(n_F-1)f_2 - (\frac{n_{d2}}{n_{d2}+1} - \frac{1}{\phi}k_1 - \frac{1}{\phi}(n_F-1)k_2)\frac{\partial f_2}{\partial k_2}}{f_2^2} - 1 = -\frac{\frac{1}{\phi}(n_F-1) + k_2\frac{\partial f_2}{\partial k_2}}{f_2} - 1$$

Thus,

$$-\begin{bmatrix}\frac{\partial g_1}{\partial k_1} & \frac{\partial g_1}{\partial k_2}\\ \frac{\partial g_2}{\partial k_1} & \frac{\partial g_2}{\partial k_2}\end{bmatrix}^{-1} = -[(\frac{\frac{1}{\phi} + k_1\frac{\partial f_1}{\partial k_1}}{f_1} + 1)(\frac{\frac{1}{\phi}(n_F-1) + k_2\frac{\partial f_2}{\partial k_2}}{f_2} + 1) - \frac{\frac{1}{\phi}(n_F-1) + k_1\frac{\partial f_1}{\partial k_2}}{f_1}\frac{\frac{1}{\phi} + k_2\frac{\partial f_2}{\partial k_1}}{f_2}]$$

$$\begin{bmatrix}-\frac{\frac{1}{\phi}(n_F-1) + k_2\frac{\partial f_2}{\partial k_2}}{f_2} - 1 & \frac{\frac{1}{\phi}(n_F-1) + k_1\frac{\partial f_1}{\partial k_2}}{f_1}\\ \frac{\frac{1}{\phi} + k_2\frac{\partial f_2}{\partial k_1}}{f_2} & -\frac{\frac{1}{\phi} + k_1\frac{\partial f_1}{\partial k_1}}{f_1} - 1\end{bmatrix}$$

$$\frac{\partial g_1}{\partial n_{d1}} = \frac{\frac{1}{(n_{di}+1)^2} + k_1\rho_1(1-\frac{1}{\phi}(n_F-1)k_2)^2\frac{1}{(n_{d1}+1)^2}}{f_1}$$

$$\frac{\partial g_2}{\partial n_{d1}} = \frac{k_2\rho_2\frac{1}{\phi^2}k_1^2\frac{1}{(n_{d1}+1)^2}}{f_2}$$

First when $\rho_1 \approx 0, \rho_2 \approx 0$, we have

$$\frac{\partial g_1}{\partial n_{d1}} \approx \frac{\frac{1}{(n_{d1}+1)^2}}{1/\phi}$$

$$\frac{\partial g_2}{\partial n_{d1}} \approx 0$$

$$-\begin{bmatrix}\frac{\partial g_1}{\partial k_1} & \frac{\partial g_1}{\partial k_2}\\ \frac{\partial g_2}{\partial k_1} & \frac{\partial g_2}{\partial k_2}\end{bmatrix}^{-1} \approx -[2n_F - (n_F-1)]\begin{bmatrix}-n_F & n_F-1\\ 1 & -2\end{bmatrix}$$

Thus

$$\frac{\partial k_1}{\partial n_{d1}} \approx \frac{[2n_F - (n_F-1)]n_F}{(n_{d1}+1)^2/\phi}$$

$$\frac{\partial k_2}{\partial n_{d1}} \approx -\frac{[2n_F - (n_F-1)]}{(n_{d1}+1)^2/\phi}$$

$$\frac{df_1}{dn_{d1}} = \frac{\partial f_1}{\partial k_1}\frac{\partial k_1}{\partial n_{d1}} + \frac{\partial f_1}{\partial k_2}\frac{\partial k_2}{\partial n_{d1}} + \frac{\partial f_1}{\partial n_{d1}}$$
$$= \rho_1 \frac{1}{(n_{d1}+1)^2}[-[\frac{1}{n_{d1}+1}(\frac{2(n_F-1)^2 k_2}{\phi} - 2(n_F-1)) + \frac{1}{n_{d2}+1}\frac{1}{\phi}2(n_F-1)k_2][2n_F-(n_F-1)]$$
$$-(1-\frac{1}{\phi}(n_F-1)k_2)^2]$$

Note as n_{d1} increases, because we let $\rho_1 \approx 0$, $\rho_2 \approx 0$, k_1/ϕ and k_2/ϕ should converge to some constant. Thus, as long as n_{d1} is sufficiently large, $\frac{df_1}{dn_{d1}}$ is negative, which implies $\mathbf{Var}[\boldsymbol{p}|\mathcal{I}_1]$ decreases as n_{d1} increases.

$$\frac{df_2}{dn_{d1}} = \frac{\partial f_2}{\partial k_1}\frac{\partial k_1}{\partial n_{d1}} + \frac{\partial f_2}{\partial k_2}\frac{\partial k_2}{\partial n_{d1}} + \frac{\partial f_2}{\partial n_{d1}}$$
$$= \rho_2[\frac{2}{n_{d2}+1}[\frac{1}{\phi}k_1 + \frac{1}{\phi}(n_F-1)k_2 - 1] + \frac{1}{n_{d1}+1}\frac{1}{\phi}2k_1]\frac{[2n_F-(n_F-1)]n_F}{(n_{d1}+1)^2}$$
$$-\rho_2[\frac{1}{n_{d2}+1}2(n_F-2)[\frac{1}{\phi}(n_F-2)k_2 - 1 + \frac{1}{\phi}k_1)] + \frac{1}{n_{d2}+1}[\frac{1}{\phi}(n_F-2)2k_2]]\frac{[2n_F-(n_F-1)]}{(n_{d1}+1)^2}$$
$$-\rho_2\frac{1}{\phi^2}k_1^2\frac{1}{(n_{d1}+1)^2}$$

If ρ_2 is small, all terms go to zero, except for $\frac{1}{n_{d2}+1}[\frac{1}{\phi}(n_F-2)2k_2]]\frac{[2n_F-(n_F-1)]}{(n_{d1}+1)^2}$. in (127), a similar positive expression is multiplied by ρ_1 to get $\frac{df_1}{dn_{d1}}$. As long as ρ_1 is sufficiently high, then $\frac{df_1}{dn_{d1}} \gg \frac{df_2}{dn_{d1}}$. □

D. Related Models: Product Innovation and Price Competition

D.1. Choosing A Location in Product Space

In the previous problem, we introduced the idea of product attributes so that a piece of data might be informative about the demand of multiple products. But we held the attributes of each product fixed. In reality, firms can choose the type of product to produce. They choose attributes. We show that the insights of the previous analysis carry over, with one small change. Data will allow a firm to choose a product that has higher-markup attributes. This makes product markups more like firm markups in the original model.

Each firm produces a single product, or bundle of products, with attributes chosen by the firm. Then the firm chooses how many units of the product or product bundle to produce. Formally, firm $i \in \{1, 2, \ldots, n_F\}$ chooses an $n \times 1$ vector a_i that describes their location in the product space, such that $\sum_j a_{ij} = 1$. As before, The jth entry of vector a_i describes how much of attribte i firm i's good contains.

The rest of the model assumptions, including consumer demand and the nature of data are the same as before. Thus, the firm's production problem is

$$max_{a_i,q_i}\mathbf{E}\left[q_i\boldsymbol{a}_i'(\tilde{\mathbf{p}}-\boldsymbol{c}_i)|\mathcal{I}_i\right] - \frac{\rho_i}{2}\mathbf{Var}\left[q_i\boldsymbol{a}_i'(\tilde{\mathbf{p}}-\boldsymbol{c}_i)|\mathcal{I}_i\right] - g(\chi_c, \boldsymbol{c}_i), \tag{128}$$

s.t. $\sum_j a_{ij} = 1$.

Just like the previous problem, prior to observing any of their data, each firm also chooses their cost vector $\boldsymbol{c}_i$. Since the data realizations are unknown in this ex-ante invetment stage, the objective is the unconditional expectation of the utility in 2

$$max_{c_i}\mathbf{E}\left[\mathbf{E}\left[q_i\boldsymbol{a}_i'(\tilde{\mathbf{p}}-\boldsymbol{c}_i)|\mathcal{I}_i\right] - \frac{\rho_i}{2}\mathbf{Var}\left[q_i\boldsymbol{a}_i'(\tilde{\mathbf{p}}-\boldsymbol{c}_i)|\mathcal{I}_i\right]\right] - g(\chi_c, \boldsymbol{c}_i). \tag{129}$$

SOLUTION Firm i's optimal production from the first order condition looks identical to the one before, except that now it is the the product of quantity and attributes that achieves this solution.

$$q_i\boldsymbol{a}_i = \left(\rho_i\mathbf{Var}[\boldsymbol{p}_i|\mathcal{I}_i] + \frac{\partial\mathbf{E}\left[\boldsymbol{p}_j|\mathcal{I}_j\right]}{\partial\boldsymbol{q}_i}\right)^{-1}(\mathbf{E}[\boldsymbol{p}_i|\mathcal{I}_i] - \boldsymbol{c}_i) \tag{130}$$

This tells us that the solution to the problem is exactly the same. In the previous problem, a firm choice produce any quantity of attributes it wanted with the right mix of products. In this problem, the firm can also choose any quantity of attributes it likes with the right quantity and product location.

The only thing that changes in this formulation of the problem is the interpretation of what constitutes a product. In the previous problem, a product had a fixed set of attributes. In this problem, a product is a fraction of the total output of the firm. Therefore the product markup here is more like what the firm markup was before. In other words, data affects the composition of a product now. Firms with data choose to produce products with higher-value attributes. This is a force that can make markups flat or increasing in data.

Proposition 6. *When firms choose attributes, product markups will increase in data, for a low enough risk aversion ρ_i.*

Proof. Comparing first-order condition (130) with original optimal choice (32), we could solve this extension model by substituting $\tilde{\boldsymbol{q}}_i$ in (32) with $q_i\boldsymbol{a}_i$ and further extend existing propositions for q_i and $\boldsymbol{a}_i$ by one-to-one mapping

$$q_i = \sum_{j=1}^{N} \tilde{\boldsymbol{q}}_{i,j} \quad \text{and} \quad \boldsymbol{a}_i = \frac{\tilde{\boldsymbol{q}}_i}{\sum_{j=1}^{N} \tilde{\boldsymbol{q}}_{i,j}} \tag{131}$$

Since firms optimize their choices in product space, the product markup is then the weighted average of attributes markups

$$M_i^p := \frac{\mathbb{E}[\boldsymbol{a}_i'\tilde{\boldsymbol{p}}_i]}{\mathbb{E}[\boldsymbol{a}_i'\tilde{\boldsymbol{c}}_i]} = \frac{\mathbb{E}[q_i\boldsymbol{a}_i'\tilde{\boldsymbol{p}}_i]}{\mathbb{E}[q_i\boldsymbol{a}_i'\tilde{\boldsymbol{c}}_i]} = \frac{\mathbb{E}[\tilde{\boldsymbol{q}}_i'\tilde{\boldsymbol{p}}_i]}{\mathbb{E}[\tilde{\boldsymbol{q}}_i'\tilde{\boldsymbol{c}}_i]} = M_i^f \tag{132}$$

This tells us that the product markups is equivalent to the firm-level markup of the original model. We already know that data boost firm-level markup with small risk aversion ρ_i (Proposition 2), thus the product markup will increase in data for a low enough risk aversion ρ_i.

This proof held marginal costs $\tilde{c}$ fixed, which corresponds to infinitely high marginal cost of adjusting c: $\chi_c \to \infty$. If we assume χ_c is sufficiently high, by continuity, the inequality will still hold.

□

This result shows why this extension is helpful for the model to match data showing flat or increasing product markups. The fact that markups had to be declining in the previous model was an artifact of the assumption that product characteristics are fixed. While that simplified the model and allowed us to focus on explaining the many other forces at play, the richer model paints a more realistic and data-consistent picture of how data, competition and markups interact.

D.2. Bertrand Competition

Many studies of markup competition use Bertrand, instead of Cournot competition. Therefore, we examine the three main effects in a Bertrand market. We find that the main results about covariance that generate the aggregation effects are similar. Cost choice is similar. But the risk effect can switch signs. Since the risk channel works against our main aggregation results (most of those results assume the price of risk sufficiently low), this strengthens the main effects.

In our model, all final goods are perfect substitutes. We know from the textbook imperfect competition models, that Bertrand with prefect substitutes results in a corner solution where the lowest cost firm captures the entire market at a price equal to the marginal cost of the second lowest cost firm, and the markup is equal to the ratio of the second lowest cost over the lowest cost. In case of a tie in the cost the marginal cost, profits are zero and there is indeterminacy in the exact allocation across firms. Showing that Bertrand competition results in lower prices than Cournot is therefore trivial.

But we know that when goods are imperfectly substitutable, Bertrand pricing is not at a corner, just as under Cournot. In this extension therefore we adjust our baseline model to include the imperfect substitutes. In the setup of our benchmark model all firms produce all goods (possibly with different weights on attributes) which are perfect substitutes. Market power therefore does not originate from the uniqueness of the goods a firm produces. Rather, market power stems from firms' heterogeneous demand shocks and production costs.

So we extend our setting, borrowing from Pellegrino (2023) where each firm produces one good only instead of all goods. First, denote the demand intercept as $\bar{\boldsymbol{p}}$ and the demand shock as $\boldsymbol{b}$. Second, we denote the similarity matrix by $\boldsymbol{\Psi}$. Therefore, we can immediately map $\psi_{ij} = 1/\phi_{ij}$ to our benchmark model where $\phi_{ij} = \phi$ which implies perfect subsitutes.[13]

We first derive a Bertrand solution and then the equivalent equilibrium conditions under Cournot. We then simulate the model to replicate the results.

[13]Pellegrino (2023) denotes the demand intercept by $\boldsymbol{b}$ and the demand similarity matrix by $\boldsymbol{I} + \boldsymbol{\Sigma}$ (note that here $\psi_{ii} = 1$ while $\sigma_{ii} = 0$ in Pellegrino's setup).

Bertrand Competition. From the inverse demand function

$$\boldsymbol{p} = \bar{\boldsymbol{p}} - \boldsymbol{\Psi}\boldsymbol{q} + \boldsymbol{b} \tag{133}$$

$$\Rightarrow \boldsymbol{q} = \boldsymbol{\Psi}^{-1}(\bar{\boldsymbol{p}} - \boldsymbol{p} + \boldsymbol{b}) \tag{134}$$

Denote $\boldsymbol{\Gamma} = \boldsymbol{\Psi}^{-1}$, so the demand function for individual firm i can be written as

$$q_i = \sum_{j=1}^{n} \gamma_{ij}(\bar{p}_j - p_j + b_j) \tag{135}$$

The expectation and variance of the quantity is given by

$$\mathbf{E}\left[q_i \mid \mathcal{I}_i\right] = \sum_{j=1}^{n} \gamma_{ij}(\bar{p}_j - p_j + K_j s_j) \tag{136}$$

$$\mathbf{Var}\left[q_i \mid \mathcal{I}_i\right] = \sum_{j=1}^{n} \gamma_{ij}^2 \mathbf{Var}\left[b_j \mid \mathcal{I}_j\right] = \sum_{j=1}^{n} \gamma_{ij}^2 (1 + \Sigma_{\epsilon_j})^{-1} \Sigma_{\epsilon_j}, \tag{137}$$

and the risk-adjusted profit function is

$$U_I = \mathbf{E}\left[q_i \mid \mathcal{I}_i\right](p_i - c_i) - \frac{\rho_i}{2} \mathbf{Var}\left[q_i \mid \mathcal{I}_i\right](p_i - c_i)^2 - g\left(\chi_c, c_i\right) \tag{138}$$

Maximization with respect to the price yields the first-order condition,

$$\frac{\partial U_i}{\partial p_i} = \mathbf{E}\left[q_i \mid \mathcal{I}_i\right] + \frac{\partial \mathbf{E}\left[q_i \mid \mathcal{I}_i\right]}{\partial p_i}(p_i - c_i) - \rho_i \mathbf{Var}\left[q_i \mid \mathcal{I}_i\right](p_i - c_i) = 0 \tag{139}$$

$$p_i = c_i + \left(\rho_i \mathbf{Var}\left[q_i \mid \mathcal{I}_i\right] - \frac{\partial \mathbf{E}\left[q_i \mid \mathcal{I}_i\right]}{\partial p_i}\right)^{-1} \mathbf{E}\left[q_i \mid \mathcal{I}_i\right] \tag{140}$$

$$p_i = c_i + \left(\rho_i \sum_{j=1}^{n} \gamma_{ij}^2 \mathbf{Var}\left[b_j \mid \mathcal{I}_j\right] + \gamma_{ii}\right)^{-1} \sum_{j=1}^{n} \gamma_{ij}(\bar{p}_j - p_j + K_j s_j) \tag{141}$$

Denote $\hat{t}_i = \left(\rho_i \sum_{j=1}^{n} \gamma_{ij}^2 \mathbf{Var}\left[b_j \mid \mathcal{I}_j\right] + \gamma_{ii}\right)^{-1}$, then

$$p_i = c_i + \hat{t}_i \sum_{j=1}^{n} \gamma_{ij}(\bar{p}_j - p_j + K_j s_j), \tag{142}$$

or equivalently in matrix notation:

$$\boldsymbol{p} = \boldsymbol{c} + \hat{\boldsymbol{T}}\boldsymbol{\Gamma}(\bar{\boldsymbol{p}} - \boldsymbol{p} + \boldsymbol{K}\boldsymbol{s}), \tag{143}$$

where $\boldsymbol{p} = [p_1, p_2, ..., p_n]'$, $\bar{\boldsymbol{p}} = [\bar{p}_1, \bar{p}_2, ..., \bar{p}_n]'$, $\boldsymbol{K}\boldsymbol{s} = [K_1 s_1, K_2 s_2, ..., K_n s_n]'$, $\boldsymbol{c} = [c_1, c_2, ..., c_n]'$ and $\hat{\boldsymbol{T}}$ is a matrix where the diagonal elements are $\hat{t}_i$, and the other elements are 0, that is

$$\hat{\boldsymbol{T}} \equiv \begin{bmatrix} \hat{t}_1 & 0 & \cdots & 0 \\ 0 & \hat{t}_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \hat{t}_n \end{bmatrix}. \tag{144}$$

Then the equilibrium prices and quantities are

$$\boldsymbol{p} = (\boldsymbol{I} + \hat{\boldsymbol{T}}\boldsymbol{\Gamma})^{-1}\boldsymbol{c} + (\boldsymbol{I} + \hat{\boldsymbol{T}}\boldsymbol{\Gamma})^{-1}\hat{\boldsymbol{T}}\boldsymbol{\Gamma}(\bar{\boldsymbol{p}} + \boldsymbol{K}\boldsymbol{s}) \tag{145}$$

$$\boldsymbol{q} = \boldsymbol{\Gamma}(\bar{\boldsymbol{p}} + \boldsymbol{b}) - \boldsymbol{\Gamma}(\boldsymbol{I} + \hat{\boldsymbol{T}}\boldsymbol{\Gamma})^{-1}\boldsymbol{c} - \boldsymbol{\Gamma}(\boldsymbol{I} + \hat{\boldsymbol{T}}\boldsymbol{\Gamma})^{-1}\hat{\boldsymbol{T}}\boldsymbol{\Gamma}(\bar{\boldsymbol{p}} + \boldsymbol{K}\boldsymbol{s}) \tag{146}$$

Lemma 7. ***Data increases price-quantity covariance*** $\partial cov(p_{ij}, q_{ij})/\partial n_{di} > 0$.

Proof: Data n_{di} enters $\boldsymbol{p}$ and $\boldsymbol{q}$ only through $\hat{\boldsymbol{T}}$ and $\boldsymbol{K}$. The terms $\bar{p}$, I and Γ are exogenous.

Step 1: Show that the diagonal elements of $\hat{\boldsymbol{T}}$ and $\boldsymbol{K}$ are increasing in n_{di}. Since Bayesian updating is the same, regardless of market structure, the Bayesian weight on the signal K, is the same as before: $K = n_{di}/(1 + n_{di})I$, which is increasing in data n_{di}. $\hat{\boldsymbol{T}}$ is also a diagonal matrix, with diagonals $\hat{t}_i$ that are decreasing in $\mathbf{Var}\left[b_j \mid \mathcal{I}_j\right]$. Since market structure does not change Bayes' law, we know from Appendix A that data reduces conditional variance (prediction errors) $\partial \mathbf{Var}\left[b_j \mid \mathcal{I}_j\right] / \partial n_{di} < 0$. Thus, the diagonals of $\hat{\boldsymbol{T}}$ are increasing as well: $\partial \hat{t}_i / \partial n_{di} > 0$.

Step 2: Larger diagonal elements of $\hat{\boldsymbol{T}}$ and $\boldsymbol{K}$ raise $cov(p_{ij}, q_{ij})$. Covariance arises from stochastic terms. There are 3 stochastic terms: s in p, and s and b in q. The b term in q is multiplied by Γ, which is exogenous. Data has no effect on that term. The s terms in both p and q are multiplied by $\hat{\boldsymbol{T}}$ and $\boldsymbol{K}$ and other positive terms: $cov(p_i, q_i) = ((\boldsymbol{I} + \hat{\boldsymbol{T}}\boldsymbol{\Gamma})^{-1}\hat{\boldsymbol{T}}\boldsymbol{\Gamma})'\boldsymbol{\Gamma}(\boldsymbol{I} + \hat{\boldsymbol{T}}\boldsymbol{\Gamma})^{-1}\boldsymbol{c} - \boldsymbol{\Gamma}(\boldsymbol{I} + \hat{\boldsymbol{T}}\boldsymbol{\Gamma})^{-1}\hat{\boldsymbol{T}}\boldsymbol{\Gamma}$. Since $(\boldsymbol{I} + \hat{\boldsymbol{T}}\boldsymbol{\Gamma})^{-1}\hat{\boldsymbol{T}}$ is a positive definite matrix, whose eigenvalues are increasing in $\hat{t}_i$, larger $\hat{t}_i$ scales up each diagonal entry of $cov(p_i, q_i)$. Similarly, larger entries of the diagonal matrix K raise each diagonal entry of $cov(p_i, q_i)$. Thus, $cov(p_{ij}, q_{ij})$ are increasing in the diagonal elements of $\hat{\boldsymbol{T}}$ and $\boldsymbol{K}$. □

Markups and profits Next, we show that changing the model of competition can change the nature of the risk effect on markups. When costs are high, risk-averse firms may price low and produce more. However, since the aggregation effects only arise when the risk channel is not too strong, this reversal does not overturn the main aggregation results of the paper.

The risk-adjusted profit in the first stage is

$$
\begin{aligned}
\mathbf{E}[U_i] =& \mathbf{E}\left[\mathbf{E}\left[q_i \mid \mathcal{I}_i\right](p_i - c_i) - \frac{\rho_i}{2}\mathbf{Var}\left[q_i \mid \mathcal{I}_i\right](p_i - c_i)^2\right] - g\left(\chi_c, c_i\right) && (147)\\
=& \mathbf{E}\left[\hat{t}_i^{-1}(p_i - c_i)^2 - \frac{\rho_i}{2}\mathbf{Var}\left[q_i \mid \mathcal{I}_i\right](p_i - c_i)^2\right] - g\left(\chi_c, c_i\right) && (148)\\
=& \mathbf{E}\left[(\gamma_{ii} + \frac{\rho_i}{2}\mathbf{Var}\left[q_i \mid \mathcal{I}_i\right])(p_i - c_i)^2\right] - g\left(\chi_c, c_i\right) && (149)\\
=& \mathbf{E}\left[\left(\gamma_{ii} + \frac{\rho_i}{2}\sum_{j=1}^{n}\gamma_{ij}^2\mathbf{Var}\left[b_j \mid \mathcal{I}_j\right]\right)(p_i - c_i)^2\right] - g\left(\chi_c, c_i\right) && (150)\\
=& \left(\gamma_{ii} + \frac{\rho_i}{2}\sum_{j=1}^{n}\gamma_{ij}^2\mathbf{Var}\left[b_j \mid \mathcal{I}_j\right]\right)\mathbf{E}\left[(p_i - c_i)^2\right] - g\left(\chi_c, c_i\right) && (151)\\
=& \left(\gamma_{ii} + \frac{\rho_i}{2}\sum_{j=1}^{n}\gamma_{ij}^2\mathbf{Var}\left[b_j \mid \mathcal{I}_j\right]\right)\left(\mathbf{E}\left[p_i - c_i\right]^2 + \mathbf{Var}\left[p_i - c_i\right]\right) - g\left(\chi_c, c_i\right) && (152)
\end{aligned}
$$

where $\mathbf{Var}[p_i - c_i] = \mathbf{Var}[p_i]$ is independent of cost choices.

Then the optimal choices of marginal cost will be

$$
\frac{\partial \mathbf{E}[U_i]}{\partial c_i} = \frac{\partial\left(\gamma_{ii} + \frac{\rho_i}{2}\sum_{j=1}^{n}\gamma_{ij}^2\mathbf{Var}\left[b_j \mid \mathcal{I}_j\right]\right)\mathbf{E}\left[p_i - c_i\right]^2}{\partial c_i} - \frac{\partial g\left(\chi_c, c_i\right)}{\partial c_i} = 0 \tag{153}
$$

Firm i's markup is defined as $M_i = \mathbf{E}[p_i]/c_i$.

$$
\boldsymbol{p} = (\boldsymbol{I} + \hat{\boldsymbol{T}}\boldsymbol{\Gamma})^{-1}\boldsymbol{c} + (\boldsymbol{I} + \hat{\boldsymbol{T}}\boldsymbol{\Gamma})^{-1}\hat{\boldsymbol{T}}\boldsymbol{\Gamma}(\bar{\boldsymbol{p}} + \boldsymbol{K}\boldsymbol{s}) \tag{154}
$$

Firm i's markup is defined as $M_i = \mathbf{E}[p_i]/c_i$. Using the Cournot price (170) and $\mathbf{E}[b_i] = 0$,

$$
\mathbf{E}[\boldsymbol{p}] = \bar{\boldsymbol{p}} - (\boldsymbol{I} + \boldsymbol{\Psi})(\boldsymbol{I} + \hat{\boldsymbol{H}}\boldsymbol{\Psi})^{-1}\hat{\boldsymbol{H}}(\bar{\boldsymbol{p}} + \boldsymbol{K}\boldsymbol{s} - \boldsymbol{c}) \equiv \bar{\boldsymbol{p}} - \boldsymbol{\Omega}\hat{\boldsymbol{H}}(\bar{\boldsymbol{p}} + \boldsymbol{K}\boldsymbol{s} - \boldsymbol{c}). \tag{155}
$$

And we can derive the equivalent condition as Proposition (1) for Cournot:

$$
\frac{\partial M_i}{\partial n_{di}} = \frac{\partial(p_i/c_i)}{\partial n_{di}} = \underbrace{\frac{c_i}{c_i^2}\frac{\partial p_i}{\partial n_{di}}}_{\text{Risk premium effect}} - \underbrace{\frac{p_i}{c_i^2}\frac{\partial c_i}{\partial n_{di}}}_{\text{Investment effect}} \tag{156}
$$

where we take the price p_i in Bertrand compared to the expected price $\mathbf{E}[p_i]$ in Cournot. That is the only difference between the two expressions.

An Equivalent Cournot Model for Comparison The inverse demand function is

$$
\boldsymbol{p} = \bar{\boldsymbol{p}} - \boldsymbol{\Psi}\boldsymbol{q} + \boldsymbol{b} \tag{157}
$$

For firm i,

$$
\begin{aligned}
p_i =& \bar{p}_i - \psi_{ii} q_i - \sum_{j \neq i} \psi_{ij} q_j + b_i \quad (\psi_{ii} = 1) && (158) \\
=& \bar{p}_i - \frac{1}{\phi_{ii}} q_i - \sum_{j \neq i} \frac{1}{\phi_{ij}} q_j + b_i && (159)
\end{aligned}
$$

The the expectation and variance of the price are

$$
\begin{aligned}
\mathbf{E}\left[p_i \mid \mathcal{I}_i\right] =& \bar{p}_i - \psi_{ii} q_i - \sum_{j \neq i} \psi_{ij} q_j + K_i s_i && (160) \\
\mathbf{Var}\left[p_i \mid \mathcal{I}_i\right] =& \mathbf{Var}\left[b_i \mid \mathcal{I}_i\right] = (1 + \Sigma_{\epsilon_i})^{-1} \Sigma_{\epsilon_i}, && (161)
\end{aligned}
$$

and the risk-adjusted profit function

$$
U_i = q_i(\mathbf{E}\left[p_i \mid \mathcal{I}_i\right] - c_i) - \frac{\rho_i}{2} \mathbf{Var}\left[p_i \mid \mathcal{I}_i\right] q_i^2 - g\left(\chi_c, c_i\right). \tag{162}
$$

The first-order condition solves

$$
\begin{aligned}
\frac{\partial U_i}{\partial q_i} &= \mathbf{E}\left[p_i \mid \mathcal{I}_i\right] - c_i + \frac{\partial \mathbf{E}\left[p_i \mid \mathcal{I}_i\right]}{\partial q_i} q_i - \rho_i \mathbf{Var}\left[b_i \mid \mathcal{I}_i\right] q_i = 0 && (163) \\
q_i &= \left(\rho_i \mathbf{Var}\left[b_i \mid \mathcal{I}_i\right] - \frac{\partial \mathbf{E}\left[p_i \mid \mathcal{I}_i\right]}{\partial q_i}\right)^{-1} \left(\mathbf{E}\left[p_i \mid \mathcal{I}_i\right] - c_i\right) && (164) \\
q_i &= \left(\rho_i \mathbf{Var}\left[b_i \mid \mathcal{I}_i\right] + \psi_{ii}\right)^{-1} \left[\bar{p}_i - \psi_{ii} q_i - \sum_{j \neq i} \psi_{ij} q_j + K_i s_i - c_i\right]. && (165)
\end{aligned}
$$

If we denote $\hat{h}_i = \left(\rho_i \mathbf{Var}\left[b_i \mid \mathcal{I}_i\right] + \psi_{ii}\right)^{-1}$, then

$$
q_i = \hat{h}_i \left(\bar{p}_i + K_i s_i - c_i - \psi_{ii} q_i - \sum_{j \neq i} \psi_{ij} q_j\right), \tag{166}
$$

or equivalently

$$
\boldsymbol{q} = \hat{\boldsymbol{H}}(\bar{\boldsymbol{p}} + \boldsymbol{K}\boldsymbol{s} - \boldsymbol{c} - \boldsymbol{\Psi}\boldsymbol{q}) \tag{167}
$$

where $\boldsymbol{q} = [q_1, q_2, ..., q_n]'$, $\bar{\boldsymbol{p}} = [\bar{p}_1, \bar{p}_2, ..., \bar{p}_n]'$, $\boldsymbol{K}\boldsymbol{s} = [K_1 s_1, K_2 s_2, ..., K_n s_n]'$, $\boldsymbol{c} = [c_1, c_2, ..., c_n]'$ and $\hat{\boldsymbol{H}}$ is a matrix where the diagonal elements are $\hat{h}_i$, and the other elements are 0, that is

$$
\hat{\boldsymbol{H}} \equiv \begin{bmatrix} \hat{h}_1 & 0 & \cdots & 0 \\ 0 & \hat{h}_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \hat{h}_n \end{bmatrix} \tag{168}
$$

Then the equilibrium price and quantity are given by

$$
\begin{aligned}
\boldsymbol{q} =& (\boldsymbol{I} + \hat{\boldsymbol{H}}\boldsymbol{\Psi})^{-1} \hat{\boldsymbol{H}}(\bar{\boldsymbol{p}} + \boldsymbol{K}\boldsymbol{s} - \boldsymbol{c}) && (169) \\
\boldsymbol{p} =& \bar{\boldsymbol{p}} - \boldsymbol{\Psi}(\boldsymbol{I} + \hat{\boldsymbol{H}}\boldsymbol{\Psi})^{-1} \hat{\boldsymbol{H}}(\bar{\boldsymbol{p}} + \boldsymbol{K}\boldsymbol{s} - \boldsymbol{c}) + \boldsymbol{b} && (170)
\end{aligned}
$$

and the risk-adjusted profit in the first stage is

$$\mathbf{E}[U_i] = \mathbf{E}\left[q_i(\mathbf{E}\left[p_i \mid \mathcal{I}_i\right] - c_i) - \frac{\rho_i}{2}\mathbf{Var}\left[p_i \mid \mathcal{I}_i\right] q_i^2\right] - g\left(\chi_c, c_i\right) \tag{171}$$

$$= \mathbf{E}\left[\hat{h}_i^{-1} q_i^2 - \frac{\rho_i}{2}\mathbf{Var}\left[b_i \mid \mathcal{I}_i\right] q_i^2\right] - g\left(\chi_c, c_i\right) \tag{172}$$

$$= \mathbf{E}\left[\left(\psi_{ii} + \frac{\rho_i}{2}\mathbf{Var}\left[b_i \mid \mathcal{I}_i\right]\right) q_i^2\right] - g\left(\chi_c, c_i\right) \tag{173}$$

$$= \left(\psi_{ii} + \frac{\rho_i}{2}\mathbf{Var}\left[b_i \mid \mathcal{I}_i\right]\right)\mathbf{E}\left[q_i^2\right] - g\left(\chi_c, c_i\right) \tag{174}$$

$$= \left(\psi_{ii} + \frac{\rho_i}{2}\mathbf{Var}\left[b_i \mid \mathcal{I}_i\right]\right)\left(\mathbf{E}\left[q_i\right]^2 + \mathbf{Var}\left[q_i\right]\right) - g\left(\chi_c, c_i\right) \tag{175}$$

where $\mathbf{Var}[q_i] = h_i^{-1}\mathbf{Cov}(p_i, q_i)$ is independent of cost choices.

Then the optimal investment choice satisfies

$$\frac{\partial \mathbf{E}[U_i]}{\partial c_i} = \frac{\partial\left(\psi_{ii} + \frac{\rho_i}{2}\mathbf{Var}\left[b_i \mid \mathcal{I}_i\right]\right)\mathbf{E}\left[q_i\right]^2}{\partial c_i} - \frac{\partial g\left(\chi_c, c_i\right)}{\partial c_i} = 0. \tag{176}$$

Firm i's markup is defined as $M_i = \mathbf{E}[p_i]/c_i$. Using (170) and $\mathbf{E}[b_i] = 0$,

$$\mathbf{E}[\boldsymbol{p}] = \bar{\boldsymbol{p}} - \boldsymbol{\Psi}(\boldsymbol{I} + \hat{\boldsymbol{H}}\boldsymbol{\Psi})^{-1}\hat{\boldsymbol{H}}(\bar{\boldsymbol{p}} + \boldsymbol{K}\boldsymbol{s} - \boldsymbol{c}) \equiv \bar{\boldsymbol{p}} - \boldsymbol{\Omega}\hat{\boldsymbol{H}}(\bar{\boldsymbol{p}} + \boldsymbol{K}\boldsymbol{s} - \boldsymbol{c}) \tag{177}$$

Then for individual firm i,

$$\mathbf{E}[p_i] = \bar{p}_i - \sum_{j=1}^{n}\omega_{ij}\hat{h}_j(\bar{p}_j + K_j s_j - c_j) \tag{178}$$

$$M_i = \mathbf{E}[p_i]/c_i \tag{179}$$

And we can write the equivalent of the equation in Proposition 1 as

$$\frac{\partial M_i}{\partial n_{di}} = \frac{\partial(\mathbf{E}[p_i]/c_i)}{\partial n_{di}} = \underbrace{\frac{c_i}{c_i^2}\frac{\partial \mathbf{E}[p_i]}{\partial n_{di}}}_{\text{Risk premium effect}} - \underbrace{\frac{\mathbf{E}[p_i]}{c_i^2}\frac{\partial c_i}{\partial n_{di}}}_{\text{Investment effect}} \tag{180}$$

Bertrand Cost Choices The optimal choices of marginal cost will be

$$\frac{\partial \mathbf{E}[U_i]}{\partial c_i} = \frac{\partial\left(\gamma_{ii} + \frac{\rho_i}{2}\sum_{j=1}^{n}\gamma_{ij}^2\mathbf{Var}\left[b_j \mid \mathcal{I}_j\right]\right)\mathbf{E}\left[p_i - c_i\right]^2}{\partial c_i} - \frac{\partial g\left(\chi_c, c_i\right)}{\partial c_i} = 0 \tag{181}$$

Assume that $g\left(\chi_c, c_i\right) = \frac{\chi_c}{2}(c_i - \bar{c})^2$.

$$\left(\gamma_{ii} + \frac{\rho_i}{2}\sum_{j=1}^{n}\gamma_{ij}^2\mathbf{Var}\left[b_j \mid \mathcal{I}_j\right]\right)\frac{\partial \mathbf{E}\left[p_i - c_i\right]^2}{\partial c_i} = \chi_c(c_i - \bar{c}) \tag{182}$$

Since $\boldsymbol{p} = (\boldsymbol{I} + \hat{\boldsymbol{T}}\boldsymbol{\Gamma})^{-1}\boldsymbol{c} + (\boldsymbol{I} + \hat{\boldsymbol{T}}\boldsymbol{\Gamma})^{-1}\hat{\boldsymbol{T}}\boldsymbol{\Gamma}(\bar{\boldsymbol{p}} + \boldsymbol{K}\boldsymbol{s}) \equiv \boldsymbol{\xi}\boldsymbol{c} + \boldsymbol{\kappa}(\bar{\boldsymbol{p}} + \boldsymbol{K}\boldsymbol{s})$,

$$p_i = \xi_{ii}c_i + \xi_{ij}c_j + \kappa_{ii}(\bar{p}_i + K_i s_i) + \kappa_{ij}(\bar{p}_j + K_j s_j) \tag{183}$$

$$p_i - c_i = (\xi_{ii} - 1)c_i + \xi_{ij}c_j + \kappa_{ii}(\bar{p}_i + K_i s_i) + \kappa_{ij}(\bar{p}_j + K_j s_j) \tag{184}$$

$$(p_i - c_i)^2 = (\xi_{ii} - 1)^2 c_i^2 + 2[\xi_{ij}c_j + \kappa_{ii}(\bar{p}_i + K_i s_i) + \kappa_{ij}(\bar{p}_j + K_j s_j)](\xi_{ii} - 1)c_i + \text{terms not related to } c_i \tag{185}$$

$$\mathbf{E}(p_i - c_i)^2 = (\xi_{ii} - 1)^2 c_i^2 + 2(\xi_{ij}c_j + \kappa_{ii}\bar{p}_i + \kappa_{ij}\bar{p}_j)(\xi_{ii} - 1)c_i + \text{terms not related to } c_i \tag{186}$$

$$\frac{\partial \mathbf{E}\left[p_i - c_i\right]^2}{\partial c_i} = 2(\xi_{ii} - 1)^2 c_i + 2(\xi_{ij}c_j + \kappa_{ii}\bar{p}_i + \kappa_{ij}\bar{p}_j)(\xi_{ii} - 1) \tag{187}$$

Then the first order condition will be

$$\left(\gamma_{ii} + \frac{\rho_i}{2}\sum_{j=1}^{n}\gamma_{ij}^2 \mathbf{Var}\left[b_j \mid \mathcal{I}_j\right]\right)\left[2(\xi_{ii}-1)^2 c_i + 2(\xi_{ij}c_j + \kappa_{ii}\bar{p}_i + \kappa_{ij}\bar{p}_j)(\xi_{ii}-1)\right] = \chi_c(c_i - \bar{c}) \tag{188}$$

Cournot Cost Choices The optimal choices of marginal cost will be

$$\frac{\partial \mathbf{E}[U_i]}{\partial c_i} = \frac{\partial\left(\psi_{ii} + \frac{\rho_i}{2}\mathbf{Var}\left[b_i \mid \mathcal{I}_i\right]\right)\mathbf{E}\left[q_i\right]^2}{\partial c_i} - \frac{\partial g\left(\chi_c, c_i\right)}{\partial c_i} = 0 \tag{189}$$

Assume that $g\left(\chi_c, c_i\right) = \frac{\chi_c}{2}(c_i - \bar{c})^2$.

$$\left(\psi_{ii} + \frac{\rho_i}{2}\mathbf{Var}\left[b_i \mid \mathcal{I}_i\right]\right)\frac{\partial \mathbf{E}\left[q_i\right]^2}{\partial c_i} = \chi_c(c_i - \bar{c}) \tag{190}$$

Since $\boldsymbol{q} = (\boldsymbol{I} + \hat{\boldsymbol{H}}\boldsymbol{\Psi})^{-1}\hat{\boldsymbol{H}}(\bar{\boldsymbol{p}} + \boldsymbol{K}\boldsymbol{s} - \boldsymbol{c}) \equiv \boldsymbol{\zeta}(\bar{\boldsymbol{p}} + \boldsymbol{K}\boldsymbol{s} - \boldsymbol{c})$,

$$q_i = \zeta_{ij}(\bar{p}_j + K_j s_j - c_j) + \zeta_{ii}(\bar{p}_i + K_i s_i - c_i) \tag{191}$$

$$q_i^2 = \zeta_{ii}^2(\bar{p}_i + K_i s_i - c_i)^2 + \zeta_{ij}^2(\bar{p}_j + K_i s_j - c_j)^2 + 2\zeta_{ii}\zeta_{ij}(\bar{p}_i + K_i s_i - c_i)(\bar{p}_j + K_i s_j - c_j) \tag{192}$$

$$\text{since } s_i \sim N(b_i, \sigma_s^2), \textit{i.i.d.} \text{ and } \mathbf{E}(b_i) \text{ in the first period is 0.} \tag{193}$$

$$\mathbf{E}q_i^2 = \zeta_{ii}^2[(\bar{p}_i - c_i)^2 + K_i^2\sigma_s^2] + \zeta_{ij}^2[(\bar{p}_j - c_j)^2 + K_j^2\sigma_s^2] + 2\zeta_{ii}\zeta_{ij}(\bar{p}_i - c_i)(\bar{p}_j - c_j) \tag{194}$$

$$\frac{\partial \mathbf{E}q_i^2}{\partial c_i} = -2\zeta_{ii}^2(\bar{p}_i - c_i) - 2\zeta_{ii}\zeta_{ij}(\bar{p}_j - c_j) \tag{195}$$

Then the first order condition will be

$$-2(\psi_{ii} + \rho_i \mathbf{Var}\left[b_i \mid \mathcal{I}_i\right])(\zeta_{ii}^2(\bar{p}_i - c_i) + \zeta_{ii}\zeta_{ij}(\bar{p}_j - c_j)) = \chi_c(c_i - \bar{c}) \tag{196}$$

Simulations. Simulations show the combined effect of the cost and risk channels. Figure 7 shows that markups and prices under Bertrand are always below those under Cournot. Because investment determines the marginal cost endogenously, marginal costs under Bertrand are higher than under Cournot. Anticipating lower profits under Bertrand, firms invest less than under Cournot. Note that more information (more draws of data) leads to lower prices but higher markups as uncertainty is reduced. The finding that Cournot pries and markups are higher than under Bertrand are robust to changes in the cost of investment. However, as Figure 8 illustrates, for high investment costs ($\chi_c = 10$) expected markups and expected prices change in opposite directions under Cournot compared to Bertrand when the number of data points increases.

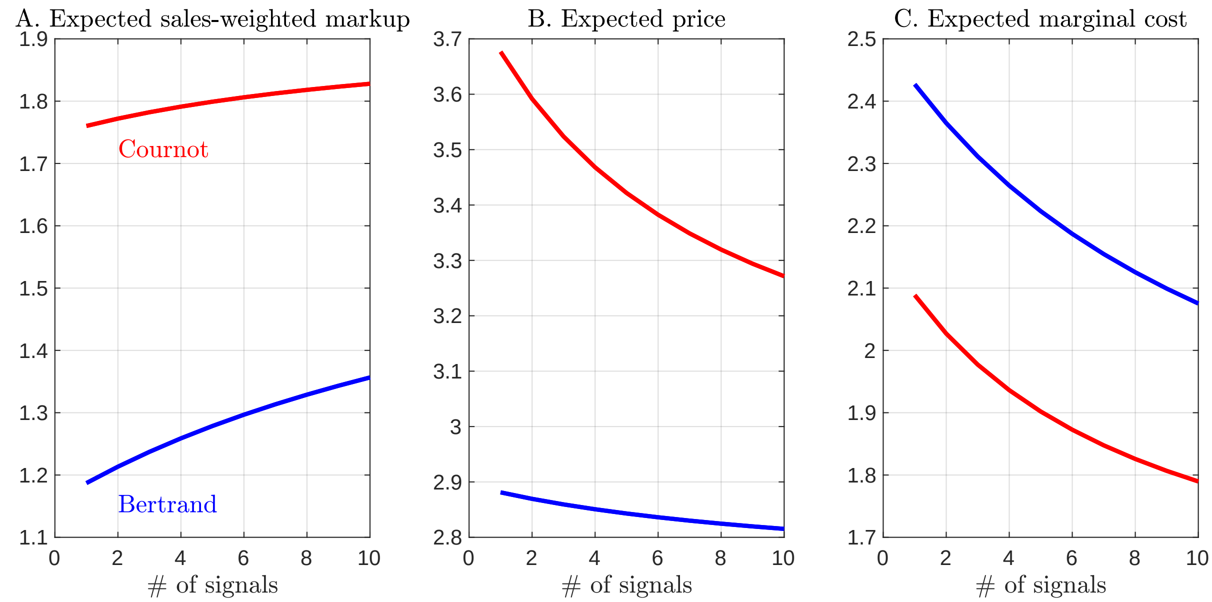


Figure 7: Cournot and Bertrand competition with investment channel dominating the risk channel.
Notes: This comparative static exercise is constructed over a duopoly example. The x-axis is the number of data points that both firms have. The investment cost function is assumed as $g(\chi_c, c_i) = \chi_c (\bar{c} - c_i)^2 / 2$ with $\chi_c = 1$ and $\bar{c} = 3$. Other parameters are: $\overline{p} = 5$, $\sigma_b = 1$, $\mu_b = 0$, $\sigma_e = 2$, $\rho_1 = \rho_2 = 1$ and $\Psi = [1, 0.5; 0.5, 1]$.

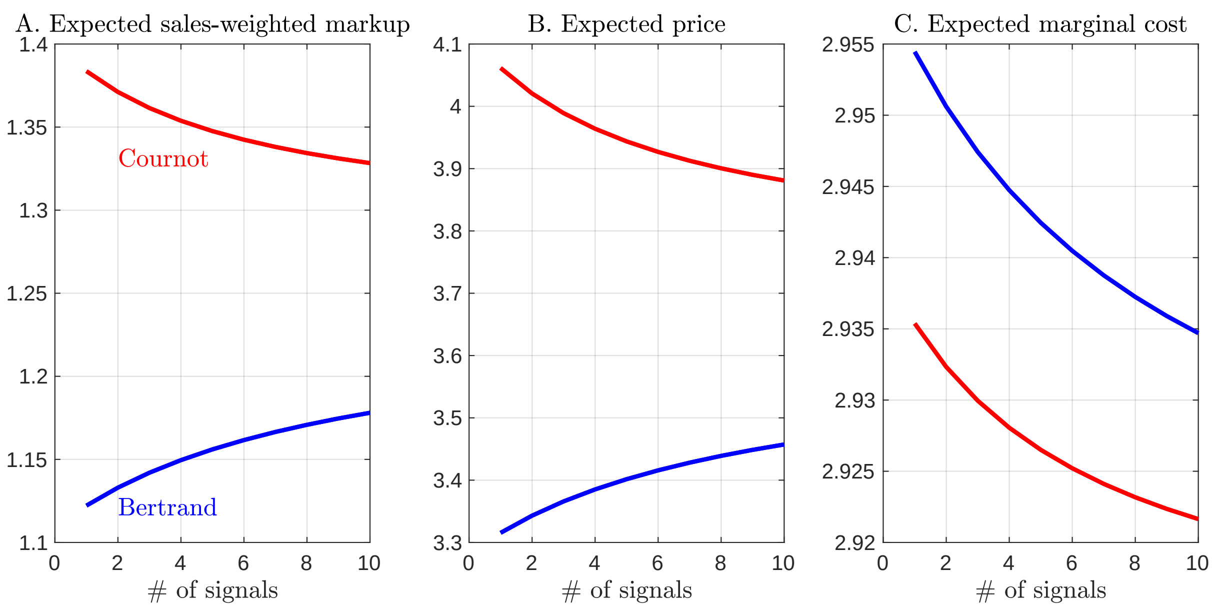


Figure 8: Cournot and Bertrand competition when investment channel matters less than risk.
Notes: This comparative static exercise is constructed over a duopoly example. The x-axis is the number of data points that both firms have. The investment cost function is assumed as $g(\chi_c, c_i) = \chi_c (\bar{c} - c_i)^2 / 2$ with $\chi_c = 10$ and $\bar{c} = 3$. Other parameters are: $\overline{p} = 5$, $\sigma_b = 1$, $\mu_b = 0$, $\sigma_e = 2$, $\rho_1 = \rho_2 = 1$ and $\Psi = [1, 0.5; 0.5, 1]$.