Data and Markups: A Macro-Finance Perspective*

Jan Eeckhout
UPF Barcelona†

Laura Veldkamp
Columbia‡

February 22, 2023

Abstract

How can we measure the extent to which data-intensive firms are using their market power? Economists typically look to markups as evidence of market power. Using a simple model with firms that price risk in their capital allocation and production decisions, we highlight the competing forces that make markups an unreliable measure of data-derived market power. Instead, we show how markups measured at different levels of aggregation reflect data and distinguish data from other intangible investments. These findings both reconcile seemingly contradictory empirical markup evidence and guide us to new ways of measuring data and its effects on markets.

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, and Bruno Pellegrino as well as seminar audiences for insightful discussions. We benefited from superb research assistance by Renjie Bao, Huasheng Nie Xiaobo Yu, and Judy Yue. Eeckhout gratefully acknowledges support from the ERC, Advanced grant 882499, and from ECO2015-67655-P. This paper originally circulated under the title, "Data and Market Power."

† jan.eeckhout@upf.edu – ICREA-BSE-CREI
‡ lv2405@columbia.edu

I Introduction

Changes in firms' market power and the sources of those changes have become the focus of intense debate. Economists point to 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). How should one measure the extent of this effect? To explore the link between data, competition, and measurable outcomes, we craft a model in which economies of scale in data induce a data-rich firm to invest in producing at a lower marginal cost and in capturing a larger market share. However, the model uncovers a rich set of interactions between data and market power measures. Data has competing effects on markups, and the trade-off between these competing effects depends on the level of aggregation at which markups are measured. The model's predictions are consistently supported by an existing empirical literature that measures markups at the product, firm, or industry level and finds different cyclical and trend behaviors at those levels. Not only does the model allow us to interpret and reconcile existing facts, it also teaches us that the difference in markups aggregated in different ways is exactly where we should be looking to see the effects of data.

To develop a tool to identify and understand the effects of data, Section II formulates a new framework, drawing on tools from multiple fields. As in macro theory, data is modeled as information; as in corporate finance, firms price risk; as in industrial organization theory, firms exploit market power. Data is digitized information. The essence of information is that it reduces uncertainty. Just like a larger data set reduces the econometrician's standard error, more data for firms reduces their forecast errors. Making future events more predictable allows the firm to make better decisions that raise expected profits, and also means the firm faces less uncertainty and less risk. While abstracting from risk is appropriate to study many questions, abstracting from risk and the price of risk when studying data removes its essential character. Since risk is central to the study of data, we adopt basic tools from corporate finance to model the ways in which firms use risk pricing to guide their production and investment decisions. Although the assumption that firms price 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).

When firms price risk in their decisions, markups conflate data and its competitive effects. A firm that prices risk requires a return to induce them to engage in risky production. This return is derived from their markup. In other words, markups compensate firms for the risk they bear. If firms use data to improve their forecasts and make their revenues more predictable and less risky, then these firms require less return to bear less risk. So, holding the size of the firm fixed, more data reduces markups. We call this the "risk channel."

However, data also makes it more profitable to grow large and develop a dominant market position. In the model, firms to choose an up-front investment, which lowers their future marginal cost of production. Because the benefits of production are unknown, this up-front investment is risky. When data lowers that risk 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. The notion that investment in cost reduction is a source of market power is in line with the view of Sutton (1991, 2001). In his language, our firms strategically use data to further differentiate themselves and thus create a dominant position. This force whereby data increases markups is what we call the "investment channel." The idea that there are socially good and bad aspects to markups, and that the balance between the two may change over time, is consistent with the evidence of Covarrubias, Gutiérrez, and Philippon (2020).

Section III shows that when a firm has more data, these two forces—the risk and investment channels—push markups in opposite directions. Therefore, markups are not a reliable indicator of data or its competitive effects. Furthermore, this trade-off between efficiency and risk compensation is similar to the trade-off between efficiency and compensation for fixed costs articulated in Weiss (2000). These two mechanisms would be difficult to distinguish based on product markups alone.

There are features of markups that allow us to identify the unique characteristics of data relative to other assets. Specifically, the growing volume of data can cause markups aggregated in different ways to diverge. The reason is that data facilitates prediction. Better predictions alter the composition of products and firms. Firms use data to adjust production. They produce more of goods that their data predicts are likely to be profitable. Of course, profitable goods are high-markup goods. Thus, even if two firms sell identical goods at identical prices with identical marginal costs, the firm with more data will be measured as a higher-markup firm because that firm uses data to skew the composition of its goods production toward higher-markup goods. Not only can data explain composition effects in markups, but some sort of information is also required 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 strategy if the firm can predict what will be profitable and what will not. Good prediction requires information.

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 creates another aggregation issue. The industry markup is likely to be higher than the average firm's markup because high-markup firms are bigger and more heavily weighted. Section III shows that the more data firms have, the more of a wedge will arise between product-level markups, firm-level markups, and industry markups measured as cost-weighted or sales-weighted aggregates. These wedges can then be used to back out the amount of data a firm has or the average amount of data held in an industry.

The fact that growing data creates wedges between various markup measures is not a mere curiosity. Such wedges exist in the data and are growing. Thus the model helps explain a curious feature of the data that has been at the heart of a debate about growing markups. From one perspective, markets are just as competitive today as in the past because good-level markups are stable (see for example Anderson, Rebelo, and Wong [2018]). Instead, growing firm-level and industry markups are evidence of declining competition (see Gutiérrez and Philippon [2016]; Furman and Orszag [2015]; Grullon, Larkin, and Michaely [2016]; De Loecker, Eeckhout, and Unger [2020]). Moreover, the distribution of markups and market shares has become more skewed, and as a result, the aggregation of markups gives rise to a different evolution of industry markups (see Hall [2018]). These facts validate our model. In turn, the model lends an economic interpretation to these facts beyond the explanation that composition effects must be at work.

Ultimately, most researchers are interested in markups because they are concerned about consumer welfare. Section IV discusses the relationship between markups, competitive outcomes, and welfare. Rising amounts of data can be good for consumers. After all, firms use data to produce more of goods that consumers want most. However, welfare may suffer when firms' data stocks become asymmetric. Our model can help to quantify that trade-off.

Another unexpected prediction of this model may help to reconcile an empirical debate about whether markups are pro- or countercyclical. This debate is central to the relevance of New Keynesian models. We find that data-intensive firms may have procyclical product markups, as in Nekarda and Ramey (2020), but countercyclical firm and industry markups (Bils 1985, 1987). In section VI, recessions are times when demand is lower on average, but also more volatile. The lower demand lowers markups. When demand is more volatile, firms that can use data to identify which product is currently in high demand can better adjust output to increase the firm's markup and profit by more than other firms. In short, higher volatility raises firm and industry markups because it creates a potential for larger composition effects. Understanding why markups measured differently have different properties allows researchers to determine which set of facts is most relevant for a given question.

A central concern about market power for digital firms is that large firms can generate lots of data, which makes them more productive and reinforces their competitive position. To understand how this dynamic fits in with our analysis, Section VII shows how our simple model could be made dynamic. When data is a by-product of firms' economic activity, this changes markups. A new term arises that pushes prices and markups downs. Firms that value data want to do more transactions to generate more data. To do more transactions, a firm must lower its price. The optimal production decision for a firm reveals that price and the marginal value of data enter as substitutes. Firms are effectively paid either with money or with their customers' data. This finding relates this way of thinking to the study of free digital goods. The idea that price does not fully capture the value of a transaction to a firm also provides one more reason that markups fail to capture market power in a data economy.

Finally, we turn to measurement. Since our contribution is to use theory to reinterpret and repurpose existing facts, we do not measure or contribute new facts. We show that if firms are increasingly using data to predict their profits, then subtle differences in how economists measure markups can deliver starkly different trends and cyclical patterns. So far, the debate surrounding these different patterns has been about the merits of measuring with micro or macro data. Our explanation—that the differences arise from firms' use of data—reframes this debate and lends new meaning to the facts. The differences no longer represent a mistake made by one group or the other, but rather an interesting and useful measure of the quantity of firms' data.

Section VIII offers guidance for how this framework might be used to measure data or the market power arising from that data. Our model teaches us that the difference between firm and product markups is a sufficient statistic for the amount of relevant data a profit-maximizing firm has about consumer demand. This would enable a sufficient statistics approach to measuring firms' data. The model could also be used for structural estimation. We discuss techniques to estimate firms' price of risk and approaches to measuring product characteristics, and we map the markup measures in our model to different empirical approaches in the markup literature.

RELATED LITERATURE Because we model data as digitized information, our tools are 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. 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 in response to information is central to our study of market power.

Existing work on the digital economy explores whether data can be a source of market power. 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. 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. De Ridder (2021) considers information technology to be something that raises fixed costs and reduces marginal costs. We do not dispute that data can be used for all of these purposes. However, at its essence, data is digitized information; information is used to reduce uncertainty or risk. The new elements we introduce are a product portfolio choice and non-indifference to risk. Both are well supported by evidence and central to our results that advance this literature by offering a new approach to data measurement and new considerations in welfare.

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. 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 explore the role of firms' data.

Empirical work on the data economy often necessarily focuses on specific markets.[1] Lambrecht and Tucker (2015) take a strategy perspective on whether data has the necessary features to confer market power. Similarly, Goldfarb and Tucker (2017) discuss the many ways in which this digital economy is transformative.

Recent work by Burstein, Carvalho, and Grassi (2020) analyzes the business cycle properties of markups. They show analytically how the sign of markup cyclicality varies with aggregation and they establish the economic importance of these markups. Our results complement these insights by proposing a specific mechanism that causes markups to fluctuate, one that is rooted in how firms use data to gain a competitive advantage.

[1] Athey, Mobius, and Pal (2017) and Athey and Gans (2010) examine media competition; Brynjolfsson, Hu, and Smith (2003) study booksellers; and Rajgopal, Srivastava, and Zhao (2021) measure digital technology firms. de Cornière and Taylor (2020) categorize uses of data as pro- or anticompetitive.

II Model

To explore the idea that data can create market power, we build a model with a few key features. 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. Allowing those goods to have correlated attributes, as in Pellegrino (2020), makes data relevant to 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. This allows us to explore if high-data firms invest to operate at a larger scale and thus grow to have more market power.

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 VII 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.

II.A Setup

FIRMS There are n_F firms, indexed by i: $i \in \{1, 2, \ldots, n_F\}$. 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.

$$U_i = \mathbf{E}\left[\pi_i | \mathcal{I}_i\right] - \frac{\rho_i}{2} \mathbf{Var}\left[\pi_i | \mathcal{I}_i\right] - g(\chi_c, \widetilde{\boldsymbol{c}}_i). \tag{1}$$

This mean-variance objective is consistent with empirical corporate finance evidence on firms' decisions (Eckbo 2008) and is a second-order approximation to a broader class of utility functions.

Firm production profit π_i depends on quantities of each good, $\mathbf{q}_i$, the market price of each good, $\mathbf{p}$, and the marginal cost of production of that good, $\boldsymbol{c}_i$:

$$\pi_i = q_i' \left(\mathbf{p} - \boldsymbol{c}_i\right). \tag{2}$$

Prior to observing any of their data, each firm chooses an up-front investment. Let $\widetilde{\boldsymbol{c}}_i$ be the vector of marginal production costs for a unit of each attribute. The up-front investment choice is modeled as a choice of $\widetilde{\boldsymbol{c}}_i$ at an investment cost $g(\chi_c, \widetilde{\boldsymbol{c}}_i)$ to maximize $E[U_i]$. The function g is strictly decreasing in $\widetilde{\boldsymbol{c}}_i$. Since lower choices of $\widetilde{\boldsymbol{c}}_i$ involve greater costs, 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 $\partial^2 g/\partial\chi_c \partial c_i < 0$. To guarantee non-negative interior marginal cost choices, one could impose $g(\chi_c, \widetilde{c}_i)$ is convex over $\widetilde{c}_i$, with $g(\chi_c, \bar{c}) = 0$ and $\lim_{\widetilde{c}\to 0} g(\chi_c, \widetilde{c}) = +\infty$. However, most of our results will not require this.

PRODUCTS AND ATTRIBUTES The product space has N attributes, indexed by $j \in \{1, 2, \ldots, N\}$. Goods, indexed by k, are combinations of attributes.

Each good $k \in \{1, 2, \ldots, N\}$ can be represented as an $N \times 1$ vector $\boldsymbol{a}_k$ of weights that good places on each attribute. The jth entry of vector $\boldsymbol{a}_k$ 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 $\boldsymbol{a}_k$'s for each good k be an $N \times N$, full-rank matrix A. For now, the mapping between attributes and products is fixed. Later, we allow firms to choose how to position their product in the product space by choosing A's.

The quantity of attributes that a firm i produces is a vector $\tilde{\mathbf{q}}_i$, with jth element $\tilde{q}_{ij}$. The attribute vector is the vector of firm i's product quantities, q_i, times the inverse attribute matrix A^{-1}:

$$\tilde{\mathbf{q}}_i = A^{-1}\mathbf{q}_i. \tag{3}$$

The marginal cost of producing a good depends on the up-front investment the firm makes and on that good's attributes. The firm's up-front investment of $g(\chi_c, \widetilde{\boldsymbol{c}}_i)$ allows it to produce each attribute j at a unit cost of $\widetilde{c}_{ij}$. The vector $\widetilde{\boldsymbol{c}}_i$ is the N-by-1 vector of all marginal production costs of firm i for each attribute. The vector $\boldsymbol{c}_i = A'\widetilde{\boldsymbol{c}}_i$ is the vector of firm i's marginal cost for each product. The cost of producing a unit of good k for firm i is therefore $\boldsymbol{c}_i = \boldsymbol{a}_k'\widetilde{\boldsymbol{c}}_i$. To keep the investment problem bounded, the investment cost function g is convex in each element $\widetilde{c}_{ij}$.

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

The price of each good depends on the attributes of a good. The price of good k is the units of each attribute a_k times the price of each attribute $\tilde{p}$:

$$p_k = \sum_{j=1}^{N} a_{jk}\tilde{p}_j. \tag{4}$$

Each attribute j has an average market price that depends on an attribute-specific constant and on the total quantity of that attribute that all firms produce:

$$\tilde{p}_j^M = \bar{p}_j - \frac{1}{\phi} \sum_{i=1}^{n_F} \tilde{q}_{ij}. \tag{5}$$

Each firm does not receive the market price for its good, but rather has a firm-specific price that depends on a firm-specific 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: $\mathbf{b}_i \sim N(0, I)$, which is i.i.d. across firms. The price a firm receives for a unit of attribute j is thus $\tilde{p}_j + b_{ij}$. We can express firm i's price in vector form as

$$\tilde{p}_i = \left[\tilde{p}_1^M, \tilde{p}_2^M, \ldots, \tilde{p}_N^M\right]' + b_i. \tag{6}$$

The price a firm receives for a unit of good k is therefore $p_k + \sum_{j=1}^N a_{jk} b_{ij}$.

INFORMATION Each firm generates n_{di} data points. Each data point is a signal about the demands for each attribute: $\tilde{\mathbf{s}}_{i,z} = \boldsymbol{b}_i + \tilde{\boldsymbol{\varepsilon}}_{i,z}$, where $\tilde{\boldsymbol{\varepsilon}}_{i,z} \sim N(\mathbf{0}, \tilde{\Sigma}_e)$ is an $N \times 1$ vector. Signal noises are uncorrelated across attributes and across firms. All firms can observe all the data generated by each firm. Of course, other firms' data is not relevant for inferring b_i. But this allows firms to know what other firms will do.

Because we are interested in how data affects competition, we will take data (n_{di} and $\tilde{\Sigma}_e$) as given. The question will be what happens to market competition and markups when we exogenously change these data conditions of some or all firms. Section VII explores what aspects of the results change when data is generated as a by-product of economic transactions.

EQUILIBRIUM

1. Each firm chooses a vector of marginal costs $\tilde{c}_i$, 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 (1).

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 (1), taking as given other firms' choices.

3. Prices clear the market for each good.

II.B Discussion of Assumptions

DATA THAT IS PUBLIC INFORMATION. The assumption that all data is public is obviously not realistic, but it is also not crucial for any of our main results. It does simplify the mathematics considerably. In a model with private signals, firms also use data to forecast what other firms will do. Data reduces risk in two ways—by predicting demand for the firm's products and by shedding light on the production decisions of other firms. Appendix C.2. shows that this strengthens the risk channel because data reduces both demand uncertainty and strategic uncertainty, prompting more production and more investment.

FIRM-SPECIFIC DEMAND SHOCKS. We also assume that shocks are firm-specific to simplify the exposition. Appendix C.1. solves the aggregate shock model and shows that all the main forces we identify here are present. The reason we relegate that model to the appendix is that the solution is complicated by firms' need to forecast what other firms know, as in Angeletos and Pavan (2007). In that setting, we can prove similar theoretical properties. But because the solutions are implicit functions, those results are less clear and thus less useful for expositing the main ideas.

FIRMS THAT PRICE RISK. The assumption that firms price risk is central to our analysis of markups. While risk pricing is novel in the markups literature, it is a bedrock principle of corporate finance and is well supported by numerous empirical studies in many domains.[2] Even if firms themselves are not risk averse, firms that take on risky projects will face a higher cost of capital.[3] So, the price-of-risk term could be interpreted as an adjustment to their expected profit. 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.

Of course, corporate finance typically assumes firms price aggregate risk. Our firms price firm-specific risk because we are exploring a market with large players where firm-specific risk is not diversifiable. Thus, the ρ_i term in (1) could capture both the price of risk and the covariance of the firm shock with market risk (the firm's beta). Furthermore, there is growing evidence that even

[2]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.

[3]While a risk price affects a firm's cost of borrowing, simply including a risky interest rate in marginal cost does not suffice. The risk of debt is far less than the risk to the enterprise value of the whole firm. Markups need to compensate the firm for risk to its debt and its equity.

idiosyncratic risk is priced, especially when firms face financial constraints (Hennesy and Whited 2007). Another reason to price the risk is that we know that the model with aggregate risk delivers similar results (Appendix C.1.). So, we can think of this is as a simplified version of that aggregate risk model. Finally, it is also possible to interpret ρ_i as the absolute risk aversion of a firm manager who is compensated with firm equity.

DATA ABOUT CONSUMER DEMAND. One might question whether data is used to forecast demand or marginal cost. Conceptually, it shouldn't matter. Firms that face risk from their cost structure should also face a higher cost of becoming more productive. 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 empiricists typically studied inverse markups, then it would be more practical to study cost uncertainty.

LINEAR DEMAND. We assume a linear demand system, which is common in the information aggregation literature.[4] Not only does the linearity assumption permit transparent results, recent work also shows that the linear setup fits the data well. Our model builds on Pellegrino (2020)'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.

However, 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.

GOODS AS BUNDLES OF ATTRIBUTES. We treat goods as collections of attributes, but this is not essential for our theoretical results. All results hold if $A = I$, in which case goods are attributes. However, the attribute structure creates correlation in demand across assets. That affects composition results and is important for measurement. To use this framework to measure a firm's data, it is crucial to recognize that information about one product can be informative about another. The correlated demand created by our attribute structure is what makes data relevant for multiple products.

[4]See Chapters 7–8 in Veldkamp (2011) for a textbook treatment.

Also, attribute-based demand is historically used in industrial organization (IO) economics because it allows researchers to predict what would happen if a new good was introduced.

NO ATTRIBUTE CHOICE. Appendix C.3. explores a model where a firm can choose the attributes of its good. The same forces are at work in that model. We choose to work with a simpler model without attribute choice to elucidate the main ideas more clearly.

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.

NO DATA CHOICE. Our main question is what the effect of data is on competition. To answer a question about the effect of data, it makes sense to take data as exogenous and explore what happens when the amount of data changes. However, future work with different objectives might investigate determinants of firms' data choices.

NO ENTRY OR EXIT. We take the number of firms as given, as in Rostek and Weretka (2012). 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.

II.C 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 c_i.

BAYESIAN UPDATING According to Bayes' law for normal variables, observing n_{di} signals, each with signal noise variance $\tilde{\Sigma}_e$, is the same as observing the average signal $\boldsymbol{s}_i = (1/n_{di}) \sum_{z=1}^{n_{di}} s_{iz} = \boldsymbol{b}_i + \boldsymbol{\varepsilon}_i$, where the variance of $\boldsymbol{\varepsilon}_i$ is $\Sigma_{\epsilon_i} = \tilde{\Sigma}_e / n_{di}$. Therefore, do a change of variable, replacing $\tilde{\Sigma}_e / n_{di}$ with Σ_{ϵ_i}. In this representation, more data points (higher n_{di}) shows up as a lower composite signal noise Σ_{ϵ_i}.

Define $\boldsymbol{K}_i$ to be the sensitivity of beliefs to the signal (also called the Kalman gain): $\boldsymbol{K}_i := (I_N + \Sigma_{\epsilon_i})^{-1}$. Then, firm i's expected value of the shock b_i can be expressed simply as $E[b_i|\mathcal{I}_i] = K_i s_i$. The expectation and variance of the pricing function (5) are

$$\begin{aligned}
\mathbf{E}\left[\tilde{\boldsymbol{p}}_i|\mathcal{I}_i\right] &= \bar{\boldsymbol{p}} + \mathbf{E}\left[\boldsymbol{b}_i|\mathcal{I}_i\right] - \frac{1}{\phi}\sum_{i'=1}^{n_F} q_{i'} \\
&= \bar{\boldsymbol{p}} + \boldsymbol{K}_i \boldsymbol{s}_i - \frac{1}{\phi}\sum_{i'=1}^{n_F} q_{i'}, \\
\mathbf{Var}\left[\tilde{\boldsymbol{p}}_i|\mathcal{I}_i\right] &= \mathbf{Var}\left[\boldsymbol{b}_i|\mathcal{I}_i\right] = \left(I_N + \Sigma_{\epsilon_i}\right)^{-1}\Sigma_{\epsilon_i}.
\end{aligned} \tag{7}$$

OPTIMAL PRODUCTION The first-order condition with respect to goods production q_i is $\partial U_i / \partial \boldsymbol{q}_i$: $\mathbf{E}\left[\boldsymbol{p}_i|\mathcal{I}_i\right] - \boldsymbol{c}_i - \frac{\partial \mathbf{E}[\boldsymbol{p}_i|\mathcal{I}_i]}{\partial \boldsymbol{q}_i}\boldsymbol{q}_i - \rho_i \mathbf{Var}\left[\boldsymbol{p}_i|\mathcal{I}_i\right]\boldsymbol{q}_i = 0$. Rearranging delivers optimal production:

$$\boldsymbol{q}_i = \left(\rho_i \mathbf{Var}\left[\boldsymbol{p}_i|\mathcal{I}_i\right] - \frac{\partial \mathbf{E}\left[\boldsymbol{p}_i|\mathcal{I}_i\right]}{\partial \boldsymbol{q}_i}\right)^{-1}\left(\mathbf{E}\left[\boldsymbol{p}_i|\mathcal{I}_i\right] - \boldsymbol{c}_i\right). \tag{8}$$

From differentiating the attribute pricing function (5), we find that the price impact of one additional unit of attribute output is

$$\frac{\partial \mathbf{E}\left[\tilde{\boldsymbol{p}}_i|\mathcal{I}_i\right]}{\partial \tilde{\boldsymbol{q}}_i} = -\frac{1}{\phi} I_N. \tag{9}$$

To simplify the problem, we can change the choice variable and have firms choose the optimal vector of attribute production $\tilde{\boldsymbol{q}}_i$. If we rewrite (8), replacing q, p, and c with attribute quantities, prices, and costs $\tilde{\boldsymbol{q}}_i$, $\tilde{\boldsymbol{p}}_i$, and $\tilde{\boldsymbol{c}}_i$, then we can substitute in the price impact (9) and conditional expectation (7) to get $\tilde{\boldsymbol{q}}_i = \hat{\boldsymbol{H}}_i\left(\overline{\boldsymbol{p}} + \boldsymbol{K}_i \boldsymbol{s}_i - \frac{1}{\phi}\sum_{i'} \tilde{\boldsymbol{q}}_{i'} - \tilde{\boldsymbol{c}}_i\right)$, where $\hat{\boldsymbol{H}}_i := \left(\rho_i \mathbf{Var}\left[\boldsymbol{b}_i|\mathcal{I}_i\right] + 1/\phi I_N\right)^{-1}$.

Next, sum production $\tilde{\boldsymbol{q}}_i$ over all firms i to get total production of each attribute $\sum_{i'} \tilde{\boldsymbol{q}}_{i'}$. This sum has a $\sum_{i'} \tilde{\boldsymbol{q}}_{i'}$ on both the left- and right-hand sides. Collect these terms and rearrange to get $\sum_{i'} \tilde{\boldsymbol{q}}_{i'} = \left(I + \frac{1}{\phi}\sum_i \hat{\boldsymbol{H}}_i\right)^{-1}\left[\sum_i \hat{\boldsymbol{H}}_i\left(\overline{\boldsymbol{p}} + \boldsymbol{K}_i \boldsymbol{s}_i - \tilde{\boldsymbol{c}}_i\right)\right]$. Substituting this total production expression for $\sum_{i'=1}^{n_F} \tilde{\boldsymbol{q}}_{i'}$ in firm i's optimal production ($\tilde{\boldsymbol{q}}_i^{*}$) yields the optimal production of each attribute by each firm i:[5]

$$\widetilde{\boldsymbol{q}}_i^* = \hat{\boldsymbol{H}}_i \left(\overline{\boldsymbol{p}} + \boldsymbol{K}_i \boldsymbol{s}_i - \widetilde{\boldsymbol{c}}_i - \left(I_N + \frac{1}{\phi} \sum_{i'} \hat{\boldsymbol{H}}_{i'} \right)^{-1} \left[\sum_{i'} \hat{\boldsymbol{H}}_{i'} \left(\overline{\boldsymbol{p}} + \boldsymbol{K}_{i'} \boldsymbol{s}_{i'} - \widetilde{\boldsymbol{c}}_{i'} \right) \right] \right).$$

Finally, the product-level optimal production function is the attribute weighting matrix A times the optimal attribute production: $\boldsymbol{q}_i^* = A \tilde{\boldsymbol{q}}_i$.

EQUILIBRIUM PRICE Substituting this aggregate quantity in the pricing function (5) yields an equilibrium average price of each attribute:

$$\tilde{\mathbf{p}}^M = \bar{p} - \left(I_N + \frac{1}{\phi} \sum_i \hat{\boldsymbol{H}}_i \right)^{-1} \left[\sum_i \hat{\boldsymbol{H}}_i \left(\overline{\boldsymbol{p}} + \boldsymbol{K}_i \boldsymbol{s}_i - \widetilde{\boldsymbol{c}}_i \right) \right]. \tag{10}$$

The average price of a good k with attribute vector a_k is then simply $p_k^M = \mathbf{a}_k' \tilde{\mathbf{p}}$, and firm i price of good k is $\mathbf{a}_k'(\tilde{\mathbf{p}}^M + \mathbf{b}_i)$.

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

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

$$\frac{\partial \mathbf{E}[U_i]}{\partial \tilde{\boldsymbol{c}}_i} = \frac{1}{2} \frac{\partial \mathbf{E}[\tilde{\boldsymbol{q}}_i]' \left(\frac{2}{\phi} I_N + \rho_i \mathbf{Var}\left[\boldsymbol{b}_i | \mathcal{I}_i \right] \right)^{-1} \mathbf{E}[\tilde{\boldsymbol{q}}]}{\partial \tilde{\boldsymbol{c}}_i} - \frac{\partial g(\chi_c, \tilde{\boldsymbol{c}}_i)}{\partial \tilde{\boldsymbol{c}}_i} = 0, \tag{11}$$

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.

III Main Results: How Data Affects Markups

We begin by exploring how more data affects a firm's choices of how much to produce and how much to invest before production. 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

[5] 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 or attribute, 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.

prices, which in turn lowers markups. More initial investment lowers marginal cost, which raises markups. This section explores that tension.

We begin by defining a product markup.

Definition 1 (Product markup)**.** *The product-level markup for product k produced by firm i is $M^p_{ik} := \mathbb{E}[\boldsymbol{p}_i(k)]/\boldsymbol{c}_i(k)$. The average product-level markup is*

$$\overline{M}^p := \frac{1}{Nn_F}\sum_{i=1}^{n_F}\sum_{k=1}^{N} M^p_{ik}. \tag{12}$$

To derive an expression for the product markup in the model, we simply divide each expected product price, using (10) and $E[\boldsymbol{b}_i] = 0$, by the marginal cost of that product, $\boldsymbol{c}_i = a_k'\widetilde{\boldsymbol{c}}_i$:

$$M^p_{ik} = \frac{1}{a_k'\widetilde{\boldsymbol{c}}_i} a_k' \left(\overline{\boldsymbol{p}} - \left(I_N + \frac{1}{\phi}\sum_{i=1}^{n_F} \hat{\boldsymbol{H}}_i \right)^{-1} \left(\sum_i \hat{\boldsymbol{H}}_i \boldsymbol{A} \left(\overline{\boldsymbol{p}} + \boldsymbol{K}_i \boldsymbol{s}_i - \widetilde{\boldsymbol{c}}_i \right) \right) \right). \tag{13}$$

Similarly, the average product markup for firm i is $\overline{M}^p_i = (1/N)\sum_{k=1}^{N} M^p_{ik}$.

What makes a markup large? Some of these causes of high markups in equation (13) are not surprising. For example, having lots of valuable attributes raises a product's markup. In the model, valuable attributes are large a_{ij}'s, especially for attributes with high expected value $\bar{p}$, relative to their cost c. Also, having fewer firms raises markups: low n_F lowers $\bar{H}$, which makes the negative term on the right smaller. This is the classic concern with concentrated markets.

Low price elasticity (low ϕ) is also a cause of high markups in (13). However, unlike many models, infinite price elasticity ($\phi = \infty$) does not eliminate the markup. In this limiting case, where firms cannot exercise market power, the markup is not zero. The markup is still positive with infinite elasticity because even if firms have no power to affect prices, they still need to be compensated for risk. The markup that remains in the infinite elasticity case is a risk premium.

Other forces arise because firms price risk. When firms are more sensitive to risk, or the price of risk in capital markets is high (high ρ), this also raises markups. They need to charge a higher markup to compensate themselves for the higher financing costs that this risk will incur. This force shows up as high ρ makes $\bar{H}$ low. When firms are very sensitive to risk, they are less sensitive to prices and cost. They won't produce more when there are small changes in profits, because they are too sensitive to the additional risk that might entail.

Finally, two forces show up in the markup formula that are affected by how much data a firm has. Those forces are risk and investment. They often compete and are at the heart of the results that follow. Therefore, we state and derive each formally.

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 1. ***Data-investment complementarity.*** *A firm with more data chooses a lower marginal cost c_i, which entails a higher cost investment and higher profitability Π_i.*

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 2. ***Higher investment raises product markups.*** *More investment (lower c_i choice) in any attribute j of good k, s.t. $a_{jk} > 0$, increases the markup of attribute j. If the markup on attribute j is less than the markup on product k ($M_{i,k} \geq M^{\tilde{p}_{i,j}}$), then this also raises the markup on good k.*

A firm that invests in producing an attribute can produce that attribute at a lower cost. If a good j does not load at all on that attribute ($a_{jk} = 0$), then the lower cost has no bearing on the cost or markup of that good. But if good j contains some of that attribute ($a_{jk} > 0$), this investment lowers the cost of producing the good. 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.

However, investment is only one channel through which data affects markups. The model teaches us that there is a second channel: data reduces the risk of production, induces more production, and thereby lowers prices and markups. We isolate this channel by holding investment (firm size) fixed. Since c_i is, of course, a choice variable which is not fixed, the correct formal statement is that this result holds when the marginal cost of adjusting investment χ is sufficiently high. Approximately, however, this parameter restriction simply serves to hold investment constant so that we can see the effect of data on output in isolation.

Lemma 3. ***Data reduces product markups (risk premium channel).*** *Holding all firms' investments fixed (χ_c sufficiently high), an increase in any firm's data about any attribute of good k reduces the markup of good k.*

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 (8) where the conditional variance in the denominator represents risk. When this variance declines, optimal production rises. More production lowers price and lowers markups. When we reduce risk with data, firms do not need as much markup compensation to be willing to produce.

This effect is always present, regardless of the level of χ_c. The restriction on χ_c is only there to isolate this channel from the investment channel, which is shut down when χ_c is sufficiently high. When χ_c is lower, 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 any attribute of good k increases the markup of good k, which loads positively on that attribute. Otherwise, an increase in any firm's data about any attribute of good k reduces the markup of good k.*

Equation (14) summarizes the effects of data on markups. The partial derivative of markups with respect to data is the difference between the risk premium effect and the investment effect.

$$\frac{\partial M^{\tilde{p}}_{i,j}}{\partial \Sigma^{-1}_{\epsilon_{i},j}} = \underbrace{\frac{\tilde{\boldsymbol{c}}_{i,j}}{\tilde{\boldsymbol{c}}^2_{i,j}} \frac{\partial \boldsymbol{D}_j}{\partial \Sigma^{-1}_{\epsilon_{i},j}}}_{\text{Risk premium effect}} - \underbrace{\frac{\boldsymbol{D}_j}{\tilde{\boldsymbol{c}}^2_{i,j}} \frac{\partial \tilde{\boldsymbol{c}}_{i,j}}{\partial \Sigma^{-1}_{\epsilon_{i},j}}}_{\text{Investment effect}} . \tag{14}$$

where

$$\boldsymbol{D}_j = \frac{\bar{\boldsymbol{p}}_j + \frac{1}{\phi} \sum_{i=1}^{n_F} \hat{\boldsymbol{H}}_{i,j} \boldsymbol{c}_{i,j}}{1 + \frac{1}{\phi} \sum_{i=1}^{n_F} \hat{\boldsymbol{H}}_{i,j}} \quad \text{and} \quad \hat{\boldsymbol{H}}_{i,j} = \left[\frac{1}{\phi} + \rho_i \left(1 + \Sigma^{-1}_{\epsilon_{i},j} \right)^{-1} \right]^{-1} . \tag{15}$$

Proposition 1 simply identifies regions of the parameter space where the first or second term of (14) dominates. High risk aversion makes the risk premium effect large. In contrast, low marginal cost of investment makes investment very responsive to data and makes the investment channel the stronger effect.

Figures 1 and 2 illustrate how the risk reduction and investment forces compete. 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 (Figure 1). 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

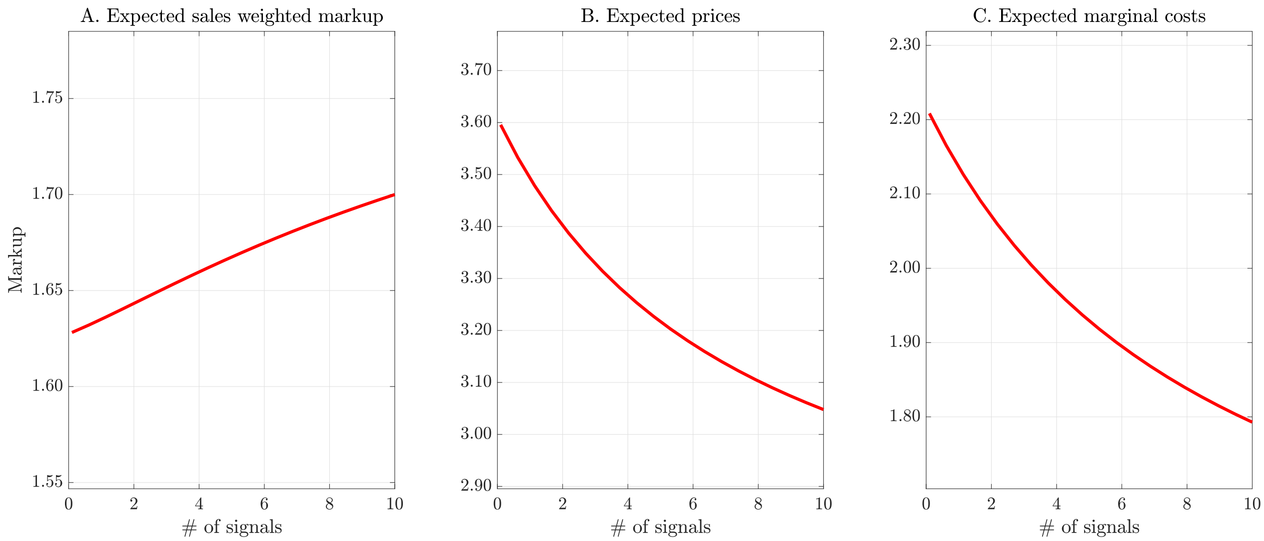


Figure 1: Data raises markups with low investment cost / price of risk

Notes: This comparative static exercise is constructed over a single-good duopoly example. The x-axis is the number of data points that both firms have. The investment cost function is assumed as $g(\chi_c, c_i) = \chi_c \left(\bar{c} - c_i\right)^2 /2$ with $\chi_c = 1$ and $\bar{c} = 3$. Other parameters are: $\overline{p} = 5$, $\phi = 1$, $\sigma_b = 1$, $\mu_b = 0$, $\sigma_e = 2$, and $\rho_1 = \rho_2 = 1$.

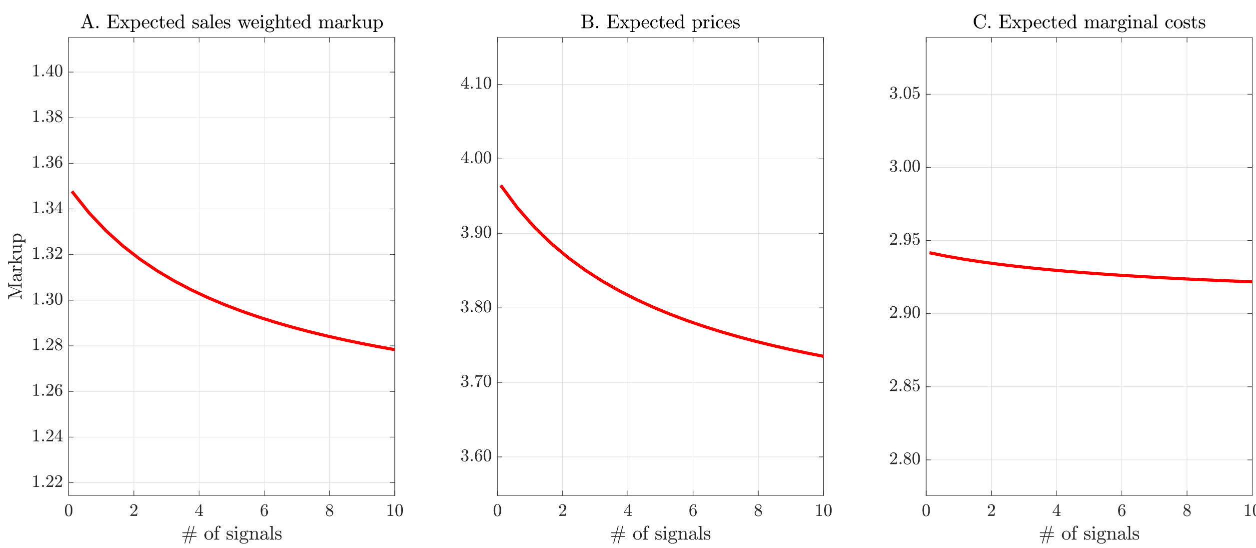


Figure 2: Data lowers markups with high investment cost / price of risk

Notes: This comparative static exercise is constructed over a single-good duopoly example. The x-axis is the number of data points that both firms have. The investment cost function is assumed as $g(\chi_c, c_i) = \chi_c \left(\bar{c} - c_i\right)^2 /2$ with $\chi_c = 10$ and $\bar{c} = 3$. Other parameters are: $\overline{p} = 5$, $\phi = 1$, $\sigma_b = 1$, $\mu_b = 0$, $\sigma_e = 2$, and $\rho_1 = \rho_2 = 1$.

and lower markups (Figure 2). Which scenario prevails depends on the strength of each force in a particular industry.

Despite the fact that markups increase in one case and decrease in the other, both results paint a rosy picture of the role of data. Even when data increases markups, it decreases price. Markups only rise because the firm could produce at a lower cost. Both results point to the efficiency-enhancing and welfare-boosting effects of data. Unfortunately, these are not the only effects data can have. The following results point out the potential problems with this rosy scenario.

IV Welfare

Typically, economists are interested in markups because they are assumed to be indicators of welfare loss or harmful market distortion. In this setting, markups perform a dual role—they are compensation for firm risk-taking and indicators of deadweight loss. This section characterizes efficient markups and welfare. We find that more data typically improves welfare, but it also makes distortions from market power more costly. When firms' stocks of data are asymmetric, exacerbating the data asymmetry can either improve welfare or harm it, depending on whether the risk or the investment effect dominates.

If firms are not compensated for the risk they bear, they will not produce. So a zero markup cannot be the efficient benchmark. Instead, we define prices to be efficient if they arise from production choices of firms that behave as if they were in a competitive market. This leads us to a new measure of market distortion, which we call the risk-adjusted markup.

Competitive firms are those who take market prices as given. In other words, they optimize as if price impact were zero: $\partial E_i[p]/\partial q_i = 0$. If we set price impact to zero in the firm's first-order condition, optimal production is

$$\boldsymbol{q}_i^{comp} = \frac{1}{\rho_i}\mathbf{Var}\left[\boldsymbol{p}_i|\mathcal{I}_i\right]^{-1}\left(\mathbf{E}\left[\boldsymbol{p}_i|\mathcal{I}_i\right] - \boldsymbol{c}_i\right). \tag{16}$$

In other words, production is the same, except that we redefine the sensitivity of production to changes in price or cost in (8) to be $H_i^{comp} = (1/\rho_i)\mathbf{Var}\left[\boldsymbol{p}_i|\mathcal{I}_i\right]^{-1}$.

The fact that market power enters only through the sensitivity term H means that in firm production (8), more market power is mathematically equivalent to increasing the conditional variance $\mathbf{Var}\left[\boldsymbol{p}_i|\mathcal{I}_i\right]$. In other words, risk mimics market power. 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.

The fact that markups reflect risk, as well as market power, suggests that measuring market power should involve controlling for risk. One such measure of market power at the product level might be

$$H_{ik}^{p} - H_{ik}^{comp}.$$

The challenge this poses is that H_{ik}^{comp} is not directly observed from firm behavior. Instead, it requires estimating a firm's data and price of risk. But using the markup wedges to measure data, as described in section VIII, makes this feasible.

WELFARE BENEFITS OF DATA When all firms get more data, this can be a Pareto improvement. Firm owners benefit because more information improves forecasts, which reduce risk that they are averse to. Also, firms with more data invest to be more efficient. On top of that, consumer surplus increases because lower production cost and more information both tighten competition among firms. The next result formalizes this logic.

Proposition 2. ***Data improves welfare.*** *If the investment cost χ_c is sufficiently high, then more data for every firm increases social welfare.*

Figure 3 illustrates this force. The upward slope of the lines tells us that welfare is increasing in the amount of data. This is true even when there is perfect competition. Even when there is no risk aversion, the ability to produce more goods to meet demand still enhances welfare.

DATA AMPLIFIES MARKET POWER COSTS Figure 3 decomposes the welfare loss into risk aversion and market power. The loss due to market power is much higher on the right, where data is abundant.

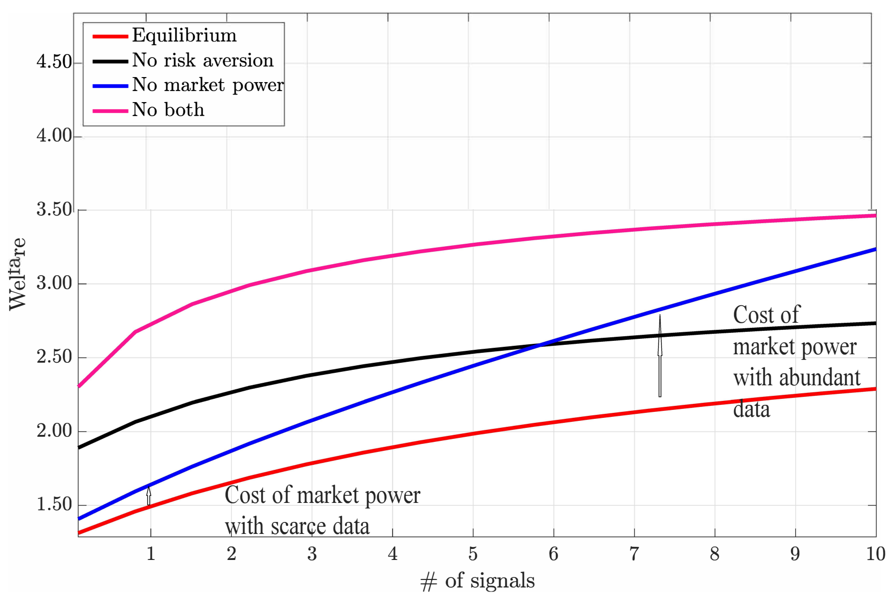


Figure 3: Welfare: Abundant data raises welfare, makes market power more costly

Notes: This counterfactual exercise is constructed over a single-good duopoly example. The x-axis is the number of data points that both firms have. The investment cost function is assumed as $g(\chi_c, c_i) = \chi_c \left(\bar{c} - c_i\right)^2 / 2$ with $\chi_c = 1$ and $\bar{c} = 3$. Other parameters are: $\overline{p} = 5$, $\phi = 1$, $\sigma_b = 1$, $\mu_b = 0$, $\sigma_e = 2$, and $\rho_1 = \rho_2 = 1$.

The reason that data makes market power more powerful can be seen in the first-order condition (8) of the firm's choice of production quantities q. The right term is expected profit per unit. That expected profit is divided by the term $\rho_i \mathbf{Var}\left[\boldsymbol{p}_i|\mathcal{I}_i\right] - \frac{\partial \mathbf{E}[\boldsymbol{p}_i|\mathcal{I}_i]}{\partial \boldsymbol{q}_i}$, which captures risk price ρ_i times risk (the conditional variance), plus the expected price impact of a trade (market power). Imagine that the product of risk price and risk is large. Then, adding some market power to this large number does not change it by much. When we divide by a large number or a slightly larger number, the answer is almost the same. Thus, when data is scarce and variance is high, market power has little effect on production.

But when data is abundant, the conditional variance is low. Lots of data makes the firm less uncertain. If the first term is small, then adding market power to it makes a big difference. Dividing by a number close to zero or a number slightly less close to zero makes a big difference. Thus, when data is abundant and risk is low, market power has an outsized effect on production choices and thus on prices and markups.

DATA ASYMMETRY So far, we have explored what happens when all firms have more data. But a key concern for market competition is the possibility that firms have highly unequal stocks of data. Next, we use our data competition framework to ask what output, prices, and markups look like when data inequality grows. Define more data asymmetry to mean adding data precision to the high-data firm in a two-firm market.

Proposition 3 (Welfare and asymmetric data)**.** *In the duopoly case, when χ_c is sufficiently large, there exists a cutoff value c^* such that,*

1. *if $\bar{c}$ (or $\underline{c}$) is greater than c^*, the social welfare is increasing in data asymmetry;*
2. *if $\bar{c}$ (or $\underline{c}$) is smaller than c^*, the social welfare is declining in data asymmetry.*

To visually illustrate this result, we consider an example with two firms. We fix the total number of data points and add data to one firm as we subtract it from second firm. Figure 4 highlights how the economy is affected by data dispersion.

What we learn is that increasing data asymmetry has two opposite welfare effects: (1) increasing market power and hence deadweight loss, and (2) lower disutility from risks because the firm with more information will produce more. When the marginal cost c_i is relatively small, a difference in data precision creates a large difference in investment and thus firm size. This force can easily enable one firm dominate the market. Therefore, a larger deadweight loss makes the welfare more likely to decline. When the marginal cost c_i is higher, the efficiency benefits prevail.

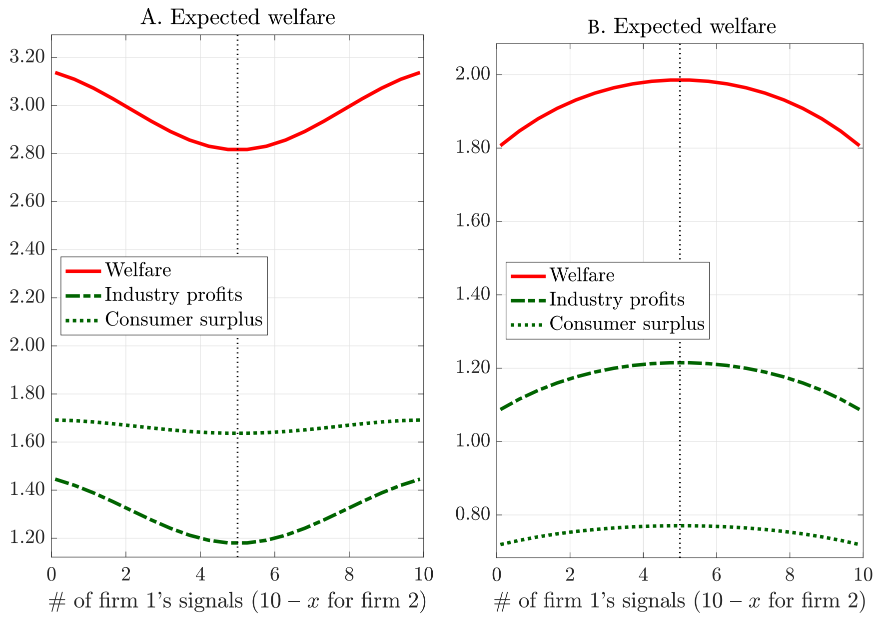


Figure 4: Data asymmetry and welfare with dominant risk channel (left) or investment channel (right).

Notes: This comparative static exercise is constructed over a single-good duopoly example. The investment cost function is assumed as $g(\chi_c, c_i) = \chi_c\,(\bar{c} - c_i)^2\,/2$. On the left, $\chi_c = 10$. On the right, $\chi_c = 1$. Other parameters are common to both plots: $\bar{c} = 3$, $\bar{p} = 5$, $\phi = 1$, $\sigma_b = 1$, $\mu_b = 0$, $\sigma_e = 2$, and $\rho_1 = \rho_2 = 1$.

While these results do not offer a simple answer or prediction about whether data is good or bad, they do provide an important input into the policy debate on data regulation. Data asymmetry and market dominance should not be seen as synonymous with welfare loss. Uncertainty is also a powerful drag on economic efficiency and on welfare. Sound data policy needs to trade off traditional market power harms against the benefits of resolving risk. This model produces a simple tool to evaluate that trade-off.

V Measuring Markups and Measuring Data

The previous analysis examined the forces that operate on product-level markups. But 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. In fact, the difference between a firm's product- and firm-level markups turns out to be a good proxy for the amount or quality of that firm's data.

These composition effects are not mere curiosities. They are also a feature of markup data.

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 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.

V.A Firm Markups

We begin by defining firm markups, exploring their relationship to product markups. In the following subsections, we will build up to the industry-level measures used by empirical researchers.

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

$$M_i^f := \frac{\mathbf{E}[\boldsymbol{q}_i'\boldsymbol{p}_i]}{\mathbf{E}[\boldsymbol{q}_i'\boldsymbol{c}_i]}. \tag{17}$$

To understand the relationship between firm markups and product markups, we can rewrite the expectation of the product of price and quantity as the product of expectations, plus a covariance term:

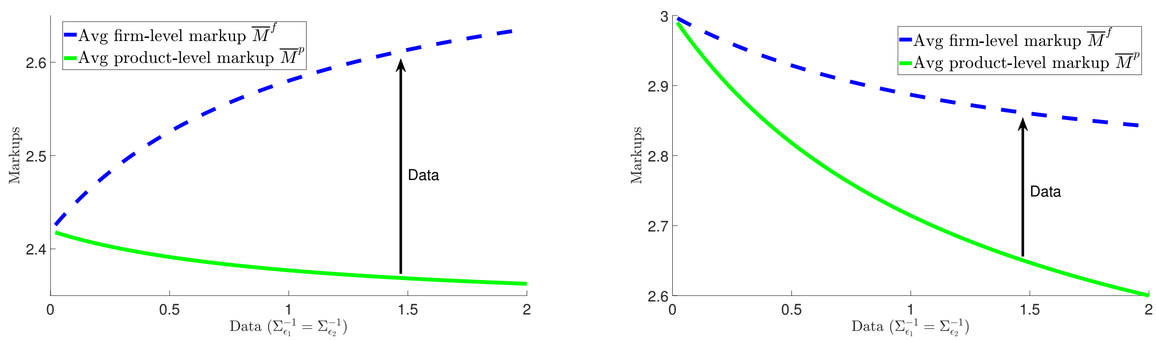
$$M_i^f = \frac{\mathbf{E}[\boldsymbol{q}_i]'\mathbf{E}[\boldsymbol{p}_i] + \mathbf{tr}\left[\mathbf{Cov}(\boldsymbol{p}_i, \boldsymbol{q}_i)\right]}{\mathbf{E}[\boldsymbol{q}_i'\boldsymbol{c}_i]} = \underbrace{\frac{\sum_{j=1}^N M_{ij}^p \boldsymbol{c}_i(j)\mathbf{E}[\boldsymbol{q}_i(j)]}{\sum_{j=1}^N \boldsymbol{c}_i(j)\mathbf{E}[\boldsymbol{q}_i(j)]}}_{\text{Cost-weighted product markups}} + \frac{\mathbf{tr}[\mathbf{Cov}\left(\boldsymbol{p}_i, \boldsymbol{q}_i\right)]}{\sum_{j=1}^N \boldsymbol{c}_i(j)\mathbf{E}[\boldsymbol{q}_i(j)]}. \tag{18}$$

The second equality just comes from using the definition of the product markup to substitute: $\mathbf{E}[\boldsymbol{p}_i] = M_i^{p\prime}\boldsymbol{c}_i$ 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 variance of prices and quantities. Firm data acts on this last term. It allows firms to produce more of goods that turn out to have high demand and thus high price.

Proposition 4. ***Data accumulation widens the wedge between product and firm markups.*** *Holding all firms' investments fixed ($(c_1, ..., c_{nF})$ given), an increase in firm i's data about any attribute increases $E[M_i^f - \overline{M}_i^p]$.*

Firm markups rise when data increases the covariance of firm's production decision q_i with the price p in (18). Without any data to predict demand, this covariance is low: without data, 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 (8), where a reduction in the conditional variance of demand makes production decisions q_i more sensitive to expected changes in price p_i. That higher sensitivity is a higher covariance.

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.



(a) Low price of risk.

(b) High price of risk.

Figure 5: More data may raise or lower markups but always causes product and firm markups to diverge. Parameters used: $\bar{p} = 5$, $\phi = 0.1$, and $A = I$. 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 5 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.

Regardless of whether product markups rise or fall as data becomes more abundant, firm-level markups rise relative to those product markups. Data allows firms to forecast which products will have high markups and to produce more of those. In fact, as we explore later, this difference between product and firm markups can be used to measure a firm's stock of data.

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 governs the difference in markups. Data changes the composition of products and firms and makes various measures of markups diverge. This is a theme that will recur as we proceed to explore markups at the industry level.

V.B Measures of Markups in an Industry

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 the markups for each firm into a single industry measure. We construct four of the most common measures here to understand how they differ. Then, we compare their theoretical predictions to empirical evidence. The model lends an interpretation to the different trends arising from the different ways empirical researchers measure industry markups.

Definition 3. *The unweighted average firm markup in an industry is*

$$\bar{M}^f := (1/N) \sum_{i=1}^{N} M_i^f. \tag{19}$$

Definition 4. *The cost-weighted markup for an industry is*

$$M^c := \sum_{i=1}^{N} w_i^c M_i^f \quad \textit{where cost weights are} \quad w_i^c = \frac{\mathbf{E}\left[\boldsymbol{q}_i' \boldsymbol{c}_i\right]}{\sum_{i=1}^{N} \mathbf{E}\left[\boldsymbol{q}_i' \boldsymbol{c}_i\right]}. \tag{20}$$

While the first definition is simply the markup of the average firm, this second definition weights larger firms more. With cost-weighted markups, larger firms are those that have larger variable costs compared with their industry competitors. The next definition also weights the markups of larger firms more. But in the sales-weighted markups, larger firms are those with larger gross revenues compared to the revenues of other firms in the same industry.

Definition 5. *The sales-weighted markup is*

$$M^s := \sum_{i=1}^{N} w_i^s M_i^f \quad \textit{where sales weights are} \quad w_i^s = \frac{\mathbf{E}\left[\boldsymbol{q}_i' \boldsymbol{p}_i\right]}{\sum_{i=1}^{N} \mathbf{E}\left[\boldsymbol{q}_i' \boldsymbol{p}_i\right]}. \tag{21}$$

Definition 6. *The industry- aggregates markup is*

$$M^{ind} := \frac{\mathbf{E}\left[\sum_{i=1}^{N} \boldsymbol{q}_i' \boldsymbol{p}_i\right]}{\mathbf{E}\left[\sum_{i=1}^{N} \boldsymbol{q}_i' \boldsymbol{c}_i\right]}. \tag{22}$$

Industry-aggregates markup is measured with data already aggregated at the industry level. It is the ratio of the total industry sales over the total industry variable cost. In theory, industry-aggregates markups are identical to cost-weighted markups:

$$M^c := \sum_{i=1}^{N} \frac{\mathbf{E}\left[\boldsymbol{q}_i' \boldsymbol{c}_i\right]}{\sum_{i=1}^{N} \mathbf{E}\left[\boldsymbol{q}_i' \boldsymbol{c}_i\right]} M_I^f = \sum_{i=1}^{N} \frac{\mathbf{E}\left[\boldsymbol{q}_i' \boldsymbol{c}_i\right]}{\sum_{i=1}^{N} \mathbf{E}\left[\boldsymbol{q}_i' \boldsymbol{c}_i\right]} \frac{\mathbf{E}\left[\boldsymbol{q}_i' \boldsymbol{p}_i\right]}{\mathbf{E}\left[\boldsymbol{q}_i' \boldsymbol{c}_i\right]} = \frac{\mathbf{E}\left[\sum_{i=1}^{N} \boldsymbol{q}_i' \boldsymbol{p}_i\right]}{\mathbf{E}\left[\sum_{i=1}^{N} \boldsymbol{q}_i' \boldsymbol{c}_i\right]} := M^{ind}. \tag{23}$$

However, in practice, with different sources of measurement error at the firm and aggregate level, each approach may deliver slightly different answers.

V.C Data and Industry Markup Measures

Our theory of data provides an explanation for the widening gap between these various markup measures. Firms that have more data can reduce uncertainty. Lower uncertainty makes larger up-front investment optimal. So, high-data firms are large firms, which are weighted more by cost weights and sales weights, relative to the unweighted firm average. As explained in the firm markup section, firms use data to skew their production toward high-markup goods, making high-data firms likely to be higher-markup firms. Thus, the measures that weight large, high-data firms more will also weight high-markup firms more, generating a higher predicted industry markup.

Proposition 5. ***Growing data increases the wedges between industry markup measures.*** *Holding all firms' investments fixed ($(c_1, ..., c_{nF})$ given) and c_i sufficiently small, an increase in firm i's data about any attribute*

a. increases the difference between cost-weighted and unweighted firm markups $E[M^c - \overline{M}^f]$.

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

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

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

Mathematically, the key to each of these results is a covariance. In the first case, the covariance is between the firm markup and the total production of a firm. Cost-weighted markups are firm markups, weighted by the firm's share of variable cost of production. If q_i is large for firms that have high markups, then the weighted average will have a higher markup than the unweighted average. This is related to data because, as discussed in the previous result, high-data firms skew their production to high-markup goods and thus have higher firm markups. High-data firms also produce more on average because data lowers their production risk. We can see this in the production first-order condition (8) where a reduction in the conditional variance reduces the denominator and makes production decisions q_i larger on average.

Economically, this is another composition or aggregation effect. 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 case (b), the key covariance is between a firm's markup and the firm's revenue. Sales-weighted industry markups are firm markups weighted by the gross revenue of the firm. Cost-weighted industry markups are firm markups weighted by the variable cost of the firm. 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.

The twist here is that "high-data firms" now means firms that have higher amounts of data than their competitors. Firms can obtain higher price relative to their cost because of information asymmetry. If all firms knew that demand would be high for an attribute and they all produced more of it, this would bring the price of that attribute back down. What we learn from this is that a divergence between sales-weighted and cost-weighted markups results from growth in cross-firm information asymmetry.

In part (c), firms' data stocks speak to the observed divergence in measures of markups using disaggregated firm data and to the measures that use total industry revenue and total industry cost. The third part of the proposition reveals that sales-weighted markups should also rise faster than markups measured on industry aggregates, if firms are accumulating more data over time. This third result follows from the second because of the theoretical equivalence between measurement using aggregates and cost-weighting the disaggregated firm markups, as shown in (23).

AN ILLUSTRATIVE EXAMPLE OF INDUSTRY MARKUP DIVERGENCE When firms choose their investment to lower their marginal cost of production, high-data firms choose to invest more. High-data firms, which we saw have higher firm-level markups, grow larger. As a result, their production accounts for a larger fraction of total production. Therefore, the higher markup of the high-data firms gets weighted more in the industry markup. Thus, investment choice amplifies the wedge between firm-level and industry markups.

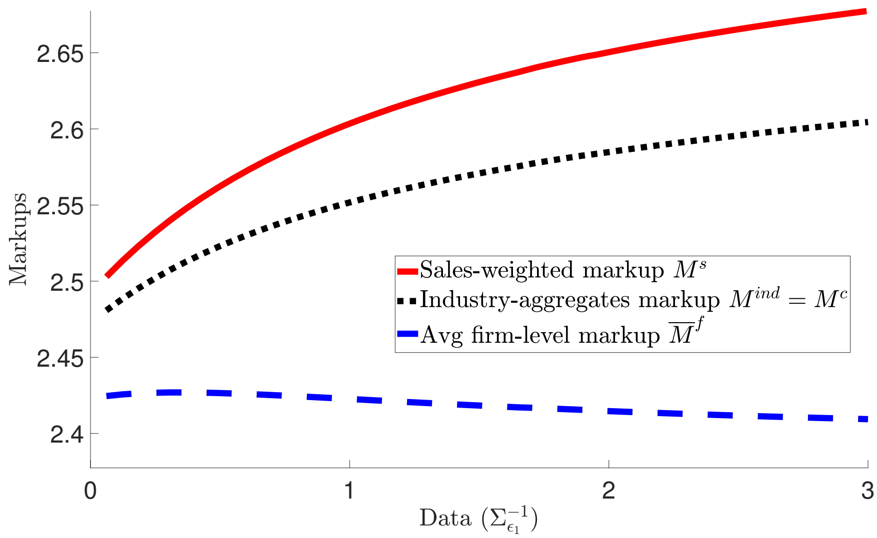


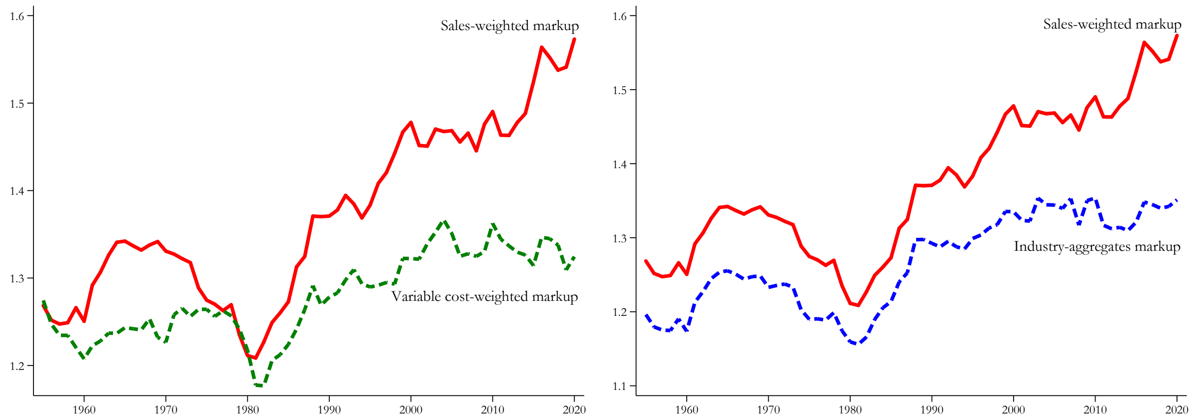
Figure 6: Data Accumulation Makes Industry Markup Measures Diverge. Investment cost function is $g(\chi_c, c_i) = \chi_c / c_i^2$, with $\chi_c = 1$. Parameters are $\bar{p} = 5$, $\rho_1 = 1$, $\rho_2 = 5$, $\phi = 0.8$, and $A = I$. Firm 1's data is measured on the x-axis. Firm 2's data is fixed at $\Sigma_{\epsilon_2}^{-1} = 1$.

In Figure 6, we see the gap between firm-level markup (blue dashed line at the bottom) and the industry markup (red solid line on top) widen relative to the previous results where that gap was much smaller. That market aggregation gap also grows as data becomes more abundant. That result suggests that as firms process more and more data, 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. Data increases markups because of investment, cross-product aggregation, and cross-firm aggregation. Data decreases markups because it induces firms to produce more (risk premium channel).

V.D 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 7a, 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 important.



(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 7: Markups Measured and Aggregated in Different Ways Diverged Over Time. Left panel is from De Loecker, Eeckhout, and Unger (2020), Figure XVI.A. Right panel is from De Loecker, Eeckhout, and Unger (2020), Figure V.

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 7b 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 7b) look remarkably similar to cost-weighted firm markups (in Figure 7a). This is due to the systematic relation between input-weighting and industry aggregates in equation (23).

VI 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 understand what is a boom or recession in the context of this model. Two relevant changes typically happen when an economy transitions from recession to boom. The first is that demand rises. The second is that the variances of demand and of output fall. Recessions are volatile, uncertain times. To formalize this new assumption, we introduce a variable *Boom* that is high in booms and low in recessions. Then, we make the level of demand procyclical and the demand variance countercyclical by assuming

$$\bar{p} = d_0 + d_1 * Boom \qquad \text{where } \; d_0, d_1 \geq 0, \tag{24}$$

$$\Sigma_b = d_2 - d_3 * Boom \qquad \text{where } \; d_2, d_3 \geq 0. \tag{25}$$

The change in the level of demand does not affect divergence, but it allows both lines to rotate. In other words, high demand in a boom regulates how countercyclical or acyclical product markups are. Falling variance in a boom is what makes the cyclical behavior of aggregate markups differ relative to product markups. The second statement is formalized in the next proposition.

Proposition 6. ***Product markups diverge from firm and industry markups when volatility rises.*** *Suppose the investment cost structure is such that firms choose identical investments ($c_i = c_j \;\; \forall\, i, j$).*

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, $\lim_{\Sigma_{b,k}\to\infty} \partial \mathbf{E}[M^f_{ij}]/\partial\Sigma_{b,j}, \partial \mathbf{E}[M^m_{ij}]/\partial\Sigma_{b,j} > 0$.

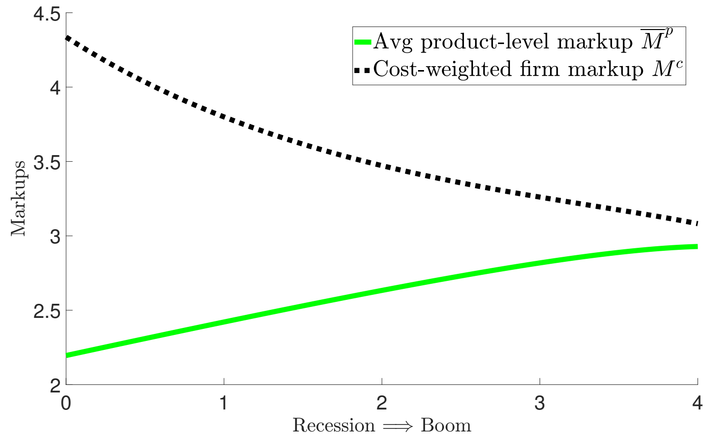


Figure 8: Procyclical product markups can coexist with countercyclical firm / industry markups. Left (right) on the x-axis represents recessions (booms), as described in (24) and (25), where $d_0 = 7/2$, $d_1 = 1/2$, $d_2 = 5$, and $d_3 = 1$. A decreasing line represents a countercyclical markup. Remaining parameters are $\rho_1 = \rho_2 = 1$, $c_1 = c_2 = 1$, $\phi = 1$, 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 8. The reason these two objects behave so differently is the covariance of demand and output. When the variance of demand rises, the covariance rises mechanically as well. The covariance of demand and 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, the volatility of recessions strengthens the composition effects that drive firm markups up but not product markups. This explains why Nekarda and Ramey (2020) found no change in markups but Bils (1985, 1987) did. Both may be right at the same time. Our model can then help researchers to think through which measure matters most for the economic question posed.

These results are for a high marginal cost of investment, which essentially holds firm size fixed. That may be a good assumption for a cyclical fluctuation. However, in the long run, investment may adjust. Appendix B.1. shows that when firms adjust investment flexibly in response to a change in demand and volatility, the effect is dampened.

VII Where Does Data Come From? A Dynamic Model

So far, the paper has taken firms' data endowments as given, because the question at hand was what effect data has on markups. However, the broader context for this question is that economists are worried about market power, which might arise because of data production. To better understand the interaction of markups and data production, we set up a model that incorporates both. Considering data production and firm growth together introduces one new effect on markups: a role for data as a means of payment. A firm's marginal benefit of a transaction is the profit and the data it generates. Revenue and data enter in a substitutable way. The 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 new role, the original insights derived from the static model survive.

Dynamic model setup Consider an economy where each firm i chooses an $n \times 1$ vector a_{it} that describes their location in the product space and a quantity q_{it} to produce. As before, firms maximize expected profit, with a price of risk adjustment as in equation (1). There are n attributes and demand shocks for each of those attributes.

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):

$$b_t = \rho b_{t-1} + \eta_{bt}, \qquad \eta_{bt} \sim iid\ N(0, \sigma_\eta I). \tag{26}$$

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 attribute is the persistent process (26), plus some transitory noise:

$$\tilde{b}_t = b_t + \epsilon_{bt} \qquad \epsilon_{bt} \sim iid\ N(0, \sigma_\epsilon I) \tag{27}$$

The fact that these demand shocks are common to all firms makes data from one firm relevant for another firm. This opens up the possibility of buying and selling data.

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 that depends on the amount produced q. Since data is about the demand for each attribute, the amount of data observed d_{it} is also proportional to the good's loading on the attribute:

$$d_{it} = q_{i,t-1} a_{i,t-1}. \tag{28}$$

This captures the idea that firms learn more about attributes they produce. If they produce cell phones, they learn about the demand for (or cost of) electronics, not about demand for food. If they want to learn about the demand for food, they need to produce something edible, or buy such data. Notice that this makes production a form of active experimentation in the product space. 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 to inform their decisions depends on data production as data purchases or sales: $D_{it} = d_{it} + m_{it}\mathcal{P}_t$ where m_{it} is the amount of data purchased by firm i at date t and $\mathcal{P}_t$ is the time-t market price per unit of data. Firms also choose an amount of data to sell 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 b_t, with precision Σ_e per signal. Firms update demand forecasts using Bayes law. Thus, when a firm obtains D_{it} units of data about each attribute, Bayes law tells the firm to average the signals to arrive at a composite signal that has precision $D_{i,t}'\Sigma_e^{-1}D_{i,t}$. Notice that Bayes' law allows us to incorporate non-integer numbers of signals. So we can proceed considering d_{it} and D_{it} to be 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 choices $\{q_{i,t}, a_{i,t}\}$ and data purchases / sales $\{m_{i,t}, l_{i,t}\}$ solve the following recursive problem:

$$V(\Omega_t) = \max_{q_{i,t},a_{i,t},m_{i,t},l_{i,t}} (P_t - c)q_{i,t}a_{i,t} + \mathcal{P}_t(l_{i,t} - m_{i,t}) + \left(\frac{1}{1+r}\right) V(\Omega_{t+1}), \tag{29}$$

where the law of motion for each firm's $\omega_{i,t}$ is given by

$$\omega_{i,t+1} = \left[\rho^2\omega_{i,t}^{-1} + \sigma_\epsilon\right]^{-1} + (n_{i,t} + m_{i,t})\sigma_\epsilon^{-2} \tag{30}$$

and the number of data points produced by the firm is $n_{i,t} = q_{i,t}a_{i,t}$.

The first-order condition for the quantity of production looks similar to (8) in the static model. Optimal production depends on risk and price impact, in the denominator, and expected profit $(p - c)$, in the numerator.

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

However, there is one new term in dynamic model: $\partial V/\partial q_i$ 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.

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 $p + \partial V / \partial q_i$. 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.

VIII Mapping Theory to Data

One reason it is important to have models that describe the relationships between quantities like data and markups is that models inform measurement. In this case, the model teaches us how to measure the amount of data a firm has and how to determine what risks that data is about. While executing the measurement is a separate paper, this section is meant to aid others who might choose to use the model as a structure for empirical analysis.

The next result shows that we can measure the amount of data a firm has by looking at the gap between average product markups and firm markups. This is analogous to looking at the alpha of a fund manager to infer how much they know.

Corollary 1. ***Markup wedges are measures of data.*** *The production-aggregation wedge $E[M_i^f - \overline{M}_i^p]$ is a monotonic function of firm i's data.*

This result is a straightforward conclusion from Proposition 4. But it is key to measurement. For many measurement exercises, an econometrician may need to know how much data a firm or a collection of firms has. This suggests a measurement approach is to look at the markups at various degrees of aggregation and use the aggregation wedge to infer a corresponding level of data.

What Is Data About? Measuring Characteristic Loadings Measuring attributes is novel in finance, but more standard in IO. One way to gauge attribute loadings is by looking at demand variance-covariance across goods and extracting principal components. The eigenvectors are loadings. There are also other orthogonal decompositions one can use. But the eigen decomposition has a nice interpretation in terms of principal components.

Another way of measuring characteristic loadings is to use the Hoberg-Phillips measure of cosine similarity from textual analysis of firms' earnings reports. This measure determines how similarly different firms describe their products to their investors.

One might think of a characteristic of a good as being its location. 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.[6]

Our predictions are consistent with the superstar firm economy of Autor et al. (2020) and the increasing span of control in Aghion et al. (2019) and Lashkari, Bauer, and Boussard (2018). The rise in firm concentration, the rise in average markups that comes from high-markup firms growing larger, and the correlation between productivity and concentration are all features of U.S. and international markets and are features of our model. Similarly, Crouzet and Eberly (2018) argue that large modern firms have high levels of intangible investment, which is correlated with having high markups. What our work adds is a mechanism—an explanation for why the accumulation of customer data can explain these trends.

Measuring the Price of Risk Measuring risk price is novel in IO, but standard in finance. A key parameter that governs the sign of many of the predictions is ρ, the price of risk. Finance has developed a whole battery of tools to determine this risk price in various ways. A common approach is to use the market prices of equities to estimate the compensation investors demand for risk in that domain and then carry the same price over to determine the price of risk that a firm faces. The argument for doing that is that the manager should be maximizing equity holders' interests. The firm's equity holders are the same agents who hold other market equities, with the

[6] 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) argues that the discrepancy stems from a mechanical relation between population size and the market definition.

same risk preferences.[7]

Distinguishing Data from Competition   Where data and market competition differ is in $cov(p, q_i)$. 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 highly correlated.

Data also enables more accurate forecasting, while market competition does not. Another approach to measuring and identifying firms' data would be to assess the accuracy of firms' forecasts.

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.[8]

Our predictions are consistent with the superstar firm economy of Autor et al. (2020) and the increasing span of control in Aghion et al. (2019) and Lashkari, Bauer, and Boussard (2018). The rise in firm concentration, the rise in average markups that comes from high-markup firms growing larger, and the correlation between productivity and concentration are all features of U.S. and international markets and are features of our model. Similarly, Crouzet and Eberly (2018) argue that large modern firms have high levels of intangible investment, which is correlated with having high markups. If a firm is a data-abundant firm, they should have high levels of intangible assets. However, since there are other intangible assets, the reverse might not be true.

IX Conclusion

We set out to explore the hypothesis that data encourages large firms to grow larger and gain market power. We constructed a new framework where firms use data to reduce uncertainty about future demand for various products. Just like managers are taught to do in MBA programs, firms' decision makers in our model make investment decisions, taking risk into account. It is this effective risk aversion that causes firms to invest more when they have more data. Data is a tool to reduce risk. With less risk from random demand, a larger investment becomes optimal. Thus high-data firms do invest more, grow larger, and exert more impact on prices.

[7] See Brealey, Myers, and Allen (2003) for a more complete explanation of the rationale and execution.

[8] 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) argues that the discrepancy stems from a mechanical relation between population size and the market definition.

But this simple story delivered some unexpected additional effects. We found that when managers price risk, markups reflect both market power and a compensation for risk. If data reduces risk by making uncertain outcomes more predictable, then it also reduces the risk premium and the markup. At the same time, 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.

Online Appendix

A. Appendix: Solution Details

ATTRIBUTE SPACE 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.

INFORMATION Each firm indexed by i has n_{di} data points, each of which is a signal of the attribute demand shock $\boldsymbol{s}_{i,j} = \boldsymbol{b_i} + \boldsymbol{\varepsilon}_{i,j}$ where $j = 1, \ldots, n_{di}$. We assume signal noises are uncorrelated and normally distributed with zero mean and precision $\Sigma_{\epsilon_{i,j}}^{-1}$. The posterior variance conditional on n_{di} signals is

$$\mathbf{Var}\left(\boldsymbol{b_i} | \{\boldsymbol{s}_{i,j}\}_{j=1}^{n_{di}}\right) = \left(\Sigma_{b_i}^{-1} + \sum_{j=1}^{n_{di}} \Sigma_{\epsilon_{i,j}}^{-1}\right)^{-1}$$

This is equivalent to a compound signal $\boldsymbol{s}_i$ with total data precision $\Sigma_{\epsilon_i}^{-1} = \sum_{j=1}^{n_{di}} \Sigma_{\epsilon_{i,j}}^{-1}$. According to Bayes's law, we have

$$\begin{aligned} \mathbf{E}\left[\tilde{\boldsymbol{p}}_i | \mathcal{I}_i\right] &= \bar{\boldsymbol{p}} + \mathbf{E}\left[\boldsymbol{b}_i | \mathcal{I}_i\right] - \frac{1}{\phi} \sum_{j=1}^{n_F} \tilde{\boldsymbol{q}}_j = \bar{\boldsymbol{p}} + \boldsymbol{K}_i \boldsymbol{s}_i - \frac{1}{\phi} \sum_{j=1}^{n_F} \tilde{\boldsymbol{q}}_j \\ \mathbf{Var}\left[\tilde{\boldsymbol{p}}_i | \mathcal{I}_i\right] &= \mathbf{Var}\left[\boldsymbol{b}_i | \mathcal{I}_i\right] = \left(\Sigma_{b_i}^{-1} + \Sigma_{\epsilon_i}^{-1}\right)^{-1} \end{aligned} \tag{32}$$

MAXIMIZING UTILITY Take first-order condition of firm's utility function, we get an expression for optimal attribute choices.

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

Differentiating the inverse demand curve $\tilde{\boldsymbol{p}}_i = \bar{\boldsymbol{p}} + \boldsymbol{b}_i - \frac{1}{\phi} \sum_{j=1}^{n_F} \tilde{\boldsymbol{q}}_j$ reveals that market power is constant:

$$\frac{\partial \mathbf{E}\left[\tilde{\boldsymbol{p}}_i | \mathcal{I}_i\right]}{\partial \tilde{\boldsymbol{q}}_i} = \frac{\partial \mathbf{E}\left[\boldsymbol{p}_i | \mathcal{I}_i\right]}{\partial \boldsymbol{q}_i} = -\frac{1}{\phi} \boldsymbol{I}_N \tag{33}$$

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 the attribute choice $\tilde{q}_i$ on both the left and the right sides of the equality. It arises on the right side because firm i's production choice $\tilde{q}_i$ affects the expected price $\mathbf{E}\left[\tilde{\boldsymbol{p}}_i | \mathcal{I}_i\right]$. Therefore, we substitute in the price and re-arrange to collect all $\tilde{q}_i$ terms and reveal the optimal production choice:

$$\begin{aligned}
\tilde{\boldsymbol{q}}_i &= \left(\rho_i \mathbf{Var}\left[\tilde{\boldsymbol{p}}_i|\mathcal{I}_i\right] - \frac{\partial \mathbf{E}\left[\tilde{\boldsymbol{p}}_i|\mathcal{I}_i\right]}{\partial \tilde{\boldsymbol{q}}_i}\right)^{-1} \left(\mathbf{E}\left[\tilde{\boldsymbol{p}}_i|\mathcal{I}_i\right] - \tilde{\boldsymbol{c}}_i\right) \\
\Rightarrow \left(\rho_i \mathbf{Var}\left[\boldsymbol{b}_i|\mathcal{I}_i\right] + \frac{1}{\phi}\boldsymbol{I}_N\right)\tilde{\boldsymbol{q}}_i &= \bar{\boldsymbol{p}} + \mathbf{E}\left[\boldsymbol{b}_i|\mathcal{I}_i\right] - \frac{1}{\phi}\sum_{j=1}^{n_F}\tilde{\boldsymbol{q}}_j - \tilde{\boldsymbol{c}}_i \\
\Rightarrow \left(\rho_i \mathbf{Var}\left[\boldsymbol{b}_i|\mathcal{I}_i\right] + \frac{2}{\phi}\boldsymbol{I}_N\right)\tilde{\boldsymbol{q}}_i &= \bar{\boldsymbol{p}} + \mathbf{E}\left[\boldsymbol{b}_i|\mathcal{I}_i\right] - \frac{1}{\phi}\sum_{j=1,j\neq i}^{n_F}\tilde{\boldsymbol{q}}_j - \tilde{\boldsymbol{c}}_i \\
\Rightarrow \tilde{\boldsymbol{q}}_i &= \left(\rho_i \mathbf{Var}\left[\boldsymbol{b}_i|\mathcal{I}_i\right] + \frac{2}{\phi}\boldsymbol{I}_N\right)^{-1}\left(\bar{\boldsymbol{p}} + \mathbf{E}\left[\boldsymbol{b}_i|\mathcal{I}_i\right] - \frac{1}{\phi}\sum_{j=1,j\neq i}^{n_F}\tilde{\boldsymbol{q}}_j - \tilde{\boldsymbol{c}}_i\right)
\end{aligned} \tag{34}$$

Define $\boldsymbol{H}_i = \left(\rho_i \mathbf{Var}\left[\boldsymbol{b}_i|\mathcal{I}_i\right] + \frac{2}{\phi}\boldsymbol{I}_N\right)^{-1}$. Using Bayes law to replace the expectation $\mathbf{E}\left[\boldsymbol{b}_i|\mathcal{I}_i\right]$ with the weighted sum of signals $\boldsymbol{K}_i\boldsymbol{s}_i$, with $\boldsymbol{K}_i = \Sigma_{b_i}(\Sigma_{b_i} + \Sigma_{\epsilon_i})^{-1}$ yields

$$\tilde{\boldsymbol{q}}_i = \boldsymbol{H}_i\left(\bar{\boldsymbol{p}} + \boldsymbol{K}_i\boldsymbol{s}_i - \frac{1}{\phi}\sum_{j=1,j\neq i}^{n_F}\tilde{\boldsymbol{q}}_j - \tilde{\boldsymbol{c}}_i\right). \tag{35}$$

SUB-GAME EQUILIBRIUM The solution above generates the best-response function given the realization of signals. We solve the sub-game Nash equilibrium by separating firm-specific terms on the left from aggregate objects on the right.

$$\begin{aligned}
&\Rightarrow \left(\boldsymbol{H}_i^{-1} - \frac{\boldsymbol{I}_N}{\phi}\right)\tilde{\boldsymbol{q}}_i = \bar{\boldsymbol{p}} + \boldsymbol{K}_i\boldsymbol{s}_i - \frac{1}{\phi}\sum_{j=1}^{n_F}\tilde{\boldsymbol{q}}_j - \tilde{\boldsymbol{c}}_i \\
&\Rightarrow \left(\boldsymbol{H}_i^{-1} - \frac{\boldsymbol{I}_N}{\phi}\right)\tilde{\boldsymbol{q}}_i - \boldsymbol{K}_i\boldsymbol{s}_i + \tilde{\boldsymbol{c}}_i \equiv \bar{\boldsymbol{p}} - \frac{1}{\phi}\sum_{j=1}^{n_F}\tilde{\boldsymbol{q}}_j \triangleq \Pi,\ \forall\, i = 1, \ldots, n_F
\end{aligned} \tag{36}$$

The right side is constant for each firm i and we denote it as Π. To solve for the equilibrium price, we re-express the optimal attribute choice in terms of the aggregate object Π and then impose consistency between firms' choices an the aggregate Π. In other words, we solve for the fixed point.

$$\begin{aligned}
\tilde{\boldsymbol{q}}_i &= \left(\boldsymbol{H}_i^{-1} - \frac{\boldsymbol{I}_N}{\phi}\right)^{-1}\left(\Pi + \boldsymbol{K}_i\boldsymbol{s}_i - \tilde{\boldsymbol{c}}_i\right) \\
\Rightarrow \Pi &= \bar{\boldsymbol{p}} - \frac{1}{\phi}\sum_{j=1}^{n_F}\tilde{\boldsymbol{q}}_j = \bar{\boldsymbol{p}} - \frac{1}{\phi}\sum_{j=1}^{n_F}\left(\boldsymbol{H}_j^{-1} - \frac{\boldsymbol{I}_N}{\phi}\right)^{-1}\left(\Pi + \boldsymbol{K}_j\boldsymbol{s}_j - \tilde{\boldsymbol{c}}_j\right) \\
\Rightarrow \left(\boldsymbol{I}_N + \frac{1}{\phi}\sum_{j=1}^{n_F}\left(\boldsymbol{H}_j^{-1} - \frac{\boldsymbol{I}_N}{\phi}\right)^{-1}\right)\Pi &= \bar{\boldsymbol{p}} - \frac{1}{\phi}\sum_{j=1}^{n_F}\left(\boldsymbol{H}_j^{-1} - \frac{\boldsymbol{I}_N}{\phi}\right)^{-1}\left(\boldsymbol{K}_j\boldsymbol{s}_j - \tilde{\boldsymbol{c}}_j\right) \\
\Rightarrow \Pi &= \left(\boldsymbol{I}_N + \frac{1}{\phi}\sum_{j=1}^{n_F}\left(\boldsymbol{H}_j^{-1} - \frac{\boldsymbol{I}_N}{\phi}\right)^{-1}\right)^{-1}\left[\bar{\boldsymbol{p}} - \frac{1}{\phi}\sum_{j=1}^{n_F}\left(\boldsymbol{H}_j^{-1} - \frac{\boldsymbol{I}_N}{\phi}\right)^{-1}\left(\boldsymbol{K}_j\boldsymbol{s}_j - \tilde{\boldsymbol{c}}_j\right)\right]
\end{aligned} \tag{37}$$

We define $\hat{\boldsymbol{H}}_i$ and $\boldsymbol{D}$ as the adjusted supply elasticity and average Π respectively.

$$\begin{aligned}\hat{\boldsymbol{H}}_i &:= \left(\boldsymbol{H}_i^{-1} - \frac{\boldsymbol{I}_N}{\phi}\right)^{-1} = \left(\frac{\boldsymbol{I}_N}{\phi} + \rho_i \mathbf{Var}\left[\boldsymbol{p}_i|\mathcal{I}_i\right]\right)^{-1} \\ \boldsymbol{D} &:= \left(\boldsymbol{I}_N + \frac{1}{\phi}\sum_{i=1}^{n_F} \hat{\boldsymbol{H}}_i\right)^{-1}\left(\bar{\boldsymbol{p}} + \frac{1}{\phi}\sum_{i=1}^{n_F}\hat{\boldsymbol{H}}_i\tilde{\boldsymbol{c}}_i\right)\end{aligned} \tag{38}$$

Finally, the equilibrium output and price are

$$\begin{aligned}\tilde{\boldsymbol{q}}_i &= \hat{\boldsymbol{H}}_i(\boldsymbol{D} - \tilde{\boldsymbol{c}}_i) + \hat{\boldsymbol{H}}_i\boldsymbol{K}_i\boldsymbol{s}_i - \frac{\hat{\boldsymbol{H}}_i}{\phi}\left(\boldsymbol{I}_N + \frac{1}{\phi}\sum_{i=1}^{n_F}\hat{\boldsymbol{H}}_i\right)^{-1}\sum_{j=1}^{n_F}\hat{\boldsymbol{H}}_j\boldsymbol{K}_j\boldsymbol{s}_j \\ \tilde{\boldsymbol{p}}_i &= \bar{\boldsymbol{p}} + \boldsymbol{b}_i - \frac{1}{\phi}\sum_{j=1}^{n_F}\tilde{\boldsymbol{q}}_j = \Pi + \boldsymbol{b}_i = \boldsymbol{D} + \boldsymbol{b}_i - \frac{1}{\phi}\left(\boldsymbol{I}_N + \frac{1}{\phi}\sum_{i=j}^{n_F}\hat{\boldsymbol{H}}_j\right)^{-1}\sum_{j=1}^{n_F}\hat{\boldsymbol{H}}_j\boldsymbol{K}_j\boldsymbol{s}_j\end{aligned} \tag{39}$$

PRICE-QUANTITY COVARIANCE A key object in our markup calculations is the co-variance between price $\tilde{\boldsymbol{p}}_i$ and quantity $\tilde{\boldsymbol{q}}_i$:

$$\begin{aligned}\mathbf{Cov}\left(\tilde{\boldsymbol{p}}_i, \tilde{\boldsymbol{q}}_i\right) = &\left(\boldsymbol{I}_N + \sum_{j=1}^{n_F}\frac{\hat{\boldsymbol{H}}_j}{\phi}\right)^{-1}\sum_{j=1}^{n_F}\hat{\boldsymbol{H}}_j\boldsymbol{K}_j\hat{\boldsymbol{H}}_j\left(\boldsymbol{I}_N + \sum_{j=1}^{n_F}\frac{\hat{\boldsymbol{H}}_j}{\phi}\right)^{-1}\frac{\hat{\boldsymbol{H}}_i}{\phi^2} + \boldsymbol{K}_i\hat{\boldsymbol{H}}_i \\ &- \left(\boldsymbol{I}_N + \sum_{j=1}^{n_F}\frac{\hat{\boldsymbol{H}}_j}{\phi}\right)^{-1}\hat{\boldsymbol{H}}_i\boldsymbol{K}_i\frac{\hat{\boldsymbol{H}}_i}{\phi} - \boldsymbol{K}_i\hat{\boldsymbol{H}}_i\left(\boldsymbol{I}_N + \sum_{j=1}^{n_F}\frac{\hat{\boldsymbol{H}}_j}{\phi}\right)^{-1}\frac{\hat{\boldsymbol{H}}_i}{\phi}\end{aligned} \tag{40}$$

PRODUCT-LEVEL MARKUP (ATTRIBUTE) The product-level markup produced by firm i is $M_{i,j}^{\tilde{p}} := \mathbf{E}[\tilde{\boldsymbol{p}}_{i,j}]/\tilde{\boldsymbol{c}}_{i,j}$. The average product-level markup on the attributes is

$$\overline{M}^{\tilde{p}} = \frac{1}{N}\frac{1}{n_F}\sum_{i=1}^{n_F}\sum_{j=1}^{N} M_{i,j}^{\tilde{p}} = \frac{1}{n_F N}\sum_{i=1}^{n_F}\sum_{j=1}^{N}\frac{\mathbf{E}[\tilde{\boldsymbol{p}}_{i,j}]}{\tilde{\boldsymbol{c}}_{i,j}} = \frac{1}{n_F N}\sum_{i=1}^{n_F}\sum_{j=1}^{N}\frac{\boldsymbol{D}_j}{\tilde{\boldsymbol{c}}_{i,j}} \tag{41}$$

We denote the posterior variance $\Sigma_{b_i} = \mathbf{Var}\left[\boldsymbol{b}_i|\mathcal{I}_i\right] = \left(I_N + \Sigma_{\epsilon_i}^{-1}\right)^{-1}$, thus

$$\frac{\partial \boldsymbol{D}_j}{\partial \Sigma_{\epsilon_i,k}^{-1}} = \delta_{jk}\frac{1}{\phi}\frac{\rho_i\hat{\boldsymbol{H}}_{i,j}^2\Sigma_{b_i,j}^2}{1 + \frac{1}{\phi}\sum_{s=1}^{n_F}\hat{\boldsymbol{H}}_{s,j}}\left(\tilde{\boldsymbol{c}}_{i,j} - \boldsymbol{D}_j\right) < 0 \Rightarrow \frac{\partial\overline{M}^{\tilde{p}}}{\partial\Sigma_{\epsilon_i,k}^{-1}} < 0 \tag{42}$$

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

$$M_i^f = \frac{\mathbf{E}[\tilde{\boldsymbol{q}}_i'\tilde{\boldsymbol{p}}_i]}{\mathbf{E}[\tilde{\boldsymbol{q}}_i'\tilde{\boldsymbol{c}}_i]} = \frac{\mathbf{E}[\tilde{\boldsymbol{q}}_i]'\mathbf{E}[\tilde{\boldsymbol{c}}] + \mathbf{trCov}\left(\tilde{\boldsymbol{p}}_i, \tilde{\boldsymbol{q}}_i\right)}{\mathbf{E}[\tilde{\boldsymbol{q}}_i'\tilde{\boldsymbol{c}}_i]} \tag{43}$$

Thus, the average firm-level markup is $\overline{M}^f = (1/n_F)\sum_{i=1}^{n_F} M_i^f$. As for the denominator, the equilibrium output increases with more data since

$$\frac{\partial \mathbf{E}\tilde{\boldsymbol{q}}_{i,j}}{\partial \Sigma_{\epsilon_i,j}^{-1}} = \rho_i \hat{\boldsymbol{H}}_{i,j}^2 \Sigma_{b_i,j}^2 \left(\boldsymbol{D}_j - \tilde{\boldsymbol{c}}_{i,j}\right) \left[1 - \frac{\frac{1}{\phi}\hat{\boldsymbol{H}}_{i,j}}{1 + \frac{1}{\phi}\sum_{k=1}^{n_F}\hat{\boldsymbol{H}}_{k,j}}\right] > 0 \tag{44}$$

Although price decreases with more data, the revenue actually benefits from it.

$$\begin{aligned}
\frac{\partial \mathbf{E}\tilde{\boldsymbol{q}}_{i,j}\mathbf{E}\tilde{\boldsymbol{p}}_{i,j}}{\partial \Sigma_{\epsilon_i,j}^{-1}} &= \rho_i \hat{\boldsymbol{H}}_{i,j}^2 \Sigma_{b_i,j}^2 \left(\boldsymbol{D}_j - \tilde{\boldsymbol{c}}_{i,j}\right) \left[\boldsymbol{D}_j \left(1 - \frac{\frac{2}{\phi}\hat{\boldsymbol{H}}_{i,j}}{1 + \frac{1}{\phi}\sum_{k=1}^{n_F}\hat{\boldsymbol{H}}_{k,j}}\right) + \frac{\frac{1}{\phi}\hat{\boldsymbol{H}}_{i,j}}{1 + \frac{1}{\phi}\sum_{k=1}^{n_F}\hat{\boldsymbol{H}}_{k,j}}\tilde{\boldsymbol{c}}_{i,j}\right] \\
&= \rho_i \hat{\boldsymbol{H}}_{i,j}^2 \Sigma_{b_i,j}^2 \left(\boldsymbol{D}_j - \tilde{\boldsymbol{c}}_{i,j}\right) \left[\boldsymbol{D}_j \frac{1 - \frac{1}{\phi}\hat{\boldsymbol{H}}_{i,j} + \frac{1}{\phi}\sum_{k=1,k\neq i}^{n_F}\hat{\boldsymbol{H}}_{k,j}}{1 + \frac{1}{\phi}\sum_{k=1}^{n_F}\hat{\boldsymbol{H}}_{k,j}} + \frac{\frac{1}{\phi}\hat{\boldsymbol{H}}_{i,j}}{1 + \frac{1}{\phi}\sum_{k=1}^{n_F}\hat{\boldsymbol{H}}_{k,j}}\tilde{\boldsymbol{c}}_{i,j}\right] > 0
\end{aligned} \tag{45}$$

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

$$M^{m,cost} := \frac{\mathbf{E}\left[\sum_{i=1}^{n_F}\tilde{\boldsymbol{q}}_i'\tilde{\boldsymbol{p}}_i\right]}{\mathbf{E}\left[\sum_{i=1}^{n_F}\tilde{\boldsymbol{q}}_i'\tilde{\boldsymbol{c}}_i\right]} = \frac{\sum_{i=1}^{n_F}\mathbf{E}\left[\tilde{\boldsymbol{q}}_i'\tilde{\boldsymbol{p}}_i\right]}{\sum_{i=1}^{n_F}\mathbf{E}\left[\tilde{\boldsymbol{q}}_i'\tilde{\boldsymbol{c}}_i\right]} = \sum_{i=1}^{n_F} w_i^{cost} M_i^f \quad \text{where} \quad w_i^{cost} = \frac{\mathbf{E}\left[\tilde{\boldsymbol{q}}_i'\tilde{\boldsymbol{c}}_i\right]}{\sum_{i=1}^{n_F}\mathbf{E}\left[\tilde{\boldsymbol{q}}_i'\tilde{\boldsymbol{c}}_i\right]}. \tag{46}$$

The weight $w_{i,j}^{cost}$ increases with more data as

$$\frac{\partial w_{i,j}^{cost}}{\partial \Sigma_{\epsilon_i,j}^{-1}} = \frac{\tilde{\boldsymbol{c}}_{i,j}}{\left(\sum_{i=1}^{n_F}\mathbf{E}\left[\tilde{\boldsymbol{q}}_{i,j}'\tilde{\boldsymbol{c}}_{i,j}\right]\right)^2}\left[\frac{\partial \mathbf{E}\tilde{\boldsymbol{q}}_{i,j}}{\partial \Sigma_{\epsilon_i,j}^{-1}}\left(\sum_{k=1,k\neq i}^{n_F}\mathbf{E}\left[\tilde{\boldsymbol{q}}_{k,j}'\tilde{\boldsymbol{c}}_{k,j}\right]\right) - \mathbf{E}\tilde{\boldsymbol{q}}_{i,j}\left(\sum_{k=1,k\neq i}^{n_F}\frac{\partial \mathbf{E}(\tilde{\boldsymbol{q}}_{k,j})}{\partial \Sigma_{\epsilon_i,j}^{-1}}\tilde{\boldsymbol{c}}_{k,j}\right)\right] > 0 \tag{47}$$

The last inequality is due to the existing results $\frac{\partial \mathbf{E}\tilde{\boldsymbol{q}}_{i,j}}{\partial \Sigma_{\epsilon_i,j}^{-1}} > 0$ and $\frac{\partial \mathbf{E}(\tilde{\boldsymbol{q}}_{i,j})}{\partial \Sigma_{\epsilon_i,j}^{-1}} = \hat{\boldsymbol{H}}_{i,j}\frac{\partial \boldsymbol{D}_j}{\partial \Sigma_{\epsilon_i,j}^{-1}} < 0$.

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

$$M^{m,sale} := \sum_{i=1}^{n_F} w_i^{sale} M_i^f = \frac{\sum_{i=1}^{n_F}\frac{\mathbf{E}^2\left[\tilde{\boldsymbol{q}}_i'\tilde{\boldsymbol{p}}_i\right]}{\mathbf{E}\left[\tilde{\boldsymbol{q}}_i'\tilde{\boldsymbol{c}}_i\right]}}{\sum_{i=1}^{n_F}\mathbf{E}\left[\tilde{\boldsymbol{q}}_i'\tilde{\boldsymbol{p}}_i\right]} \quad \text{where} \quad w_i^{sale} = \frac{\mathbf{E}\left[\tilde{\boldsymbol{q}}_i'\tilde{\boldsymbol{p}}_i\right]}{\sum_{i=1}^{n_F}\mathbf{E}\left[\tilde{\boldsymbol{q}}_i'\tilde{\boldsymbol{p}}_i\right]}. \tag{48}$$

EXPECTED UTILITY To solve for the firms' cost choices, we need to solve for expected utility of each firm. We start with the expected profits. According to equation (36), we have

$$\begin{aligned}
\mathbf{E}\left[\tilde{\boldsymbol{q}}_i'\left(\tilde{\boldsymbol{p}}_i - \tilde{\boldsymbol{c}}_i\right)\right] &= \mathbf{E}\left[\tilde{\boldsymbol{q}}_i'\left(\mathbf{E}\left[\tilde{\boldsymbol{p}}_i|\mathcal{I}_i\right] - \tilde{\boldsymbol{c}}_i\right)\right] \\
&= \mathbf{E}\left[\tilde{\boldsymbol{q}}_i'\hat{\boldsymbol{H}}_i^{-1}\tilde{\boldsymbol{q}}_i\right]
\end{aligned} \tag{49}$$

The full expected utility is not conditional on the firm's signals because firms choose cost before signals are observed. This utility could be expressed as expected profit minus the price of risk.

Substituting in the expected profit expression above, we get

$$
\begin{aligned}
\mathbf{E}[U_i] &= \mathbf{E}\left[\tilde{\boldsymbol{q}}_i'\left(\tilde{\boldsymbol{p}}_i-\tilde{\boldsymbol{c}}_i\right)\right]-\frac{\rho_i}{2}\mathbf{E}\left[\tilde{\boldsymbol{q}}_i'\mathbf{Var}\left[\tilde{\boldsymbol{p}}_i|\mathcal{I}_i\right]\tilde{\boldsymbol{q}}_i\right]-g(\chi_c,\tilde{\boldsymbol{c}}_i)\\
&= \mathbf{E}\left[\tilde{\boldsymbol{q}}_i'\left(\hat{\boldsymbol{H}}_i^{-1}-\frac{\rho_i}{2}\mathbf{Var}\left[\tilde{\boldsymbol{p}}_i|\mathcal{I}_i\right]\right)\tilde{\boldsymbol{q}}_i\right]-g(\chi_c,\tilde{\boldsymbol{c}}_i)\\
&= \mathbf{E}\left[\tilde{\boldsymbol{q}}_i'\left(\frac{I_N}{\phi}+\frac{\rho_i}{2}\mathbf{Var}\left[\tilde{\boldsymbol{p}}_i|\mathcal{I}_i\right]\right)\tilde{\boldsymbol{q}}_i\right]-g(\chi_c,\tilde{\boldsymbol{c}}_i)\\
&= \frac{1}{2}\mathbf{E}\left[\tilde{\boldsymbol{q}}_i'\boldsymbol{H}_i^{-1}\tilde{\boldsymbol{q}}_i\right]-g(\chi_c,\tilde{\boldsymbol{c}}_i)\\
&= \frac{1}{2}\left(\mathbf{E}[\tilde{\boldsymbol{q}}_i]'\boldsymbol{H}_i^{-1}\mathbf{E}[\tilde{\boldsymbol{q}}_i]+\mathbf{tr}\left(\boldsymbol{H}_i^{-1}\mathbf{Var}[\tilde{\boldsymbol{q}}_i]\right)\right)-g(\chi_c,\tilde{\boldsymbol{c}}_i)
\end{aligned}
\tag{50}
$$

where $\mathbf{Var}[\tilde{\boldsymbol{q}}_i]=\boldsymbol{H}_i^{-1}\mathbf{Cov}\left(\tilde{\boldsymbol{p}}_i,\tilde{\boldsymbol{q}}_i\right)$ is independent of cost choices.

OPTIMAL CHOICES OF MARGINAL COST The first and second order condition for the optimal marginal cost choice $\boldsymbol{c}_i$ is

$$
\begin{aligned}
\frac{\partial\mathbf{E}[U_i]}{\partial\tilde{\boldsymbol{c}}_i} &= \frac{1}{2}\frac{\partial\mathbf{E}[\tilde{\boldsymbol{q}}_i]'\boldsymbol{H}_i^{-1}\mathbf{E}[\tilde{\boldsymbol{q}}_i]}{\partial\tilde{\boldsymbol{c}}_i}-\frac{\partial g(\chi_c,\tilde{\boldsymbol{c}}_i)}{\partial\tilde{\boldsymbol{c}}_i}=0\\
\frac{\partial^2\mathbf{E}[U_i]}{\partial\tilde{\boldsymbol{c}}_i\partial\tilde{\boldsymbol{c}}_i'} &= \frac{1}{2}\frac{\partial^2\mathbf{E}[\tilde{\boldsymbol{q}}_i]'\boldsymbol{H}_i^{-1}\mathbf{E}[\tilde{\boldsymbol{q}}_i]}{\partial\tilde{\boldsymbol{c}}_i\partial\tilde{\boldsymbol{c}}_i'}-\frac{\partial^2 g(\chi_c,\tilde{\boldsymbol{c}}_i)}{\partial\tilde{\boldsymbol{c}}_i\partial\tilde{\boldsymbol{c}}_i'}\ \text{ is negative semi-definite}
\end{aligned}
\tag{51}
$$

Assuming diagonal signal noise Σ_{ϵ_i}, we have $\boldsymbol{H}_{i,j}^{-1}=\frac{2}{\phi}+\rho_i\Sigma_{b_i,j}$ and $\Sigma_{b_{i,j}}=\left(1+\Sigma_{\epsilon_{i,j}}^{-1}\right)^{-1}$. Thus the FOC and SOC could be written as

$$
\begin{aligned}
\frac{\partial\mathbf{E}[U_i]}{\partial\tilde{\boldsymbol{c}}_i} &= \frac{1}{2}\frac{\partial}{\partial\tilde{\boldsymbol{c}}_i}\left(\sum_{s=1}^{N}\left(\boldsymbol{D}_s-\tilde{\boldsymbol{c}}_{i,s}\right)^2\hat{\boldsymbol{H}}_{i,s}^2\boldsymbol{H}_{i,s}^{-1}\right)-\frac{\partial g(\chi_c,\tilde{\boldsymbol{c}}_i)}{\partial\tilde{\boldsymbol{c}}_i}=0\\
\frac{\partial\mathbf{E}[U_i]}{\partial\tilde{\boldsymbol{c}}_{i,j}} &= \left(\boldsymbol{D}_j-\tilde{\boldsymbol{c}}_{i,j}\right)\hat{\boldsymbol{H}}_{i,j}^2\boldsymbol{H}_{i,j}^{-1}\left[\frac{\hat{\boldsymbol{H}}_{i,j}\phi^{-1}}{1+\frac{1}{\phi}\sum_{s=1}^{n_F}\hat{\boldsymbol{H}}_{s,j}}-1\right]-\frac{\partial g(\chi_c,\tilde{\boldsymbol{c}}_i)}{\partial\tilde{\boldsymbol{c}}_{i,j}}=0\\
\frac{\partial^2\mathbf{E}[U_i]}{\partial\tilde{\boldsymbol{c}}_{i,j}\partial\tilde{\boldsymbol{c}}_{i,k}} &= \delta_{jk}\hat{\boldsymbol{H}}_{i,j}^2\boldsymbol{H}_{i,j}^{-1}\left[\frac{\hat{\boldsymbol{H}}_{i,j}\phi^{-1}}{1+\frac{1}{\phi}\sum_{s=1}^{n_F}\hat{\boldsymbol{H}}_{s,j}}-1\right]\frac{\hat{\boldsymbol{H}}_{i,j}\phi^{-1}}{1+\frac{1}{\phi}\sum_{s=1}^{n_F}\hat{\boldsymbol{H}}_{s,j}}-\frac{\partial g(\chi_c,\tilde{\boldsymbol{c}}_i)}{\partial\tilde{\boldsymbol{c}}_{i,j}\partial\tilde{\boldsymbol{c}}_{i,k}}
\end{aligned}
\tag{52}
$$

since

$$
\boldsymbol{D}_j=\frac{\bar{\boldsymbol{p}}_j+\frac{1}{\phi}\sum_{s=1}^{n_F}\hat{\boldsymbol{H}}_{s,j}\tilde{\boldsymbol{c}}_{s,j}}{1+\frac{1}{\phi}\sum_{s=1}^{n_F}\hat{\boldsymbol{H}}_{s,j}}\quad\text{and}\quad\frac{\partial\boldsymbol{D}_j}{\partial\tilde{\boldsymbol{c}}_{i,k}}=\delta_{jk}\frac{\hat{\boldsymbol{H}}_{i,j}\phi^{-1}}{1+\frac{1}{\phi}\sum_{s=1}^{n_F}\hat{\boldsymbol{H}}_{s,j}}
\tag{53}
$$

B. Proofs and Auxiliary Results

Proof of Lemma 1: Data-Investment Complementarity

Proof. Starting from the expected utility, we have

$$
\mathbb{E}\left[U_i\right]=\frac{1}{2}\left[\mathbb{E}\left[\widetilde{\boldsymbol{q}}_i\right]'\boldsymbol{H}_i^{-1}\mathbb{E}\left[\widetilde{\boldsymbol{q}}_i\right]+\mathrm{tr}\left(\boldsymbol{H}_i^{-1}\mathbb{V}\left[\widetilde{\boldsymbol{q}}_i\right]\right)\right]-g\left(\chi_c,\widetilde{\boldsymbol{c}}_i\right)
\tag{54}
$$

To show this complementarity between information and costs, we first differentiate $\mathbb{E}\left[U_i\right]$ with regarding to marginal cost. Here, $\tilde{c}_{ij}$ denotes firm i's marginal cost of producing attribute j. $\widehat{\boldsymbol{H}}_{ij}$ 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. Then,

$$\begin{aligned}\frac{\partial \mathbb{E}\left[U_i\right]}{\partial \widetilde{c}_{ij}} &= \frac{\partial}{\partial \widetilde{c}_{ij}}\left\{\frac{1}{2}\left[\mathbb{E}\left[\widetilde{\boldsymbol{q}}_i\right]' \boldsymbol{H}_i^{-1}\mathbb{E}\left[\widetilde{\boldsymbol{q}}_i\right] + \operatorname{tr}\left(\boldsymbol{H}_i^{-1}\mathbb{V}\left[\widetilde{\boldsymbol{q}}_i\right]\right)\right] - g\left(\chi_c, \widetilde{\boldsymbol{c}}_i\right)\right\} \\ &= -\left(\widehat{\boldsymbol{H}}_{ij}\right)^2 \boldsymbol{H}_{ij}^{-1}\left\{\frac{\left[\phi\left(\overline{\boldsymbol{p}} - \widetilde{c}_{ij}\right) + \sum_{s\neq i}^{n_F}\widehat{H}_{sj}\left(\widetilde{c}_{sj} - \widetilde{c}_{ij}\right)\right]\left(\phi + \sum_{s\neq i}\widehat{H}_{sj}\right)}{\left(\phi + \sum_{i=1}^{n_F}\widehat{\boldsymbol{H}}_{ij}\right)^2}\right\} - \frac{\partial g\left(\chi_c, \widetilde{\boldsymbol{c}}_i\right)}{\partial \widetilde{c}_i}\end{aligned} \tag{55}$$

Denote $\mathbb{V} = \mathbb{V}\left[\boldsymbol{b}_{ij}|\mathcal{I}\right]$. Then, the second order cross derivative is:

$$\begin{aligned}\frac{\partial^2 \mathbb{E}\left[U_i\right]}{\partial \widetilde{c}_{ij}\partial \mathbb{V}\left[\boldsymbol{b}_{ij}|\mathcal{I}\right]} &= \frac{\partial}{\partial \mathbb{V}}\left\{-\left(\widehat{\boldsymbol{H}}_{ij}\right)^2 \boldsymbol{H}_{ij}^{-1}\left\{\frac{\left[\phi\left(\overline{\boldsymbol{p}} - \widetilde{c}_{ij}\right) + \sum_{s\neq i}^{n_F}\widehat{H}_{sj}\left(\widetilde{c}_{sj} - \widetilde{c}_{ij}\right)\right]\left(\phi + \sum_{s\neq i}\widehat{H}_{sj}\right)}{\left(\phi + \sum_{i=1}^{n_F}\widehat{\boldsymbol{H}}_{ij}\right)^2}\right\} - \frac{\partial g\left(\chi_c, \widetilde{\boldsymbol{c}}_i\right)}{\partial \widetilde{c}_i}\right\} \\ &= \underbrace{-\left\{\left[\phi\left(\overline{\boldsymbol{p}} - \widetilde{c}_{ij}\right) + \sum_{s\neq i}^{n_F}\widehat{H}_{sj}\left(\widetilde{c}_{sj} - \widetilde{c}_{ij}\right)\right]\left(\phi + \sum_{s\neq i}\widehat{H}_{sj}\right)\right\}}_{\text{A negative term } (-)} \\ &\qquad \times \frac{\partial}{\partial \mathbb{V}_i}\left\{\frac{\frac{2}{\phi} + \rho_i\mathbb{V}_i}{\left(\phi + \sum_{s=1}^{n_F}\left(\frac{1}{\phi} + \rho_s\mathbb{V}_s\right)^{-1}\right)^2\left(\frac{1}{\phi} + \rho_i\mathbb{V}_i\right)^2}\right\} \\ &= (-) \times \underbrace{\left\{-\frac{\left(\phi + \sum_{s\neq i}^{n_F}\left(\frac{1}{\phi} + \rho_s\mathbb{V}_s\right)^{-1}\right)\left(\frac{1}{\phi} + \rho_i\mathbb{V}_i\right) + 1 + \frac{2}{\phi}\sum_{s\neq i}^{n_F}\left(\frac{1}{\phi} + \rho_s\mathbb{V}_s\right)^{-1}}{\left(\phi + \sum_{s=1}^{n_F}\left(\frac{1}{\phi} + \rho_s\mathbb{V}_s\right)^{-1}\right)^3\left(\frac{1}{\phi} + \rho_i\mathbb{V}_i\right)^3}\right\}}_{\text{Negative, i.e., } <0} > 0 \\ &= (-) \times \underbrace{\frac{\partial}{\partial \mathbb{V}_i}\left\{\frac{1}{\left(2 + \phi\rho_i\mathbb{V}_i + \sum_{s\neq i}^{n_F}\frac{\frac{1}{\phi} + \rho_i\mathbb{V}_i}{\frac{1}{\phi} + \rho_s\mathbb{V}_s}\right)\left(\phi + \sum_{s\neq i}^{n_F}\left[\frac{1}{\frac{1}{\phi} + \rho_s\mathbb{V}_s}\left(1 - \frac{1}{2 + \phi\rho_i\mathbb{V}_i}\right)\right]\right)}\right\}}_{\text{Decreasing on } \mathbb{V}_i, \text{ i.e., } <0} > 0\end{aligned} \tag{56}$$

Hence, we get $\frac{\partial^2 \mathbb{E}[U_i]}{\partial \widetilde{c}_{ij}\partial \mathbb{V}} > 0$, which means the marginal benefit from reducing costs is higher (more negative) when firms have better information (lower variance).

□

Proof of Lemma 2: Greater investment raises a firm's product markup.

Proof. More investment would lower marginal cost $\tilde{\boldsymbol{c}}_{i,j}$ and its derivative is

$$\frac{\partial M^{\tilde{p}}_{i,j}}{\partial \tilde{\boldsymbol{c}}_{i,j}} = \frac{\frac{\partial \boldsymbol{D}_j}{\partial \tilde{c}_{i,j}}\tilde{\boldsymbol{c}}_{i,j} - \boldsymbol{D}_j}{\tilde{\boldsymbol{c}}^2_{i,j}} = \frac{\frac{1}{\phi}\hat{\boldsymbol{H}}_{i,j}\tilde{\boldsymbol{c}}_{i,j} - \bar{\boldsymbol{p}}_j - \frac{1}{\phi}\sum_{s=1}^{n_F}\hat{\boldsymbol{H}}_{s,j}\tilde{\boldsymbol{c}}_{s,j}}{\tilde{\boldsymbol{c}}^2_{i,j}\left(1+\frac{1}{\phi}\sum_{s=1}^{n_F}\hat{\boldsymbol{H}}_{s,j}\right)} = -\frac{\bar{\boldsymbol{p}}_j + \frac{1}{\phi}\sum_{s=1,s\neq i}^{n_F}\hat{\boldsymbol{H}}_{s,j}\tilde{\boldsymbol{c}}_{s,j}}{\tilde{\boldsymbol{c}}^2_{i,j}\left(1+\frac{1}{\phi}\sum_{s=1}^{n_F}\hat{\boldsymbol{H}}_{s,j}\right)} \leq 0 \tag{57}$$

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

To link it to the product level, let's look at the product k that used $A_{kj} > 0$ attribute j.

$$\frac{\partial M_{i,k}}{\partial \tilde{\boldsymbol{c}}_{i,j}} = \frac{\partial}{\partial \tilde{\boldsymbol{c}}_{i,j}} \frac{\sum_j A_{k,j}\mathbb{E}[\tilde{p}_{i,j}]}{\sum_j A_{kj}\tilde{\boldsymbol{c}}_{i,j}} \tag{58}$$

$$= \frac{[\sum_j A_{kj}\tilde{\boldsymbol{c}}_{i,j}]A_{kj}\frac{\partial}{\partial \tilde{c}_{i,j}}\mathbb{E}[\tilde{p}_{i,j}] - [\sum_j A_{k,j}\mathbb{E}[\tilde{p}_{i,j}]]A_{kj}}{[\sum_j A_{kj}\tilde{\boldsymbol{c}}_{i,j}]^2} \tag{59}$$

$$= \frac{A_{kj}}{\sum_j A_{kj}\tilde{\boldsymbol{c}}_{i,j}}\left[\frac{\partial}{\partial \tilde{\boldsymbol{c}}_{i,j}}\mathbb{E}[\tilde{p}_{i,j}] - M_{i,k}\right]. \tag{60}$$

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

$$\frac{\partial M^{\tilde{p}}_{i,j}}{\partial \tilde{\boldsymbol{c}}_{i,j}} = \frac{1}{\tilde{\boldsymbol{c}}_{i,j}}\left[\frac{\partial}{\partial \tilde{\boldsymbol{c}}_{i,j}}\mathbb{E}[\tilde{p}_{i,j}] - M_{i,k}\right] \leq 0. \tag{61}$$

Therefore, (60) is negative if the markup on product k is greater than the markup on attribute j: $M_{i,k} \geq M^{\tilde{p}_{i,j}}$. □

Comment: To see why not every attribute markup increase raises the product markup, consider a numerical example where a product uses 99% of an attribute with a price 101 and cost 100 and 1% of an attribute that costs 1 and has a price 5. The product markup is

$$\frac{99\% \cdot 101 + 1\% \cdot 5}{99\% \cdot 100 + 1\% \cdot 1} \approx 1.10403 \tag{62}$$

Now suppose we decrease the cost of product 2 to 0.9 and the price to 4.6 (note the attribute markup increases from 5 to 5.11)

$$\frac{99\% \cdot 101 + 1\% \cdot 4.6}{99\% \cdot 100 + 1\% \cdot 0.9} \approx 1.10373 < 1.10403 \tag{63}$$

Therefore, we proved that lowering the cost of a low-markup attribute can increases the product markup but not necessarily a high-markup attribute.

Proof of Lemma 3: (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 $\mathbf{E}(p_i/c_i) = \mathbf{E}(p_i)/c_i$ 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 $M^{\tilde{p}}_{i,j} := \mathbf{E}[\tilde{\boldsymbol{p}}_{i,j}]/\boldsymbol{c}_{i,j}$. The average markup on the attributes is

$$\overline{M}^{\tilde{p}} = \frac{1}{N}\frac{1}{n_F}\sum_{i=1}^{n_F}\sum_{j=1}^{N} M_{i,j}^{\tilde{p}} = \frac{1}{n_F N}\sum_{i=1}^{n_F}\sum_{j=1}^{N}\frac{\mathbf{E}[\tilde{\boldsymbol{p}}_{i,j}]}{\boldsymbol{c}_{i,j}} = \frac{1}{n_F N}\sum_{i=1}^{n_F}\sum_{j=1}^{N}\frac{\boldsymbol{D}_j}{\boldsymbol{c}_{i,j}} \tag{64}$$

We denote the posterior variance $\Sigma_{b_i} = \mathbf{Var}\left[\boldsymbol{b}_i|\mathcal{I}_i\right] = \left(I_N + \Sigma_{\epsilon_i}^{-1}\right)^{-1}$. The j^{th} term of equilibrium price $\boldsymbol{D}_j$ is

$$\boldsymbol{D}_j = \frac{\bar{\boldsymbol{p}}_j + \frac{1}{\phi}\sum_{i=1}^{n_F}\hat{\boldsymbol{H}}_{i,j}\boldsymbol{c}_{i,j}}{1 + \frac{1}{\phi}\sum_{i=1}^{n_F}\hat{\boldsymbol{H}}_{i,j}} \quad \text{where} \quad \hat{\boldsymbol{H}}_{i,j} = \left[\frac{1}{\phi} + \rho_i\left(1 + \Sigma_{\epsilon_{i,j}}^{-1}\right)^{-1}\right]^{-1} \tag{65}$$

The positive output means $\boldsymbol{D} \geq \boldsymbol{c}_i$, thus

$$\frac{\partial \boldsymbol{D}_j}{\partial \Sigma_{\epsilon_{i,k}}^{-1}} = \delta_{jk}\frac{1}{\phi}\frac{\rho_i \hat{\boldsymbol{H}}_{i,j}^2 \Sigma_{b_i,j}^2}{1 + \frac{1}{\phi}\sum_{s=1}^{n_F}\hat{\boldsymbol{H}}_{s,j}}\left(\boldsymbol{c}_{i,j} - \boldsymbol{D}_j\right) < 0 \Rightarrow \frac{\partial \overline{M}^{\tilde{p}}}{\partial \Sigma_{\epsilon_i,k}^{-1}} < 0 \tag{66}$$

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 $M_{i,j}^{\tilde{p}} = \mathbf{E}[\tilde{\boldsymbol{p}}_{i,j}]/\tilde{\boldsymbol{c}}_{i,j} = \boldsymbol{D}_j/\tilde{\boldsymbol{c}}_{i,j}$. Its partial derivative to data is

$$\frac{\partial M_{i,j}^{\tilde{p}}}{\partial \Sigma_{\epsilon_i,j}^{-1}} = \frac{\frac{\partial \boldsymbol{D}_j}{\partial \Sigma_{\epsilon_i,j}^{-1}}\tilde{\boldsymbol{c}}_{i,j} - \boldsymbol{D}_j\frac{\partial \tilde{\boldsymbol{c}}_{i,j}}{\partial \Sigma_{\epsilon_i,j}^{-1}}}{\tilde{\boldsymbol{c}}_{i,j}^2} \tag{67}$$

According to (42) and Lemma 1, we have

$$\frac{\partial \boldsymbol{D}_j}{\partial \Sigma_{\epsilon_i,j}^{-1}} = \frac{1}{\phi}\frac{\rho_i \hat{\boldsymbol{H}}_{i,j}^2 \Sigma_{b_i,j}^2}{1 + \frac{1}{\phi}\sum_{s=1}^{n_F}\hat{\boldsymbol{H}}_{s,j}}\left(\tilde{\boldsymbol{c}}_{i,j} - \boldsymbol{D}_j\right) < 0 \quad \text{and} \quad \frac{\partial \tilde{\boldsymbol{c}}_i}{\partial \Sigma_{\epsilon_i}} = \left(-\frac{\partial^2 \mathbf{E}[U_i]}{\partial \tilde{\boldsymbol{c}}_i \partial \tilde{\boldsymbol{c}}_i'}\right)^{-1}\Lambda, \quad \frac{\partial \tilde{\boldsymbol{c}}_{i,j}}{\partial \Sigma_{\epsilon_i,j}^{-1}} \leq 0 \tag{68}$$

If marginal cost $\tilde{\boldsymbol{c}}_{i,j}$ or price of risk ρ_i is sufficiently low, the second term in the numerator $-\boldsymbol{D}_j\frac{\partial \tilde{\boldsymbol{c}}_{i,j}}{\partial \Sigma_{\epsilon_i,j}^{-1}} > 0$ dominates the marginal effect, thus increasing product markups.

□

WELFARE PRELIMINARIES In this section, we work out the components of welfare, as preliminaries to the two welfare results that follow. We start with firms' profits:

$$\mathbb{E}\left[U_i\right] = \frac{1}{2}\left[\mathbb{E}\left[\widetilde{\boldsymbol{q}}_i\right]' \boldsymbol{H}_i^{-1}\mathbb{E}\left[\widetilde{\boldsymbol{q}}_i\right] + \boldsymbol{H}_i^{-1}\mathbb{V}\left[\widetilde{\boldsymbol{q}}_i\right]\right] \tag{69}$$

$$\text{where} \qquad \widetilde{\boldsymbol{q}}_i = \widehat{\boldsymbol{H}}_i\left(\boldsymbol{D}-\overline{\boldsymbol{c}}\right) + \widehat{\boldsymbol{H}}_i\boldsymbol{K}_i\boldsymbol{s}_i - \frac{\widehat{\boldsymbol{H}}_i}{\phi}\left(1+\frac{1}{\phi}\sum_{i=1}^{n_F}\widehat{\boldsymbol{H}}_i\right)^{-1}\sum_{j=1}^{n_F}\widehat{\boldsymbol{H}}_j\boldsymbol{K}_j\boldsymbol{s}_j \tag{70}$$

Combining these, we have:

$$\mathbb{E}\left[U_i\right] = \frac{1}{2}\left(\boldsymbol{D}-\overline{\boldsymbol{c}}\right)^\top \widehat{\boldsymbol{H}}_i\boldsymbol{H}_i^{-1}\widehat{\boldsymbol{H}}_i\left(\boldsymbol{D}-\overline{\boldsymbol{c}}\right) + \frac{1}{2}\left(\phi I + \sum_{i=1}^{n_F}\widehat{\boldsymbol{H}}_i\right)^\top \widehat{\boldsymbol{H}}_i H_i^{-1}\widehat{\boldsymbol{H}}_i\left(\phi I + \sum_{i=1}^{n_F}\widehat{\boldsymbol{H}}_i\right) \tag{71}$$

$$\times\left[\left(\phi I_N + \sum_j \widehat{\boldsymbol{H}}_j\right)^\top K_i^\top\left(I_N+\Sigma_{\varepsilon,i}\right)K_i\left(\phi I_N + \sum_j\widehat{\boldsymbol{H}}_j\right) + \sum_{j\neq i}\widehat{\boldsymbol{H}}_j\boldsymbol{K}_j^\top\left(I_N+\Sigma_{\varepsilon,j}\right)\boldsymbol{K}_j\widehat{\boldsymbol{H}}_j\right] \tag{72}$$

Consumer surplus:

$$\mathbb{E}CS = \left\{\frac{\left(\mathbb{E}\left[\widetilde{Q}\right]\right)^2 + \mathbb{V}\left[\widetilde{Q}\right]}{2\phi}\right\} \tag{73}$$

$$= \frac{\left(\boldsymbol{D}-\overline{\boldsymbol{c}}\right)^\top\left(\widehat{\boldsymbol{H}}_1+\widehat{\boldsymbol{H}}_2\right)^\top\left(\widehat{\boldsymbol{H}}_1+\widehat{\boldsymbol{H}}_2\right)\left(\boldsymbol{D}-\overline{\boldsymbol{c}}\right)}{2\phi} + \frac{\phi}{2}\left(\phi I + \widehat{\boldsymbol{H}}_1+\widehat{\boldsymbol{H}}_2\right)^{-1} \tag{74}$$

$$\times\left[\left(\widehat{\boldsymbol{H}}_1\boldsymbol{K}_1\right)^\top\left(1+\Sigma_{\varepsilon,1}\right)\left(\widehat{\boldsymbol{H}}_1\boldsymbol{K}_1\right) + \left(\widehat{\boldsymbol{H}}_2\boldsymbol{K}_2\right)^\top\left(1+\Sigma_{\varepsilon,2}\right)\left(\widehat{\boldsymbol{H}}_2\boldsymbol{K}_2\right)\right] \tag{75}$$

Welfare is thus:

$$\begin{aligned}\mathbb{W} = \frac{1}{2}\left(\phi I + \widehat{\boldsymbol{H}}_1+\widehat{\boldsymbol{H}}_2\right)^{-1}&\left\{\phi^2\left(\overline{\boldsymbol{p}}-\overline{\boldsymbol{c}}\right)^\top\left[\widehat{\boldsymbol{H}}_1 H_1^{-1}\widehat{\boldsymbol{H}}_1 + \widehat{\boldsymbol{H}}_2 H_2^{-1}\widehat{\boldsymbol{H}}_2 + \frac{\left(\widehat{\boldsymbol{H}}_1+\widehat{\boldsymbol{H}}_2\right)^\top\left(\widehat{\boldsymbol{H}}_1+\widehat{\boldsymbol{H}}_2\right)}{\phi}\right]\left(\overline{\boldsymbol{p}}-\overline{\boldsymbol{c}}\right)\right.\\ &+\frac{\widehat{\boldsymbol{H}}_1^2}{1+\Sigma_{\varepsilon,1}}\left[\left(\phi I + \widehat{\boldsymbol{H}}_2\right)H_1^{-1}\left(\phi I+\widehat{\boldsymbol{H}}_2\right) + \widehat{\boldsymbol{H}}_2 H_2^{-1}\widehat{\boldsymbol{H}}_2 + \phi I\right] + \\ &\left.\frac{\widehat{\boldsymbol{H}}_2^2}{1+\Sigma_{\varepsilon,2}}\left[\left(\phi I + \widehat{\boldsymbol{H}}_1\right)H_2^{-1}\left(\phi I+\widehat{\boldsymbol{H}}_1\right) + \widehat{\boldsymbol{H}}_1 H_1^{-1}\widehat{\boldsymbol{H}}_1 + \phi I\right]\right\}\left(\phi I + \widehat{\boldsymbol{H}}_1+\widehat{\boldsymbol{H}}_2\right)^{-1}\end{aligned} \tag{76}$$

where

$$
\begin{aligned}
\boldsymbol{D} &= \left(1 + \frac{1}{\phi}\sum_{i=1}^{n_F} \widehat{\boldsymbol{H}}_i\right)^{-1}\left(\overline{\boldsymbol{p}} + \frac{1}{\phi}\sum_{i=1}^{n_F} \widehat{H}_i \overline{\boldsymbol{c}}\right) && (77)\\
\boldsymbol{H}_i &= \left(\frac{2}{\phi} + \rho_i \mathbb{V}\left[b_i|\mathcal{I}_i\right]\right)^{-1} && (78)\\
\boldsymbol{K}_i &= \left(1+\Sigma_{\varepsilon,i}\right)^{-1} && (79)\\
\widehat{\boldsymbol{H}}_i &= \left(\frac{1}{\phi} + \rho_i \mathbb{V}\left[b_i|\mathcal{I}_i\right]\right)^{-1} && (80)\\
\mathbb{V}\left[b_i|\mathcal{I}_i\right] &= \frac{\Sigma_{\varepsilon,i}}{1+\Sigma_{\varepsilon,i}} && (81)\\
\Sigma_{\varepsilon,i} &= \widetilde{\Sigma}/n_{di} && (82)
\end{aligned}
$$

Proof of Proposition 2: Welfare For simplicity, we denote $n_{d,1} = n_{d,2} = x$. In this case, we can get rid of the i subscripts in (76), which gives us:

$$
\begin{aligned}
\mathbb{W} = {} & \frac{\left(x+\widetilde{\Sigma}\right)^2}{\left[\left(x+\widetilde{\Sigma}+\phi\rho_i\widetilde{\Sigma}\right)+2\left(x+\widetilde{\Sigma}\right)\right]^2}\left\{\phi\left(\frac{2x+\left(2+\phi\rho_i\right)\widetilde{\Sigma}}{x+\widetilde{\Sigma}}+2\right)\left(\overline{\boldsymbol{p}}-\overline{\boldsymbol{c}}\right)^2\right.\\
& \left.+\frac{x}{x+\widetilde{\Sigma}}\left[\frac{\phi\left[2x+\left(2+\phi\rho_i\right)\widetilde{\Sigma}\right]^3}{\left(x+\widetilde{\Sigma}+\phi\rho_i\widetilde{\Sigma}\right)^2\left(x+\widetilde{\Sigma}\right)}+\frac{\phi\left(x+\widetilde{\Sigma}\right)\left[2x+\left(2+\phi\rho_i\right)\widetilde{\Sigma}\right]}{\left(x+\widetilde{\Sigma}+\phi\rho_i\widetilde{\Sigma}\right)^2}+\phi\right]\right\}
\end{aligned} \tag{83}
$$

Denote $y := x + \widetilde{\Sigma}$, we have

$$
\mathbb{W} = \frac{\phi y}{\left[\left(y+\phi\rho_i\widetilde{\Sigma}\right)+2y\right]^2}\left\{\left(\phi\rho_i\widetilde{\Sigma}+4y\right)\left(\overline{\boldsymbol{p}}-\overline{\boldsymbol{c}}\right)^2+\left(y-\widetilde{\Sigma}\right)\left[\frac{\left[2y+\phi\rho_i\widetilde{\Sigma}\right]^3}{y\left(y+\phi\rho_i\widetilde{\Sigma}\right)^2}+\frac{y\left[2y+\phi\rho_i\widetilde{\Sigma}\right]}{\left(y+\phi\rho_i\widetilde{\Sigma}\right)^2}+1\right]\right\} \tag{84}
$$

Finally, the derivative is:

$$
\begin{aligned}
\frac{\partial\mathbb{W}}{\partial x} = \frac{\partial\mathbb{W}}{\partial y} & > \frac{\phi^2\rho\widetilde{\Sigma}\left(5y+\phi\rho_i\widetilde{\Sigma}\right)\left(\overline{\boldsymbol{p}}-\overline{\boldsymbol{c}}\right)^2}{\left(3y+\phi\rho_i\widetilde{\Sigma}\right)^3}+\frac{\phi\widetilde{\Sigma}}{\left[3y+\phi\rho_i\widetilde{\Sigma}\right]^2}+\phi\frac{2y\left[5y^2+4\phi\rho_i\widetilde{\Sigma}y+\left(\phi\rho_i\widetilde{\Sigma}\right)^2\right]}{\left[3y+\phi\rho_i\widetilde{\Sigma}\right]^2\left(y+\phi\rho_i\widetilde{\Sigma}\right)^2}\\
& \quad + \frac{\phi^2\rho_i\widetilde{\Sigma}\left\{2.5y\left[5y^2+4\phi\rho_i\widetilde{\Sigma}y+\left(\phi\rho_i\widetilde{\Sigma}\right)^2\right]+\left(y+2\phi\rho_i\widetilde{\Sigma}\right)\left[\frac{1y^2+4\phi\rho_i\widetilde{\Sigma}y+\left(\phi\rho_i\widetilde{\Sigma}\right)^2}{2}\right]\right\}}{\left[3y+\phi\rho_i\widetilde{\Sigma}\right]^3\left(y+\phi\rho_i\widetilde{\Sigma}\right)^2}\\
& > 0
\end{aligned} \tag{85}
$$

This proof holds for the case with $\chi_c \to +\infty$ or 0. We can extend the results for sufficiently large (small) χ_c by using continuity. Thus, the welfare is increasing in the number of data points.

Proof of Proposition 3: Welfare with Asymmetric Firms Consider the case where firms have different numbers of data points.
Typically, we write $N_1 = n_{d,1} = N - x$ and $N_2 = n_{d,2} = N + x$.
Also, denote $\widetilde{N} = N + \widetilde{\Sigma}$
Then, we have:

$$\mathbb{E}CS = \frac{\phi\left(\widehat{\boldsymbol{H}}_1 + \widehat{\boldsymbol{H}}_2\right)^2}{2\left(\phi + \widehat{\boldsymbol{H}}_1 + \widehat{\boldsymbol{H}}_2\right)^2}\left[(\overline{\boldsymbol{p}} - \overline{\boldsymbol{c}})\right]^2 + \frac{\phi}{2\left(\phi + \widehat{\boldsymbol{H}}_1 + \widehat{\boldsymbol{H}}_2\right)^2}\left[\frac{\left(\widehat{\boldsymbol{H}}_1\right)^2}{1+\Sigma_{\varepsilon,1}} + \frac{\left(\widehat{\boldsymbol{H}}_2\right)^2}{1+\Sigma_{\varepsilon,2}}\right]$$

$$\begin{aligned}\mathbb{E}\left[\Pi\right] = {} & \phi\frac{\left\{10\phi\rho\widetilde{\Sigma}\left[\left(N+\widetilde{\Sigma}\right)^2 - x^2\right] + 2\left(\phi\rho\widetilde{\Sigma}\right)^2\left(4N + 4\widetilde{\Sigma} + \phi\rho\widetilde{\Sigma}\right)\right\}\left(N+\widetilde{\Sigma}\right) + 4\left[\left(N+\widetilde{\Sigma}\right)^2 - x^2\right]^2}{2\left[3\left(N^2-x^2\right) + \left(3+2\phi\rho\right)\widetilde{\Sigma}\left(2N\right) + \left(1+\phi\rho\right)\left(3+\phi\rho\right)\widetilde{\Sigma}^2\right]^2}\left(\overline{\boldsymbol{p}} - \overline{\boldsymbol{c}}\right)^2 \\ & + \phi\frac{\left\{5\left[\left(N+\widetilde{\Sigma}\right)^2 - x^2\right] + 2\phi\rho\widetilde{\Sigma}\left(2N + 2\widetilde{\Sigma}\right) + \left(\phi\rho\widetilde{\Sigma}\right)^2\right\}\left[4\left(N^2-x^2\right) + \left(2+\phi\rho\right)\widetilde{\Sigma}\left(2N\right)\right]}{2\left[3\left(N^2-x^2\right) + \left(3+2\phi\rho\right)\widetilde{\Sigma}\left(2N\right) + \left(1+\phi\rho\right)\left(3+\phi\rho\right)\widetilde{\Sigma}^2\right]^2}\end{aligned} \tag{86}$$

Notice that here we have a component $(\overline{\boldsymbol{p}} - \overline{\boldsymbol{c}})^2$ that divides the expression into two parts
The first term (with $(\overline{\boldsymbol{p}} - \overline{\boldsymbol{c}})^2$)

$$\mathbb{W}_{(1)} = \phi\left(\overline{\boldsymbol{p}} - \overline{\boldsymbol{c}}\right)^2\frac{\left[\left(\widetilde{N}^2 - x^2\right) + \phi\rho\widetilde{\Sigma}\widetilde{N}\right]\left[4\left(\widetilde{N}^2 - x^2\right) + 5\phi\rho\widetilde{\Sigma}\widetilde{N}\right] + \widetilde{N}\left(\phi\rho\widetilde{\Sigma}\right)^2\left(\widetilde{N} + \phi\rho\widetilde{\Sigma}\right)}{\left[3\left(\widetilde{N}^2 - x^2\right) + \left(\phi\rho\widetilde{\Sigma}\right)\left(4\widetilde{N} + \phi\rho\widetilde{\Sigma}\right)\right]^2} \tag{87}$$

Denote $y := \left(\widetilde{N}^2 - x^2\right)$

$$\mathbb{W}_{(1)} = \phi\left(\overline{\boldsymbol{p}} - \overline{\boldsymbol{c}}\right)^2\frac{\left(y + \phi\rho\widetilde{\Sigma}\widetilde{N}\right)\left(4y + 5\phi\rho\widetilde{\Sigma}\widetilde{N}\right) + \widetilde{N}\left(\phi\rho\widetilde{\Sigma}\right)^2\left(\widetilde{N} + \phi\rho\widetilde{\Sigma}\right)}{\left[3y + \left(\phi\rho\widetilde{\Sigma}\right)\left(4\widetilde{N} + \phi\rho\widetilde{\Sigma}\right)\right]^2} \tag{88}$$

The derivative is

$$\begin{aligned}\frac{\partial\mathbb{W}_{(1)}}{\partial x} &= \frac{\partial\mathbb{W}_{(1)}}{\partial y}\frac{\partial y}{\partial x} &(89)\\ &= -2\phi\left(\overline{\boldsymbol{p}} - \overline{\boldsymbol{c}}\right)^2 x\frac{\phi\rho\widetilde{\Sigma}\widetilde{N}\left(5y\right) + \left(\phi\rho\widetilde{\Sigma}\right)^2\left(8y + 3\phi\rho\widetilde{\Sigma}\widetilde{N}\right)}{\left[3y + \left(\phi\rho\widetilde{\Sigma}\right)\left(4\widetilde{N} + \phi\rho\widetilde{\Sigma}\right)\right]^3} &(90)\\ &< 0 &(91)\end{aligned}$$

Therefore, the first part is decreasing over the data asymmetry. The second term:

$$
\begin{aligned}
\mathbb{W}_{(2)} = \phi & \frac{\left\{\left[5\widetilde{N}^2 + \phi\rho\widetilde{\Sigma}\left(4\widetilde{N} + \phi\rho\widetilde{\Sigma}\right)\right]\left(2\widetilde{N} + \phi\rho\widetilde{\Sigma}\right) + \widetilde{N}\left(\widetilde{N} + \phi\rho\widetilde{\Sigma}\right)^2\right\}\left(\widetilde{N} - \widetilde{\Sigma}\right)}{\left[3\left(\widetilde{N}^2 - x^2\right) + \left(\phi\rho\widetilde{\Sigma}\right)\left(4\widetilde{N} + \phi\rho\widetilde{\Sigma}\right)\right]^2} \\
& + \phi\frac{x^2\left[11x^2 - 14\widetilde{N}^2 + 7\left(1 - \phi\rho\right)\widetilde{\Sigma}\widetilde{N} + \left(3 - \phi\rho\right)\phi\rho\widetilde{\Sigma}^2\right]}{\left[3\left(\widetilde{N}^2 - x^2\right) + \left(\phi\rho\widetilde{\Sigma}\right)\left(4\widetilde{N} + \phi\rho\widetilde{\Sigma}\right)\right]^2}
\end{aligned}
\tag{92}
$$

The denominator is decreasing over x. The numerator is increasing on x. Therefore, the second term is increasing when the data asymmetry expands. This proof holds for the case with $\chi_c \to +\infty$ or 0. We can extend the results for sufficiently large (small) χ_c by using continuity.

Proof of Proposition 4: The firm-level markup wedge increases in data.

Proof. Firm-level markup for firm i is M_i^f is defined as

$$
M_i^f = \frac{\mathbf{E}[\tilde{\boldsymbol{q}}_i'\tilde{\boldsymbol{p}}_i]}{\mathbf{E}[\tilde{\boldsymbol{q}}_i'\tilde{\boldsymbol{c}}_i]} = \frac{\mathbf{E}[\tilde{\boldsymbol{q}}_i]'\mathbf{E}[\tilde{\boldsymbol{p}}_i] + \mathbf{trCov}\left(\tilde{\boldsymbol{p}}_i, \tilde{\boldsymbol{q}}_i\right)}{\mathbf{E}[\tilde{\boldsymbol{q}}_i'\tilde{\boldsymbol{c}}_i]} = \frac{\sum_{l=1}^N \mathbf{E}[\tilde{\boldsymbol{q}}_{i,l}]\mathbf{E}[\tilde{\boldsymbol{p}}_{i,l}] + \sum_{l=1}^N \mathbf{Cov}_{i,l}}{\sum_{l=1}^N \mathbf{E}[\tilde{\boldsymbol{q}}_{i,l}]\tilde{\boldsymbol{c}}_{i,l}} \tag{93}
$$

where $\mathbf{Cov}_{i,l}$ is the l^{th} diagonal value of the price-quantity covariance matrix $cov(\tilde{p}_i, \tilde{q}_i)$. From (40) this is

$$
\mathbf{Cov}_{i,l} = \frac{\hat{\boldsymbol{H}}_{i,l}}{\left(1 + \sum_{s=1}^{n_F}\frac{1}{\phi}\hat{\boldsymbol{H}}_{s,l}\right)^2}\left[\sum_{s=1,s\neq i}^{n_F}\left(\frac{\hat{\boldsymbol{H}}_{s,l}}{\phi}\right)^2\boldsymbol{K}_{s,l} + \boldsymbol{K}_{i,l}\left(1 + \sum_{s=1,s\neq i}^{n_F}\frac{\hat{\boldsymbol{H}}_{s,l}}{\phi}\right)^2\right] \tag{94}
$$

where K_{il} is firm i's Bayesian updating weight on the signal about attribute l.

Taking partial derivative of the Kalman gain $\boldsymbol{K}_i = (I_N + \Sigma_{\epsilon_i})^{-1}$, with respect to $\Sigma_{\epsilon_k,l}^{-1}$ yields

$$
\frac{\partial \boldsymbol{K}_{i,j}}{\partial \Sigma_{\epsilon,k,l}^{-1}} = \delta_{ki}\delta_{jl}(1 + \Sigma_{\epsilon,i,j})^{-2}\Sigma_{\epsilon,i,j}^{-2} \tag{95}
$$

Recall that $\hat{\boldsymbol{H}}_{k,l} = (\phi^{-1} + \rho_k(\Sigma_{b_k,l}^{-1} + \Sigma_{\epsilon,k,l}^{-1})^{-1}))^{-1}$. This implies

$$
\frac{\partial \hat{\boldsymbol{H}}_{k,l}}{\partial \Sigma_{\epsilon,i,j}^{-1}} = \delta_{lj}\delta_{ki}\hat{\boldsymbol{H}}_{k,l}^2\rho_k(\Sigma_{b_k,l}^{-1} + \Sigma_{\epsilon_k,l}^{-1})^{-2} \tag{96}
$$

Similarly, using $\mathbb{E}\left[\tilde{\boldsymbol{p}}_{k,l}\right] = \boldsymbol{D}_l = \frac{\bar{p}_l + \frac{1}{\phi}\sum_{i'=1}^{n_F} \hat{\boldsymbol{H}}_{i',l}\tilde{\boldsymbol{c}}_{i'l}}{1+\frac{1}{\phi}\sum_{i'=1}^{n_F} \hat{\boldsymbol{H}}_{i',l}}$, we obtain

$$\frac{\partial \mathbb{E}\left[\tilde{\boldsymbol{p}}_{k,l}\right]}{\partial \Sigma_{\epsilon_i,j}^{-1}} = \frac{(1+\frac{1}{\phi}\sum_{i'=1}^{n_F} \hat{\boldsymbol{H}}_{i',l})\frac{1}{\phi}\sum_{i'=1}^{n_F} \frac{\partial \hat{\boldsymbol{H}}_{i',l}}{\partial \Sigma_{\epsilon_i,j}^{-1}}\tilde{\boldsymbol{c}}_{i'l} - (\bar{p}_l + \frac{1}{\phi}\sum_{i'=1}^{n_F} \hat{\boldsymbol{H}}_{i',l}\tilde{\boldsymbol{c}}_{il})\frac{1}{\phi}\sum_{i'=1}^{n_F} \frac{\partial \hat{\boldsymbol{H}}_{i',l}}{\partial \Sigma_{\epsilon_i,j}^{-1}}}{(1+\frac{1}{\phi}\sum_{i'=1}^{n_F} \hat{\boldsymbol{H}}_{i',l})^2} \tag{97}$$

$$= \delta_{jl}\frac{(1+\frac{1}{\phi}\sum_{i'=1}^{n_F} \hat{\boldsymbol{H}}_{i',l})\frac{1}{\phi}\hat{\boldsymbol{H}}_{i,l}^2\rho_i(\Sigma_{b_i,l}^{-1}+\Sigma_{\epsilon_i,l}^{-1})^{-2}\tilde{\boldsymbol{c}}_{il} - (\bar{p}_l + \frac{1}{\phi}\sum_{i'=1}^{n_F} \hat{\boldsymbol{H}}_{i',l}\tilde{\boldsymbol{c}}_{i'l})\frac{1}{\phi}\hat{\boldsymbol{H}}_{i,l}^2\rho_i(\Sigma_{b_i,l}^{-1}+\Sigma_{\epsilon_i,l}^{-1})^{-2}}{(1+\frac{1}{\phi}\sum_{i'=1}^{n_F} \hat{\boldsymbol{H}}_{i',l})^2} \tag{98}$$

$$= \delta_{jl}\frac{1}{\phi}\hat{\boldsymbol{H}}_{i,l}^2\rho_i(\Sigma_{b_i,l}^{-1}+\Sigma_{\epsilon_i,l}^{-1})^{-2}\frac{(1+\frac{1}{\phi}\sum_{i'=1}^{n_F} \hat{\boldsymbol{H}}_{i',l})\tilde{\boldsymbol{c}}_{il} - (\bar{p}_l + \frac{1}{\phi}\sum_{i'=1}^{n_F} \hat{\boldsymbol{H}}_{i',l}\tilde{\boldsymbol{c}}_{i'l})}{(1+\frac{1}{\phi}\sum_{i'=1}^{n_F} \hat{\boldsymbol{H}}_{i',l})^2} \tag{99}$$

$$= \delta_{jl}\frac{1}{\phi}\hat{\boldsymbol{H}}_{i,l}^2\rho_i(\Sigma_{b_i,l}^{-1}+\Sigma_{\epsilon_i,l}^{-1})^{-2}\frac{(1+\frac{1}{\phi}\sum_{i'=1}^{n_F} \hat{\boldsymbol{H}}_{i',l})\tilde{\boldsymbol{c}}_{il} - (\bar{p}_l + \frac{1}{\phi}\sum_{i'=1}^{n_F} \hat{\boldsymbol{H}}_{i',l}\tilde{\boldsymbol{c}}_{i'l})}{(1+\frac{1}{\phi}\sum_{i'=1}^{n_F} \hat{\boldsymbol{H}}_{i',l})^2} \tag{100}$$

$$= \delta_{jl}\frac{\hat{\boldsymbol{H}}_{i,l}^2\rho_i(\Sigma_{b_i,l}^{-1}+\Sigma_{\epsilon_i,l}^{-1})^{-2}}{\phi+\sum_{i'=1}^{n_F} \hat{\boldsymbol{H}}_{i',l}}(\tilde{\boldsymbol{c}}_{i,l} - \boldsymbol{D}_l) \tag{101}$$

Similarly, for the expected quantity produced $\mathbb{E}\left[\tilde{\boldsymbol{q}}_{k,l}\right] = \hat{\boldsymbol{H}}_{k,l}(\boldsymbol{D}_l - \tilde{\boldsymbol{c}}_{k,l})$:

$$\frac{\partial \mathbb{E}\left[\tilde{\boldsymbol{q}}_{k,l}\right]}{\partial \Sigma_{\epsilon_i,j}^{-1}} = \frac{\partial \hat{\boldsymbol{H}}_{k,l}}{\partial \Sigma_{\epsilon_i,j}^{-1}}(\boldsymbol{D}_l - \tilde{\boldsymbol{c}}_{k,l}) + \hat{\boldsymbol{H}}_{k,l}\frac{\partial \mathbb{E}\left[\tilde{\boldsymbol{p}}_{k,l}\right]}{\partial \Sigma_{\epsilon_i,j}^{-1}} \tag{102}$$

$$= \delta_{lj}\delta_{ki}\hat{\boldsymbol{H}}_{k,l}^2\rho_k(\Sigma_{b_k,l}^{-1}+\Sigma_{\epsilon_k,l}^{-1})^{-2}(\boldsymbol{D}_l - \tilde{\boldsymbol{c}}_{k,l}) + \delta_{jl}\hat{\boldsymbol{H}}_{k,l}\frac{\hat{\boldsymbol{H}}_{i,l}^2\rho_i(\Sigma_{b_i,l}^{-1}+\Sigma_{\epsilon_i,l}^{-1})^{-2}}{\phi+\sum_{i'=1}^{n_F} \hat{\boldsymbol{H}}_{i',l}}(\tilde{\boldsymbol{c}}_{i,l} - \boldsymbol{D}_l) \tag{103}$$

$$= \delta_{jl}\hat{\boldsymbol{H}}_{i,l}^2\rho_i(\Sigma_{b_i,l}^{-1}+\Sigma_{\epsilon_i,l}^{-1})^{-2}(\boldsymbol{D}_l - \tilde{\boldsymbol{c}}_{i,l})\left(\delta_{ki} - \frac{\hat{\boldsymbol{H}}_{k,l}}{\phi+\sum_{i'=1}^{n_F} \hat{\boldsymbol{H}}_{i',l}}\right) \tag{104}$$

Thus the derivative of numerator is

$$\frac{\partial \mathbf{E}\tilde{\boldsymbol{q}}_{i,l}\mathbf{E}\tilde{\boldsymbol{p}}_{i,l}}{\partial \Sigma_{\epsilon_i,j}^{-1}} = \delta_{jl}\rho_i\hat{\boldsymbol{H}}_{i,j}^2\Sigma_{b_i,j}^2\left(\boldsymbol{D}_j - \tilde{\boldsymbol{c}}_{i,j}\right)\left(\frac{1-\frac{1}{\phi}\hat{\boldsymbol{H}}_{i,j}+\frac{1}{\phi}\sum_{s=1,s\neq i}^{n_F}\hat{\boldsymbol{H}}_{s,j}}{1+\frac{1}{\phi}\sum_{s=1}^{n_F}\hat{\boldsymbol{H}}_{s,j}}\boldsymbol{D}_j + \frac{\frac{1}{\phi}\hat{\boldsymbol{H}}_{i,j}\tilde{\boldsymbol{c}}_{i,j}}{1+\frac{1}{\phi}\sum_{s=1}^{n_F}\hat{\boldsymbol{H}}_{s,j}}\right) \geq 0$$

$$\frac{\partial \mathbf{Cov}_{i,l}}{\partial \Sigma_{\epsilon_i,j}^{-1}} = \delta_{jl}\hat{\boldsymbol{H}}_{i,j}\Sigma_{b_i,j}^2\left(\frac{\rho_i\left(1-\frac{1}{\phi}\hat{\boldsymbol{H}}_{i,j}+\frac{1}{\phi}\sum_{s=1,s\neq i}^{n_F}\hat{\boldsymbol{H}}_{s,j}\right)}{\left(1+\sum_{s=1}^{n_F}\frac{1}{\phi}\hat{\boldsymbol{H}}_{s,j}\right)}\mathbf{Cov}_{i,j} + \frac{\left(1+\sum_{s=1,s\neq i}^{n_F}\frac{1}{\phi}\hat{\boldsymbol{H}}_{s,j}\right)^2}{\left(1+\sum_{s=1}^{n_F}\frac{1}{\phi}\hat{\boldsymbol{H}}_{s,j}\right)^2}\right) \geq 0 \tag{105}$$

Since the covariance term is the difference between the firm markup and the average product markup, this proves that that difference, the firm-level markup wedge in increasing in the firm's data. Moreover, firm-level markup increases with more data with small price of risk ρ_i since

$$\frac{\partial M_i^f}{\partial \Sigma_{\epsilon_i,j}^{-1}} = \frac{\left(\frac{\partial \mathbf{E}\tilde{\boldsymbol{q}}_{i,j}\mathbf{E}\tilde{\boldsymbol{p}}_{i,j}}{\partial \Sigma_{\epsilon_i,j}^{-1}} + \frac{\partial \mathbf{Cov}_{i,j}}{\partial \Sigma_{\epsilon_i,j}^{-1}}\right)}{\sum_{l=1}^N \mathbf{E}\tilde{\boldsymbol{q}}_{i,l}\tilde{\boldsymbol{c}}_{i,l}} - \frac{\left(\sum_{l=1}^N \mathbf{E}[\tilde{\boldsymbol{q}}_{i,l}\tilde{\boldsymbol{p}}_{i,l}]\right)\frac{\partial \mathbf{E}\tilde{\boldsymbol{q}}_{i,j}}{\partial \Sigma_{\epsilon_i,j}^{-1}}\tilde{\boldsymbol{c}}_{i,j}}{\left(\sum_{l=1}^N \mathbf{E}\tilde{\boldsymbol{q}}_{i,l}\tilde{\boldsymbol{c}}_{i,l}\right)^2} \quad \text{and} \lim_{\rho_i \to 0} \frac{\partial M_i^f}{\partial \Sigma_{\epsilon_i,j}^{-1}} = \frac{\hat{\boldsymbol{H}}_{i,j}\Sigma_{b_i,j}^2 \frac{\left(1+\sum_{s=1,s\neq i}^{n_F}\frac{1}{\phi}\hat{\boldsymbol{H}}_{s,j}\right)^2}{\left(1+\sum_{s=1}^{n_F}\frac{1}{\phi}\hat{\boldsymbol{H}}_{s,j}\right)^2}}{\sum_{l=1}^N \mathbf{E}\tilde{\boldsymbol{q}}_{i,l}\tilde{\boldsymbol{c}}_{i,l}} > 0 \tag{106}$$

We prove the negative first order derivative for fixed choices of cost $\tilde{\boldsymbol{c}}_i$, which corresponds to infinite 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, by continuity, the inequality will still hold. □

Proof of Proposition 5a: 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

$$w_i^{cost} = \frac{\mathbf{E}\left[\tilde{\boldsymbol{q}}_i'\tilde{\boldsymbol{c}}_i\right]}{\sum_{k=1}^{n_F}\mathbf{E}\left[\tilde{\boldsymbol{q}}_k'\tilde{\boldsymbol{c}}_k\right]} = \frac{\sum_{l=1}^N \mathbf{E}[\tilde{\boldsymbol{q}}_{i,l}]\tilde{\boldsymbol{c}}_{i,l}}{\sum_{k=1}^{n_F}\sum_{l=1}^N \mathbf{E}[\tilde{\boldsymbol{q}}_{k,l}]\tilde{\boldsymbol{c}}_{k,l}} \tag{107}$$

This weight is increasing in data for the firm i since

$$\begin{aligned}\frac{\partial w_i^{cost}}{\partial \Sigma_{\epsilon_i,j}^{-1}} &= \frac{\frac{\partial \mathbf{E}[\tilde{\boldsymbol{q}}_{i,j}]}{\partial \Sigma_{\epsilon_i,j}^{-1}}\tilde{\boldsymbol{c}}_{i,j}\left(\sum_{k=1,k\neq i}^{n_F}\mathbf{E}\left[\tilde{\boldsymbol{q}}_k'\tilde{\boldsymbol{c}}_k\right]\right) - \mathbf{E}\left[\tilde{\boldsymbol{q}}_i'\tilde{\boldsymbol{c}}_i\right]\sum_{k=1,k\neq i}^{n_F}\tilde{\boldsymbol{c}}_{k,j}\frac{\partial \mathbf{E}[\tilde{\boldsymbol{q}}_{k,j}]}{\partial \Sigma_{\epsilon_i,j}^{-1}}}{\left(\sum_{k=1}^{n_F}\mathbf{E}\left[\tilde{\boldsymbol{q}}_k'\tilde{\boldsymbol{c}}_k\right]\right)^2} \\ &= \rho_i \hat{\boldsymbol{H}}_{i,j}^2\Sigma_{b_i,j}^2\left(\boldsymbol{D}_j - \tilde{\boldsymbol{c}}_{i,j}\right)\left[\frac{\tilde{\boldsymbol{c}}_{i,j}\left(\sum_{k=1,k\neq i}^{n_F}\mathbf{E}\left[\tilde{\boldsymbol{q}}_k'\tilde{\boldsymbol{c}}_k\right]\right)}{\left(\sum_{k=1}^{n_F}\mathbf{E}\left[\tilde{\boldsymbol{q}}_k'\tilde{\boldsymbol{c}}_k\right]\right)^2}\frac{1+\sum_{s\neq i,s=1}^{n_F}\frac{\hat{\boldsymbol{H}}_{s,j}}{\phi}}{1+\sum_{s=1}^{n_F}\frac{\hat{\boldsymbol{H}}_{s,j}}{\phi}} + \frac{\mathbf{E}\left[\tilde{\boldsymbol{q}}_i'\tilde{\boldsymbol{c}}_i\right]\sum_{k=1,k\neq i}^{n_F}\frac{\frac{1}{\phi}\hat{\boldsymbol{H}}_{k,j}\tilde{\boldsymbol{c}}_{k,j}}{1+\frac{1}{\phi}\sum_{s=1}^{n_F}\hat{\boldsymbol{H}}_{s,j}}}{\left(\sum_{k=1}^{n_F}\mathbf{E}\left[\tilde{\boldsymbol{q}}_k'\tilde{\boldsymbol{c}}_k\right]\right)^2}\right] \geq 0\end{aligned} \tag{108}$$

This inequality indicates that high-data firms produce more on average and have larger impacts on cost-weighted industry markup. Furthermore, firm-level markup increases in data if cost is small enough and $N > 1$ since

$$\frac{\partial M_i^f}{\partial \Sigma_{\epsilon_i,j}^{-1}} = \frac{\left(\frac{\partial \mathbb{E}\tilde{\boldsymbol{q}}_{i,j}\mathbb{E}\tilde{\boldsymbol{p}}_{i,j}}{\partial \Sigma_{\epsilon_i,j}^{-1}} + \frac{\partial \mathbf{Cov}_{i,j}}{\partial \Sigma_{\epsilon_i,j}^{-1}}\right)}{\sum_{l=1}^N \mathbf{E}\tilde{\boldsymbol{q}}_{i,l}\tilde{\boldsymbol{c}}_{i,l}} - \frac{\left(\sum_{l=1}^N \mathbb{E}[\tilde{\boldsymbol{q}}_{i,l}\tilde{\boldsymbol{p}}_{i,l}]\right)\frac{\partial \mathbb{E}\tilde{\boldsymbol{q}}_{i,j}}{\partial \Sigma_{\epsilon_i,j}^{-1}}\tilde{\boldsymbol{c}}_{i,j}}{\left(\sum_{l=1}^N \mathbb{E}\tilde{\boldsymbol{q}}_{i,l}\tilde{\boldsymbol{c}}_{i,l}\right)^2} \Rightarrow \lim_{\tilde{\boldsymbol{c}}_{i,j}\to 0}\frac{\partial M_i^f}{\partial \Sigma_{\epsilon_i,j}^{-1}} = \frac{\left(\frac{\partial \mathbb{E}\tilde{\boldsymbol{q}}_{i,j}\mathbb{E}\tilde{\boldsymbol{p}}_{i,j}}{\partial \Sigma_{\epsilon_i,j}^{-1}} + \frac{\partial \mathbf{Cov}_{i,j}}{\partial \Sigma_{\epsilon_i,j}^{-1}}\right)}{\sum_{l\neq j}^N \mathbb{E}\tilde{\boldsymbol{q}}_{i,l}\tilde{\boldsymbol{c}}_{i,l}} > 0 \tag{109}$$

These two forces intensify each other and drive up the industry-level markups compared to unweighted firm-level markups, leading to increasing wedge between these two markups. This proof holds for fixed choices of cost $\tilde{\boldsymbol{c}}_i$, which corresponds to infinite high marginal cost $\chi_c \to \infty$. The inequality still holds for large enough χ_c by using continuity. □

Proof of Proposition 5b: Sales weighted vs cost-weighted markup Notice that the wedge between the sales- and cost-weighted markups is

$$M^{m,sales} - M^{m} = \sum_{i=1}^{N} \left(\underbrace{\frac{\mathbb{E}\left[\tilde{\boldsymbol{q}}_i' \tilde{\boldsymbol{p}}_i\right]}{\sum_{i=1}^{N} \mathbb{E}\left[\tilde{\boldsymbol{q}}_i' \tilde{\boldsymbol{p}}_i\right]}}_{w_i^{sales}} - \underbrace{\frac{\mathbb{E}\left[\tilde{\boldsymbol{q}}_i' \tilde{\boldsymbol{c}}_i\right]}{\sum_{i=1}^{N} \mathbb{E}\left[\tilde{\boldsymbol{q}}_i' \tilde{\boldsymbol{c}}_i\right]}}_{w_i^{m}} \right) \frac{\mathbb{E}\left[\tilde{\boldsymbol{q}}_i' \tilde{\boldsymbol{p}}_i\right]}{\mathbb{E}\left[\tilde{\boldsymbol{q}}_i' \tilde{\boldsymbol{c}}_i\right]} \tag{110}$$

When firms are ex ante identical, this wedge is zero $M^{m,sales} - M^{m}$.

To see how data $\Sigma_{\epsilon_i,j}^{-1}$ affects the wedge, let's first take a loot at how it affects the difference between the sales weight and the cost weight of and firm k:

$$\begin{aligned}
\frac{\partial}{\partial \Sigma_{\epsilon_i,j}^{-1}} (w_k^{sales} - w_k^{m}) = & \; w_k^{sales} \left(\frac{1}{\mathbb{E}\left[\tilde{\boldsymbol{q}}_k' \tilde{\boldsymbol{p}}_k\right]} \frac{\partial \mathbb{E}\left[\tilde{\boldsymbol{q}}_k' \tilde{\boldsymbol{p}}_k\right]}{\partial \Sigma_{\epsilon_i,j}^{-1}} - \frac{1}{\sum_{k'} \mathbb{E}\left[\tilde{\boldsymbol{q}}_{k'}' \tilde{\boldsymbol{p}}_{k'}\right]} \sum_{k'} \frac{\partial \mathbb{E}\left[\tilde{\boldsymbol{q}}_{k'}' \tilde{\boldsymbol{p}}_{k'}\right]}{\partial \Sigma_{\epsilon_i,j}^{-1}} \right) && (111)\\
& - w_k^{m} \left(\frac{1}{\mathbb{E}\left[\tilde{\boldsymbol{q}}_k' \tilde{\boldsymbol{c}}_k\right]} \frac{\partial \mathbb{E}\left[\tilde{\boldsymbol{q}}_k' \tilde{\boldsymbol{c}}_k\right]}{\partial \Sigma_{\epsilon_i,j}^{-1}} - \frac{1}{\sum_{k'} \mathbb{E}\left[\tilde{\boldsymbol{q}}_{k'}' \tilde{\boldsymbol{c}}_{k'}\right]} \sum_{k'} \frac{\partial \mathbb{E}\left[\tilde{\boldsymbol{q}}_{k'}' \tilde{\boldsymbol{c}}_{k'}\right]}{\partial \Sigma_{\epsilon_i,j}^{-1}} \right) && (112)\\
= & \; w_k^{sales} \left(\frac{1 - w_k^{sales}}{\mathbb{E}\left[\tilde{\boldsymbol{q}}_k' \tilde{\boldsymbol{p}}_k\right]} \frac{\partial \mathbb{E}\left[\tilde{\boldsymbol{q}}_k' \tilde{\boldsymbol{p}}_k\right]}{\partial \Sigma_{\epsilon_i,j}^{-1}} - \frac{1}{\sum_{k'} \mathbb{E}\left[\tilde{\boldsymbol{q}}_{k'}' \tilde{\boldsymbol{p}}_{k'}\right]} \sum_{k' \neq k} \frac{\partial \mathbb{E}\left[\tilde{\boldsymbol{q}}_{k'}' \tilde{\boldsymbol{p}}_{k'}\right]}{\partial \Sigma_{\epsilon_i,j}^{-1}} \right) && (113)\\
& - w_k^{m} \left(\frac{1 - w_k^{m}}{\mathbb{E}\left[\tilde{\boldsymbol{q}}_k' \tilde{\boldsymbol{c}}_k\right]} \frac{\partial \mathbb{E}\left[\tilde{\boldsymbol{q}}_k' \tilde{\boldsymbol{c}}_k\right]}{\partial \Sigma_{\epsilon_i,j}^{-1}} - \frac{1}{\sum_{k'} \mathbb{E}\left[\tilde{\boldsymbol{q}}_{k'}' \tilde{\boldsymbol{c}}_{k'}\right]} \sum_{k' \neq k} \frac{\partial \mathbb{E}\left[\tilde{\boldsymbol{q}}_{k'}' \tilde{\boldsymbol{c}}_{k'}\right]}{\partial \Sigma_{\epsilon_i,j}^{-1}} \right) && (114)\\
& && (115)
\end{aligned}$$

Using the assumptions that firms are ex ante identical, we have $w_k^{sales} = w_k^{m} = \frac{1}{n_F}$ and $\mathbb{E}\left[\tilde{\boldsymbol{q}}_k' \tilde{\boldsymbol{p}}_k\right] = \mathbb{E}\left[\tilde{\boldsymbol{q}}_i' \tilde{\boldsymbol{p}}_i\right], \forall k, i$ and the effect of information on the weights can be simplified to

$$\begin{aligned}
\frac{\partial}{\partial \Sigma_{\epsilon_i,j}^{-1}} (w_k^{sales} - w_k^{m}) = & \; \frac{1}{n_F} \left(\frac{1}{\mathbb{E}\left[\tilde{\boldsymbol{q}}_k' \tilde{\boldsymbol{p}}_k\right]} \frac{\partial \mathbb{E}\left[\tilde{\boldsymbol{q}}_k' \tilde{\boldsymbol{p}}_k\right]}{\partial \Sigma_{\epsilon_i,j}^{-1}} - \frac{n_F - 1}{n_F \mathbb{E}\left[\tilde{\boldsymbol{q}}_{k'}' \tilde{\boldsymbol{p}}_{k'}\right]} \frac{\partial \mathbb{E}\left[\tilde{\boldsymbol{q}}_{k'}' \tilde{\boldsymbol{p}}_{k'}\right]}{\partial \Sigma_{\epsilon_i,j}^{-1}} - \frac{1}{n_F \mathbb{E}\left[\tilde{\boldsymbol{q}}_{k'}' \tilde{\boldsymbol{p}}_{k'}\right]} \frac{\partial \mathbb{E}\left[\tilde{\boldsymbol{q}}_i' \tilde{\boldsymbol{p}}_i\right]}{\partial \Sigma_{\epsilon_i,j}^{-1}} \right) && (116)\\
& - \frac{1}{n_F} \left(\frac{1}{\mathbb{E}\left[\tilde{\boldsymbol{q}}_k' \tilde{\boldsymbol{c}}_k\right]} \frac{\partial \mathbb{E}\left[\tilde{\boldsymbol{q}}_k' \tilde{\boldsymbol{c}}_k\right]}{\partial \Sigma_{\epsilon_i,j}^{-1}} - \frac{n_F - 1}{n_F \mathbb{E}\left[\tilde{\boldsymbol{q}}_{k'}' \tilde{\boldsymbol{c}}_{k'}\right]} \frac{\partial \mathbb{E}\left[\tilde{\boldsymbol{q}}_{k'}' \tilde{\boldsymbol{c}}_{k'}\right]}{\partial \Sigma_{\epsilon_i,j}^{-1}} - \frac{1}{n_F \mathbb{E}\left[\tilde{\boldsymbol{q}}_{k'}' \tilde{\boldsymbol{c}}_{k'}\right]} \frac{\partial \mathbb{E}\left[\tilde{\boldsymbol{q}}_i' \tilde{\boldsymbol{c}}_i\right]}{\partial \Sigma_{\epsilon_i,j}^{-1}} \right) && (117)\\
= & \; \frac{1}{n_F^2 \mathbb{E}\left[\tilde{\boldsymbol{q}}_k' \tilde{\boldsymbol{p}}_k\right]} \left(\frac{\partial \mathbb{E}\left[\tilde{\boldsymbol{q}}_k' \tilde{\boldsymbol{p}}_k\right]}{\partial \Sigma_{\epsilon_i,j}^{-1}} - \frac{\partial \mathbb{E}\left[\tilde{\boldsymbol{q}}_i' \tilde{\boldsymbol{p}}_i\right]}{\partial \Sigma_{\epsilon_i,j}^{-1}} \right) && (118)\\
& - \frac{1}{n_F^2 \mathbb{E}\left[\tilde{\boldsymbol{q}}_k' \tilde{\boldsymbol{c}}_k\right]} \left(\frac{\partial \mathbb{E}\left[\tilde{\boldsymbol{q}}_k' \tilde{\boldsymbol{c}}_k\right]}{\partial \Sigma_{\epsilon_i,j}^{-1}} - \frac{\partial \mathbb{E}\left[\tilde{\boldsymbol{q}}_i' \tilde{\boldsymbol{c}}_i\right]}{\partial \Sigma_{\epsilon_i,j}^{-1}} \right), \forall k \neq i && (119)\\
& && (120)
\end{aligned}$$

and similarly for firm i itself

$$\frac{\partial}{\partial \Sigma_{\epsilon_i,j}^{-1}}(w_i^{sales} - w_i^m) = \frac{n_F - 1}{n_F^2 \mathbb{E}\left[\tilde{\boldsymbol{q}}_k' \tilde{\boldsymbol{p}}_k\right]} \left(\frac{\partial \mathbb{E}\left[\tilde{\boldsymbol{q}}_i' \tilde{\boldsymbol{p}}_i\right]}{\partial \Sigma_{\epsilon_i,j}^{-1}} - \frac{\partial \mathbb{E}\left[\tilde{\boldsymbol{q}}_k' \tilde{\boldsymbol{p}}_k\right]}{\partial \Sigma_{\epsilon_i,j}^{-1}} \right) \tag{121}$$

$$- \frac{n_F - 1}{n_F^2 \mathbb{E}\left[\tilde{\boldsymbol{q}}_k' \tilde{\boldsymbol{c}}_k\right]} \left(\frac{\partial \mathbb{E}\left[\tilde{\boldsymbol{q}}_i' \tilde{\boldsymbol{c}}_i\right]}{\partial \Sigma_{\epsilon_i,j}^{-1}} - \frac{\partial \mathbb{E}\left[\tilde{\boldsymbol{q}}_k' \tilde{\boldsymbol{c}}_k\right]}{\partial \Sigma_{\epsilon_i,j}^{-1}} \right) \tag{122}$$

where k is any firm different from i.

Notice that the condition that the wedge in the weights widens amount to showing that the elasticity in the different of sales in higher than that of the cost

$$\frac{1}{\mathbb{E}\left[\tilde{\boldsymbol{q}}_k' \tilde{\boldsymbol{p}}_k\right]} \left(\frac{\partial \mathbb{E}\left[\tilde{\boldsymbol{q}}_i' \tilde{\boldsymbol{p}}_i\right]}{\partial \Sigma_{\epsilon_i,j}^{-1}} - \frac{\partial \mathbb{E}\left[\tilde{\boldsymbol{q}}_k' \tilde{\boldsymbol{p}}_k\right]}{\partial \Sigma_{\epsilon_i,j}^{-1}} \right) \geq \frac{1}{\mathbb{E}\left[\tilde{\boldsymbol{q}}_k' \tilde{\boldsymbol{c}}_k\right]} \left(\frac{\partial \mathbb{E}\left[\tilde{\boldsymbol{q}}_i' \tilde{\boldsymbol{c}}_i\right]}{\partial \Sigma_{\epsilon_i,j}^{-1}} - \frac{\partial \mathbb{E}\left[\tilde{\boldsymbol{q}}_k' \tilde{\boldsymbol{c}}_k\right]}{\partial \Sigma_{\epsilon_i,j}^{-1}} \right) \tag{123}$$

which is equivalent to the information

$$\frac{\partial}{\partial \Sigma_{\epsilon_i,j}^{-1}} \frac{\mathbb{E}\left[\tilde{\boldsymbol{q}}_i' \tilde{\boldsymbol{p}}_i\right]}{\mathbb{E}\left[\tilde{\boldsymbol{q}}_i' \tilde{\boldsymbol{c}}_i\right]} \geq \frac{\partial}{\partial \Sigma_{\epsilon_i,j}^{-1}} \frac{\mathbb{E}\left[\tilde{\boldsymbol{q}}_k' \tilde{\boldsymbol{p}}_k\right]}{\mathbb{E}\left[\tilde{\boldsymbol{q}}_k' \tilde{\boldsymbol{c}}_k\right]} \tag{124}$$

This means if difference in weights $w_i^{m,sales} - w_i^m$ turns positive whenever it increases the markup of firm i relatively more than other firms. Therefore, the wedge between the sales weighted markup and the cost-weighted markup is always weakly increasing in $\Sigma_{\epsilon_i,j}^{-1}$. And it is strict if the information of firm i affects the markup of firm i differently from that of firm k, which is generically true as the information firm i to concentrate more on high-markup products while the opposite for the other firms.

Indeed, this result reflect the fact that the wedge between the sales-weighted markup and the cost-weighted markup is always non-negative and it is zero if and only if all firms are symmetric. To see this point, notice we can write the wedge as

$$M^{m,sales} - M^m = \frac{\sum_{i=1}^N \mathbb{E}\left[\tilde{\boldsymbol{q}}_i' \tilde{\boldsymbol{c}}_i\right] \sum_{i=1}^N \frac{\mathbb{E}\left[\tilde{\boldsymbol{q}}_i' \tilde{\boldsymbol{p}}_i\right]^2}{\mathbb{E}\left[\tilde{\boldsymbol{q}}_i' \tilde{\boldsymbol{c}}_i\right]} - \left(\sum_{i=1}^N \mathbb{E}\left[\tilde{\boldsymbol{q}}_i' \tilde{\boldsymbol{p}}_i\right] \right)^2}{\sum_{i=1}^N \mathbb{E}\left[\tilde{\boldsymbol{q}}_i' \tilde{\boldsymbol{p}}_i\right] \sum_{i=1}^N \mathbb{E}\left[\tilde{\boldsymbol{q}}_i' \tilde{\boldsymbol{c}}_i\right]} \tag{125}$$

Recall Cauchy-Schwarz inequality $\left(\sum_{i=1}^N u_i v_i\right)^2 \leq \left(\sum_{i=1}^N u_i^2\right)\left(\sum_{i=1}^N v_i^2\right)$. Let $u_i = \sqrt{\mathbb{E}\left[\tilde{\boldsymbol{q}}_i' \tilde{\boldsymbol{c}}_i\right]}$ $v_i = \sqrt{\frac{\mathbb{E}\left[\tilde{\boldsymbol{q}}_i' \tilde{\boldsymbol{p}}_i\right]^2}{\mathbb{E}\left[\tilde{\boldsymbol{q}}_i' \tilde{\boldsymbol{c}}_i\right]}}$. The Cauchy-Schwarz inequality says [9]

$$\sum_{i=1}^N \mathbb{E}\left[\tilde{\boldsymbol{q}}_i' \tilde{\boldsymbol{c}}_i\right] \sum_{i=1}^N \frac{\mathbb{E}\left[\tilde{\boldsymbol{q}}_i' \tilde{\boldsymbol{p}}_i\right]^2}{\mathbb{E}\left[\tilde{\boldsymbol{q}}_i' \tilde{\boldsymbol{c}}_i\right]} - \left(\sum_{i=1}^N \mathbb{E}\left[\tilde{\boldsymbol{q}}_i' \tilde{\boldsymbol{p}}_i\right] \right)^2 \geq 0 \tag{126}$$

and the equality holds if and only if all firms have the same markup.

Intuitively, a high-markup firm has higher sales relative to it's costs so, it has a higher sales-weight than its cost weights. Similarly, low-markup firm tends to have lower sales-weight than cost-weight but is out-weighted by the high-markup firms. The wedge achieves the minimum 0

[9] This special case is also referred to as Sedrakyan's inequality, Bergström's inequality, Engel's form, the T2 lemma, or Titu's lemma.

when all firms are symmetric and it gets larger as the information brings more asymmetry in the production.

Proof of Proposition 5c: Sales-weighted vs. industry aggregates markup The reason this corollary follows directly from Proposition 5b, 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:

$$M^{ag} := \frac{\mathbf{E}\left[\sum_{i=1}^{N} \boldsymbol{q}_i' \boldsymbol{p}_i\right]}{\mathbf{E}\left[\sum_{i=1}^{N} \boldsymbol{q}_i' \boldsymbol{c}_i\right]} = \frac{\sum_{i=1}^{N} \mathbf{E}\left[\boldsymbol{q}_i' \boldsymbol{p}_i\right]}{\sum_{i=1}^{N} \mathbf{E}\left[\boldsymbol{q}_i' \boldsymbol{c}_i\right]} = \sum_{i=1}^{N} w_i^m M_i^f = M^m \quad \text{where} \quad w_i^c = \frac{\mathbf{E}\left[\boldsymbol{q}_i' \boldsymbol{c}_i\right]}{\sum_{i=1}^{N} \mathbf{E}\left[\boldsymbol{q}_i' \boldsymbol{c}_i\right]}. \tag{127}$$

B.1. Cyclical Markups

Proof of proposition 6 Part a: product markups are increasing in demand variance and converge to a constant.

Proof. According to the definition of $\hat{\boldsymbol{H}}_i$, we have

$$\begin{aligned} &\hat{\boldsymbol{H}}_i = \left(\frac{\boldsymbol{I}_N}{\phi} + \rho_i \mathbf{Var}(\tilde{\boldsymbol{p}}_i|\mathcal{I}_i)\right)^{-1} \quad \text{and} \quad \mathbf{Var}(\tilde{\boldsymbol{p}}_i|\mathcal{I}_i) = \left(\Sigma_b^{-1} + \Sigma_{\epsilon_i}^{-1}\right)^{-1} \\ &\Rightarrow \lim_{\Sigma_b \to \infty} \mathbf{Var}(\tilde{\boldsymbol{p}}_i|\mathcal{I}_i) = \Sigma_{\epsilon_i}, \quad \tilde{\hat{\boldsymbol{H}}}_i := \lim_{\Sigma_b \to \infty} \hat{\boldsymbol{H}}_i = \left(\frac{\boldsymbol{I}_N}{\phi} + \rho_i \Sigma_{\epsilon_i}\right)^{-1} \end{aligned} \tag{128}$$

The equilibrium price is given by

$$\mathbf{E}\left[\tilde{\boldsymbol{p}}_i\right] = \boldsymbol{D} = \left(\boldsymbol{I}_N + \frac{1}{\phi}\sum_{i=1}^{n_F} \hat{\boldsymbol{H}}_i\right)^{-1} \left(\bar{\boldsymbol{p}} + \frac{1}{\phi}\sum_{i=1}^{n_F} \hat{\boldsymbol{H}}_i \boldsymbol{c}_i\right) \tag{129}$$

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

$$\begin{aligned} \tilde{\boldsymbol{p}} := \lim_{\Sigma_b \to \infty} \mathbf{E}\left[\tilde{\boldsymbol{p}}_i\right] &= \left(\boldsymbol{I}_N + \frac{1}{\phi}\sum_{i=1}^{n_F} \lim_{\Sigma_b \to \infty} \hat{\boldsymbol{H}}_i\right)^{-1} \left(\bar{\boldsymbol{p}} + \frac{1}{\phi}\sum_{i=1}^{n_F} \lim_{\Sigma_b \to \infty} \hat{\boldsymbol{H}}_i \boldsymbol{c}_i\right) \\ &= \left[\boldsymbol{I}_N + \sum_{i=1}^{n_F} \left(\boldsymbol{I}_N + \phi\rho_i \Sigma_{\epsilon_i}\right)^{-1}\right]^{-1} \left[\bar{\boldsymbol{p}} + \sum_{i=1}^{n_F} \boldsymbol{c}_i \left(\boldsymbol{I}_N + \phi\rho_i \Sigma_{\epsilon_i}\right)^{-1}\right] \end{aligned} \tag{130}$$

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.

$$\boldsymbol{q}_i = \boldsymbol{A}\tilde{\boldsymbol{q}}_i \quad \text{and} \quad \boldsymbol{p}_i = \boldsymbol{A}\tilde{\boldsymbol{p}}_i \Rightarrow \overline{M}^p = \frac{1}{N}\frac{1}{n_F}\sum_{i=1}^{n_F}\sum_{j=1}^{N} \frac{\left(\boldsymbol{A}\mathbf{E}\left[\tilde{\boldsymbol{p}}_i\right]\right)_j}{\left(\boldsymbol{A}\boldsymbol{c}_i\right)_j} \quad \text{converges.} \tag{131}$$

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

j is

$$\frac{\partial \mathbf{E}[\tilde{\boldsymbol{p}}_{i,j}]}{\partial \Sigma_{b,j}} = \frac{(\bar{\boldsymbol{c}}_j - \bar{\boldsymbol{p}}_j)\frac{1}{\phi}\sum_{i=1}^{n_F}\frac{\partial \hat{\boldsymbol{H}}_{i,j}}{\partial \Sigma_{b,j}}}{\left(1+\frac{1}{\phi}\sum_{i=1}^{n_F}\hat{\boldsymbol{H}}_{i,j}\right)^2} \quad \text{and} \quad \frac{\partial \hat{\boldsymbol{H}}_{i,j}}{\partial \Sigma_{b,j}} = -\frac{\hat{\boldsymbol{H}}_{i,j}^2 \rho_i \Sigma_{\epsilon_i,j}^2}{\left(\Sigma_{b.j}+\Sigma_{\epsilon_i,j}\right)^2} \leq 0 \tag{132}$$

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. First, We will show that the trace of the covariance $\mathbf{tr}[\mathbf{Cov}\,(\tilde{\boldsymbol{p}}_i, \tilde{\boldsymbol{q}}_i)]$ is always positive.

$$\begin{aligned}\mathbf{Cov}\,(\tilde{\boldsymbol{p}}_i, \tilde{\boldsymbol{q}}_i) = &\left(\boldsymbol{I}_N + \sum_{j=1}^{n_F}\frac{\hat{\boldsymbol{H}}_j}{\phi}\right)^{-1}\sum_{j=1}^{n_F}\hat{\boldsymbol{H}}_j\mathbf{Var}(\boldsymbol{K}_j\boldsymbol{s}_j)\hat{\boldsymbol{H}}_j\left(\boldsymbol{I}_N + \sum_{j=1}^{n_F}\frac{\hat{\boldsymbol{H}}_j}{\phi}\right)^{-1}\frac{\hat{\boldsymbol{H}}_i}{\phi^2} + \mathbf{Var}(\boldsymbol{K}_i\boldsymbol{s}_i)\hat{\boldsymbol{H}}_i \\ &-\left(\boldsymbol{I}_N + \sum_{j=1}^{n_F}\frac{\hat{\boldsymbol{H}}_j}{\phi}\right)^{-1}\hat{\boldsymbol{H}}_i\mathbf{Var}(\boldsymbol{K}_i\boldsymbol{s}_i)\frac{\hat{\boldsymbol{H}}_i}{\phi} - \mathbf{Var}(\boldsymbol{K}_i\boldsymbol{s}_i)\hat{\boldsymbol{H}}_i\left(\boldsymbol{I}_N + \sum_{j=1}^{n_F}\frac{\hat{\boldsymbol{H}}_j}{\phi}\right)^{-1}\frac{\hat{\boldsymbol{H}}_i}{\phi}\end{aligned} \tag{133}$$

Denote $\boldsymbol{Y}$ the sum of price impacts $\boldsymbol{Y} = \boldsymbol{I}_N + \sum_{j=1}^{n_F}\frac{\hat{\boldsymbol{H}}_j}{\phi}$. The trace could be written as

$$\begin{aligned}&\mathbf{tr}[\mathbf{Cov}\,(\tilde{\boldsymbol{p}}_i, \tilde{\boldsymbol{q}}_i)] \\ =&\mathbf{tr}\left[\boldsymbol{Y}^{-1}\sum_{j=1}^{n_F}\hat{\boldsymbol{H}}_j\mathbf{Var}(\boldsymbol{K}_j\boldsymbol{s}_j)\hat{\boldsymbol{H}}_j\boldsymbol{Y}^{-1}\frac{\hat{\boldsymbol{H}}_i}{\phi^2}\right] + \mathbf{tr}\left[\mathbf{Var}(\boldsymbol{K}_i\boldsymbol{s}_i)\hat{\boldsymbol{H}}_i\right] - \mathbf{tr}\left[\boldsymbol{Y}^{-1}\hat{\boldsymbol{H}}_i\mathbf{Var}(\boldsymbol{K}_i\boldsymbol{s}_i)\frac{\hat{\boldsymbol{H}}_i}{\phi}\right] - \mathbf{tr}\left[\mathbf{Var}(\boldsymbol{K}_i\boldsymbol{s}_i)\hat{\boldsymbol{H}}_i\boldsymbol{Y}^{-1}\frac{\hat{\boldsymbol{H}}_i}{\phi}\right] \\ \geq&\mathbf{tr}\left[\boldsymbol{Y}^{-1}\hat{\boldsymbol{H}}_i\mathbf{Var}(\boldsymbol{K}_i\boldsymbol{s}_i)\hat{\boldsymbol{H}}_i\boldsymbol{Y}^{-1}\frac{\hat{\boldsymbol{H}}_i}{\phi^2}\right] + \mathbf{tr}\left[\mathbf{Var}(\boldsymbol{K}_i\boldsymbol{s}_i)\hat{\boldsymbol{H}}_i\right] - \mathbf{tr}\left[\boldsymbol{Y}^{-1}\hat{\boldsymbol{H}}_i\mathbf{Var}(\boldsymbol{K}_i\boldsymbol{s}_i)\frac{\hat{\boldsymbol{H}}_i}{\phi}\right] - \mathbf{tr}\left[\mathbf{Var}(\boldsymbol{K}_i\boldsymbol{s}_i)\hat{\boldsymbol{H}}_i\boldsymbol{Y}^{-1}\frac{\hat{\boldsymbol{H}}_i}{\phi}\right] \\ =&\mathbf{tr}\left[\mathbf{Var}(\boldsymbol{K}_i\boldsymbol{s}_i)\hat{\boldsymbol{H}}_i\boldsymbol{Y}^{-1}\frac{\hat{\boldsymbol{H}}_i}{\phi^2}\boldsymbol{Y}^{-1}\hat{\boldsymbol{H}}_i\right] + \mathbf{tr}\left[\mathbf{Var}(\boldsymbol{K}_i\boldsymbol{s}_i)\hat{\boldsymbol{H}}_i\right] - \mathbf{tr}\left[\mathbf{Var}(\boldsymbol{K}_i\boldsymbol{s}_i)\frac{\hat{\boldsymbol{H}}_i}{\phi}\boldsymbol{Y}^{-1}\hat{\boldsymbol{H}}_i\right] - \mathbf{tr}\left[\mathbf{Var}(\boldsymbol{K}_i\boldsymbol{s}_i)\hat{\boldsymbol{H}}_i\boldsymbol{Y}^{-1}\frac{\hat{\boldsymbol{H}}_i}{\phi}\right] \\ =&\phi\mathbf{tr}\left[\mathbf{Var}(\boldsymbol{K}_i\boldsymbol{s}_i)\left(\frac{\hat{\boldsymbol{H}}_i}{\phi}\boldsymbol{Y}^{-1}\frac{\hat{\boldsymbol{H}}_i}{\phi}\boldsymbol{Y}^{-1}\frac{\hat{\boldsymbol{H}}_i}{\phi} + \frac{\hat{\boldsymbol{H}}_i}{\phi} - \frac{\hat{\boldsymbol{H}}_i}{\phi}\boldsymbol{Y}^{-1}\frac{\hat{\boldsymbol{H}}_i}{\phi} - \frac{\hat{\boldsymbol{H}}_i}{\phi}\boldsymbol{Y}^{-1}\frac{\hat{\boldsymbol{H}}_i}{\phi}\right)\right] \\ =&\phi\mathbf{tr}\left[\mathbf{Var}(\boldsymbol{K}_i\boldsymbol{s}_i)\frac{\hat{\boldsymbol{H}}_i}{\phi}\left(\boldsymbol{Y}^{-1}\frac{\hat{\boldsymbol{H}}_i}{\phi} - \boldsymbol{I}_N\right)^2\right] \geq 0\end{aligned} \tag{134}$$

We denote $\boldsymbol{x}_i = \frac{\hat{\boldsymbol{H}}_i}{\phi}$ and $Z_i = \mathbf{Var}(\boldsymbol{K}_i\boldsymbol{s}_i) = \frac{\Sigma_b^2}{\Sigma_b+\Sigma_i}$ and consider diagonal shock and signal variance. $\boldsymbol{x}_i$, $\boldsymbol{Y}$ and Z_i are diagonal under our assumption. The covariance matrix is simplified as

$$\mathbf{Cov}\,(\tilde{\boldsymbol{p}}_i, \tilde{\boldsymbol{q}}_i) = \phi\left[\boldsymbol{Y}^{-1}\sum_{j=1}^{n_F}\boldsymbol{x}_j Z_j \boldsymbol{x}_j \boldsymbol{Y}^{-1}\boldsymbol{x}_i + Z_i\boldsymbol{x}_i - \boldsymbol{Y}^{-1}\boldsymbol{x}_i Z_i \boldsymbol{x}_i - Z_i \boldsymbol{x}_i \boldsymbol{Y}^{-1}\boldsymbol{x}_i\right] \tag{135}$$

The covariance matrix is also diagonal and denote the k^{th} diagonal $\mathbf{Cov}_{i,k}$. Subscript k refers to the

k^{th} diagonal value.

$$\begin{aligned}
\mathbf{Cov}_{i,k} &\coloneqq \mathbf{Cov}\left(\boldsymbol{p}_{i,k}, \tilde{\boldsymbol{q}}_{i,k}\right) \\
&= \phi\left[\boldsymbol{Y}_k^{-1}\sum_{j=1}^{n_F}\boldsymbol{x}_{j,k}Z_{j,k}\boldsymbol{x}_{j,k}\boldsymbol{Y}_k^{-1}\boldsymbol{x}_{i,k} + Z_i\boldsymbol{x}_{i,k} - \boldsymbol{Y}_k^{-1}\boldsymbol{x}_{i,k}Z_{i,k}\boldsymbol{x}_{i,k} - Z_{i,k}\boldsymbol{x}_{i,k}\boldsymbol{Y}_k^{-1}\boldsymbol{x}_{i,k}\right] \\
&= \phi\frac{\boldsymbol{x}_{i,k}}{\boldsymbol{Y}_k^2}\left[\sum_{j\neq i,j=1}^{n_F}\boldsymbol{x}_{j,k}^2 Z_{j,k} + Z_{i,k}\left(\boldsymbol{x}_{i,k} - \boldsymbol{Y}_k\right)^2\right]
\end{aligned} \tag{136}$$

The limiting behavioral for all variables are

$$\begin{aligned}
\lim_{\Sigma_{b,k}\to\infty}\boldsymbol{x}_{i,k} &= \left(1+\phi\rho_i\Sigma_{\epsilon_i,k}\right)^{-1} \\
\lim_{\Sigma_{b,k}\to\infty}\boldsymbol{Y}_k &= 1+\sum_{j=1}^{n_F}\lim_{\Sigma_{b,k}\to\infty}\boldsymbol{x}_{j,k} = 1+\sum_{j=1}^{n_F}\left(1+\phi\rho_j\Sigma_{\epsilon_j,k}\right)^{-1} \\
\lim_{\Sigma_{b,k}\to\infty}\frac{Z_{i,k}}{\Sigma_{b,k}} &= \lim_{\Sigma_{b,k}\to\infty}\frac{\frac{\Sigma_{b,k}^2}{\Sigma_{b,k}+\Sigma_{i,k}}}{\Sigma_{b,k}} = 1
\end{aligned} \tag{137}$$

The ratio of covariance to shock variance converges as

$$\lim_{\Sigma_{b,k}\to\infty}\frac{\mathbf{Cov}_{i,k}}{\Sigma_{b,k}} = \frac{\phi\left(1+\phi\rho_i\Sigma_{\epsilon_i,k}\right)^{-1}\left[\sum_{j\neq i,j=1}^{n_F}\left(1+\phi\rho_j\Sigma_{\epsilon_j,k}\right)^{-2} + \left(1+\sum_{j=1,j\neq i}^{n_F}\left(1+\phi\rho_j\Sigma_{\epsilon_j,k}\right)^{-1}\right)^2\right]}{\left(1+\sum_{j=1}^{n_F}\left(1+\phi\rho_j\Sigma_{\epsilon_j,k}\right)^{-1}\right)^2} \tag{138}$$

we have

$$\begin{aligned}
M_i^f &= \frac{\mathbf{E}[\tilde{\boldsymbol{q}}_i'\tilde{\boldsymbol{p}}_i]}{\mathbf{E}[\tilde{\boldsymbol{q}}_i'\boldsymbol{c}_i]} = \frac{\mathbf{E}[\tilde{\boldsymbol{q}}_i]'\mathbf{E}[\boldsymbol{p}] + \mathbf{tr}\left[\mathbf{Cov}(\tilde{\boldsymbol{p}}_i, \tilde{\boldsymbol{q}}_i)\right]}{\mathbf{E}[\tilde{\boldsymbol{q}}_i'\boldsymbol{c}_i]} \\
&= \frac{\sum_{j=1}^N(\mathbf{E}(\tilde{\boldsymbol{p}}_{i,j}) - \boldsymbol{c}_{i,j})\mathbf{E}(\tilde{\boldsymbol{p}}_{i,j})\hat{\boldsymbol{H}}_{i,j} + \sum_{j=1}^N\mathbf{Cov}_{i,j}}{\sum_{j=1}^N(\mathbf{E}(\tilde{\boldsymbol{p}}_{i,j}) - \boldsymbol{c}_{i,j})\boldsymbol{c}_{i,j}\hat{\boldsymbol{H}}_{i,j}}
\end{aligned} \tag{139}$$

We assume the diagonal values of shock variance are the same, so the asymptote of M_i^f is

$$\begin{aligned}
\alpha_i &\coloneqq \lim_{\Sigma_b\to\infty}\frac{M_i^f}{\Sigma_b} = \frac{\sum_{j=1}^N\lim_{\Sigma_b\to\infty}\frac{\mathbf{Cov}_{i,j}}{\Sigma_{b,k}}}{\sum_{j=1}^N(\tilde{\boldsymbol{p}}_j - \boldsymbol{c}_{i,j})\boldsymbol{c}_{i,j}\tilde{\hat{\boldsymbol{H}}}_{i,j}} > 0 \\
\gamma_i &\coloneqq \lim_{\Sigma_b\to\infty}\left(M_i^f - \alpha_i\Sigma_b\right) = \frac{\sum_{j=1}^N(\tilde{\boldsymbol{p}}_j - \boldsymbol{c}_{i,j})\tilde{\boldsymbol{p}}_j\tilde{\hat{\boldsymbol{H}}}_{i,j} + \widetilde{\mathbf{Cov}}_{i,j}}{\sum_{j=1}^N(\tilde{\boldsymbol{p}}_j - \boldsymbol{c}_{i,j})\boldsymbol{c}_{i,j}\tilde{\hat{\boldsymbol{H}}}_{i,j}}
\end{aligned} \tag{140}$$

where the difference $\widetilde{\mathbf{Cov}}_i$ is definted as

$$\widetilde{\mathbf{Cov}}_i := \lim_{\Sigma_b \to \infty} \left(\mathbf{Cov}_{i,k} - \left(\lim_{\Sigma_b \to \infty} \frac{\mathbf{Cov}_{i,k}}{\Sigma_b} \right) \Sigma_b \right)$$
$$= -\frac{\phi \left(1 + \phi \rho_i \Sigma_{\epsilon_i,k}\right)^{-1} \left[\sum_{j \neq i, j=1}^{n_F} \left(1 + \phi \rho_j \Sigma_{\epsilon_j,k}\right)^{-2} \Sigma_{\epsilon_j,k} + \left(1 + \sum_{j=1, j \neq i}^{n_F} \left(1 + \phi \rho_j \Sigma_{\epsilon_j,k}\right)^{-1}\right)^2 \Sigma_{\epsilon_i,k} \right]}{\left(1 + \sum_{j=1}^{n_F} \left(1 + \phi \rho_j \Sigma_{\epsilon_j,k}\right)^{-1}\right)^2} \tag{141}$$

The average firm-level markup $\overline{M}^f = (1/n_F) \sum_{i=1}^{n_F} M_i^f$ approaches $\sum_{i=1}^{n_F} \frac{\alpha_i}{n_F} \Sigma_b + \sum_{i=1}^{n_F} \frac{\gamma_i}{n_F}$ in the long run. The economy-level markup is $M^m = \sum_{i=1}^{n_F} w^{\boldsymbol{H}_i} M_i^f$ with $w^{\boldsymbol{H}_i} = \frac{\mathbf{E}\left[\tilde{\boldsymbol{q}}_i' \boldsymbol{c}_i\right]}{\sum_{i=1}^{n_F} \mathbf{E}\left[\tilde{\boldsymbol{q}}_i' \boldsymbol{c}_i\right]}$. The weight $w^{\boldsymbol{H}_i}$ converges to w_i as shock variance goes to infinity, implying an asymptote of economy-level markup.

$$w_i := \lim_{\Sigma_b \to \infty} w^{\boldsymbol{H}_i} = \frac{\sum_{j=1}^{N} (\tilde{\boldsymbol{p}}_j - \boldsymbol{c}_{i,j}) \boldsymbol{c}_{i,j} \tilde{\tilde{\boldsymbol{H}}}_{i,j}}{\sum_{i=1}^{n_F} \sum_{j=1}^{N} (\tilde{\boldsymbol{p}}_j - \boldsymbol{c}_{i,j}) \boldsymbol{c}_{i,j} \tilde{\tilde{\boldsymbol{H}}}_{i,j}} \Rightarrow M^m \text{ approaches } \sum_{i=1}^{n_F} w_i \alpha_i \Sigma_b + \sum_{i=1}^{n_F} w_i \gamma_i \tag{142}$$

Finally, the derivative of each component of covariance is

$$\begin{aligned}
\frac{\partial \boldsymbol{x}_{i,k}}{\partial \Sigma_{b,k}} &= -\phi \rho_i \boldsymbol{x}_{i,k}^2 \left(\frac{\Sigma_{\epsilon_i,k}}{\Sigma_{b,k} + \Sigma_{\epsilon_i,k}} \right)^2 = -\frac{\boldsymbol{x}_{i,k}(1 - \boldsymbol{x}_{i,k}) \Sigma_{\epsilon_i,k}}{\Sigma_{b,k}(\Sigma_{b,k} + \Sigma_{\epsilon_i,k})} \\
\frac{\partial \boldsymbol{Y}_k}{\partial \Sigma_{b,k}} &= \sum_{j=1}^{n_F} \frac{\partial \boldsymbol{x}_{j,k}}{\partial \Sigma_{b,k}} = -\sum_{j=1}^{n_F} \frac{\boldsymbol{x}_{j,k}(1 - \boldsymbol{x}_{j,k}) \Sigma_{\epsilon_j,k}}{\Sigma_{b,k}(\Sigma_{b,k} + \Sigma_{\epsilon_j,k})} \\
\frac{\partial Z_{i,k}}{\partial \Sigma_{b,k}} &= \frac{\Sigma_{b,k}(\Sigma_{b,k} + 2\Sigma_{\epsilon_i,k})}{\left(\Sigma_{b,k} + \Sigma_{\epsilon_i,k}\right)^2} = \frac{Z_{i,k}(\Sigma_{b,k} + 2\Sigma_{\epsilon_i,k})}{\Sigma_{b,k} \left(\Sigma_{b,k} + \Sigma_{\epsilon_i,k}\right)} \\
\frac{\partial \frac{\boldsymbol{x}_{i,k}}{\boldsymbol{Y}_k^2}}{\partial \Sigma_{b,k}} &= \frac{\boldsymbol{x}_{i,k}}{\Sigma_{b,k} \boldsymbol{Y}_k^2} \left[\frac{2}{\boldsymbol{Y}_k} \sum_{j=1}^{n_F} \frac{\boldsymbol{x}_{j,k}(1 - \boldsymbol{x}_{j,k}) \Sigma_{\epsilon_j,k}}{\Sigma_{b,k} + \Sigma_{\epsilon_j,k}} - \frac{(1 - \boldsymbol{x}_{i,k}) \Sigma_{\epsilon_i,k}}{\Sigma_{b,k} + \Sigma_{\epsilon_i,k}} \right] \\
\frac{\partial \sum_{j \neq i, j=1}^{n_F} \boldsymbol{x}_{j,k}^2 Z_{j,k}}{\partial \Sigma_{b,k}} &= \sum_{j \neq i, j=1}^{n_F} \frac{\boldsymbol{x}_{j,k}^2 Z_{j,k}}{(\Sigma_{b,k} + \Sigma_{\epsilon_j,k}) \Sigma_{b,k}} \left[\Sigma_{b,k} + 2\Sigma_{\epsilon_j,k} - 2(1 - \boldsymbol{x}_{j,k}) \Sigma_{\epsilon_j,k} \right] \\
\frac{\partial Z_{i,k} \left(\boldsymbol{x}_{i,k} - \boldsymbol{Y}_k\right)^2}{\partial \Sigma_{b,k}} &= \frac{Z_{i,k} \left(\boldsymbol{x}_{i,k} - \boldsymbol{Y}_k\right)^2 \left(\Sigma_{b,k} + 2\Sigma_{\epsilon_i,k}\right)}{(\Sigma_{b,k} + 2\Sigma_{\epsilon_i,k}) \Sigma_{b,k}} - 2 \left(\boldsymbol{Y}_k - \boldsymbol{x}_{i,k}\right) \frac{Z_{i,k}}{\Sigma_{b,k}} \sum_{j \neq i, j=1}^{n_F} \boldsymbol{x}_{j,k}(1 - \boldsymbol{x}_{j,k}) \frac{\Sigma_{\epsilon_j,k}}{\Sigma_{b,k} + \Sigma_{\epsilon_j,k}}
\end{aligned} \tag{143}$$

So the derivative of covariance $\mathbf{Cov}_{i,k}$ could be decomposed into two parts

$$\frac{\partial \mathbf{Cov}_{i,k}}{\partial \Sigma_{b,k}} = \phi \frac{\boldsymbol{x}_{i,k}}{\boldsymbol{Y}_k^2} \left[\boldsymbol{G}_1 + \boldsymbol{G}_2 \right] \tag{144}$$

where

$$
\begin{aligned}
\boldsymbol{G}_1 &:= Z_{i,k}\left(\boldsymbol{x}_{i,k}-\boldsymbol{Y}_k\right)^2 \frac{\Sigma_{b,k}+(1+\boldsymbol{x}_{i,k})\Sigma_{\epsilon_i,k}}{\Sigma_{b,k}+\Sigma_{\epsilon_i,k}} - 2Z_{i,k}\left(\boldsymbol{Y}_k-\boldsymbol{x}_{i,k}\right)\frac{\boldsymbol{x}_{i,k}}{\boldsymbol{Y}_k}\sum_{j\neq i,j=1}^{n_F}\frac{\boldsymbol{x}_{j,k}(1-\boldsymbol{x}_{j,k})\Sigma_{\epsilon_j,k}}{\Sigma_{b,k}+\Sigma_{\epsilon_j,k}} \\
\boldsymbol{G}_2 &:= \frac{\Sigma_{b,k}+\boldsymbol{x}_{i,k}\Sigma_{\epsilon_i,k}}{\Sigma_{b,k}+\Sigma_{\epsilon_i,k}}\sum_{j\neq i,j=1}^{n_F}\boldsymbol{x}_{j,k}^2 Z_{j,k} + \frac{2}{\boldsymbol{Y}_k}\sum_{j=1}^{n_F}\frac{\boldsymbol{x}_{j,k}(1-\boldsymbol{x}_{j,k})\Sigma_{\epsilon_j,k}}{\Sigma_{b,k}+\Sigma_{\epsilon_j,k}}\sum_{j\neq i,j=1}^{n_F}\boldsymbol{x}_{j,k}^2 Z_{j,k} \\
&\qquad + \sum_{j\neq i,j=1}^{n_F}\boldsymbol{x}_{j,k}^2 Z_{j,k}\frac{\Sigma_{\epsilon_j,k}}{\Sigma_{b,k}+\Sigma_{\epsilon_j,k}}\left[1-2(1-\boldsymbol{x}_{j,k})\right]
\end{aligned} \tag{145}
$$

We can prove that $\boldsymbol{G}_1$ is always positive

$$
\boldsymbol{G}_1 \geq 0 \Leftrightarrow Z_{i,k}\left(\boldsymbol{x}_{i,k}-\boldsymbol{Y}_k\right)^2 \frac{\Sigma_{b,k}+(1+\boldsymbol{x}_{i,k})\Sigma_{\epsilon_i,k}}{\Sigma_{b,k}+\Sigma_{\epsilon_i,k}} \geq 2Z_{i,k}\left(\boldsymbol{Y}_k-\boldsymbol{x}_{i,k}\right)\frac{\boldsymbol{x}_{i,k}}{\boldsymbol{Y}_k}\sum_{j\neq i,j=1}^{n_F}\frac{\boldsymbol{x}_{j,k}(1-\boldsymbol{x}_{j,k})\Sigma_{\epsilon_j,k}}{\Sigma_{b,k}+\Sigma_{\epsilon_j,k}} \tag{146}
$$

Since $\boldsymbol{Y}_k \geq 1+\boldsymbol{x}_{i,k}$ and $0 \leq \boldsymbol{x}_{i,k} \leq 1$, we have

$$
\begin{aligned}
Z_{i,k}\left(\boldsymbol{x}_{i,k}-\boldsymbol{Y}_k\right)^2 \frac{\Sigma_{b,k}+(1+\boldsymbol{x}_{i,k})\Sigma_{\epsilon_i,k}}{\Sigma_{b,k}+\Sigma_{\epsilon_i,k}} &\geq Z_{i,k}\left(\boldsymbol{Y}_k-\boldsymbol{x}_{i,k}\right)^2 \\
&\geq Z_{i,k}\left(\boldsymbol{Y}_k-\boldsymbol{x}_{i,k}\right)\frac{2\boldsymbol{x}_{i,k}}{\boldsymbol{Y}_k}\left(1+\sum_{j\neq i,j=1}^{n_F}\boldsymbol{x}_{j,k}\right) \\
&\geq Z_{i,k}\left(\boldsymbol{Y}_k-\boldsymbol{x}_{i,k}\right)\frac{2\boldsymbol{x}_{i,k}}{\boldsymbol{Y}_k}\sum_{j\neq i,j=1}^{n_F}\frac{\boldsymbol{x}_{j,k}(1-\boldsymbol{x}_{j,k})\Sigma_{\epsilon_j,k}}{\Sigma_{b,k}+\Sigma_{\epsilon_j,k}}
\end{aligned} \tag{147}
$$

As for the $\boldsymbol{G}_2$, large shock variance ($\Sigma_{b,k} \geq \Sigma_{\epsilon_j,k}, \forall j$) guarantees its positivity since

$$
\begin{aligned}
&\Sigma_{b,k} \geq \Sigma_{\epsilon_j,k} \Rightarrow \frac{\Sigma_{b,k}}{\Sigma_{b,k}+\Sigma_{\epsilon_i,k}} \geq \frac{\Sigma_{\epsilon_j,k}}{\Sigma_{b,k}+\Sigma_{\epsilon_j,k}} \\
\Rightarrow &\frac{\Sigma_{b,k}+\boldsymbol{x}_{i,k}\Sigma_{\epsilon_i,k}}{\Sigma_{b,k}+\Sigma_{\epsilon_i,k}}\sum_{j\neq i,j=1}^{n_F}\boldsymbol{x}_{j,k}^2 Z_{j,k} \geq \sum_{j\neq i,j=1}^{n_F}\boldsymbol{x}_{j,k}^2 Z_{j,k}\frac{\Sigma_{\epsilon_j,k}}{\Sigma_{b,k}+\Sigma_{\epsilon_j,k}} \\
\Rightarrow &\boldsymbol{G}_2 \geq \sum_{j\neq i,j=1}^{n_F}\boldsymbol{x}_{j,k}^2 Z_{j,k}\frac{\Sigma_{\epsilon_j,k}}{\Sigma_{b,k}+\Sigma_{\epsilon_j,k}}\left[2-2(1-\boldsymbol{x}_{j,k})\right] \geq 0
\end{aligned} \tag{148}
$$

So the derivative of covariance $\mathbf{Cov}_{i,k}$ is positive when shock variance is large enough.

This proof held marginal costs $\tilde{c}$ fixed. If we assume the marginal cost of adjusting c is sufficiently high, by continuity, the inequality will still hold. □

CYCLICAL MARKUPS WITH EFFICIENT INVESTMENT The trade-off between the risk premium and motivation effect still exists here. When the variances of shocks increase, the firm has a tendency to charge a higher price in order to compensate for the increasing risk. On the other hand, they will become less willing to invest, which leads to higher production costs and thus drive markups down. As we show here, depending on the parameters, the aggregated markups can both increase or decrease with the economic cycle.

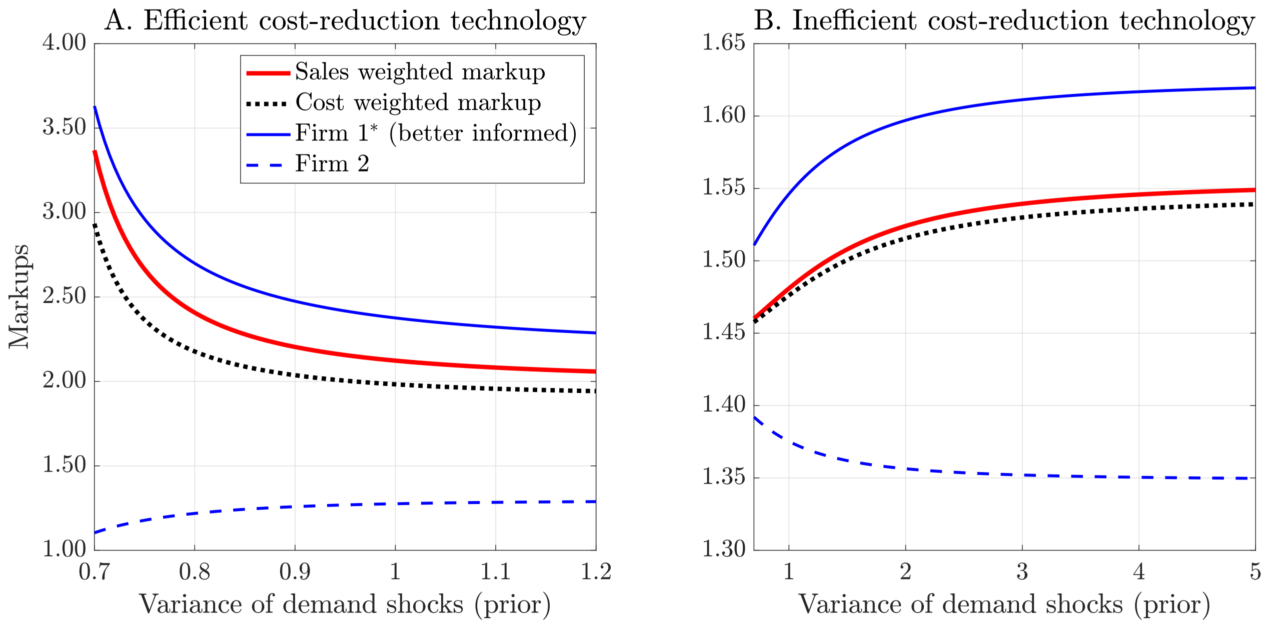


Figure 9: Comparative static: markups and economic cycles, $\rho = 1$

Notes: These two panels depict how markups change with the variance of demand shock. For tractability, the weighted markups are weighted by expected sales (costs) over expected markups. At firm level, firm 1 has eight data points while firm 2 only observes two. The parameter for investment, χ_c, is 1 and 1.5 for the two panels, respectively. Moreover, $\rho_1 = \rho_2 = 1$.

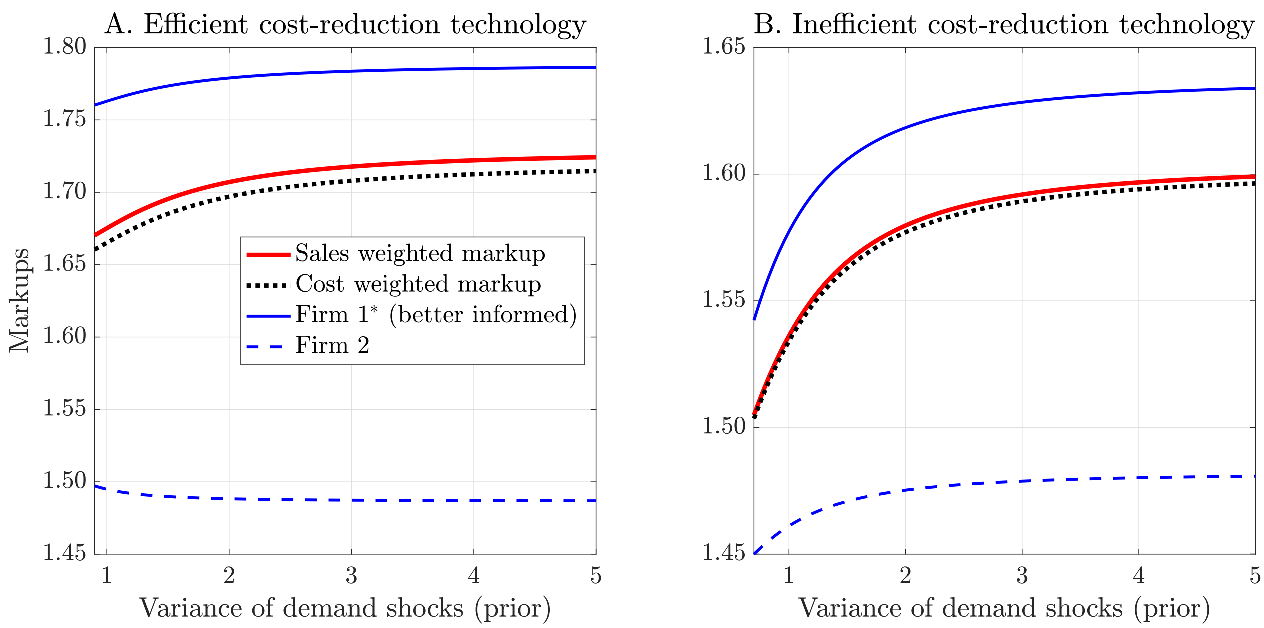


Figure 10: Comparative static: markups and economic cycles, $\rho = 4$

Notes: These two panels depict how markups change with the variance of demand shock. For tractability, the weighted markups are weighted by expected sales (costs) over expected markups. At firm level, firm 1 has eight data points while firm 2 only observes two. The parameter for investment, χ_c, is 1 and 1.5 for the two panels, respectively. Moreover, $\rho_1 = \rho_2 = 4$.

C. Solutions to Alternative Models

C.1. A Model with Aggregate Demand Shocks

Our results can be explained more clearly in a model with firm-specific demand. However, none of the results is dependent on the firm-specific nature of the shocks. In this appendix, we setup, solve and analyze a model where shocks affect the demand for attributes. These shocks affect all firms whose product load on these attributes. Signals are about the aggregate vector of attribute demand shocks. The new complication in this model is that the solution is not explicit. The solution is characterized by a set of $n_F + 3$ equations in $n_F + 3$ unknowns.

Changes to model setup.

DEMAND The first order condition for demand is a linear combination of $\boldsymbol{b}$ and price $\boldsymbol{p}$, with a constant term $\bar{\boldsymbol{p}}$

$$\frac{1}{\phi}\sum_{i=1}^{n_F}\tilde{\boldsymbol{q}}_i = \bar{\boldsymbol{p}} + \boldsymbol{b} - \boldsymbol{p} \tag{149}$$

INFORMATION Each firm sees a private signal $\boldsymbol{s}_i$ is standard normal and $\boldsymbol{s}_i = \boldsymbol{b} + \boldsymbol{\varepsilon}_i$ where the variance of $\boldsymbol{b}$ and $\boldsymbol{\varepsilon}_i$ are Σ_b and $\Sigma_{\epsilon_i} = \tilde{\Sigma}_{\epsilon_i}/n_{di}$ respectively.

SOLUTION Each Firm has the same mean-variance objective. Its first-order condition with respect to $\tilde{\boldsymbol{q}}_i$ is

$$\tilde{\boldsymbol{q}}_i = \left(\rho_i \mathbf{Var}\left[\boldsymbol{p}_i|\mathcal{I}_i\right] - \frac{\partial \mathbf{E}\left[\boldsymbol{p}_i|\mathcal{I}_i\right]}{\partial \tilde{\boldsymbol{q}}_i}\right)^{-1}\left(\mathbf{E}\left[\boldsymbol{p}_i|\mathcal{I}_i\right] - \boldsymbol{c}_i\right) \tag{150}$$

From differentiating the pricing function (149), we find that the price impact of one additional unit of attribute output is

$$\frac{\partial \mathbf{E}\left[\boldsymbol{p}_i|\mathcal{I}_i\right]}{\partial \tilde{\boldsymbol{q}}_i} = -\frac{1}{\phi}\boldsymbol{I}_N \;\Rightarrow\; \hat{\boldsymbol{H}}_i \equiv \left(\rho_i \mathbf{Var}\left[\boldsymbol{p}_i|\mathcal{I}_i\right] + \frac{\boldsymbol{I}_N}{\phi}\right)^{-1} \tag{151}$$

So the optimal production is $\tilde{\boldsymbol{q}}_i = \hat{\boldsymbol{H}}_i\left(\mathbf{E}\left[\boldsymbol{p}_i|\mathcal{I}_i\right] - \boldsymbol{c}_i\right)$

BAYESIAN UPDATING 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 $\boldsymbol{D}$, $\boldsymbol{F}$ and $\{\boldsymbol{h}_i\}_{i=1,\ldots,n_F}$.

$$\boldsymbol{p} = \boldsymbol{D} + \boldsymbol{F}\boldsymbol{b} + \sum_{i=1}^{n_F}\boldsymbol{h}_i\boldsymbol{\varepsilon}_i \tag{152}$$

Since firm i could only observe $\boldsymbol{s}_i$, its expectation of the price is

$$\mathbf{E}[\boldsymbol{p}|\boldsymbol{s}_i] = \boldsymbol{D} + \boldsymbol{\beta}_i\boldsymbol{s}_i \;\text{ where }\; \boldsymbol{\beta}_i = \mathbf{Cov}(\boldsymbol{p},\boldsymbol{s}_i)\mathbf{Var}(\boldsymbol{s}_i)^{-1} \tag{153}$$

The variance of price forecast error is

$$\mathbf{Var}[\boldsymbol{p}|\boldsymbol{s}_i] = \mathbf{Var}(\boldsymbol{p}) - \mathbf{Cov}(\boldsymbol{p},\boldsymbol{s}_i)\mathbf{Var}(\boldsymbol{s}_i)^{-1}\mathbf{Cov}(\boldsymbol{p},\boldsymbol{s}_i)' \tag{154}$$

The optimal production is

$$\begin{aligned} \tilde{\boldsymbol{q}}_i &= \hat{\boldsymbol{H}}_i(\boldsymbol{D} - \boldsymbol{c}_i + \boldsymbol{\beta}_i \boldsymbol{s}_i) \\ \mathbf{E}[\tilde{\boldsymbol{q}}_i] &= \hat{\boldsymbol{H}}_i(\boldsymbol{D} - \boldsymbol{c}_i) \end{aligned} \tag{155}$$

SOLUTION According to the total demand function, we could match the coefficients

$$\begin{aligned} & \frac{1}{\phi}\sum_{i=1}^{n_F} \tilde{\boldsymbol{q}}_i = \bar{\boldsymbol{p}} + \boldsymbol{b} - \boldsymbol{p} \\ \Rightarrow & \frac{1}{\phi}\sum_{i=1}^{n_F} \hat{\boldsymbol{H}}_i(\boldsymbol{D} - \boldsymbol{c}_i + \boldsymbol{\beta}_i \boldsymbol{s}_i) = \bar{\boldsymbol{p}} + \boldsymbol{b} - \left(\boldsymbol{D} + \boldsymbol{F}\boldsymbol{b} + \sum_{i=1}^{n_F} \boldsymbol{h}_i \boldsymbol{\varepsilon}_i\right) \\ \Rightarrow & \left(\boldsymbol{F} - \boldsymbol{I}_N + \frac{1}{\phi}\sum_{i=1}^{n_F} \hat{\boldsymbol{H}}_i \boldsymbol{\beta}_i\right)\boldsymbol{b} + \sum_{i=1}^{n_F}\left(\boldsymbol{h}_i + \frac{1}{\phi}\hat{\boldsymbol{H}}_i\boldsymbol{\beta}_i\right)\boldsymbol{\varepsilon}_i + \left(\boldsymbol{I}_N + \frac{1}{\phi}\sum_{i=1}^{n_F}\hat{\boldsymbol{H}}_i\right)\boldsymbol{D} - \bar{\boldsymbol{p}} - \frac{1}{\phi}\sum_{i=1}^{n_F}\hat{\boldsymbol{H}}_i\boldsymbol{c}_i = 0 \end{aligned} \tag{156}$$

So the coefficients must satisfy

$$\begin{aligned} \boldsymbol{F} &= \boldsymbol{I}_N - \frac{1}{\phi}\sum_{i=1}^{n_F}\hat{\boldsymbol{H}}_i\boldsymbol{\beta}_i \\ \boldsymbol{h}_i &= -\frac{1}{\phi}\hat{\boldsymbol{H}}_i\boldsymbol{\beta}_i, \;\; \forall\, i = 1, \ldots, n_F \\ \boldsymbol{D} &= \left(\boldsymbol{I}_N + \frac{1}{\phi}\sum_{i=1}^{n_F}\hat{\boldsymbol{H}}_i\right)^{-1}\left(\bar{\boldsymbol{p}} + \frac{1}{\phi}\sum_{i=1}^{n_F}\hat{\boldsymbol{H}}_i\boldsymbol{c}_i\right) \end{aligned} \tag{157}$$

where the $\boldsymbol{\beta}_i$ and $\hat{\boldsymbol{H}}_i$ are endogenously determined by $\boldsymbol{F}$ and $\{\boldsymbol{h}_j\}_{j=1,\ldots,n_F}$

$$\begin{aligned} \boldsymbol{\beta}_i &= (\boldsymbol{F}\Sigma_b + \boldsymbol{h}_i\Sigma_{\epsilon_i})\left(\Sigma_b + \Sigma_{\epsilon_i}\right)^{-1} \\ \hat{\boldsymbol{H}}_i &= \left[\rho_i\left(\boldsymbol{F}\Sigma_b\boldsymbol{F}' + \sum_{i=1}^{n_F}\boldsymbol{h}_i^2\Sigma_{\epsilon_i} - (\boldsymbol{F}\Sigma_b + \boldsymbol{h}_i\Sigma_{\epsilon_i})\left(\Sigma_b + \Sigma_{\epsilon_i}\right)^{-1}(\boldsymbol{F}\Sigma_b + \boldsymbol{h}_i\Sigma_{\epsilon_i})'\right) + \frac{\boldsymbol{I}_N}{\phi}\right]^{-1} \end{aligned} \tag{158}$$

PRODUCT-LEVEL MARKUP The product-level markup for product k produced by firm i is $M^p_{ik} := \mathbf{E}[\boldsymbol{p}_i(j)]/\boldsymbol{c}_i(j)$. The average product-level markup is

$$\overline{M}^p = \frac{1}{N}\frac{1}{n_F}\sum_{i=1}^{n_F}\sum_{j=1}^{N} M^p_{ij} = \frac{1}{N}\frac{1}{n_F}\sum_{i=1}^{n_F}\sum_{j=1}^{N}\frac{\boldsymbol{D}(j)}{\boldsymbol{c}_i(j)} \tag{159}$$

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

$$\begin{aligned} M^f_i &= \frac{\mathbf{E}[\tilde{\boldsymbol{q}}_i'\boldsymbol{p}]}{\mathbf{E}[\tilde{\boldsymbol{q}}_i'\boldsymbol{c}_i]} = \frac{\mathbf{E}[\tilde{\boldsymbol{q}}_i]'\mathbf{E}[\boldsymbol{p}] + \mathbf{tr}\left[\mathbf{Cov}(\boldsymbol{p}_i, \tilde{\boldsymbol{q}}_i)\right]}{\mathbf{E}[\tilde{\boldsymbol{q}}_i'\boldsymbol{c}_i]} \\ &= \frac{(\boldsymbol{D} - \boldsymbol{c}_i)'\hat{\boldsymbol{H}}_i\boldsymbol{D} + \mathbf{tr}\left(\hat{\boldsymbol{H}}_i\boldsymbol{\beta}_i\mathbf{Var}(\boldsymbol{s}_i)\boldsymbol{\beta}_i'\right)}{(\boldsymbol{D} - \boldsymbol{c}_i)'\hat{\boldsymbol{H}}_i\boldsymbol{c}_i} > \frac{(\boldsymbol{D} - \boldsymbol{c}_i)'\hat{\boldsymbol{H}}_i\boldsymbol{D}}{(\boldsymbol{D} - \boldsymbol{c}_i)'\hat{\boldsymbol{H}}_i\boldsymbol{c}_i} \end{aligned} \tag{160}$$

Thus, the average firm-level markup is $\overline{M}^f = (1/n_F)\sum_{i=1}^{n_F} M^f_i$.

ECONOMY-LEVEL MARKUP The industry markup is

$$M^m := \frac{\mathbf{E}\left[\sum_{i=1}^{n_F} \tilde{\boldsymbol{q}}_i' \boldsymbol{p}_i\right]}{\mathbf{E}\left[\sum_{i=1}^{n_F} \tilde{\boldsymbol{q}}_i' \boldsymbol{c}_i\right]} = \frac{\sum_{i=1}^{n_F} \mathbf{E}\left[\tilde{\boldsymbol{q}}_i' \boldsymbol{p}_i\right]}{\sum_{i=1}^{n_F} \mathbf{E}\left[\tilde{\boldsymbol{q}}_i' \boldsymbol{c}_i\right]} = \sum_{i=1}^{n_F} w^{\boldsymbol{H}_i} M_i^f \quad \text{where} \quad w^{\boldsymbol{H}_i} = \frac{\mathbf{E}\left[\tilde{\boldsymbol{q}}_i' \boldsymbol{c}_i\right]}{\sum_{i=1}^{n_F} \mathbf{E}\left[\tilde{\boldsymbol{q}}_i' \boldsymbol{c}_i\right]}. \tag{161}$$

Aggregate Demand Model: Cyclical Markup Fluctuations

Proposition 7. *The product-level markup converges as shock variance tends to infinity given identical risk aversion and signal precision across all firms.*

Proof. Define $M_i = \left(\Sigma_b + \Sigma_{\epsilon_i}\left(I_N + \frac{\hat{\boldsymbol{H}}_i}{\phi}\right)\right)^{-1}$, we have $\lim_{\Sigma_b \to \infty} M_i \Sigma_b = I_N$. The unknown coefficients could be expressed in $\hat{\boldsymbol{H}}_i$ and M_i.

$$\begin{aligned}
\boldsymbol{\beta}_i &= \frac{\Sigma_b M_i}{I_N + \sum_{j=1}^{n_F} \Sigma_b M_j \frac{\hat{\boldsymbol{H}}_j}{\phi}} \\
\boldsymbol{h}_i &= -\boldsymbol{\beta}_i \frac{\hat{\boldsymbol{H}}_i}{\phi} = -\frac{\Sigma_b M_i \frac{\hat{\boldsymbol{H}}_i}{\phi}}{I_N + \sum_{j=1}^{n_F} \Sigma_b M_j \frac{\hat{\boldsymbol{H}}_j}{\phi}} \\
\boldsymbol{F} &= I_N + \sum_{j=1}^{n_F} \boldsymbol{h}_i = \frac{1}{1 + \sum_{j=1}^{n_F} \Sigma_b M_j \frac{\hat{\boldsymbol{H}}_j}{\phi}}
\end{aligned} \tag{162}$$

The price impact $\hat{\boldsymbol{H}}_i$ satisfy following system of equations

$$\left(\frac{\hat{\boldsymbol{H}}_i}{\phi}\right)^{-1} = I_N + \rho_i \phi \left(\boldsymbol{F}^2 \Sigma_b + \sum_{i=1}^{n_F} \boldsymbol{h}_i^2 \Sigma_{\epsilon_i} - \boldsymbol{\beta}_i^2 (\Sigma_b + \Sigma_{\epsilon_i})\right) \tag{163}$$

By symmetry, all firm choose the same impact function $\hat{\boldsymbol{H}}_i$, thus

$$\begin{aligned}
\left(\frac{\hat{\boldsymbol{H}}_{i,k}}{\phi}\right)^{-1} &= 1 + \rho_i \phi \frac{\left(\Sigma_{\epsilon_{i,k}} + \Sigma_{b,k} + \Sigma_{\epsilon_{i,k}} \frac{\hat{\boldsymbol{H}}_{i,k}}{\phi}\right)^2 \Sigma_{b,k} + n_F \frac{\hat{\boldsymbol{H}}_{i,k}^2}{\phi^2} \Sigma_{b,k}^2 \Sigma_{\epsilon_{i,k}} - \Sigma_{b,k}^2 (\Sigma_{b,k} + \Sigma_{\epsilon_{i,k}})}{\left(\Sigma_{b,k} + \Sigma_{\epsilon_{i,k}} + (\Sigma_{\epsilon_{i,k}} + n_F \Sigma_{b,k}) \frac{\hat{\boldsymbol{H}}_{i,k}}{\phi}\right)^2} \\
&= 1 + \rho_i \phi \frac{\left(1 + \frac{\hat{\boldsymbol{H}}_{i,k}}{\phi}\right)\left(2 + \left(1 + \frac{\hat{\boldsymbol{H}}_{i,k}}{\phi}\right) \frac{\Sigma_{\epsilon_{i,k}}}{\Sigma_{b,k}}\right) + n_F \frac{\hat{\boldsymbol{H}}_{i,k}^2}{\phi^2} \Sigma_{\epsilon_{i,k}} - \Sigma_{\epsilon_{i,k}}}{\left(1 + \frac{\Sigma_{\epsilon_{i,k}}}{\Sigma_{b,k}} + (\frac{\Sigma_{\epsilon_{i,k}}}{\Sigma_{b,k}} + n_F) \frac{\hat{\boldsymbol{H}}_{i,k}}{\phi}\right)^2}
\end{aligned} \tag{164}$$

This is a cubic equation for $\hat{\boldsymbol{H}}_{i,k}$ and has explicit solution. Moreover, the solution is convergent since all coefficients converge as shock variance $\Sigma_{b,k}$ goes to infinity. Another observation is that $\boldsymbol{F}^2 \Sigma_b - \boldsymbol{\beta}_i^2 \Sigma_b$ is bounded since

$$(\boldsymbol{F}^2 - \boldsymbol{\beta}_i^2)\Sigma_b = \left(1 + \sum_{j=1}^{n_F} \Sigma_b M_j \frac{\hat{\boldsymbol{H}}_j}{\phi}\right)^{-2} \frac{2\Sigma_{\epsilon_i}\left(1 + \frac{\hat{\boldsymbol{H}}_i}{\phi}\right) + \frac{\left(\Sigma_{\epsilon_i}\left(1 + \frac{\hat{\boldsymbol{H}}_i}{\phi}\right)\right)^2}{\Sigma_b}}{\left(1 + \frac{\Sigma_{\epsilon_i}}{\Sigma_b}\left(1 + \frac{\hat{\boldsymbol{H}}_i}{\phi}\right)\right)^2} \tag{165}$$

So the RHS of equation (163) is bounded and $\hat{\boldsymbol{H}}_i$ is positive in the limit. The product-level markup is clearly convergent because $\lim_{\Sigma_b \to \infty} \mathbf{E}\left[\boldsymbol{p}_i\right] = \lim_{\Sigma_b \to \infty} \left(\boldsymbol{I}_N + \frac{1}{\phi} \sum_{i=1}^{n_F} \hat{\boldsymbol{H}}_i\right)^{-1} \left(\bar{\boldsymbol{p}} + \frac{1}{\phi} \sum_{i=1}^{n_F} \hat{\boldsymbol{H}}_i \boldsymbol{c}_i\right)$ exists.

This proof held marginal costs $\tilde{c}$ fixed. If we assume the marginal cost of adjusting c is sufficiently high, by continuity, the inequality will still hold. □

Proposition 8. *The firm-level and economy-level markups are strictly increasing in demand shock variance, if the shock variance is large enough. In the high-variance limit, markups approach linear asymptotes.*

Proof. The covariance term is $\boldsymbol{\beta}_i \mathbf{Var}(\boldsymbol{s}_i)\boldsymbol{\beta}_i' \hat{\boldsymbol{H}}_i$ and we have

$$
\begin{aligned}
M_i^f &= \frac{\mathbf{E}[\tilde{\boldsymbol{q}}_i' \boldsymbol{p}_i]}{\mathbf{E}[\tilde{\boldsymbol{q}}_i' \boldsymbol{c}_i]} = \frac{\mathbf{E}[\tilde{\boldsymbol{q}}_i]' \mathbf{E}[\boldsymbol{p}] + \mathbf{tr}\left[\mathbf{Cov}(\boldsymbol{p}_i, \tilde{\boldsymbol{q}}_i)\right]}{\mathbf{E}[\tilde{\boldsymbol{q}}_i' \boldsymbol{c}_i]} \\
&= \frac{\sum_{j=1}^N (\mathbf{E}(\boldsymbol{p}_{i,j}) - \boldsymbol{c}_{i,j}) \mathbf{E}(\boldsymbol{p}_{i,j}) \hat{\boldsymbol{H}}_{i,j} + \sum_{j=1}^N \hat{\boldsymbol{H}}_{i,j} \frac{\boldsymbol{\beta}_{i,j}^2 \Sigma_{b,j}^2}{\Sigma_{b,j} + \Sigma_{\epsilon_i,j}}}{\sum_{j=1}^N (\mathbf{E}(\boldsymbol{p}_{i,j}) - \boldsymbol{c}_{i,j}) \boldsymbol{c}_{i,j} \hat{\boldsymbol{H}}_{i,j}}
\end{aligned} \tag{166}
$$

The β_i converges as $\lim_{\Sigma_b \to \infty} M_i \Sigma_b = 1$. The asymptote for M_i^f is

$$
\alpha_i := \lim_{\Sigma_b \to \infty} \frac{M_i^f}{\Sigma_b} = \frac{\sum_{j=1}^N \tilde{\hat{\boldsymbol{H}}}_{i,j} \tilde{\boldsymbol{\beta}}_j^2}{\sum_{j=1}^N (\tilde{\boldsymbol{p}}_j - \boldsymbol{c}_{i,j}) \boldsymbol{c}_{i,j} \tilde{\hat{\boldsymbol{H}}}_{i,j}}, \quad \gamma_i := \lim_{\Sigma_b \to \infty} \left(M_i^f - \alpha_i \Sigma_b \right) = \frac{\sum_{j=1}^N \left((\tilde{\boldsymbol{p}}_j - \boldsymbol{c}_{i,j}) \tilde{\boldsymbol{p}}_j - \tilde{\boldsymbol{\beta}}_j^2 \Sigma_{\epsilon_i,j} \right) \tilde{\hat{\boldsymbol{H}}}_{i,j}}{\sum_{j=1}^N (\tilde{\boldsymbol{p}}_j - \boldsymbol{c}_{i,j}) \boldsymbol{c}_{i,j} \tilde{\hat{\boldsymbol{H}}}_{i,j}} \tag{167}
$$

Where $\lim_{\Sigma_b \to \infty} \hat{\boldsymbol{H}}_i = \tilde{\hat{\boldsymbol{H}}}_i$ and $\lim_{\Sigma_b \to \infty} \boldsymbol{\beta}_i = \tilde{\boldsymbol{\beta}}_i = \left(I_N + \sum_{i=1}^{n_F} \tilde{\hat{\boldsymbol{H}}}_i \right)^{-1}$. The average firm-level markup $\overline{M}^f = (1/n_F) \sum_{i=1}^{n_F} M_i^f$ approaches $\sum_{i=1}^{n_F} \frac{\alpha_i}{n_F} \Sigma_b + \sum_{i=1}^{n_F} \frac{\gamma_i}{n_F}$ in the long run. The economy-level markup is $M^m = \sum_{i=1}^{n_F} w^{\boldsymbol{H}_i} M_i^f$ with $w^{\boldsymbol{H}_i} = \frac{\mathbf{E}\left[\tilde{\boldsymbol{q}}_i' \boldsymbol{c}_i\right]}{\sum_{i=1}^{n_F} \mathbf{E}\left[\tilde{\boldsymbol{q}}_i' \boldsymbol{c}_i\right]}$. The weight $w^{\boldsymbol{H}_i}$ converges to w_i as shock variance goes to infinity.

$$
w_i := \lim_{\Sigma_b \to \infty} w^{\boldsymbol{H}_i} = \frac{\sum_{j=1}^N (\tilde{\boldsymbol{p}}_j - \boldsymbol{c}_{i,j}) \boldsymbol{c}_{i,j} \tilde{\hat{\boldsymbol{H}}}_{i,j}}{\sum_{i=1}^{n_F} \sum_{j=1}^N (\tilde{\boldsymbol{p}}_j - \boldsymbol{c}_{i,j}) \boldsymbol{c}_{i,j} \tilde{\hat{\boldsymbol{H}}}_{i,j}} \Rightarrow M^m \text{ approaches } \sum_{i=1}^{n_F} w_i \alpha_i \Sigma_b + \sum_{i=1}^{n_F} w_i \gamma_i \tag{168}
$$

This proof held marginal costs $\tilde{c}$ fixed. If we assume the marginal cost of adjusting c is sufficiently high, by continuity, the inequality will still hold. □

C.2. A Model with Data as Private Information

For simplicity, we assumed that all firms see the signals of all other firms in the economy. In this appendix we solve a model with signals that are privately observed by one firm only. We compare the solution in the private and public signal models and find modest differences.

The only change to the setup of the main model is the information set. Firm i observes only the n_{di} data points generated by firm i, not the data produced by other firms. This is equivalent to conditioning expectations on the composite signal $\tilde{\mathbf{s}}_i$.

The first-order condition for firms still holds given their beliefs and strategies adopted by other firms. We denote the conditional expectation $\mathbf{E}_i(\cdot) = \mathbf{E}(\cdot|\mathcal{I}_i)$ for firm i. The inverse demand function is given by

$$\begin{aligned}
\boldsymbol{p}_i &= \bar{\boldsymbol{p}} + \boldsymbol{b}_i - \frac{1}{\phi}\sum_{j=1}^{n_F}\tilde{\boldsymbol{q}}_j \\
\Rightarrow \mathbf{E}\left[\boldsymbol{p}_i|\mathcal{I}_i\right] &= \bar{\boldsymbol{p}} + \mathbf{E}\left[\boldsymbol{b}_i|\mathcal{I}_i\right] - \frac{1}{\phi}\sum_{j=1}^{n_F}\mathbf{E}\left[\tilde{\boldsymbol{q}}_j|\mathcal{I}_i\right] \\
\Rightarrow \mathbf{E}_i\boldsymbol{p}_i &= \bar{\boldsymbol{p}} + \mathbf{E}_i\boldsymbol{b}_i - \frac{1}{\phi}\sum_{j=1}^{n_F}\mathbf{E}_i\tilde{\boldsymbol{q}}_j
\end{aligned} \tag{169}$$

So the optimal output in the incomplete information setup is

$$\begin{aligned}
\tilde{\boldsymbol{q}}_i &= \left(\rho_i\mathbf{Var}\left[\boldsymbol{p}_i|\mathcal{I}_i\right] - \frac{\partial\mathbf{E}\left[\boldsymbol{p}_i|\mathcal{I}_i\right]}{\partial\tilde{\boldsymbol{q}}_i}\right)^{-1}\left(\mathbf{E}\left[\boldsymbol{p}_i|\mathcal{I}_i\right] - \boldsymbol{c}_i\right) \\
\Rightarrow &\left(\rho_i\mathbf{Var}\left[\boldsymbol{p}_i|\mathcal{I}_i\right] - \frac{\partial\mathbf{E}\left[\boldsymbol{p}_i|\mathcal{I}_i\right]}{\partial\tilde{\boldsymbol{q}}_i}\right)\tilde{\boldsymbol{q}}_i = \mathbf{E}\left[\boldsymbol{p}_i|\mathcal{I}_i\right] - \boldsymbol{c}_i \\
\Rightarrow &\left(\rho_i\mathbf{Var}\left[\boldsymbol{p}_i|\mathcal{I}_i\right] + \frac{1}{\phi}\boldsymbol{I}_N\right)\tilde{\boldsymbol{q}}_i = \bar{\boldsymbol{p}} + \mathbf{E}_i\boldsymbol{b}_i - \frac{1}{\phi}\sum_{j=1}^{n_F}\mathbf{E}_i\tilde{\boldsymbol{q}}_j - \boldsymbol{c}_i \\
\Rightarrow &\left(\rho_i\mathbf{Var}\left[\boldsymbol{p}_i|\mathcal{I}_i\right] + \frac{2}{\phi}\boldsymbol{I}_N\right)\tilde{\boldsymbol{q}}_i = \bar{\boldsymbol{p}} + \mathbf{E}_i\boldsymbol{b}_i - \frac{1}{\phi}\sum_{j=1,j\neq i}^{n_F}\mathbf{E}_i\tilde{\boldsymbol{q}}_j - \boldsymbol{c}_i \\
\Rightarrow &\tilde{\boldsymbol{q}}_i = \boldsymbol{H}_i\left(\bar{\boldsymbol{p}} + \mathbf{E}_i\boldsymbol{b}_i - \boldsymbol{c}_i - \frac{1}{\phi}\sum_{j=1,j\neq i}^{n_F}\mathbf{E}_i\tilde{\boldsymbol{q}}_j\right),\ \forall\, i = 1,\ldots,n_F
\end{aligned} \tag{170}$$

Linear Equilibrium We first solve for a linear equilibrium in which optimal output is a linear function of signal. Suppose that the each firm follows a linear strategy of the form

$$\tilde{\boldsymbol{q}}_i = \boldsymbol{\alpha}_i + \boldsymbol{\gamma}_i\mathbf{E}_i\boldsymbol{b}_i = \boldsymbol{\alpha}_i + \boldsymbol{\gamma}_i\boldsymbol{K}_i\boldsymbol{s}_i \tag{171}$$

Then the optimal action function (170) across all firms is

$$\begin{aligned}
\tilde{\boldsymbol{q}}_i &= \boldsymbol{H}_i\left(\bar{\boldsymbol{p}} + \mathbf{E}_i\boldsymbol{b}_i - \boldsymbol{c}_i - \frac{1}{\phi}\sum_{j=1,j\neq i}^{n_F}\mathbf{E}_i\tilde{\boldsymbol{q}}_j\right) \\
\Rightarrow \boldsymbol{\alpha}_i + \boldsymbol{\gamma}_i\mathbf{E}_i\boldsymbol{b}_i &= \boldsymbol{H}_i\left(\bar{\boldsymbol{p}} + \mathbf{E}_i\boldsymbol{b}_i - \boldsymbol{c}_i - \frac{1}{\phi}\sum_{j=1,j\neq i}^{n_F}\mathbf{E}_i\left(\boldsymbol{\alpha}_j + \boldsymbol{\gamma}_j\mathbf{E}_j\boldsymbol{b}_j\right)\right) \\
\Rightarrow \boldsymbol{\alpha}_i + \boldsymbol{\gamma}_i\mathbf{E}_i\boldsymbol{b}_i &= \boldsymbol{H}_i\left(\bar{\boldsymbol{p}} + \mathbf{E}_i\boldsymbol{b}_i - \boldsymbol{c}_i - \frac{1}{\phi}\sum_{j=1,j\neq i}^{n_F}\boldsymbol{\alpha}_j\right) \\
\Rightarrow \boldsymbol{\alpha}_i - \boldsymbol{H}_i&\left(\bar{\boldsymbol{p}} - \boldsymbol{c}_i - \frac{1}{\phi}\sum_{j=1,j\neq i}^{n_F}\boldsymbol{\alpha}_j\right) + \left(\boldsymbol{\gamma}_i - \boldsymbol{H}_i\right)\mathbf{E}_i\boldsymbol{b}_i = 0,\ \forall\, i = 1,\ldots,n_F
\end{aligned} \tag{172}$$

Since the last equation holds for arbitrary $\mathbf{E}_i \boldsymbol{b}_i$, we must have

$$\begin{aligned}
\boldsymbol{\alpha}_i &= \boldsymbol{H}_i \left(\bar{\boldsymbol{p}} - \boldsymbol{c}_i - \frac{1}{\phi} \sum_{j=1, j\neq i}^{n_F} \boldsymbol{\alpha}_j \right) \\
\gamma_i &= \boldsymbol{H}_i = \hat{\boldsymbol{H}}_i \left(\boldsymbol{I}_N + \frac{1}{\phi} \hat{\boldsymbol{H}}_i \right)^{-1}
\end{aligned} \tag{173}$$

From the first equation we can solve for $\boldsymbol{\alpha}_i$ (similar to section 8)

$$\begin{aligned}
\boldsymbol{\alpha}_i &= \left(\boldsymbol{H}_i^{-1} - \frac{\boldsymbol{I}_N}{\phi} \right)^{-1} \left[\left(\boldsymbol{I}_N + \frac{1}{\phi} \sum_{j=1}^{n_F} \left(\boldsymbol{H}_j^{-1} - \frac{\boldsymbol{I}_N}{\phi} \right)^{-1} \right)^{-1} \left(\bar{\boldsymbol{p}} + \frac{1}{\phi} \sum_{j=1}^{n_F} \left(\boldsymbol{H}_j^{-1} - \frac{\boldsymbol{I}_N}{\phi} \right)^{-1} \boldsymbol{c}_j \right) - \boldsymbol{c}_i \right] \\
&= \hat{\boldsymbol{H}}_i \left[\left(\boldsymbol{I}_N + \frac{1}{\phi} \sum_{j=1}^{n_F} \hat{\boldsymbol{H}}_j \right)^{-1} \left(\bar{\boldsymbol{p}} + \frac{1}{\phi} \sum_{j=1}^{n_F} \hat{\boldsymbol{H}}_j \boldsymbol{c}_j \right) - \boldsymbol{c}_i \right]
\end{aligned} \tag{174}$$

Finally the equilibrium output $\tilde{\boldsymbol{q}}_i$ is given by

$$\begin{aligned}
\tilde{\boldsymbol{q}}_i &= \hat{\boldsymbol{H}}_i \left[\left(\boldsymbol{I}_N + \frac{1}{\phi} \sum_{j=1}^{n_F} \hat{\boldsymbol{H}}_j \right)^{-1} \left(\bar{\boldsymbol{p}} + \frac{1}{\phi} \sum_{j=1}^{n_F} \hat{\boldsymbol{H}}_j \boldsymbol{c}_j \right) - \boldsymbol{c}_i \right] + \boldsymbol{H}_i \mathbf{E}_i \boldsymbol{b}_i \\
&= \hat{\boldsymbol{H}}_i \left(\boldsymbol{D} - \boldsymbol{c}_i \right) + \hat{\boldsymbol{H}}_i \left(\boldsymbol{I}_N + \frac{1}{\phi} \hat{\boldsymbol{H}}_i \right)^{-1} \boldsymbol{K}_i \boldsymbol{s}_i
\end{aligned} \tag{175}$$

The equilibrium price and output are

$$\begin{aligned}
\mathbf{E}(\tilde{\boldsymbol{q}}_i) &= \hat{\boldsymbol{H}}_i \left(\boldsymbol{D} - \boldsymbol{c}_i \right) \\
\boldsymbol{p}_i &= \boldsymbol{D} + \boldsymbol{b}_i - \frac{1}{\phi} \sum_{j=1}^{n_F} \hat{\boldsymbol{H}}_j \left(\boldsymbol{I}_N + \frac{1}{\phi} \hat{\boldsymbol{H}}_j \right)^{-1} \boldsymbol{K}_j \boldsymbol{s}_j \Rightarrow \mathbf{E}(\boldsymbol{p}_i) = \boldsymbol{D}
\end{aligned} \tag{176}$$

In the case where there are two firms, we can prove that this equilibrium exists and is unique. Proof available on request. We omit it here for now because it is lengthy.

PRODUCT-LEVEL MARKUP The product-level markup for product k produced by firm i is $M_{ik}^p := \mathbf{E}[\boldsymbol{p}_i(j)]/\boldsymbol{c}_i(j)$. The average product-level markup is

$$\overline{M}^p = \frac{1}{N} \frac{1}{n_F} \sum_{i=1}^{n_F} \sum_{j=1}^{N} M_{ij}^p = \frac{1}{N} \frac{1}{n_F} \sum_{i=1}^{n_F} \sum_{j=1}^{N} \frac{\boldsymbol{D}(j)}{\boldsymbol{c}_i(j)} \tag{177}$$

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

$$\begin{aligned}
M_i^f &= \frac{\mathbf{E}[\tilde{\boldsymbol{q}}_i' \boldsymbol{p}]}{\mathbf{E}[\tilde{\boldsymbol{q}}_i' \boldsymbol{c}_i]} = \frac{\mathbf{E}[\tilde{\boldsymbol{q}}_i]' \mathbf{E}[\boldsymbol{p}] + \mathbf{tr}\left[\mathbf{Cov}(\boldsymbol{p}_i, \tilde{\boldsymbol{q}}_i)\right]}{\mathbf{E}[\tilde{\boldsymbol{q}}_i' \boldsymbol{c}_i]} \\
&= \frac{(\boldsymbol{D} - \boldsymbol{c}_i)' \hat{\boldsymbol{H}}_i \boldsymbol{D} + \mathbf{tr}\left(\left(\boldsymbol{I}_N - \frac{\boldsymbol{H}_i}{\phi} \right) \boldsymbol{K}_i \mathbf{Var}(\boldsymbol{s}_i) \boldsymbol{K}_i' \boldsymbol{H}_i \right)}{(\boldsymbol{D} - \boldsymbol{c}_i)' \hat{\boldsymbol{H}}_i \boldsymbol{c}_i} > \frac{(\boldsymbol{D} - \boldsymbol{c}_i)' \hat{\boldsymbol{H}}_i \boldsymbol{D}}{(\boldsymbol{D} - \boldsymbol{c}_i)' \hat{\boldsymbol{H}}_i \boldsymbol{c}_i}
\end{aligned} \tag{178}$$

Thus, the average firm-level markup is $\overline{M}^f = (1/n_F)\sum_{i=1}^{n_F} M_i^f$.

INDUSTRY MARKUP The industry markup is

$$M^m := \frac{\mathbf{E}\left[\sum_{i=1}^{n_F} \tilde{\boldsymbol{q}}_i' \boldsymbol{p}_i\right]}{\mathbf{E}\left[\sum_{i=1}^{n_F} \tilde{\boldsymbol{q}}_i' \boldsymbol{c}_i\right]} = \frac{\sum_{i=1}^{n_F} \mathbf{E}\left[\tilde{\boldsymbol{q}}_i' \boldsymbol{p}_i\right]}{\sum_{i=1}^{n_F} \mathbf{E}\left[\tilde{\boldsymbol{q}}_i' \boldsymbol{c}_i\right]} = \sum_{i=1}^{n_F} w^{\boldsymbol{H}_i} M_i^f \quad \text{where} \quad w^{\boldsymbol{H}_i} = \frac{\mathbf{E}\left[\tilde{\boldsymbol{q}}_i' \boldsymbol{c}_i\right]}{\sum_{i=1}^{n_F} \mathbf{E}\left[\tilde{\boldsymbol{q}}_i' \boldsymbol{c}_i\right]}. \tag{179}$$

Private Information Model: Cyclical Markup Behavior

Proposition 9. *The product-level markup converges as shock variance goes to infinity given identical risk aversion and signal precision across all firms.*

Proof. We analyze an economy consisted of identical firms with diagonal firm and shock variance matrices. The price impact $\hat{\boldsymbol{H}}_i$ satisfies the following equation

$$\begin{aligned}
\hat{\boldsymbol{H}}_i &= \left[\rho_i \mathbf{Var}(\boldsymbol{b}_i|\mathcal{I}_i) + \frac{\boldsymbol{I}_N}{\phi} + \rho_i(n_F-1)\frac{\hat{\boldsymbol{H}}_i}{\phi}\left(\boldsymbol{I}_N + \frac{1}{\phi}\hat{\boldsymbol{H}}_i\right)^{-1} \mathbf{Var}(\boldsymbol{K}_i\boldsymbol{s}_i)\left(\boldsymbol{I}_N + \frac{1}{\phi}\hat{\boldsymbol{H}}_i\right)^{-1}\frac{\hat{\boldsymbol{H}}_i}{\phi}\right]^{-1} \\
\Rightarrow \hat{\boldsymbol{H}}_{i,k}^{-1} &= \rho_i \frac{\Sigma_{b,k}\Sigma_{\epsilon_i,k}}{\Sigma_{b,k}+\Sigma_{\epsilon_i,k}} + \frac{1}{\phi} + \rho_i(n_F-1)\left(\frac{\hat{\boldsymbol{H}}_{i,k}}{\phi+\hat{\boldsymbol{H}}_{i,k}}\right)^2 \frac{\Sigma_{b,k}^2}{\Sigma_{b,k}+\Sigma_{\epsilon_i,k}},\ \forall k = 1,\ldots,N
\end{aligned} \tag{180}$$

Taking derivative with respect to $\Sigma_{b,k}$ for both sides, we have

$$\begin{aligned}
-\hat{\boldsymbol{H}}_{i,k}^{-2}\frac{\partial \hat{\boldsymbol{H}}_{i,k}}{\partial \Sigma_{b,k}} &= \frac{\rho_i \Sigma_{\epsilon_i,k}^2}{\Sigma_{b,k}+\Sigma_{\epsilon_i,k}} + \rho_i(n_F-1)\left[\frac{\Sigma_{b,k}^2}{\Sigma_{b,k}+\Sigma_{\epsilon_i,k}}\frac{2\phi\hat{\boldsymbol{H}}_{i,k}}{(\hat{\boldsymbol{H}}_{i,k}+\phi)^3}\frac{\partial \hat{\boldsymbol{H}}_{i,k}}{\partial \Sigma_{b,k}} + \left(\frac{\hat{\boldsymbol{H}}_{i,k}}{\phi+\hat{\boldsymbol{H}}_{i,k}}\right)^2 \frac{\Sigma_{b,k}(\Sigma_{b,k}+2\Sigma_{\epsilon_i,k})}{\Sigma_{b,k}+\Sigma_{\epsilon_i,k}}\right] \\
-\left(\hat{\boldsymbol{H}}_{i,k}^{-2} + \rho_i(n_F-1)\frac{\Sigma_{b,k}^2}{\Sigma_{b,k}+\Sigma_{\epsilon_i,k}}\frac{2\phi\hat{\boldsymbol{H}}_{i,k}}{(\hat{\boldsymbol{H}}_{i,k}+\phi)^3}\right)\frac{\partial \hat{\boldsymbol{H}}_{i,k}}{\partial \Sigma_{b,k}} &= \frac{\rho_i \Sigma_{\epsilon_i,k}^2}{\Sigma_{b,k}+\Sigma_{\epsilon_i,k}} + \rho_i(n_F-1)\left(\frac{\hat{\boldsymbol{H}}_{i,k}}{\phi+\hat{\boldsymbol{H}}_{i,k}}\right)^2 \frac{\Sigma_{b,k}(\Sigma_{b,k}+2\Sigma_{\epsilon_i,k})}{\Sigma_{b,k}+\Sigma_{\epsilon_i,k}}
\end{aligned} \tag{181}$$

The derivative $\frac{\partial \hat{\boldsymbol{H}}_{i,k}}{\partial \Sigma_{b,k}}$ is clearly negative, implying convergent $\hat{\boldsymbol{H}}_{i,k}$ (decreasing and non-negative) as shock variance goes to infinity. Furthermore, $\hat{\boldsymbol{H}}_{i,k}$ must converges to zero, otherwise the RHS of equation (180) is unbounded while the LHS is bounded. The product-level markup $\overline{M}^p$ is convergent:

$$\overline{M}^p = \frac{1}{N}\frac{1}{n_F}\sum_{i=1}^{n_F}\sum_{j=1}^{N} M_{ij}^p = \frac{1}{N}\frac{1}{n_F}\sum_{i=1}^{n_F}\sum_{j=1}^{N}\frac{\mathbf{E}(\boldsymbol{p}_{i,j})}{\boldsymbol{c}_{i,j}} \quad \text{and} \quad \lim_{\Sigma_b\to\infty}\overline{M}^p = \frac{1}{N}\frac{1}{n_F}\sum_{i=1}^{n_F}\sum_{j=1}^{N}\frac{\bar{\boldsymbol{p}}_j}{\boldsymbol{c}_{i,j}} \tag{182}$$

since $\mathbf{E}\left[\boldsymbol{p}_i\right] = \boldsymbol{D} = \left(\boldsymbol{I}_N + \frac{1}{\phi}\sum_{i=1}^{n_F}\hat{\boldsymbol{H}}_i\right)^{-1}\left(\bar{\boldsymbol{p}} + \frac{1}{\phi}\sum_{i=1}^{n_F}\hat{\boldsymbol{H}}_i\boldsymbol{c}_i\right)$ and $\lim_{\Sigma_b\to\infty}\mathbf{E}\left[\boldsymbol{p}_i\right] = \bar{\boldsymbol{p}}$

□

Proposition 10. *The firm-level and economy-level markups are strictly increasing if the shock variance is large enough, and approach their linear asymptotes.*

Proof. The firm-level markup for firm i is the quantity-weighted prices divided by quantity-weighted costs:

$$
\begin{aligned}
M_i^f &= \frac{\mathbf{E}[\tilde{\boldsymbol{q}}_i' \boldsymbol{p}_i]}{\mathbf{E}[\tilde{\boldsymbol{q}}_i' \boldsymbol{c}_i]} = \frac{\mathbf{E}[\tilde{\boldsymbol{q}}_i]'\mathbf{E}[\boldsymbol{p}] + \mathbf{tr}\left[\mathbf{Cov}(\boldsymbol{p}_i, \tilde{\boldsymbol{q}}_i)\right]}{\mathbf{E}[\tilde{\boldsymbol{q}}_i' \boldsymbol{c}_i]} \\
&= \frac{(\boldsymbol{D} - \boldsymbol{c}_i)' \hat{\boldsymbol{H}}_i \boldsymbol{D} + \mathbf{tr}\left(\left(\boldsymbol{I}_N + \frac{1}{\phi}\hat{\boldsymbol{H}}_i\right)^{-1} \mathbf{Var}(\boldsymbol{K}_i \boldsymbol{s}_i) \left(\boldsymbol{I}_N + \frac{1}{\phi}\hat{\boldsymbol{H}}_i\right)^{-1} \hat{\boldsymbol{H}}_i\right)}{(\boldsymbol{D} - \boldsymbol{c}_i)' \hat{\boldsymbol{H}}_i \boldsymbol{c}_i}
\end{aligned}
\tag{183}
$$

We further assume all the products have same shock Σ_b and signal variance, thus ensuring the same diagonal values of $\hat{\boldsymbol{H}}_i$. The M_i^f can be simplified as

$$
M_i^f = \frac{\sum_{j=1}^N (\boldsymbol{D}_j - \boldsymbol{c}_{i,j})\boldsymbol{D}_j + \sum_{j=1}^N \left(1 + \frac{1}{\phi}\hat{\boldsymbol{H}}_{i,j}\right)^{-2} \frac{\Sigma_{b,j}^2}{\Sigma_{b,j} + \Sigma_{\epsilon_i,j}}}{\sum_{j=1}^N (\boldsymbol{D}_j - \boldsymbol{c}_{i,j})\boldsymbol{c}_{i,j}}
\tag{184}
$$

M_i^f also admits an asymptote as

$$
\alpha_i := \lim_{\Sigma_b \to \infty} \frac{M_i^f}{\Sigma_b} = \frac{N}{\sum_{j=1}^N (\bar{\boldsymbol{p}}_j - \boldsymbol{c}_{i,j})\boldsymbol{c}_{i,j}} \quad \text{and} \quad \gamma_i := \lim_{\Sigma_b \to \infty} \left(M_i^f - \alpha_i \Sigma_b\right) = \frac{\sum_{j=1}^N \left((\bar{\boldsymbol{p}}_j - \boldsymbol{c}_{i,j})\bar{\boldsymbol{p}}_j - \Sigma_{\epsilon_i,j}\right)}{\sum_{j=1}^N (\bar{\boldsymbol{p}}_j - \boldsymbol{c}_{i,j})\boldsymbol{c}_{i,j}}
\tag{185}
$$

The average firm-level markup $\overline{M}^f = (1/n_F)\sum_{i=1}^{n_F} M_i^f$ approaches $\sum_{i=1}^{n_F} \frac{\alpha_i}{n_F}\Sigma_b + \sum_{i=1}^{n_F} \frac{\gamma_i}{n_F}$ in the long run. The economy-level markup is $M^m = \sum_{i=1}^{n_F} w^{\boldsymbol{H}_i} M_i^f$ with $w^{\boldsymbol{H}_i} = \frac{\mathbf{E}\left[\tilde{\boldsymbol{q}}_i' \boldsymbol{c}_i\right]}{\sum_{i=1}^{n_F} \mathbf{E}\left[\tilde{\boldsymbol{q}}_i' \boldsymbol{c}_i\right]}$. The weight $w^{\boldsymbol{H}_i}$ converges to w_i as shock variance goes to infinity.

$$
w_i := \lim_{\Sigma_b \to \infty} w^{\boldsymbol{H}_i} = \frac{\sum_{j=1}^N (\bar{\boldsymbol{p}}_j - \boldsymbol{c}_{i,j})\boldsymbol{c}_{i,j}}{\sum_{i=1}^{n_F}\sum_{j=1}^N (\bar{\boldsymbol{p}}_j - \boldsymbol{c}_{i,j})\boldsymbol{c}_{i,j}} \Rightarrow M^m \text{ approaches } \sum_{i=1}^{n_F} w_i \alpha_i \Sigma_b + \sum_{i=1}^{n_F} w_i \gamma_i
\tag{186}
$$

□

C.3. 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

$$
max_{a_i, q_i} \mathbf{E}\left[q_i \boldsymbol{a}_i' \left(\tilde{\mathbf{p}} - \boldsymbol{c}_i\right) | \mathcal{I}_i\right] - \frac{\rho_i}{2} \mathbf{Var}\left[q_i \boldsymbol{a}_i' \left(\tilde{\mathbf{p}} - \boldsymbol{c}_i\right) | \mathcal{I}_i\right] - g(\chi_c, \boldsymbol{c}_i),
\tag{187}
$$

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 1

$$max_{c_i}\mathbf{E}\left[\ \mathbf{E}\left[q_i\boldsymbol{a}_i'\left(\tilde{\mathbf{p}}-\boldsymbol{c}_i\right)|\mathcal{I}_i\right]-\frac{\rho_i}{2}\mathbf{Var}\left[q_i\boldsymbol{a}_i'\left(\tilde{\mathbf{p}}-\boldsymbol{c}_i\right)|\mathcal{I}_i\right]\right]-g(\chi_c,\boldsymbol{c}_i). \tag{188}$$

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.

$$q_i\boldsymbol{a}_i=\left(\rho_i\mathbf{Var}\left[\boldsymbol{p}_i|\mathcal{I}_i\right]+\frac{\partial\mathbf{E}\left[\boldsymbol{p}_j|\mathcal{I}_j\right]}{\partial\boldsymbol{q}_i}\right)^{-1}\left(\mathbf{E}\left[\boldsymbol{p}_i|\mathcal{I}_i\right]-\boldsymbol{c}_i\right) \tag{189}$$

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

Proof. Comparing first-order condition (189) with original optimal choice (34), we could solve this extension model by substituting $\tilde{\boldsymbol{q}}_i$ in (34) with $q_i\boldsymbol{a}_i$ and further extend existing propositions for q_i and $\boldsymbol{a}_i$ by one-to-one mapping

$$q_i=\sum_{j=1}^{N}\tilde{\boldsymbol{q}}_{i,j}\quad\text{and}\quad\boldsymbol{a}_i=\frac{\tilde{\boldsymbol{q}}_i}{\sum_{j=1}^{N}\tilde{\boldsymbol{q}}_{i,j}} \tag{190}$$

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

$$M_i^p:=\frac{\mathbb{E}[\boldsymbol{a}_i'\tilde{\boldsymbol{p}}_i]}{\mathbb{E}[\boldsymbol{a}_i'\tilde{\boldsymbol{c}}_i]}=\frac{\mathbb{E}[q_i\boldsymbol{a}_i'\tilde{\boldsymbol{p}}_i]}{\mathbb{E}[q_i\boldsymbol{a}_i'\tilde{\boldsymbol{c}}_i]}=\frac{\mathbb{E}[\tilde{\boldsymbol{q}}_i'\tilde{\boldsymbol{p}}_i]}{\mathbb{E}[\tilde{\boldsymbol{q}}_i'\tilde{\boldsymbol{c}}_i]}=M_i^f \tag{191}$$

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 4), 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.

References

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

Aghion, Philippe, Antonin Bergeaud, Timo Boppart, Peter Klenow, and Huiyu Li, "A Theory of Falling Growth and Rising Rents," NBER Working Paper 26448, 2019.

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.

Athey, Susan, and Joshua S. Gans, "The Impact of Targeting Technology on Advertising Markets and Media Competition," *American Economic Review*, 100 (2010), 608–613.

Athey, Susan, Mark Mobius, and Jeno Pal, "The Impact of News Aggregators on Internet News Consumption: the Case of Localization," Stanford Technical Report 3353, 2017.

Autor, David, David Dorn, Lawrence F Katz, Christina Patterson, and John Van Reenen, "The Fall of the Labor Share and the Rise of Superstar Firms," *The Quarterly Journal of Economics*, 135 (2020), 645–709.

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.

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.

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

Brynjolfsson, Eric, Y Hu, and M Smith, “Consumer Surplus in the Digital Economy: Estimating the Value of Increased Product Variety at Online Booksellers,” *Management Science*, 49 (2003), 1580–1596.

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

de Cornière, Alexandre, and Greg Taylor, “Data and Competition: a General Framework with Applications to Mergers, Market Structure, and Privacy Policy,” TSE Working Paper 1076, 2020.

Covarrubias, Matias, Germán Gutiérrez, and Thomas Philippon, “From Good to Bad Concentration? US Industries over the Past 30 Years,” *NBER Macroeconomics Annual*, 34 (2020), 1–46.

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.

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.

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

Goldfarb, Avi, and Catherine E. Tucker, "Digital Economics," NBER Working Paper 23684, 2017.

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

Gutiérrez, Germán, and Thomas Philippon, "Investment-less Growth: An Empirical Investigation," NBER Working Paper 22897, 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.

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

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

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

Lambrecht, Anja, and Catherine E. Tucker, "Can Big Data Protect a Firm from Competition?", mimeoTechnical report, 2015.

Lashkari, Danial, Arthur Bauer, and Jocelyn Boussard, "Information Technology and Returns to Scale," Boston College Mimeo, 2018.

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.

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," WRDS Research Paper, 2020.

Rajgopal, Shivaram, Anup Srivastava, and Rong Zhao, "Do Digital Technology Firms Earn Excess Profits? An Alternative Perspective," mimeoTechnical report, 2021.

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).

Veldkamp, Laura, *Information Choice in Macroeconomics and Finance* (Princeton, NJ: Princeton University Press, 2011).

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

Weiss, Christoph, "Mark-ups, Industry Structure and the Business Cycle," *Applied Economics Letters*, 7 (2000), 189–191.

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