Data and Market Power*

Jan Eeckhout
UPF Barcelona†

Laura Veldkamp
Columbia‡

August 5, 2021

Abstract

Might firms' use of data create market power? To explore this hypothesis, 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 larger scale. However, the model uncovers much richer interactions between data, welfare and market power. Data affects risk, firm size and the composition of the goods firm produce, all of which affect markups. The tradeoff between these forces depends on the level of aggregation at which markups are measured. Empirical researchers who measure markups at the product level, firm level or industry level come to different conclusions about trends and cyclical fluctuations in markups. Our results reconcile and re-interpret these facts. The divergence between product, firm and industry markups can be a sign that firms are using data efficiently to produce the goods consumers want most.

Keywords. Market Power. Data. Risk. Technological Change. Market Structure. Endogenous Markups.

JEL. C6. D4. D5. L1.

*We thank Huasheng Nie and Renjie Bao for excellent research assistance. We also thank Ariel Pakes and Hugo Hopenhayn for useful discussions that led to this project. Eeckhout gratefully acknowledges support from the ERC, Advanced grant 882499, and from ECO2015-67655-P.Preliminary. Please do not circulate.

† jan.eeckhout@upf.edu – ICREA-GSE-CREi

‡ lv2405@columbia.edu

Changes in firms' market power and the sources of those changes have become the focus of intense debate. Economists point to 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). To explore this hypothesis, 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 larger scale. As a result, the data-rich firm exerts more market power. However, a simple model that embodies this logic uncovers a much richer set of interactions between data and market power measures. We find that data has competing effects on markups and that the tradeoff between these competing effects depends on the level of aggregation at which markups are measured. The results help to resolve a dispute in the empirical literature between researchers who measure markups at the product, firm or industry level and come to different conclusions about the cyclicality of markups and trends in competition.

In order to explore the hypothesis that increasing returns to scale in information or data is creating market power, Section 1 formulates a new framework, where data is modeled as information. The essence of information is that it is something that can reduce uncertainty or risk. For such risk-reduction to matter, we need a setting with firms that price risk, or are averse to risk. When data helps firms resolve their risk, risk-averse firms are emboldened to invest more and grow larger. This is the assumption that will link data and firm size, which is a key part of the data competition hypothesis that we set out to explore. Assuming that firms price risk is unusual in the firm competition literature, but has a long history in corporate finance (Brealey et al., 2003). For decades, every major MBA program has taught future firm managers to evaluate investments, accounting for the price of risk.

When firms price risk and data is used to formulate more accurate forecasts that avoid risk, then data makes risky investment less costly for a firm. In our setting, firms can make 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 that the firm grows larger, produces more at lower marginal cost and earns higher markups. This force whereby data increases markups is what we call the "investment channel." The idea that data makes firms large and profitable is standard logic in the debate about data and competition.

The model teaches us that investment is not the only force linking data and markups. When data reduces risk, this also encourages firms to produce more. Increasing production pushes prices and markups down. This downward force of data on markups is what we call the "risk channel." Section 2 derives the investment effect and the risk premium effect, shows that they move markups in opposite directions and derives conditions under which each force dominates.

An auxiliary prediction of the model is that the growing volume of data can cause product markups to fall, while firm and industry markups rise. The reason is that data affects 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. In other words, markups measured at different levels of aggregation have composition effects, and data creates and strengthens those composition effects. Section 2 shows that the more data firms acquire, the more of a wedge will arise between product-level markups and firm-level markups. When firms' data pushes up the firm markup, relative to the product markup, we call that a "firm-level aggregation" effect.

Data also causes the markup of an average firm and its industry to diverge. When firms can make an investment that lowers their marginal cost of production, 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 therefore are weighted more heavily in the industry markup. When growing data widens the gap between industry and firm markups, this "industry aggregation" effect of data captures classic concerns about competition. 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.

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 Gutierrez and Philippon (2016), Furman and Orszag (2015), Grullon et al. (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)). Our analysis finds some truth in each view, but concludes that none of the markup measures alone is sufficient to draw welfare conclusions. Data does grow market power and create distortions. But it also allows firms to operate more efficiently, by producing more of the goods that consumers want most.

Another unexpected prediction of this model may help to reconcile an empirical debate about whether markups are pro- or counter-cyclical. This debate is central to the relevance of New Keynesian models. We find that data-intensive firms may have procyclical product markups, as in Ramey and Nakarda (2020), but counter-cyclical firm and industry markups Bils (1985, 1987). In section 5, 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 adjust output more to increase the firm's markup and profit by more. In short, higher volatility raises firm and industry markups because it creates a potential for larger composition effects. Understanding why markups measured at different levels of aggregation have different cyclical properties allows researchers to determine which set of facts is most relevant for a given question.

Ultimately, most researchers are interested in markups because they are concerned about consumer welfare. Section 3 discusses the relationship between markups, competitive outcomes and welfare. Rising amounts of data can be good consumers. After all, firms use data to produce more of goods that consumers want most. Welfare suffers when firms' data stocks become asymmetric. When data-rich firms get more data, they grow larger, exert more market power and can harm consumers.

Our model primarily contributes new thinking about competition in the digital age. It also offers new approaches to measurement. Section 6 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 difference between the firm and product markup is a sufficient statistic for the amount of relevant data a profit-maximizing firm has about consumer demand. This would enable a reduced-form approach to measuring the amount of data a firm has on the data asymmetry in an industry. The model could also be used for structural estimation. For this purpose, we also discuss different approaches to measuring hedonic product characteristics, techniques to estimate firms' price of risk, and we map the markup measures in our model to different empirical approaches in the markup literature. Both appoaches could be helpful to study firm competition, or a host of other questions related to firms' ownership of and use of consumer data.

Related literature Existing work on the digital economy does explore whether data can be a source of market power. In Kirpalani and Philippon (2020), data enables directed two-sided search.

Acemoglu et al. (2021) and Bergemann and Bonatti (2019) model data as information and explore whether data markets are efficient. Ichihashi (2020) show how firms can use consumer data to price discriminate, while Liang and Madsen (2021) explore the use of data in labor markets. 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. Empirical work on the data economy often, necessarily focuses on specific markets.[1] None of these analyses explore data's effect on firms' risk, or considers product markups at various levels of aggregation.

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 Edmond, Midrigan, and Xu (2019) and De Loecker, Eeckhout, and Mongey (2021)). Those models offer tools to evaluate the welfare consequences of markups. Our approach differs because we explore data specifically.

Our predictions are consistent with the superstar firm economy, as described in Autor et al. (2020). 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 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.

Rossi-Hansberg, Sarte, and Trachter (2018) have pointed out yet another divergence in measures of market power, namely between local and national markets. However, that difference in measures of market power is not expressed in markups but in concentration indices such as HHI (Herfindahl-Hirschman Index). Expressed in markups, there is no divergence between market power in local and national markets.[2]

Our approach is related to work on banking competition (Vives and Ye, 2021). Because the banking literature has long been informed by finance, it is more common to model banks who price risk. However, there are also relevant differences. Information is used by banks, not to forecast demand, but to forecast loan repayment. Banks don't choose which loans to produce or

[1] Athey et al. (2017); Athey (2010) examine media competition, Brynjolfsson et al. (2003) study booksellers and Rajgopal et al. (2021) measure digital technology firms. de Cornière and Taylor (2020) categorize uses of data as pro- or anti-competitive.

[2] Benkard, Yurukoglu, and Zhang (2021) argue that HHI is defined over the market where consumers are located, whereas data used to measure HHI is based on the location of production, which leads to misleading and inconsistent findings when aggregating. Eeckhout (2020) shows the discrepancy stems from a mechanical relation between population size and the market definition.

what their marginal cost of banking will be. While important features differ, many of the efficiency benefits and competition concerns surrounding data are similar.

1 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. Second, data is used to resolve this uncertainty. Data is informative about what demand will be. Third, firms face a cost of bearing risk. This price of risk is what links data and uncertainty to investment. Fourth, in order to explore the relationship between data and the composition of the goods a firm produces, it is useful if production is something deterministic that firms choose. Therefore, we model firms that choose quantities of multiple goods. Allowing those goods to have correlated attributes, as in Pelligrino (2020), makes data relevant to multiple goods. Finally, since the data competition hypothesis was 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 explore these features in a static model because dynamics are not essential to elucidate the mechanisms we consider. However, such a model could easily repeated.

1.1 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, vector $\mathbf{q}_i$, to maximize risk-adjusted profit, where the price of risk is ρ_i.

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

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. This is investment is modeled as a choice of marginal cost vector $\boldsymbol{c}_i$, at an investment cost $g(\chi_c, \widetilde{\boldsymbol{c}}_i)$, to maximize $E[U_i]$. We interpret lower choices of $\boldsymbol{c}_i$ as larger investments.

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 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 a matrix A. For now, the mapping between attributes and products is fixed. In fact, one could set $A = I$, equate attributes and products and most of the results would be unchanged. But this mapping allows us to consider that data might be relevant for multiple products. Later, we will allow firms to choose how to position their product in the product space.

The marginal cost of producing a good depends on the up-front investment the firm makes and on the 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 eac 'h 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 $\tilde{p}$:

$$p_k = \sum_{j=1}^{N} a_{jk}\tilde{p}_j \tag{3}$$

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

$$\tilde{p}_j = \bar{p}_j - \frac{1}{\phi}\sum_{i=1}^{N}\tilde{q}_{ij} \tag{4}$$

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}:

$$\tilde{\mathbf{q}}_i = A^{-1}\mathbf{q}_i \tag{5}$$

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 $\tilde{p}_j + b_{ij}$. 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: $\tilde{\mathbf{s}}_i = \boldsymbol{b}_i + \tilde{\boldsymbol{\varepsilon}}_i$, where $\tilde{\boldsymbol{\varepsilon}}_i \sim N(\mathbf{0}, \tilde{\Sigma}_{\epsilon_i})$ 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 $\tilde{\Sigma}_{\epsilon_i}$) 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.

Equilibrium

1. Each firm sequentially chooses a vector of marginal costs $\boldsymbol{c}_i$, 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 (1.1).

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

3. Prices clear the market for each good.

1.2 Discussion of assumptions

Data that is public information. The assumption that all data is public is obviously not realistic. It is also not crucial for any of our main results. It does simplify the mathematics considerably. One interpretation of this assumption is that firms can choose output conditional on the average price. However, public signals are not essential. In a model with private signals, firms also used data to forecast what other firms will do. Data reduces risk in two ways – about the firm's demand and about the production decisions of other firms. Appendix C.6 shows that similar results arise because data still reduces uncertainty, which prompts 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 an implicit solution to a set of non-linear equations. We can prove theoretical properties using the implicit function theorem. But they are less clear and thus less useful for expositing the ideas we wish to convey.

Firms that price risk. What is essential is the assumption that firms price risk. Even if firms themselves are not risk-averse, firms that take on risky projects will face higher cost of capital. So the price of risk term could be interpreted as an adjustment to their expected profit. Standard MBA curricula typically teach managers to set their price of risk ρ to match the risk-premium on an equity index like the S&P 500. The idea is that if a firm gets less return per unit of risk than this, the firm would be better off not investing in production and instead investing the firm's cash in a market portfolio of equity.

Typically, firms only price aggregate risk. However, we are studying markets with a small number of large firms. Firm-specific risk is not idiosyncratic. It is not easily diversifiable. Therefore, this risk should be priced. In addition, there is growing evidence that managers do price idiosyncratic risk, especially when firms face financial constraints (Whited and HENNESSY, 2007). Also, since we know that the model with aggregate risk delivers similar results (Appendix C.1), 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, whose risk cannot be fully hedged. In that case, ρ_i might vary by firm.

Data about consumer demand. Finally, 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 capital. 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.

Goods as bundles of attributes. 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 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.

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

No attribute choice. Appendix C.9 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 to elucidate the main ideas more clearly.

No data choice. Our main question is what the effect of data is on competition. To answer that question, it makes sense to take data as exogenous and ask what happens when the amount of data changes. However, future work with different objectives might investigate determinants of firms data choices.

1.3 Solution

We solve the model by backwards induction, starting with the quantity choices, for given production costs 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 $\tilde{\Sigma}_{\epsilon_i}$ is the same as observing the average signal $\boldsymbol{s}_i = (1/n_{di}) \sum_{m=1}^{n_{di}} s_{im} = \boldsymbol{b}_i + \boldsymbol{\varepsilon}_i$, where the variance of $\boldsymbol{\varepsilon}_i$ is $\Sigma_{\epsilon_i} = \tilde{\Sigma}_{\epsilon_i}/n_{di}$. Therefore, we proceed as if each firm observes the one composite signal with lower signal noise. 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 (4) are

$$\begin{aligned}
\mathbf{E}\left[\boldsymbol{p}_i|\mathcal{I}_i\right] &= \bar{\boldsymbol{p}} + \mathbf{E}\left[\boldsymbol{A}\boldsymbol{b}_i|\mathcal{I}_i\right] - \frac{1}{\phi}\left(\boldsymbol{q}_i + \boldsymbol{q}_j\right) \\
&= \bar{\boldsymbol{p}} + \boldsymbol{A}\boldsymbol{K}_i\boldsymbol{s}_i - \frac{1}{\phi}\left(\boldsymbol{q}_i + \boldsymbol{q}_j\right) \\
\mathbf{Var}\left[\boldsymbol{p}_i|\mathcal{I}_i\right] &= \boldsymbol{A}\mathbf{Var}\left[\boldsymbol{b}_i|\mathcal{I}_i\right]\boldsymbol{A}' \\
&= \boldsymbol{A}\left(I_N + \Sigma_{\epsilon_i}\right)^{-1}\Sigma_{\epsilon_i}\boldsymbol{A}'
\end{aligned} \tag{6}$$

Optimal Production The first-order condition with respect to 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:

$$\boldsymbol{q}_i = \left(\rho_i \mathbf{Var}\left[\boldsymbol{p}_i|\mathcal{I}_i\right] - \frac{\partial \mathbf{E}\left[\boldsymbol{p}_i|\mathcal{I}_i\right]}{\partial \boldsymbol{q}_i}\right)^{-1}\left(\mathbf{E}\left[\boldsymbol{p}_i|\mathcal{I}_i\right] - \boldsymbol{c}_i\right) \tag{7}$$

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

$$\frac{\partial \mathbf{E}\left[\tilde{\boldsymbol{p}}_i|\mathcal{I}_i\right]}{\partial \tilde{\boldsymbol{q}}_i} = -\frac{1}{\phi} I_N. \tag{8}$$

Substituting the conditional expectation (6) and price impact (8) into the first-order condition (7) yields

$$\boldsymbol{q}_i^* = \hat{\boldsymbol{H}}_i \left(\boldsymbol{A}\overline{\boldsymbol{p}} + \boldsymbol{A}\boldsymbol{K}_i \boldsymbol{s}_i - \frac{1}{\phi}\sum_j \boldsymbol{q}_j - \boldsymbol{c}_i \right) \tag{9}$$

where $\hat{\boldsymbol{H}}_i := \left(\frac{1}{\phi} I_N + \rho_i \mathbf{Var}\left[\boldsymbol{b}_i|\mathcal{I}_i\right]\right)^{-1}$..

Note that (9) has firm i's optimal quantity both on the left and the right-hand sides. Collecting q_i terms on the left and writing in terms of attribute quantities yields the optimal production choices of each firm i, as a function of the production choices of all other firms:

$$\tilde{\boldsymbol{q}}_i = \boldsymbol{A}^{-1}\boldsymbol{H}_i\boldsymbol{A} \left(\bar{\boldsymbol{p}} + \boldsymbol{K}_i \boldsymbol{s}_i - \frac{1}{\phi}\sum_{j=1, j\neq i}^{nF} \tilde{\boldsymbol{q}}_j - \widetilde{\boldsymbol{c}}_i \right) \tag{10}$$

$$\text{where} \qquad \boldsymbol{H}_i := \left(\frac{2}{\phi} I_N + \rho_i \mathbf{Var}\left[\boldsymbol{b}_i|\mathcal{I}_i\right]\right)^{-1} \tag{11}$$

is the sensitivity of production to a change in the expected price. This sensitivity or supply elasticity captured by $\boldsymbol{H}_i$ will play a key role in equilibrium markups.

To express the solution in terms of model primitives, we compute total output and the subgame perfect output of each firm, in terms of the signals they observe. If we sum production (9) over all firms i, we get total production of each attribute: $\left(I + \frac{1}{\phi}\sum_i \boldsymbol{A}^{-1}\hat{\boldsymbol{H}}_i\boldsymbol{A}\right)^{-1} \left[\sum_i \boldsymbol{A}^{-1}\hat{\boldsymbol{H}}_i\boldsymbol{A}\left(\overline{\boldsymbol{p}} + \boldsymbol{K}_i\boldsymbol{s}_i - \widetilde{\boldsymbol{c}}_i\right)\right]$. Substituting this total production expression for $\sum_{i=1}^{n_F} \tilde{\boldsymbol{q}}_i$ in firm i's optimal production (9) yields the optimal production of each attribute by each firm i:[3]

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

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

[3]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.

Equilibrium Price Substituting this aggregate quantity in the pricing function (4) yields an equilibrium average price of each attribute:

$$\tilde{p} = \bar{p} - \left(\phi I + \sum_i \boldsymbol{A}^{-1}\hat{\boldsymbol{H}}_i\boldsymbol{A}\right)^{-1}\left[\sum_i \boldsymbol{A}^{-1}\hat{\boldsymbol{H}}_i\boldsymbol{A}\left(\overline{\boldsymbol{p}} + \boldsymbol{K}_i\boldsymbol{s}_i - \widetilde{\boldsymbol{c}}_i\right)\right]. \quad (12)$$

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

Optimal Investment Choices We begin with a firm that moves last, taking all other firms' investment choices as given. Then, we explore the choice of the previous movers. Firm i chooses cost $\boldsymbol{c}_i$ to maximize its unconditional expected utility $\mathbf{E}[U_i]$.

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

$$\frac{\partial \mathbf{E}[U_i]}{\partial \boldsymbol{c}_i} = \frac{1}{2}\frac{\partial \mathbf{E}[\boldsymbol{q}_i]'\boldsymbol{H}_i^{-1}\mathbf{E}[\boldsymbol{q}_i]}{\partial \boldsymbol{c}_i} - \frac{\partial g(\chi_c, \boldsymbol{c}_i)}{\partial \boldsymbol{c}_i} = 0 \quad (13)$$

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.

2 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 lower price, which lowers markups. More initial investment lowers marginal cost, which raises markups. This section explores that tension.

We begin by defining a product markup.

Definition 2.1 (Product 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

$$\overline{M}^p := \frac{1}{2N}\sum_{i=1}^{2}\sum_{j=1}^{N} M_{ij}^p. \quad (14)$$

To derive an expression for the product markup in the model, we simply divide each product price, using (12), by the marginal cost of that product, $a_k'c_i$:

$$M_{ik}^p = \frac{1}{a_k' c_i} a_k'(\bar{p} + b) - \frac{1}{\phi} a_k'(I + \bar{H})^{-1} \left(\frac{1}{c_i} \bar{H} \bar{p} + \sum_i \frac{E[b|\mathcal{I}_i] - c_j}{c_i} \right) \tag{15}$$

where $\bar{H} := \sum_{i=1}^{n_F} H_i$. Similarly, the average product markup for firm i is $\overline{M}_i^p = (1/N) \sum_{k=1}^{N} M_{ik}^p$.

What makes a markup large? Equation (2) reveals many forces that explain the results that follow. Some of these forces 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. Another unsurprising force that raises markups is high price sensitivity to changes in aggregate supply: Low ϕ makes H low, which makes the negative term on the right smaller. 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.

Other forces arise because firms price risk. When firms are more sensitive to risk, or the price of risk in capital markets is high, 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. These two forces often compete and are at the heart of the results that follow. Therefore, we state and drive 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.

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

The proofs of this and all further results are in Appendix B. Data both increases the expected revenue of a firm, by allowing it to produce more in states where the price will be high. It also reduces the uncertainty around that investment and lowers the risk of the firm. Both of these effects make the marginal benefit of production and the marginal benefit of investment higher.

What this means is that high-data firms invest more and grow larger. As the next result shows, it is also a channel through which data increases product markups.

Proposition 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 good k.

A firm that invests in producing an attribute can produce that attribute at a lower cost. If a good 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 j. But if the good contains some of that attribute ($a_{jk} > 0$), this investment lowers the cost of producing good j. 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 with the marginal cost of adjusting investment χ is sufficiently high. But approximately, this parameter restriction simply serves to hold investment constant, so that we can see the effect of data on output in isolation.

Proposition 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 (1.3) 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 4. 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 reduces 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 (16) 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.

$$\frac{\partial M_{i,j}^{\tilde{p}}}{\partial \Sigma_{\epsilon_i,j}^{-1}} = \underbrace{\frac{\tilde{\boldsymbol{c}}_{i,j}}{\tilde{\boldsymbol{c}}_{i,j}^2} \frac{\partial \boldsymbol{D}_j}{\partial \Sigma_{\epsilon_i,j}^{-1}}}_{\text{Risk premium effect}} - \underbrace{\frac{\boldsymbol{D}_j}{\tilde{\boldsymbol{c}}_{i,j}^2} \frac{\partial \tilde{\boldsymbol{c}}_{i,j}}{\partial \Sigma_{\epsilon_i,j}^{-1}}}_{\text{Investment effect}} \tag{16}$$

Proposition 4 simply identifies regions of the parameter space where the first or second term of (16) 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.

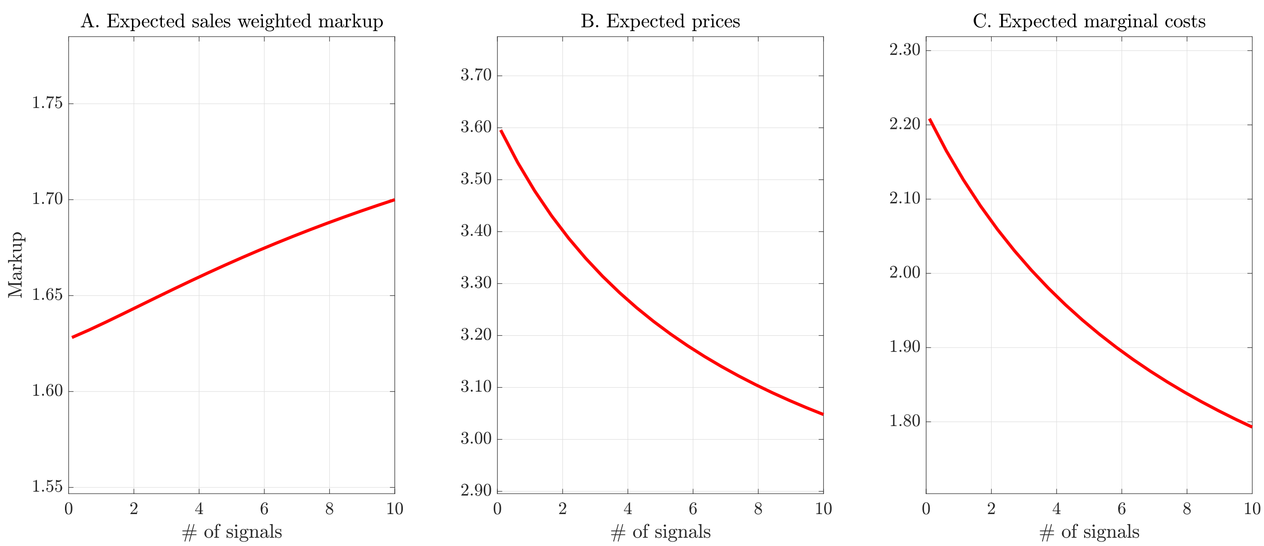

Figure 1: Data raises markups, with low investment cost / price of risk

Notes: This comparative static exercise is constructed over single-good duopoly example. The x-axis is the number of data points that both firms have. The investment cost function is assumed as $g(\chi_c, c_i) = \chi_c (\bar{c} - c_i)^2 / 2$ 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 $\rho_1 = \rho_2 = 1$.

Figures 1 and 2 illustrate how the risk reduction and investment forces compete. When firm investments greatly decrease marginal cost (low χ), 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

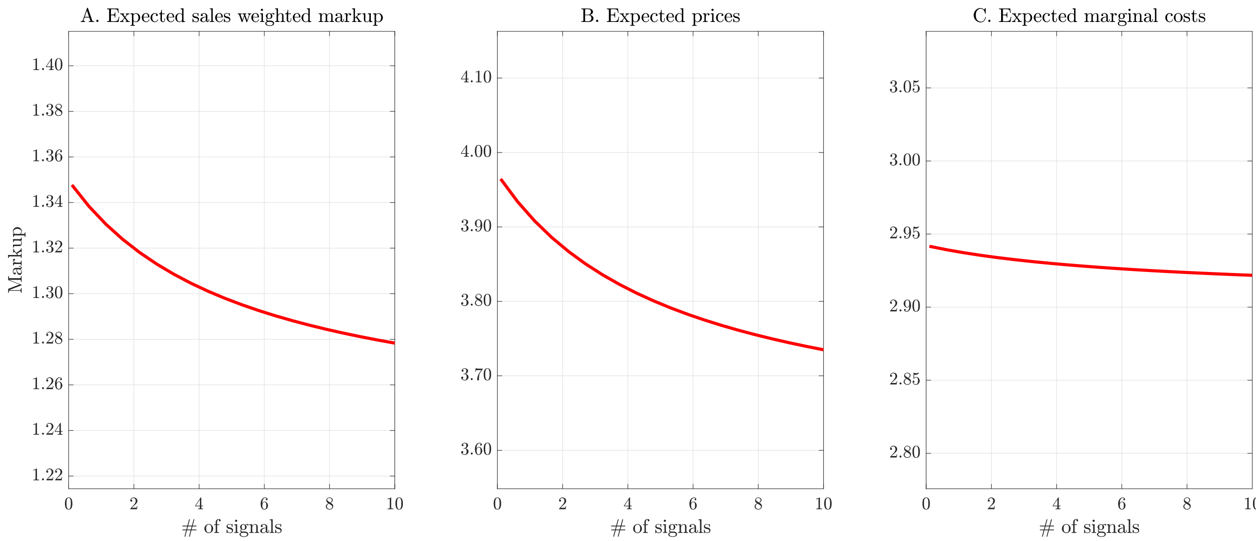

Figure 2: Data lowers markups, with high investment cost / price of risk

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

on more risk with a large production plan. By producing more, data-rich firms drive prices down 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.

3 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 investment effect dominates.

If firms are not compensated for the risk they bear, they will not produce. So 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

$$\boldsymbol{q}_i^{comp} = \frac{1}{\rho_i}\mathbf{Var}\left[\boldsymbol{p}_i|\mathcal{I}_i\right]^{-1}\left(\mathbf{E}\left[\boldsymbol{p}_i|\mathcal{I}_i\right] - \boldsymbol{c}_i\right) \tag{17}$$

In other words, production in is the same, except that we redefine the sensitivity of production to changes in price or cost in (1.3) 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 (1.3), 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 mimcs 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.[4]

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 6 makes this feasible.

Welfare benefits of data. When all firms get more data, this can be a Pareto improvement. Firms earn more utility because more information directly reduces 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 the competition among firms. The next result formalizes this logic.

Proposition 5. Data improves welfare. If the investment cost χ_c is sufficiently high, then more data for every firm increases social welfare.

[4] Later, when we define firm- and product-level markups, the competitive benchmark takes the same form: Simply replace $\bar{H}$ with $\bar{H}^{comp} = \sum_i H_i^{comp}$.

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 averseion, 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.

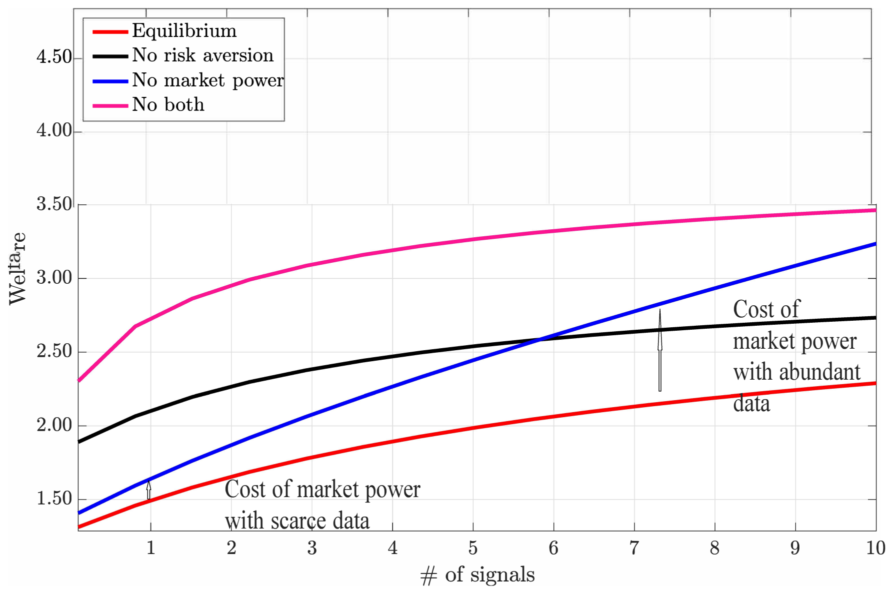

Figure 3: Welfare: Abundant data raises welfare, makes market power more costly

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

The reason that data makes market power more powerful can be seen in the first order condition (7) 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 $\rho_i \mathbf{Var}[\boldsymbol{p}_i|\mathcal{I}_i] - \frac{\partial \mathbf{E}[\boldsymbol{p}_i|\mathcal{I}_i]}{\partial \boldsymbol{q}_i}$, 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 scare 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 6 (Welfare and asymmetric data)**.** In the duopoly case, when χ_c is sufficiently large or small, there exists a cutoff value c^* 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. This highlights how the economy is affected by data dispersion.

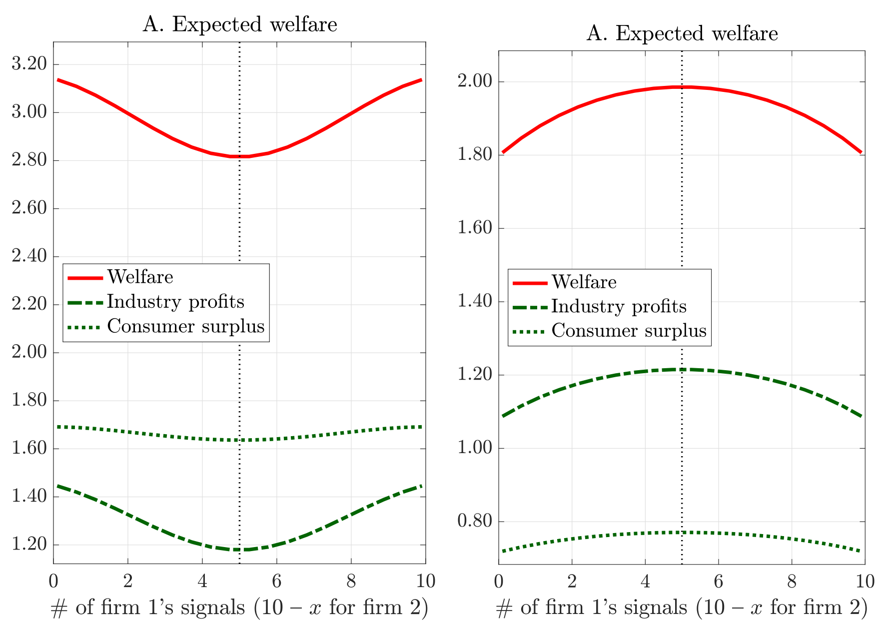

Figure 4: Data asymmetry and welfare, with dominant risk channel (left) or investment channel (right).

Notes: This comparative static exercise is constructed over 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 $\rho_1 = \rho_2 = 1$.

The left panel of Figure 4 shows that unequal distribution of data can be good for welfare. When the risk channel dominates (χ_c is large), the firm with more data, produces more. In doing so, it lowers the market price. The ability of data to incentivize production by reducing risk overwhelms the desire of the firm to reduce production to earn monopoly rents.

This trade-off shifts when the investment is efficient, i.e., χ_c is small. In the right panel of Figure 4, welfare is maximized when data is evenly distributed across firms. Here, the key mechanism is about market power. As the economy gets more asymmetric, the large, high-data firm invests more and grows larger. It has a larger markup from exploiting its market power. Higher markups create deadweight loss and hence lower welfare.

What we learn is that increasing data asymmetry has two opposite welfare effects: (1) increasing market power and hence deadweight loss; (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.

4 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 aggregate 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 et al. (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 asset. 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 explaning why firms' used of predictive data can generate such statistical patterns.

We begin by defining the aggregate markup measures and relating them to the objects in our model.

Definition 4.1 (Firm-level markup)**.** The firm-level markup for firm i is the quantity-weighted prices divided by quantity-weighted costs:

$$M_i^f := \frac{\mathbf{E}[\boldsymbol{q}_i'\boldsymbol{p}_i]}{\mathbf{E}[\boldsymbol{q}_i'\boldsymbol{c}_i]} = \frac{\mathbf{E}[\boldsymbol{q}_i]'\mathbf{E}[\boldsymbol{p}_i] + \mathbf{tr}\left[\mathbf{Cov}(\boldsymbol{p}_i, \boldsymbol{q}_i)\right]}{\mathbf{E}[\boldsymbol{q}_i'\boldsymbol{c}_i]} = \frac{\sum_{j=1}^N M_{ij}^p \boldsymbol{c}_i(j)\mathbf{E}[\boldsymbol{q}_i(j)] + \mathbf{tr}[\mathbf{Cov}\left(\boldsymbol{p}_i, \boldsymbol{q}_i\right)]}{\sum_{j=1}^N \boldsymbol{c}_i(j)\mathbf{E}[\boldsymbol{q}_i(j)]} \tag{18}$$

Similarly, the average firm-level markup is $\overline{M}^f = (1/n_F)\sum_{i=1}^{nF} M_i^f$. One interesting thing here is that the first part is cost-weighted markup:

$$M_i^f = \underbrace{\frac{\sum_{j=1}^N M_{ij}^p \boldsymbol{c}_i(j)\mathbf{E}[\boldsymbol{q}_i(j)]}{\sum_{j=1}^N \boldsymbol{c}_i(j)\mathbf{E}[\boldsymbol{q}_i(j)]}}_{\text{Cost-weighted markup}} + \frac{\mathbf{tr}[\mathbf{Cov}\left(\boldsymbol{p}_i, \boldsymbol{q}_i\right)]}{\sum_{j=1}^N \boldsymbol{c}_i(j)\mathbf{E}[\boldsymbol{q}_i(j)]} \tag{19}$$

Definition 4.2 (industry markup)**.** The industry markup is

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

Proposition 7. Data creates a wedge between product and firm markups. Holding all firms' invetments fixed ($(c_1, ..., c_{nF})$ given), an increase in firm i's data about any attribute increases $E[M_i^f - \overline{M}_i^p]$.

Mathematically, the key to this result is the covariance term in the firm level markup equation (18). It is the covariance of the price p with the firm's production decision q_i. Data increases this covariance. Without any data to predict demand, this covariance is low. The relationship between data and covariance shows up in the production first order condition (7) where a reduction in the conditional variance makes production decisions q_i more sensitive to expected changes in price p_i. Data also makes those expected changes in price covary more with actual price changes.

Economically, this is simply a composition effect at work. Firm markups are weighted averages of product markups, where highly-produced products are weighted more. Firms use data to produce more of high-price, high-markup goods. That is how firms maximize profit. Data allows the firm to skew the composition of their products in the direction of high-markup goods. So data increases the composition effect and makes firm markups larger than the average product markup of that firm.

Proposition 8. Data creates a wedge between firm and industry markups. Holding all firms' invetments fixed ($(c_1, ..., c_{nF})$ given), an increase in firm i's data about any attribute increases $E[M_i^m - \overline{M}^f]$.

Again, the key to this result is a covariance. This time the covariance is between the firm markup and the total production of a firm. industry markups are firm markups, weighted by the total output of the firm. 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 the in the production first order condition (7) 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.

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 decrease markups because it induces firms to produce more (risk-premium channel).

Data creates the product-firm markup wedge. To build intuition for the mechanisms at work, we explore a numerical example with inefficient investment

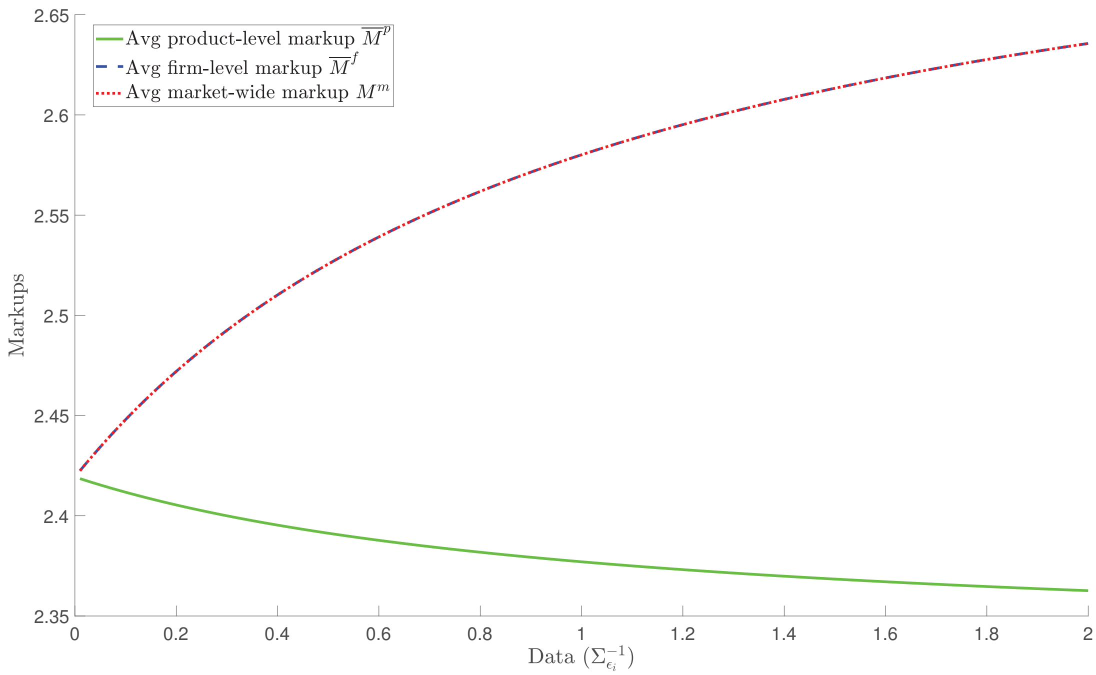

Figure 5: When all firms have more data, product markups fall, while firm and industry markups rise. Parameters used: $\bar{p} = 5$, $\rho_1 = \rho_2 = 1$, $\phi = 0.1$, $A = I$. Firm marginal costs are not chosen here. They are fixed as $c_1 = c_2 = 1$.

Figure 5 illustrates the competing effects data has on product and firm/industry markups. The product-level markup falls as data rises. In this case, both firms are being endowed with more data points. 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.

At the same time, the firm-level and industry markups are rising. Data allows firms to forecast which products will be high-markup products and to produce more of those. After all, producing more high-markup products raises a firm's profits. When firms shift the composition of their production to high-markup products, the average markup, averaged across all units of goods the firm produces rises. This is just a composition effect.

In this instance, both firms are symmetric. Because of this symmetry, the firm-level markup is the same for both firms. Therefore, the average firm markup is the same as the industry markup. We see this in Figure 5 as the blue and red dashed line lying on top of each other.

It is not always the case that product-level markups and firm/market-level markups move in opposite directions. For example, if the demand elasticity parameter were higher, $\phi = 1$, instead of $\phi = 0.1$, then the risk aversion effect would dominate and all three markups would decline as firms accumulated more data. What is general is the gap between product and firm/market level markups that arises when firms use data to change to composition of the goods they produce and skew it toward high-markup products.

From the previous example, we saw that when firms are perfectly symmetric, the industry markup is the same as the firm markup because it is a weighted average of markups that are all the same. However, when asymmetry is broken and some firms choose to be larger than others, industry aggregation matters.

Investment-data complementarity creates the industry markup wedge When firms choose their investment , to lower their marginal cost of production, high-data firms choose to invest more. This is the investment-data complementarity. This means that high-data firms, which we saw have higher firm-level markups, grow larger. 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.

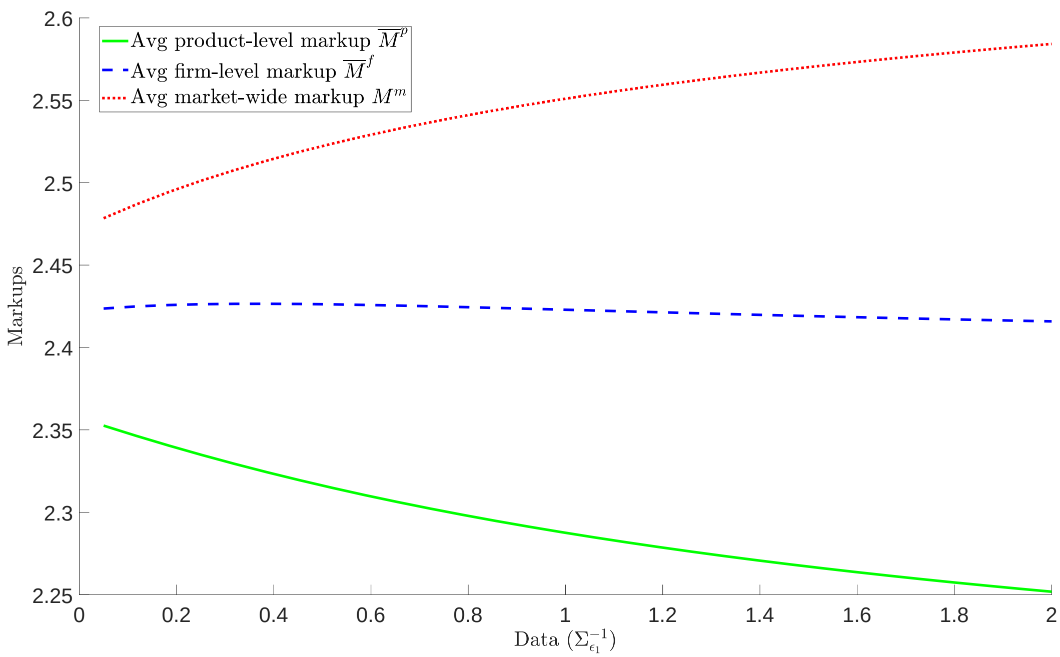

Figure 6: Firm Size Choice Makes All Three Markups Diverge. Investment cost function is $g(\chi_c, c_i) = \chi_c/c_i^2$, with $\chi_c = 1$. Parameters are $\bar{p} = 5$, $\rho_1 = 1$, $\rho_2 = 5$, $\phi = 0.8$ and $A = I$. Firm 1's data is measured on the x-axis. Firm 2's data is fixed at $\Sigma_{\epsilon_2}^{-1}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

In Figure 6, we see the gap between firm-level markup (blue dashed line in middle) and the industry markup (red dotted 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, that the differences between markups, measured at various levels of aggregation, will continue to grow.

5 Cyclicality of Markups

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

In order 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 variance of demand and of output falls. Recessions are volatile, uncertain times. In the results that follow, we investigate the effect of the change in volatility and show that it can explain the conflicting empirical findings.

Proposition 9. 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 asymptote to a function increasing in variance, $\lim_{\Sigma_{b,k} \to \infty} \partial \mathbf{E}[M^f_{ij}]/\partial \Sigma_{b,j}, \partial \mathbf{E}[M^m_{ij}]/\partial \Sigma_{b,j} > 0$.

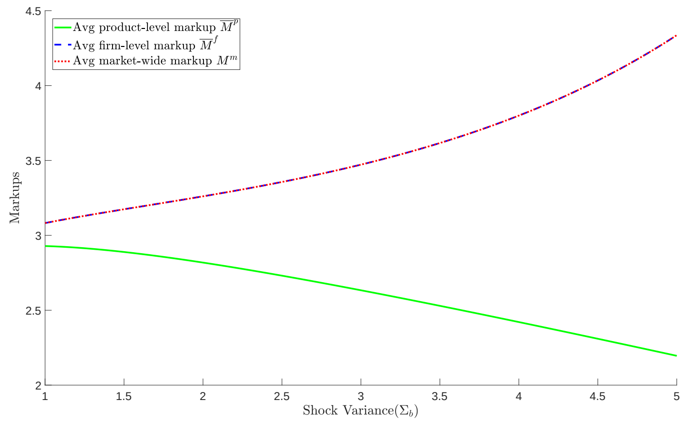

Figure 7: Procyclical product markups can co-exist with counter-cyclical firm / industry markups. Right on the x-axis represents recessions, with high shock variance and low average demand. Left is booms. An increasing line therefore represents a counter-cyclical markup. Demand and variance are inversely linked through $\bar{p} = 6 - 0.5\Sigma_b$. Remaining parameters are $\rho_1 = \rho_2 = 1$,$c_1 = c_2 = 1$, $\phi = 1$, $\Sigma_{\epsilon_1} = \Sigma_{\epsilon_2} = 1$.

The co-existence of a pro-cyclical product markup and a counter-cyclical firm or industry markup is illustrated in Figure 7. The reason these two objects behave so differently is the co-variance 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 is why the model can explain why Ramey found no change in markups, but Bils 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.

6 Mapping Theory to Data

One of the reasons that it is important to have models the 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 what risks that data is about.

Measure data by looking at the spread bt 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. The cross-firm aggregation wedge $E[M_i^m - \overline{M}^f]$ is a monotonic function of a weighted average of all firms' data precision.

This result is a straightforward conclusion from propositions 7 and 8. 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 suggest 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 Eigen-vectors are loadings. There are also other orthogonal decompositions one cas 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 desribe their products to their investors.

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 compenstaion 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 same risk preferences. (See Brealey and Myers, 2008 for a more complete explanation of the rationale and execution.)

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, whereas market competition does not. Another approach to measuring and identifying firms' data would be to asses the accuracy of firms' forecasts.

7 Conclusion

We set out to explore the hypothesis that data encouraged 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, the firm 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.

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 is 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 shifts good toward high markup goods, which raises markups. the tug-of-war between risk reduction and the composition effects induced by data plays out different 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. Out model suggests an 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.

References

ACEMOGLU, D., A. MAKHDOUMI, A. MALEKIAN, AND A. OZDAGLAR (2021): "Too Much Data: Prices and Inefficiencies in Data Markets," MIT working paper.

ANDERSON, E., S. REBELO, AND A. WONG (2018): "Markups across space and time," Tech. rep., National Bureau of Economic Research.

ATHEY, S. (2010): "The impact of targeting technology on advertising markets and media competition," *American Economic Review*, 100, 608–13.

ATHEY, S., M. MOBIUS, AND P. J. (2017): "The impact of news aggregators on internet news consumption: The case of localization," Technical Report 3353.

AUTOR, D., D. DORN, L. F. KATZ, C. PATTERSON, AND J. VAN REENEN (2020): "The Fall of the Labor Share and the Rise of Superstar Firms*," *The Quarterly Journal of Economics*, 135, 645–709.

BENKARD, L., A. YURUKOGLU, AND A. L. ZHANG (2021): "Concentration in Product Markets," Tech. rep., Stanford University.

BERGEMANN, D. AND A. BONATTI (2019): "The Economics of Social Data: An Introduction," Cowles Foundation Discussion Papers 2171, Cowles Foundation for Research in Economics, Yale University.

BILS, M. (1987): "The Cyclical Behavior of Marginal Cost and Price," *American Economic Review*, 77, 838–55.

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

BREALEY, R., S. MYERS, AND F. ALLEN (2003): *Principles of Corporate Finance*, New York: McGraw-Hill, 7th ed.

BRYNJOLFSSON, E., Y. HU, AND M. SMITH (2003): "Consumer surplus in the digital economy: Estimating the value of increased product variety at online booksellers," *Management Science*, 49, 1580–1596.

CROUZET, N. AND J. EBERLY (2018): "Understanding Weak Capital Investment: the Role of Market Concentration and Intangibles," *Jackson Hole Economic Policy Symposium*.

de Cornière, A. and G. Taylor (2020): "Data and Competition: a General Framework with Applications to Mergers, Market Structure, and Privacy Policy," Tech. Rep. 20-1076, TSE Working Papers.

De Loecker, J., J. Eeckhout, and S. Mongey (2021): "Quantifying Market Power and Business Dynamism in the Macroeconomy," Mimeo.

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

Edmond, C., V. Midrigan, and D. Y. Xu (2019): "How costly are markups?" Tech. rep., National Bureau of Economic Research.

Eeckhout, J. (2020): "Comment on: Diverging Trends in National and Local Concentration," in *NBER Macroeconomics Annual 2020, Volume 35*, NBER.

Furman, J. and P. Orszag (2015): "A Firm-Level Perspective on the Role of Rents in the Rise in Inequality," Speech.

Goldfarb, A. and C. Tucker (2017): "Digital Economics," NBER Working Paper No. w23684.

Grullon, G., Y. Larkin, and R. Michaely (2016): "Are US Industries Becoming More Concentrated?" *Unpublished Working Paper*.

Gutierrez, G. and T. Philippon (2016): "Investment-less Growth: An Empirical Investigation," NBER Working Paper 22897, National Bureau of Economic Research.

Hall, R. E. (2018): "New Evidence on the Markup of Prices over Marginal Costs and the Role of Mega-Firms in the US Economy," Tech. rep., National Bureau of Economic Research.

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

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

Kirpalani, R. and T. Philippon (2020): "Data Sharing and Market Power with Two-Sided Platforms," NBER Working Paper 28023, National Bureau of Economic Research.

Lambrecht, A. and C. Tucker (2015): "Can Big Data Protect a Firm from Competition?" Mimeo.

LIANG, A. AND E. MADSEN (2021): "Data and Incentives," Working Paper.

PELLIGRINO, B. (2020): "Product Differentiation and Oligopoly: a Network Approach," Mimeo.

RAJGOPAL, S., A. SRIVASTAVA, AND R. ZHAO (2021): "Rajgopal, Shivaram and Srivastava, Anup and Zhao, Rong, Do Digital Technology Firms Earn Excess Profits? An Alternative Perspective," SSRN paper 3739037.

RAMEY, V. AND C. NAKARDA (2020): "The Cyclical Behavior of the Price-Cost Markup," Nber working paper 19099.

ROSSI-HANSBERG, E., P.-D. SARTE, AND N. TRACHTER (2018): "Diverging trends in national and local concentration," Tech. rep., National Bureau of Economic Research.

SUTTON, J. (1991): *Sunk costs and market structure: Price competition, advertising, and the evolution of concentration*, MIT press.

——— (2001): *Technology and market structure: theory and history*, MIT press.

VIVES, X. AND Z. YE (2021): "Information Technology and Bank Competition," Iese working paper, IESE.

WHITED, T. AND C. HENNESSY (2007): "How Costly Is External Financing? Evidence From a Structural Estimation," *Journal of Finance*, 62, 1705–1745.

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

$$\begin{aligned} \mathbf{E}\left[\tilde{\boldsymbol{p}}_i | \mathcal{I}_i\right] &= \bar{\boldsymbol{p}} + \mathbf{E}\left[\boldsymbol{b}_i | \mathcal{I}_i\right] - \frac{1}{\phi}\sum_{j=1}^{n_F} \tilde{\boldsymbol{q}}_j = \bar{\boldsymbol{p}} + \boldsymbol{K}_i \boldsymbol{s}_i - \frac{1}{\phi}\sum_{j=1}^{n_F} \tilde{\boldsymbol{q}}_j \\ \mathbf{Var}\left[\tilde{\boldsymbol{p}}_i | \mathcal{I}_i\right] &= \mathbf{Var}\left[\boldsymbol{b}_i | \mathcal{I}_i\right] = \left(\Sigma_{b_i}^{-1} + \Sigma_{\epsilon_i}^{-1}\right)^{-1} \end{aligned} \tag{21}$$

Maximizing Utility Take first-order condition of firm's utility function, we have

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

It is straightforward to see that market power is constant given $\tilde{\boldsymbol{p}}_i = \bar{\boldsymbol{p}} + \boldsymbol{b}_i - \frac{1}{\phi}\sum_{j=1}^{n_F} \tilde{\boldsymbol{q}}_j$:

$$\frac{\partial \mathbf{E}\left[\tilde{\boldsymbol{p}}_i | \mathcal{I}_i\right]}{\partial \tilde{\boldsymbol{q}}_i} = -\frac{1}{\phi}\boldsymbol{I}_N \tag{22}$$

Substituting this constant market power into the first order condition for optimal output yields the optimal attribute production,

$$\begin{aligned} \tilde{\boldsymbol{q}}_i &= \left(\rho_i \mathbf{Var}\left[\tilde{\boldsymbol{p}}_i | \mathcal{I}_i\right] - \frac{\partial \mathbf{E}\left[\tilde{\boldsymbol{p}}_i | \mathcal{I}_i\right]}{\partial \tilde{\boldsymbol{q}}_i}\right)^{-1} \left(\mathbf{E}\left[\tilde{\boldsymbol{p}}_i | \mathcal{I}_i\right] - \tilde{\boldsymbol{c}}_i\right) \\ \Rightarrow \left(\rho_i \mathbf{Var}\left[\boldsymbol{b}_i | \mathcal{I}_i\right] + \frac{1}{\phi}\boldsymbol{I}_N\right)\tilde{\boldsymbol{q}}_i &= \bar{\boldsymbol{p}} + \mathbf{E}\left[\boldsymbol{b}_i | \mathcal{I}_i\right] - \frac{1}{\phi}\sum_{j=1}^{n_F} \tilde{\boldsymbol{q}}_j - \tilde{\boldsymbol{c}}_i \\ \Rightarrow \left(\rho_i \mathbf{Var}\left[\boldsymbol{b}_i | \mathcal{I}_i\right] + \frac{2}{\phi}\boldsymbol{I}_N\right)\tilde{\boldsymbol{q}}_i &= \bar{\boldsymbol{p}} + \mathbf{E}\left[\boldsymbol{b}_i | \mathcal{I}_i\right] - \frac{1}{\phi}\sum_{j=1, j\neq i}^{n_F} \tilde{\boldsymbol{q}}_j - \tilde{\boldsymbol{c}}_i \\ \Rightarrow \tilde{\boldsymbol{q}}_i &= \left(\rho_i \mathbf{Var}\left[\boldsymbol{b}_i | \mathcal{I}_i\right] + \frac{2}{\phi}\boldsymbol{I}_N\right)^{-1} \left(\bar{\boldsymbol{p}} + \mathbf{E}\left[\boldsymbol{b}_i | \mathcal{I}_i\right] - \frac{1}{\phi}\sum_{j=1, j\neq i}^{n_F} \tilde{\boldsymbol{q}}_j - \tilde{\boldsymbol{c}}_i\right) \end{aligned} \tag{23}$$

Recall that $\boldsymbol{H}_i = \left(\rho_i \mathbf{Var}\left[\boldsymbol{b}_i | \mathcal{I}_i\right] + \frac{2}{\phi}\boldsymbol{I}_N\right)^{-1}$. Substituting in $\mathbf{Var}\left[\boldsymbol{b}_i | \mathcal{I}_i\right]$ and $\boldsymbol{K}_i = \Sigma_{b_i}(\Sigma_{b_i} + \Sigma_{\epsilon_i})$ yields the solution in (23).

Sub-game equilibrium The FOC above generates the best-response function given the realization of signals. We could solve the sub-game Nash equilibrium by

$$
\begin{aligned}
&\tilde{\boldsymbol{q}}_i = \boldsymbol{H}_i\left(\bar{\boldsymbol{p}} + \boldsymbol{K}_i\boldsymbol{s}_i - \frac{1}{\phi}\sum_{j=1,j\neq i}^{n_F}\tilde{\boldsymbol{q}}_j - \tilde{\boldsymbol{c}}_i\right)\\
&\Rightarrow \left(\boldsymbol{H}_i^{-1} - \frac{\boldsymbol{I}_N}{\phi}\right)\tilde{\boldsymbol{q}}_i = \bar{\boldsymbol{p}} + \boldsymbol{K}_i\boldsymbol{s}_i - \frac{1}{\phi}\sum_{j=1}^{n_F}\tilde{\boldsymbol{q}}_j - \tilde{\boldsymbol{c}}_i\\
&\Rightarrow \left(\boldsymbol{H}_i^{-1} - \frac{\boldsymbol{I}_N}{\phi}\right)\tilde{\boldsymbol{q}}_i - \boldsymbol{K}_i\boldsymbol{s}_i + \tilde{\boldsymbol{c}}_i \equiv \bar{\boldsymbol{p}} - \frac{1}{\phi}\sum_{j=1}^{n_F}\tilde{\boldsymbol{q}}_j \triangleq \Pi,\ \forall\, i = 1,\ldots,n_F
\end{aligned}
\tag{24}
$$

The RHS is constant for each firm i and we denote it as Π. We have

$$
\begin{aligned}
&\tilde{\boldsymbol{q}}_i = \left(\boldsymbol{H}_i^{-1} - \frac{\boldsymbol{I}_N}{\phi}\right)^{-1}(\Pi + \boldsymbol{K}_i\boldsymbol{s}_i - \tilde{\boldsymbol{c}}_i)\\
&\Rightarrow \Pi = \bar{\boldsymbol{p}} - \frac{1}{\phi}\sum_{j=1}^{n_F}\tilde{\boldsymbol{q}}_j = \bar{\boldsymbol{p}} - \frac{1}{\phi}\sum_{j=1}^{n_F}\left(\boldsymbol{H}_j^{-1} - \frac{\boldsymbol{I}_N}{\phi}\right)^{-1}(\Pi + \boldsymbol{K}_j\boldsymbol{s}_j - \tilde{\boldsymbol{c}}_j)\\
&\Rightarrow \left(\boldsymbol{I}_N + \frac{1}{\phi}\sum_{j=1}^{n_F}\left(\boldsymbol{H}_j^{-1} - \frac{\boldsymbol{I}_N}{\phi}\right)^{-1}\right)\Pi = \bar{\boldsymbol{p}} - \frac{1}{\phi}\sum_{j=1}^{n_F}\left(\boldsymbol{H}_j^{-1} - \frac{\boldsymbol{I}_N}{\phi}\right)^{-1}(\boldsymbol{K}_j\boldsymbol{s}_j - \tilde{\boldsymbol{c}}_j)\\
&\Rightarrow \Pi = \left(\boldsymbol{I}_N + \frac{1}{\phi}\sum_{j=1}^{n_F}\left(\boldsymbol{H}_j^{-1} - \frac{\boldsymbol{I}_N}{\phi}\right)^{-1}\right)^{-1}\left[\bar{\boldsymbol{p}} - \frac{1}{\phi}\sum_{j=1}^{n_F}\left(\boldsymbol{H}_j^{-1} - \frac{\boldsymbol{I}_N}{\phi}\right)^{-1}(\boldsymbol{K}_j\boldsymbol{s}_j - \tilde{\boldsymbol{c}}_j)\right]
\end{aligned}
\tag{25}
$$

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

$$
\begin{aligned}
\hat{\boldsymbol{H}}_i &:= \left(\boldsymbol{H}_i^{-1} - \frac{\boldsymbol{I}_N}{\phi}\right)^{-1} = \left(\frac{\boldsymbol{I}_N}{\phi} + \rho_i\mathbf{Var}\left[\boldsymbol{p}_i|\mathcal{I}_i\right]\right)^{-1}\\
\boldsymbol{D} &:= \left(\boldsymbol{I}_N + \frac{1}{\phi}\sum_{i=1}^{n_F}\hat{\boldsymbol{H}}_i\right)^{-1}\left(\bar{\boldsymbol{p}} + \frac{1}{\phi}\sum_{i=1}^{n_F}\hat{\boldsymbol{H}}_i\tilde{\boldsymbol{c}}_i\right)
\end{aligned}
\tag{26}
$$

Finally, the equilibrium output and price are

$$
\begin{aligned}
\tilde{\boldsymbol{q}}_i &= \hat{\boldsymbol{H}}_i(\boldsymbol{D} - \tilde{\boldsymbol{c}}_i) + \hat{\boldsymbol{H}}_i\boldsymbol{K}_i\boldsymbol{s}_i - \frac{\hat{\boldsymbol{H}}_i}{\phi}\left(\boldsymbol{I}_N + \frac{1}{\phi}\sum_{i=1}^{n_F}\hat{\boldsymbol{H}}_i\right)^{-1}\sum_{j=1}^{n_F}\hat{\boldsymbol{H}}_j\boldsymbol{K}_j\boldsymbol{s}_j\\
\tilde{\boldsymbol{p}}_i &= \bar{\boldsymbol{p}} + \boldsymbol{b}_i - \frac{1}{\phi}\sum_{j=1}^{n_F}\tilde{\boldsymbol{q}}_j = \Pi + \boldsymbol{b}_i = \boldsymbol{D} + \boldsymbol{b}_i - \frac{1}{\phi}\left(\boldsymbol{I}_N + \frac{1}{\phi}\sum_{i=j}^{n_F}\hat{\boldsymbol{H}}_j\right)^{-1}\sum_{j=1}^{n_F}\hat{\boldsymbol{H}}_j\boldsymbol{K}_j\boldsymbol{s}_j
\end{aligned}
\tag{27}
$$

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

$$
\overline{M}^{\tilde{p}} = \frac{1}{N}\frac{1}{n_F}\sum_{i=1}^{n_F}\sum_{j=1}^{N}M_{i,j}^{\tilde{p}} = \frac{1}{n_F N}\sum_{i=1}^{n_F}\sum_{j=1}^{N}\frac{\mathbf{E}[\tilde{\boldsymbol{p}}_{i,j}]}{\tilde{\boldsymbol{c}}_{i,j}} = \frac{1}{n_F N}\sum_{i=1}^{n_F}\sum_{j=1}^{N}\frac{\boldsymbol{D}_j}{\tilde{\boldsymbol{c}}_{i,j}}
\tag{28}
$$

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

$$\frac{\partial \boldsymbol{D}_j}{\partial \Sigma_{\epsilon_i,k}^{-1}} = \delta_{jk}\frac{1}{\phi}\frac{\rho_i \hat{\boldsymbol{H}}_{i,j}^2 \Sigma_{b_i,j}^2}{1+\frac{1}{\phi}\sum_{s=1}^{n_F}\hat{\boldsymbol{H}}_{s,j}}\left(\tilde{\boldsymbol{c}}_{i,j}-\boldsymbol{D}_j\right) < 0 \Rightarrow \frac{\partial \overline{M}^{\tilde{p}}}{\partial \Sigma_{\epsilon_i,k}^{-1}} < 0 \tag{29}$$

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

$$M_i^f = \frac{\mathbf{E}[\tilde{\boldsymbol{q}}_i'\tilde{\boldsymbol{p}}_i]}{\mathbf{E}[\tilde{\boldsymbol{q}}_i'\tilde{\boldsymbol{c}}_i]} = \frac{\mathbf{E}[\tilde{\boldsymbol{q}}_i]'\mathbf{E}[\tilde{\boldsymbol{c}}] + \mathbf{trCov}\left(\tilde{\boldsymbol{p}}_i, \tilde{\boldsymbol{q}}_i\right)}{\mathbf{E}[\tilde{\boldsymbol{q}}_i'\tilde{\boldsymbol{c}}_i]} \tag{30}$$

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

$$\frac{\partial \mathbf{E}\tilde{\boldsymbol{q}}_{i,j}}{\partial \Sigma_{\epsilon_i,j}^{-1}} = \rho_i \hat{\boldsymbol{H}}_{i,j}^2 \Sigma_{b_i,j}^2\left(\boldsymbol{D}_j - \tilde{\boldsymbol{c}}_{i,j}\right)\left[1 - \frac{\frac{1}{\phi}\hat{\boldsymbol{H}}_{i,j}}{1+\frac{1}{\phi}\sum_{k=1}^{n_F}\hat{\boldsymbol{H}}_{k,j}}\right] > 0 \tag{31}$$

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

$$\begin{aligned}\frac{\partial \mathbf{E}\tilde{\boldsymbol{q}}_{i,j}\mathbf{E}\tilde{\boldsymbol{p}}_{i,j}}{\partial \Sigma_{\epsilon_i,j}^{-1}} &= \rho_i \hat{\boldsymbol{H}}_{i,j}^2 \Sigma_{b_i,j}^2\left(\boldsymbol{D}_j - \tilde{\boldsymbol{c}}_{i,j}\right)\left[\boldsymbol{D}_j\left(1 - \frac{\frac{2}{\phi}\hat{\boldsymbol{H}}_{i,j}}{1+\frac{1}{\phi}\sum_{k=1}^{n_F}\hat{\boldsymbol{H}}_{k,j}}\right) + \frac{\frac{1}{\phi}\hat{\boldsymbol{H}}_{i,j}}{1+\frac{1}{\phi}\sum_{k=1}^{n_F}\hat{\boldsymbol{H}}_{k,j}}\tilde{\boldsymbol{c}}_{i,j}\right] \\ &= \rho_i \hat{\boldsymbol{H}}_{i,j}^2 \Sigma_{b_i,j}^2\left(\boldsymbol{D}_j - \tilde{\boldsymbol{c}}_{i,j}\right)\left[\boldsymbol{D}_j\frac{1-\frac{1}{\phi}\hat{\boldsymbol{H}}_{i,j}+\frac{1}{\phi}\sum_{k=1,k\neq i}^{n_F}\hat{\boldsymbol{H}}_{k,j}}{1+\frac{1}{\phi}\sum_{k=1}^{n_F}\hat{\boldsymbol{H}}_{k,j}} + \frac{\frac{1}{\phi}\hat{\boldsymbol{H}}_{i,j}}{1+\frac{1}{\phi}\sum_{k=1}^{n_F}\hat{\boldsymbol{H}}_{k,j}}\tilde{\boldsymbol{c}}_{i,j}\right] > 0\end{aligned} \tag{32}$$

Cost-weighted industry markup The industry markup weighted by cost is

$$M^{m,cost} := \frac{\mathbf{E}\left[\sum_{i=1}^{n_F}\tilde{\boldsymbol{q}}_i'\tilde{\boldsymbol{p}}_i\right]}{\mathbf{E}\left[\sum_{i=1}^{n_F}\tilde{\boldsymbol{q}}_i'\tilde{\boldsymbol{c}}_i\right]} = \frac{\sum_{i=1}^{n_F}\mathbf{E}\left[\tilde{\boldsymbol{q}}_i'\tilde{\boldsymbol{p}}_i\right]}{\sum_{i=1}^{n_F}\mathbf{E}\left[\tilde{\boldsymbol{q}}_i'\tilde{\boldsymbol{c}}_i\right]} = \sum_{i=1}^{n_F} w_i^{cost} M_i^f \quad \text{where} \quad w_i^{cost} = \frac{\mathbf{E}\left[\tilde{\boldsymbol{q}}_i'\tilde{\boldsymbol{c}}_i\right]}{\sum_{i=1}^{n_F}\mathbf{E}\left[\tilde{\boldsymbol{q}}_i'\tilde{\boldsymbol{c}}_i\right]}. \tag{33}$$

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

$$\frac{\partial w_{i,j}^{cost}}{\partial \Sigma_{\epsilon_i,j}^{-1}} = \frac{\tilde{\boldsymbol{c}}_{i,j}}{\left(\sum_{i=1}^{n_F}\mathbf{E}\left[\tilde{\boldsymbol{q}}_{i,j}'\tilde{\boldsymbol{c}}_{i,j}\right]\right)^2}\left[\frac{\partial \mathbf{E}\tilde{\boldsymbol{q}}_{i,j}}{\partial \Sigma_{\epsilon_i,j}^{-1}}\left(\sum_{k=1,k\neq i}^{n_F}\mathbf{E}\left[\tilde{\boldsymbol{q}}_{k,j}'\tilde{\boldsymbol{c}}_{k,j}\right]\right) - \mathbf{E}\tilde{\boldsymbol{q}}_{i,j}\left(\sum_{k=1,k\neq i}^{n_F}\frac{\partial \mathbf{E}(\tilde{\boldsymbol{q}}_{k,j})}{\partial \Sigma_{\epsilon_i,j}^{-1}}\tilde{\boldsymbol{c}}_{k,j}\right)\right] > 0 \tag{34}$$

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

$$M^{m,sale} := \sum_{i=1}^{n_F} w_i^{sale} M_i^f = \frac{\sum_{i=1}^{n_F}\frac{\mathbf{E}^2\left[\tilde{\boldsymbol{q}}_i'\tilde{\boldsymbol{p}}_i\right]}{\mathbf{E}\left[\tilde{\boldsymbol{q}}_i'\tilde{\boldsymbol{c}}_i\right]}}{\sum_{i=1}^{n_F}\mathbf{E}\left[\tilde{\boldsymbol{q}}_i'\tilde{\boldsymbol{p}}_i\right]} \quad \text{where} \quad w_i^{sale} = \frac{\mathbf{E}\left[\tilde{\boldsymbol{q}}_i'\tilde{\boldsymbol{p}}_i\right]}{\sum_{i=1}^{n_F}\mathbf{E}\left[\tilde{\boldsymbol{q}}_i'\tilde{\boldsymbol{p}}_i\right]}. \tag{35}$$

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

$$\begin{aligned}\mathbf{Cov}\left(\tilde{\boldsymbol{p}}_i, \tilde{\boldsymbol{q}}_i\right) = &\left(\boldsymbol{I}_N + \sum_{j=1}^{n_F} \frac{\hat{\boldsymbol{H}}_j}{\phi}\right)^{-1} \sum_{j=1}^{n_F} \hat{\boldsymbol{H}}_j \boldsymbol{K}_j \hat{\boldsymbol{H}}_j \left(\boldsymbol{I}_N + \sum_{j=1}^{n_F} \frac{\hat{\boldsymbol{H}}_j}{\phi}\right)^{-1} \frac{\hat{\boldsymbol{H}}_i}{\phi^2} + \boldsymbol{K}_i \hat{\boldsymbol{H}}_i \\ &- \left(\boldsymbol{I}_N + \sum_{j=1}^{n_F} \frac{\hat{\boldsymbol{H}}_j}{\phi}\right)^{-1} \hat{\boldsymbol{H}}_i \boldsymbol{K}_i \frac{\hat{\boldsymbol{H}}_i}{\phi} - \boldsymbol{K}_i \hat{\boldsymbol{H}}_i \left(\boldsymbol{I}_N + \sum_{j=1}^{n_F} \frac{\hat{\boldsymbol{H}}_j}{\phi}\right)^{-1} \frac{\hat{\boldsymbol{H}}_i}{\phi}\end{aligned} \tag{36}$$

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 (24), we have

$$\begin{aligned}\mathbf{E}\left[\tilde{\boldsymbol{q}}_i' \left(\tilde{\boldsymbol{p}}_i - \tilde{\boldsymbol{c}}_i\right)\right] &= \mathbf{E}\left[\tilde{\boldsymbol{q}}_i' \left(\mathbf{E}\left[\tilde{\boldsymbol{p}}_i | \mathcal{I}_i\right] - \tilde{\boldsymbol{c}}_i\right)\right] \\ &= \mathbf{E}\left[\tilde{\boldsymbol{q}}_i' \hat{\boldsymbol{H}}_i^{-1} \tilde{\boldsymbol{q}}_i\right]\end{aligned} \tag{37}$$

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

$$\begin{aligned}\mathbf{E}[U_i] &= \mathbf{E}\left[\tilde{\boldsymbol{q}}_i' \left(\tilde{\boldsymbol{p}}_i - \tilde{\boldsymbol{c}}_i\right)\right] - \frac{\rho_i}{2} \mathbf{E}\left[\tilde{\boldsymbol{q}}_i' \mathbf{Var}\left[\tilde{\boldsymbol{p}}_i | \mathcal{I}_i\right] \tilde{\boldsymbol{q}}_i\right] - g(\chi_c, \tilde{\boldsymbol{c}}_i) \\ &= \mathbf{E}\left[\tilde{\boldsymbol{q}}_i' \left(\hat{\boldsymbol{H}}_i^{-1} - \frac{\rho_i}{2} \mathbf{Var}\left[\tilde{\boldsymbol{p}}_i | \mathcal{I}_i\right]\right) \tilde{\boldsymbol{q}}_i\right] - g(\chi_c, \tilde{\boldsymbol{c}}_i) \\ &= \mathbf{E}\left[\tilde{\boldsymbol{q}}_i' \left(\frac{I_N}{\phi} + \frac{\rho_i}{2} \mathbf{Var}\left[\tilde{\boldsymbol{p}}_i | \mathcal{I}_i\right]\right) \tilde{\boldsymbol{q}}_i\right] - g(\chi_c, \tilde{\boldsymbol{c}}_i) \\ &= \frac{1}{2} \mathbf{E}\left[\tilde{\boldsymbol{q}}_i' \boldsymbol{H}_i^{-1} \tilde{\boldsymbol{q}}_i\right] - g(\chi_c, \tilde{\boldsymbol{c}}_i) \\ &= \frac{1}{2} \left(\mathbf{E}[\tilde{\boldsymbol{q}}_i]' \boldsymbol{H}_i^{-1} \mathbf{E}[\tilde{\boldsymbol{q}}_i] + \mathbf{tr}\left(\boldsymbol{H}_i^{-1} \mathbf{Var}[\tilde{\boldsymbol{q}}_i]\right)\right) - g(\chi_c, \tilde{\boldsymbol{c}}_i)\end{aligned} \tag{38}$$

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

$$\begin{aligned}\frac{\partial \mathbf{E}[U_i]}{\partial \tilde{\boldsymbol{c}}_i} &= \frac{1}{2} \frac{\partial \mathbf{E}[\tilde{\boldsymbol{q}}_i]' \boldsymbol{H}_i^{-1} \mathbf{E}[\tilde{\boldsymbol{q}}_i]}{\partial \tilde{\boldsymbol{c}}_i} - \frac{\partial g(\chi_c, \tilde{\boldsymbol{c}}_i)}{\partial \tilde{\boldsymbol{c}}_i} = 0 \\ \frac{\partial^2 \mathbf{E}[U_i]}{\partial \tilde{\boldsymbol{c}}_i \partial \tilde{\boldsymbol{c}}_i'} &= \frac{1}{2} \frac{\partial^2 \mathbf{E}[\tilde{\boldsymbol{q}}_i]' \boldsymbol{H}_i^{-1} \mathbf{E}[\tilde{\boldsymbol{q}}_i]}{\partial \tilde{\boldsymbol{c}}_i \partial \tilde{\boldsymbol{c}}_i'} - \frac{\partial^2 g(\chi_c, \tilde{\boldsymbol{c}}_i)}{\partial \tilde{\boldsymbol{c}}_i \partial \tilde{\boldsymbol{c}}_i'} \text{ is negative semi-definite}\end{aligned} \tag{39}$$

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

$$\begin{aligned}
\frac{\partial \mathbf{E}[U_i]}{\partial \tilde{\boldsymbol{c}}_i} &= \frac{1}{2}\frac{\partial}{\partial \tilde{\boldsymbol{c}}_i}\left(\sum_{s=1}^{N}\left(\boldsymbol{D}_s - \tilde{\boldsymbol{c}}_{i,s}\right)^2 \hat{\boldsymbol{H}}_{i,s}^2 \boldsymbol{H}_{i,s}^{-1}\right) - \frac{\partial g(\chi_c, \tilde{\boldsymbol{c}}_i)}{\partial \tilde{\boldsymbol{c}}_i} = 0 \\
\Rightarrow \frac{\partial \mathbf{E}[U_i]}{\partial \tilde{\boldsymbol{c}}_{i,j}} &= \left(\boldsymbol{D}_j - \tilde{\boldsymbol{c}}_{i,j}\right)\hat{\boldsymbol{H}}_{i,j}^2 \boldsymbol{H}_{i,j}^{-1}\left[\frac{\hat{\boldsymbol{H}}_{i,j}}{1 + \frac{1}{\phi}\sum_{s=1}^{n_F}\hat{\boldsymbol{H}}_{s,j}} - 1\right] - \frac{\partial g(\chi_c, \tilde{\boldsymbol{c}}_i)}{\partial \tilde{\boldsymbol{c}}_{i,j}} = 0 \\
\Rightarrow \frac{\partial^2 \mathbf{E}[U_i]}{\partial \tilde{\boldsymbol{c}}_{i,j}\partial \tilde{\boldsymbol{c}}_{i,k}} &= \delta_{jk}\hat{\boldsymbol{H}}_{i,j}^2 \boldsymbol{H}_{i,j}^{-1}\left[\frac{\hat{\boldsymbol{H}}_{i,j}}{1 + \frac{1}{\phi}\sum_{s=1}^{n_F}\hat{\boldsymbol{H}}_{s,j}} - 1\right]^2 - \frac{\partial g(\chi_c, \tilde{\boldsymbol{c}}_i)}{\partial \tilde{\boldsymbol{c}}_{i,j}\partial \tilde{\boldsymbol{c}}_{i,k}}
\end{aligned} \tag{40}$$

since

$$\boldsymbol{D}_j = \frac{\bar{\boldsymbol{p}}_j + \frac{1}{\phi}\sum_{s=1}^{n_F}\hat{\boldsymbol{H}}_{s,j}\tilde{\boldsymbol{c}}_{s,j}}{1 + \frac{1}{\phi}\sum_{s=1}^{n_F}\hat{\boldsymbol{H}}_{s,j}} \quad \text{and} \quad \frac{\partial \boldsymbol{D}_j}{\partial \tilde{\boldsymbol{c}}_{i,k}} = \delta_{jk}\frac{\hat{\boldsymbol{H}}_{i,j}}{1 + \frac{1}{\phi}\sum_{s=1}^{n_F}\hat{\boldsymbol{H}}_{s,j}} \tag{41}$$

B Proofs and Auxiliary Results

Proof of Proposition 1: Data-Investment Complementarity

Proof. Starting from the expected utility, we have

$$\mathbb{E}\left[U_i\right] = \frac{1}{2}\left[\mathbb{E}\left[\widetilde{\boldsymbol{q}}_i\right]' \boldsymbol{H}_i^{-1}\mathbb{E}\left[\widetilde{\boldsymbol{q}}_i\right] + \operatorname{tr}\left(\boldsymbol{H}_i^{-1}\mathbb{V}\left[\widetilde{\boldsymbol{q}}_i\right]\right)\right] - g\left(\chi_c, \widetilde{\boldsymbol{c}}_i\right) \tag{42}$$

To show this complementarity between information and costs, we first differentiate $\mathbb{E}\left[U_i\right]$ with regarding to marginal cost:

$$\begin{aligned}
\frac{\partial \mathbb{E}\left[U_i\right]}{\partial \widetilde{c}_{ij}} &= \frac{\partial}{\partial \widetilde{c}_{ij}}\left\{\frac{1}{2}\left[\mathbb{E}\left[\widetilde{\boldsymbol{q}}_i\right]' \boldsymbol{H}_i^{-1}\mathbb{E}\left[\widetilde{\boldsymbol{q}}_i\right] + \operatorname{tr}\left(\boldsymbol{H}_i^{-1}\mathbb{V}\left[\widetilde{\boldsymbol{q}}_i\right]\right)\right] - g\left(\chi_c, \widetilde{\boldsymbol{c}}_i\right)\right\} \\
&= -\left(\widehat{\boldsymbol{H}}_{ij}\right)^2 \boldsymbol{H}_{ij}^{-1}\left\{\frac{\left[\phi\left(\overline{\boldsymbol{p}} - \widetilde{c}_{ij}\right) + \sum_{s\neq i}^{n_F}\widehat{H}_{sj}\left(\widetilde{c}_{sj} - \widetilde{c}_{ij}\right)\right]\left(\phi + \sum_{s\neq i}\widehat{H}_{sj}\right)}{\left(\phi + \sum_{i=1}^{n_F}\widehat{\boldsymbol{H}}_{ij}\right)^2}\right\} - \frac{\partial g\left(\chi_c, \widetilde{\boldsymbol{c}}_i\right)}{\partial \widetilde{c}_i}
\end{aligned} \tag{43}$$

The second order cross derivative is: (denote $\mathbb{V} = \mathbb{V}\left[\boldsymbol{b}_{ij}|\mathcal{I}\right]$)

$$\begin{aligned}
\frac{\partial^2 \mathbb{E}\left[U_i\right]}{\partial \widetilde{c}_{ij}\partial \mathbb{V}\left[\boldsymbol{b}_{ij}|\mathcal{I}\right]} &= \frac{\partial}{\partial \mathbb{V}}\left\{-\left(\widehat{\boldsymbol{H}}_{ij}\right)^2 \boldsymbol{H}_{ij}^{-1}\left\{\frac{\left[\phi\left(\overline{\boldsymbol{p}}-\widetilde{c}_{ij}\right)+\sum_{s\neq i}^{n_F}\widehat{H}_{sj}\left(\widetilde{c}_{sj}-\widetilde{c}_{ij}\right)\right]\left(\phi+\sum_{s\neq i}\widehat{H}_{sj}\right)}{\left(\phi+\sum_{i=1}^{n_F}\widehat{\boldsymbol{H}}_{ij}\right)^2}\right\}-\frac{\partial g\left(\chi_c,\widetilde{\boldsymbol{c}}_i\right)}{\partial \widetilde{c}_i}\right\} \\
&= \underbrace{-\left\{\left[\phi\left(\overline{\boldsymbol{p}}-\widetilde{c}_{ij}\right)+\sum_{s\neq i}^{n_F}\widehat{H}_{sj}\left(\widetilde{c}_{sj}-\widetilde{c}_{ij}\right)\right]\left(\phi+\sum_{s\neq i}\widehat{H}_{sj}\right)\right\}}_{\text{A negative term } (-)} \\
&\qquad \times \frac{\partial}{\partial \mathbb{V}_i}\left\{\frac{\frac{2}{\phi}+\rho_i\mathbb{V}_i}{\left(\phi+\sum_{s=1}^{n_F}\left(\frac{1}{\phi}+\rho_s\mathbb{V}_s\right)^{-1}\right)^2\left(\frac{1}{\phi}+\rho_i\mathbb{V}_i\right)^2}\right\} \\
&= (-)\times \underbrace{\frac{\partial}{\partial \mathbb{V}_i}\left\{\frac{1}{\left(2+\phi\rho_i\mathbb{V}_i+\sum_{s\neq i}^{n_F}\frac{\frac{1}{\phi}+\rho_i\mathbb{V}_i}{\frac{1}{\phi}+\rho_s\mathbb{V}_s}\right)\left(\phi+\sum_{s\neq i}^{n_F}\left[\frac{1}{\frac{1}{\phi}+\rho_s\mathbb{V}_s}\left(1-\frac{1}{2+\phi\rho_i\mathbb{V}_i}\right)\right]\right)}\right\}}_{\text{Decreasing on } \mathbb{V}_i\text{, i.e., }<0} > 0
\end{aligned} \tag{44}$$

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 Proposition ??: Greater investment raises a firm's product markup.

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

$$\frac{\partial M_{i,j}^{\tilde{p}}}{\partial \boldsymbol{c}_{i,j}} = \frac{\frac{\partial \boldsymbol{D}_j}{\partial \boldsymbol{c}_{i,j}}\boldsymbol{c}_{i,j}-\boldsymbol{D}_j}{\boldsymbol{c}_{i,j}^2} = \frac{\frac{1}{\phi}\hat{\boldsymbol{H}}_{i,j}\boldsymbol{c}_{i,j}-\bar{\boldsymbol{p}}_j-\frac{1}{\phi}\sum_{s=1}^{n_F}\hat{\boldsymbol{H}}_{s,j}\boldsymbol{c}_{s,j}}{\boldsymbol{c}_{i,j}^2\left(1+\frac{1}{\phi}\sum_{s=1}^{n_F}\hat{\boldsymbol{H}}_{s,j}\right)} = -\frac{\bar{\boldsymbol{p}}_j+\frac{1}{\phi}\sum_{s=1,s\neq i}^{n_F}\hat{\boldsymbol{H}}_{s,j}\boldsymbol{c}_{s,j}}{\boldsymbol{c}_{i,j}^2\left(1+\frac{1}{\phi}\sum_{s=1}^{n_F}\hat{\boldsymbol{H}}_{s,j}\right)} \leq 0 \tag{45}$$

The negative derivative confirms that more investment leads to higher product-level markup. □

Proof of Proposition 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 $\mathbf{E}\left(p_i/c_i\right) = \mathbf{E}\left(p_i\right)/c_i$ is decreasing in data.

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

$$\overline{M}^{\tilde{p}} = \frac{1}{N}\frac{1}{n_F}\sum_{i=1}^{n_F}\sum_{j=1}^{N}M_{i,j}^{\tilde{p}} = \frac{1}{n_F N}\sum_{i=1}^{n_F}\sum_{j=1}^{N}\frac{\mathbf{E}[\tilde{\boldsymbol{p}}_{i,j}]}{\boldsymbol{c}_{i,j}} = \frac{1}{n_F N}\sum_{i=1}^{n_F}\sum_{j=1}^{N}\frac{\boldsymbol{D}_j}{\boldsymbol{c}_{i,j}} \tag{46}$$

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

$$\boldsymbol{D}_j = \frac{\bar{\boldsymbol{p}}_j + \frac{1}{\phi}\sum_{i=1}^{n_F}\hat{\boldsymbol{H}}_{i,j}\boldsymbol{c}_{i,j}}{1+\frac{1}{\phi}\sum_{i=1}^{n_F}\hat{\boldsymbol{H}}_{i,j}} \quad \text{where} \quad \hat{\boldsymbol{H}}_{i,j} = \left[\frac{1}{\phi} + \rho_i\left(1+\Sigma_{\epsilon_{i,j}}^{-1}\right)^{-1}\right]^{-1} \tag{47}$$

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

$$\frac{\partial \boldsymbol{D}_j}{\partial \Sigma_{\epsilon_i,k}^{-1}} = \delta_{jk}\frac{1}{\phi}\frac{\rho_i \hat{\boldsymbol{H}}_{i,j}^2 \Sigma_{b_i,j}^2}{1+\frac{1}{\phi}\sum_{s=1}^{n_F}\hat{\boldsymbol{H}}_{s,j}}\left(\boldsymbol{c}_{i,j} - \boldsymbol{D}_j\right) < 0 \Rightarrow \frac{\partial \overline{M}^{\tilde{p}}}{\partial \Sigma_{\epsilon_i,k}^{-1}} < 0 \tag{48}$$

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.
□

□

Proof of Proposition 4: 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

$$\frac{\partial M_{i,j}^{\tilde{p}}}{\partial \Sigma_{\epsilon_{i,j}}^{-1}} = \frac{\frac{\partial \boldsymbol{D}_j}{\partial \Sigma_{\epsilon_{i,j}}^{-1}}\tilde{\boldsymbol{c}}_{i,j} - \boldsymbol{D}_j\frac{\partial \tilde{\boldsymbol{c}}_{i,j}}{\partial \Sigma_{\epsilon_{i,j}}^{-1}}}{\tilde{\boldsymbol{c}}_{i,j}^2} \tag{49}$$

According to (29) and (**??**), we have

$$\frac{\partial \boldsymbol{D}_j}{\partial \Sigma_{\epsilon_{i,j}}^{-1}} = \frac{1}{\phi}\frac{\rho_i \hat{\boldsymbol{H}}_{i,j}^2 \Sigma_{b_i,j}^2}{1+\frac{1}{\phi}\sum_{s=1}^{n_F}\hat{\boldsymbol{H}}_{s,j}}\left(\tilde{\boldsymbol{c}}_{i,j} - \boldsymbol{D}_j\right) < 0 \quad \text{and} \quad \frac{\partial \tilde{\boldsymbol{c}}_i}{\partial \Sigma_{\epsilon_i}} = \left(-\frac{\partial^2 \mathbf{E}[U_i]}{\partial \tilde{\boldsymbol{c}}_i \partial \tilde{\boldsymbol{c}}_i'}\right)^{-1}\Lambda, \quad \frac{\partial \tilde{\boldsymbol{c}}_{i,j}}{\partial \Sigma_{\epsilon_{i,j}}^{-1}} \leq 0 \tag{50}$$

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.

□

Proof of Proposition ??: Markups increase in data asymmetry.

Proof. Data asymmetry means adding a unit of total data precision for firm m and subtracting the same amount of precision from firm n. Using result (ADD EQN REF), and taking the difference for the two firms' effects, we get

$$\begin{aligned}
\frac{\partial \boldsymbol{D}_j}{\partial \Sigma_{\epsilon_m,k}^{-1}} &= \delta_{jk}\frac{1}{\phi}\frac{\rho_m \hat{\boldsymbol{H}}_{m,j}^2 \Sigma_{b_m,j}^2}{1+\frac{1}{\phi}\sum_{i=1}^{n_F}\hat{\boldsymbol{H}}_{i,j}}\left(\boldsymbol{c}_{m,j} - \boldsymbol{D}_j\right) = -\delta_{jk}\frac{1}{\phi}\frac{\frac{\partial \ln \hat{\boldsymbol{H}}_{m,j}}{\partial \Sigma_{\epsilon_{m,j}}^{-1}}\mathbf{E}(\tilde{\boldsymbol{q}}_{m,j})}{1+\frac{1}{\phi}\sum_{i=1}^{n_F}\hat{\boldsymbol{H}}_{i,j}} \\
\frac{\partial \boldsymbol{D}_j}{\partial \Sigma_{\epsilon_n,k}^{-1}} &= \delta_{jk}\frac{1}{\phi}\frac{\rho_n \hat{\boldsymbol{H}}_{n,j}^2 \Sigma_{b_n,j}^2}{1+\frac{1}{\phi}\sum_{i=1}^{n_F}\hat{\boldsymbol{H}}_{i,j}}\left(\boldsymbol{c}_{n,j} - \boldsymbol{D}_j\right) = -\delta_{jk}\frac{1}{\phi}\frac{\frac{\partial \ln \hat{\boldsymbol{H}}_{n,j}}{\partial \Sigma_{\epsilon_{n,j}}^{-1}}\mathbf{E}(\tilde{\boldsymbol{q}}_{n,j})}{1+\frac{1}{\phi}\sum_{i=1}^{n_F}\hat{\boldsymbol{H}}_{i,j}}
\end{aligned} \tag{51}$$

So the combined effect on the j^{th} price is

$$d\boldsymbol{D}_j = \frac{\partial \boldsymbol{D}_j}{\partial \Sigma^{-1}_{\epsilon_m,k}} d\Sigma^{-1}_{\epsilon_m,k} + \frac{\partial \boldsymbol{D}_j}{\partial \Sigma^{-1}_{\epsilon_n,k}} d\Sigma^{-1}_{\epsilon_n,k} \approx \frac{\partial \boldsymbol{D}_j}{\partial \Sigma^{-1}_{\epsilon_m,k}} \Delta\Sigma^{-1}_{\epsilon_m,k} - \frac{\partial \boldsymbol{D}_j}{\partial \Sigma^{-1}_{\epsilon_n,k}} \Delta\Sigma^{-1}_{\epsilon_n,k} \tag{52}$$

The last linear approximation governs how the product-level markup reacts to the data asymmetry, in general we have

$$\frac{\rho_m \hat{\boldsymbol{H}}^2_{m,j} \Sigma^2_{b_m,j} \left(\boldsymbol{D}_j - \boldsymbol{c}_{m,j}\right)}{\rho_n \hat{\boldsymbol{H}}^2_{n,j} \Sigma^2_{b_n,j} \left(\boldsymbol{D}_j - \boldsymbol{c}_{n,j}\right)} \geq (\leq) \frac{\Delta\Sigma^{-1}_{\epsilon_m,k}}{\Delta\Sigma^{-1}_{\epsilon_n,k}} \Leftrightarrow \text{decreasing (increasing) product-level markup} \tag{53}$$

Suppose the linear constraint of precisions ($\sum_i \Sigma^{-1}_{\epsilon_i,j} \leq \tilde{K}_j$) is binding given utility increasing in data. The ratio of the $\Delta\Sigma^{-1}_{\epsilon_m,k}$ to $\Delta\Sigma^{-1}_{\epsilon_n,k}$ is a constant (equal to one), and the product-level markup is increasing if LHS of (53) is smaller than one. In other words, increasing product-level markup implies

$$\rho_m \hat{\boldsymbol{H}}^2_{m,j} \Sigma^2_{b_m,j} \left(\boldsymbol{D}_j - \boldsymbol{c}_{m,j}\right) \leq \rho_n \hat{\boldsymbol{H}}^2_{n,j} \Sigma^2_{b_n,j} \left(\boldsymbol{D}_j - \boldsymbol{c}_{n,j}\right) \tag{54}$$

Since coefficients $\hat{\boldsymbol{H}}$ and Σ^2_b are independent of marginal cost $\boldsymbol{c}$, we have

$$\boldsymbol{D}_j - \boldsymbol{c}_{m,j} = \frac{\bar{\boldsymbol{p}}_j + \frac{1}{\phi}\sum_{i=1}^{n_F} \hat{\boldsymbol{H}}_{i,j}\boldsymbol{c}_{i,j} - \boldsymbol{c}_{m,j} - \frac{1}{\phi}\sum_{i=1}^{n_F} \hat{\boldsymbol{H}}_{i,j}\boldsymbol{c}_{m,j}}{1 + \frac{1}{\phi}\sum_{i=1}^{n_F} \hat{\boldsymbol{H}}_{i,j}} \tag{55}$$

Thus, the equation (54) is equivalent to the the following condition

$$\begin{aligned} &\rho_m \hat{\boldsymbol{H}}^2_{m,j} \Sigma^2_{b_m,j} \left(\boldsymbol{D}_j - \boldsymbol{c}_{m,j}\right) \leq \rho_n \hat{\boldsymbol{H}}^2_{n,j} \Sigma^2_{b_n,j} \left(\boldsymbol{D}_j - \boldsymbol{c}_{n,j}\right) \\ \Leftrightarrow\ &\boldsymbol{c}_{m,j} \leq \frac{(1-K)\hat{\boldsymbol{H}}_{n,j} + K\left(\phi + \sum_{i=1}^{n_F} \hat{\boldsymbol{H}}_{i,j}\right)}{K\hat{\boldsymbol{H}}_{m,j} + (1-K)\left(\phi + \sum_{i=1}^{n_F} \hat{\boldsymbol{H}}_{i,j}\right)} \boldsymbol{c}_{n,j} \quad \text{where} \quad K := \frac{\rho_n \hat{\boldsymbol{H}}^2_{n,j} \Sigma^2_{b_n,j}}{\rho_m \hat{\boldsymbol{H}}^2_{m,j} \Sigma^2_{b_m,j}} \end{aligned} \tag{56}$$

It's clear that if the marginal cost $\boldsymbol{c}_{m,j}$ is small enough, the product-level markup will increase. Moreover, the coefficient before $\left(\boldsymbol{D}_j - \boldsymbol{c}_{m,j}\right)$ converges to zero as

$$\begin{aligned} &\rho_m \hat{\boldsymbol{H}}^2_{m,j} \Sigma^2_{b_m,j} = \left(\frac{2}{\Sigma_{b_m,j}\phi} + \rho_m + \frac{1}{\rho_m \phi^2 \Sigma^2_{b_m,j}}\right)^{-1} = \left(\rho_m + \frac{2}{\phi}\left(1 + \Sigma^{-1}_{\epsilon_m,j}\right) + \frac{1}{\rho_m \phi^2}\left(1 + \Sigma^{-1}_{\epsilon_m,j}\right)^2\right)^{-1} \\ \Rightarrow &\lim_{\Sigma^{-1}_{\epsilon_m,j} \to \infty} \rho_m \hat{\boldsymbol{H}}^2_{m,j} \Sigma^2_{b_m,j} = \lim_{\rho_m \to 0} \rho_m \hat{\boldsymbol{H}}^2_{m,j} \Sigma^2_{b_m,j} = 0 \end{aligned} \tag{57}$$

So the inequality (54) holds if the risk aversion is small enough or data precision is large enough. □

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:

$$\mathbb{E}\left[U_i\right] = \frac{1}{2}\left[\mathbb{E}\left[\widetilde{\boldsymbol{q}}_i\right]' \boldsymbol{H}_i^{-1}\mathbb{E}\left[\widetilde{\boldsymbol{q}}_i\right] + \boldsymbol{H}_i^{-1}\mathbb{V}\left[\widetilde{\boldsymbol{q}}_i\right]\right] \tag{58}$$

$$\text{where} \qquad \widetilde{\boldsymbol{q}}_i = \widehat{\boldsymbol{H}}_i\left(\boldsymbol{D}-\overline{\boldsymbol{c}}\right) + \widehat{\boldsymbol{H}}_i\boldsymbol{K}_i\boldsymbol{s}_i - \frac{\widehat{\boldsymbol{H}}_i}{\phi}\left(1+\frac{1}{\phi}\sum_{i=1}^{n_F}\widehat{\boldsymbol{H}}_i\right)^{-1}\sum_{j=1}^{n_F}\widehat{\boldsymbol{H}}_j\boldsymbol{K}_j\boldsymbol{s}_j \tag{59}$$

Combining these, we have:

$$\mathbb{E}\left[U_i\right] = \frac{\widehat{\boldsymbol{H}}_i^2}{2H_i}\left(\boldsymbol{D}-\overline{\boldsymbol{c}}\right)^2 + \frac{\widehat{\boldsymbol{H}}_i^2}{2H_i\left(\phi+\sum_{i=1}^{n_F}\widehat{\boldsymbol{H}}_i\right)^2}\left[\left(\phi+\widehat{\boldsymbol{H}}_j\right)^2 K_i^2\left(1+\Sigma_{\varepsilon,i}\right) + \left(\widehat{\boldsymbol{H}}_j\right)^2\boldsymbol{K}_j^2\left(1+\Sigma_{\varepsilon,j}\right)\right] \tag{60}$$

Consumer surplus:

$$\mathbb{E}CS = \left\{\frac{\left(\mathbb{E}\left[\widetilde{Q}\right]\right)^2 + \mathbb{V}\left[\widetilde{Q}\right]}{2\phi}\right\} \tag{61}$$

$$= \frac{\left(\widehat{\boldsymbol{H}}_1+\widehat{\boldsymbol{H}}_2\right)^2\left(\boldsymbol{D}-\overline{\boldsymbol{c}}\right)^2}{2\phi} + \frac{\phi}{2\left(\phi+\widehat{\boldsymbol{H}}_1+\widehat{\boldsymbol{H}}_2\right)^2}\left[\left(\widehat{\boldsymbol{H}}_1\boldsymbol{K}_1\right)^2\left(1+\Sigma_{\varepsilon,1}\right) + \left(\widehat{\boldsymbol{H}}_2\boldsymbol{K}_2\right)^2\left(1+\Sigma_{\varepsilon,2}\right)\right] \tag{62}$$

Welfare is thus:

$$\mathbb{W} = \frac{1}{2\left(\phi+\widehat{\boldsymbol{H}}_1+\widehat{\boldsymbol{H}}_2\right)^2}\left\{\left[\frac{\widehat{\boldsymbol{H}}_1^2}{H_1}+\frac{\widehat{\boldsymbol{H}}_2^2}{H_2}+\frac{\left(\widehat{\boldsymbol{H}}_1+\widehat{\boldsymbol{H}}_2\right)^2}{\phi}\right]\phi^2\left(\overline{\boldsymbol{p}}-\overline{\boldsymbol{c}}\right)^2 + \frac{\widehat{\boldsymbol{H}}_1^2}{1+\Sigma_{\varepsilon,1}}\left[\frac{\left(\phi+\widehat{\boldsymbol{H}}_2\right)^2}{H_1}+\frac{\widehat{\boldsymbol{H}}_2^2}{H_2}+\phi\right] + \frac{\widehat{\boldsymbol{H}}_2^2}{1+\Sigma_{\varepsilon,2}}\left[\frac{\left(\phi+\widehat{\boldsymbol{H}}\right.}{H_2}\right.\right. \tag{63}$$

where

$$\boldsymbol{D} = \left(1+\frac{1}{\phi}\sum_{i=1}^{n_F}\widehat{\boldsymbol{H}}_i\right)^{-1}\left(\overline{\boldsymbol{p}}+\frac{1}{\phi}\sum_{i=1}^{n_F}\widehat{H}_i\overline{\boldsymbol{c}}\right) \tag{64}$$

$$\boldsymbol{H}_i = \left(\frac{2}{\phi}+\rho_i\mathbb{V}\left[b_i|\mathcal{I}_i\right]\right)^{-1} \tag{65}$$

$$\boldsymbol{K}_i = \left(1+\Sigma_{\varepsilon,i}\right)^{-1} \tag{66}$$

$$\widehat{\boldsymbol{H}}_i = \left(\frac{1}{\phi}+\rho_i\mathbb{V}\left[b_i|\mathcal{I}_i\right]\right)^{-1} \tag{67}$$

$$\mathbb{V}\left[b_i|\mathcal{I}_i\right] = \frac{\Sigma_{\varepsilon,i}}{1+\Sigma_{\varepsilon,i}} \tag{68}$$

$$\Sigma_{\varepsilon,i} = \widetilde{\Sigma}/n_{di} \tag{69}$$

Proof of Proposition 5 : For simplicity, we denote $n_{d,1} = n_{d,2} = x$. In this case, we can get rid of the i subscripts in (63), which gives us:

$$\mathbb{W} = \frac{\left(x+\widetilde{\Sigma}\right)^2}{\left[\left(x+\widetilde{\Sigma}+\phi\rho_i\widetilde{\Sigma}\right)+2\left(x+\widetilde{\Sigma}\right)\right]^2}\left\{\phi\left(\frac{2x+(2+\phi\rho_i)\,\widetilde{\Sigma}}{x+\widetilde{\Sigma}}+2\right)(\overline{\boldsymbol{p}}-\overline{\boldsymbol{c}})^2+\frac{x}{x+\widetilde{\Sigma}}\left[\frac{\phi\left[2x+(2+\phi\rho_i)\,\widetilde{\Sigma}\right]^3}{\left(x+\widetilde{\Sigma}+\phi\rho_i\widetilde{\Sigma}\right)^2\left(x+\widetilde{\Sigma}\right)}\right.\right. \tag{70}$$

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

$$\mathbb{W} = \frac{\phi y}{\left[\left(y+\phi\rho_i\widetilde{\Sigma}\right)+2y\right]^2}\left\{\left(\phi\rho_i\widetilde{\Sigma}+4y\right)(\overline{\boldsymbol{p}}-\overline{\boldsymbol{c}})^2+\left(y-\widetilde{\Sigma}\right)\left[\frac{\left[2y+\phi\rho_i\widetilde{\Sigma}\right]^3}{y\left(y+\phi\rho_i\widetilde{\Sigma}\right)^2}+\frac{y\left[2y+\phi\rho_i\widetilde{\Sigma}\right]}{\left(y+\phi\rho_i\widetilde{\Sigma}\right)^2}+1\right]\right\} \tag{71}$$

Finally, the derivative is:

$$\frac{\partial\mathbb{W}}{\partial x} = \frac{\partial\mathbb{W}}{\partial y} > \frac{\phi^2\rho\widetilde{\Sigma}\left(5y+\phi\rho_i\widetilde{\Sigma}\right)(\overline{\boldsymbol{p}}-\overline{\boldsymbol{c}})^2}{\left(3y+\phi\rho_i\widetilde{\Sigma}\right)^3}+\frac{\phi\widetilde{\Sigma}}{\left[3y+\phi\rho_i\widetilde{\Sigma}\right]^2}+\phi\frac{2y\left[5y^2+4\phi\rho_i\widetilde{\Sigma}y+\left(\phi\rho_i\widetilde{\Sigma}\right)^2\right]}{\left[3y+\phi\rho_i\widetilde{\Sigma}\right]^2\left(y+\phi\rho_i\widetilde{\Sigma}\right)^2} \tag{72}$$

$$+\frac{\phi^2\rho_i\widetilde{\Sigma}\left\{2.5y\left[5y^2+4\phi\rho_i\widetilde{\Sigma}y+\left(\phi\rho_i\widetilde{\Sigma}\right)^2\right]+\left(y+2\phi\rho_i\widetilde{\Sigma}\right)\left[\frac{1y^2+4\phi\rho_i\widetilde{\Sigma}y+\left(\phi\rho_i\widetilde{\Sigma}\right)^2}{2}\right]\right\}}{\left[3y+\phi\rho_i\widetilde{\Sigma}\right]^3\left(y+\phi\rho_i\widetilde{\Sigma}\right)^2} \tag{73}$$

$$> 0 \tag{74}$$

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 6 : 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:

$$\mathbb{E}CS = \frac{\phi\left(\widehat{\boldsymbol{H}}_1+\widehat{\boldsymbol{H}}_2\right)^2}{2\left(\phi+\widehat{\boldsymbol{H}}_1+\widehat{\boldsymbol{H}}_2\right)^2}\left[(\overline{\boldsymbol{p}}-\overline{\boldsymbol{c}})\right]^2+\frac{\phi}{2\left(\phi+\widehat{\boldsymbol{H}}_1+\widehat{\boldsymbol{H}}_2\right)^2}\left[\frac{\left(\widehat{\boldsymbol{H}}_1\right)^2}{1+\Sigma_{\varepsilon,1}}+\frac{\left(\widehat{\boldsymbol{H}}_2\right)^2}{1+\Sigma_{\varepsilon,2}}\right] \tag{75}$$

$$\mathbb{E}\left[\Pi\right] = \phi\frac{10\phi\rho\widetilde{\Sigma}\left[\left(N+\widetilde{\Sigma}\right)^2-x^2\right]\left(N+\widetilde{\Sigma}\right)+4\left[\left(N+\widetilde{\Sigma}\right)^2-x^2\right]^2+2\left(\phi\rho\widetilde{\Sigma}\right)^2\left(N+\widetilde{\Sigma}\right)\left(4N+4\widetilde{\Sigma}+\phi\rho\widetilde{\Sigma}\right)}{2\left[3\left(N^2-x^2\right)+\left(3+2\phi\rho\right)\widetilde{\Sigma}\left(2N\right)+\left(1+\phi\rho\right)\left(3+\phi\rho\right)\widetilde{\Sigma}^2\right]^2}\left(\overline{\boldsymbol{p}}-\right. \tag{76}$$

$$+\phi\frac{\left[5\left[\left(N+\widetilde{\Sigma}\right)^2-x^2\right]+2\phi\rho\widetilde{\Sigma}\left(2N+2\widetilde{\Sigma}\right)+\left(\phi\rho\widetilde{\Sigma}\right)^2\right]\left[4\left(N^2-x^2\right)+\left(2+\phi\rho\right)\widetilde{\Sigma}\left(2N\right)\right]}{2\left[3\left(N^2-x^2\right)+\left(3+2\phi\rho\right)\widetilde{\Sigma}\left(2N\right)+\left(1+\phi\rho\right)\left(3+\phi\rho\right)\widetilde{\Sigma}^2\right]^2} \tag{77}$$

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 $(\overline{\boldsymbol{p}}-\overline{\boldsymbol{c}})^2$)

$$\mathbb{W}_{(1)} = \phi\left(\overline{\boldsymbol{p}}-\overline{\boldsymbol{c}}\right)^2\frac{\left[\left(\widetilde{N}^2-x^2\right)+\phi\rho\widetilde{\Sigma}\widetilde{N}\right]\left[4\left(\widetilde{N}^2-x^2\right)+5\phi\rho\widetilde{\Sigma}\widetilde{N}\right]+\widetilde{N}\left(\phi\rho\widetilde{\Sigma}\right)^2\left(\widetilde{N}+\phi\rho\widetilde{\Sigma}\right)}{\left[3\left(\widetilde{N}^2-x^2\right)+\left(\phi\rho\widetilde{\Sigma}\right)\left(4\widetilde{N}+\phi\rho\widetilde{\Sigma}\right)\right]^2} \tag{78}$$

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

$$= \phi\left(\overline{\boldsymbol{p}}-\overline{\boldsymbol{c}}\right)^2\frac{\left(y+\phi\rho\widetilde{\Sigma}\widetilde{N}\right)\left(4y+5\phi\rho\widetilde{\Sigma}\widetilde{N}\right)+\widetilde{N}\left(\phi\rho\widetilde{\Sigma}\right)^2\left(\widetilde{N}+\phi\rho\widetilde{\Sigma}\right)}{\left[3y+\left(\phi\rho\widetilde{\Sigma}\right)\left(4\widetilde{N}+\phi\rho\widetilde{\Sigma}\right)\right]^2} \tag{79}$$

The derivative is

$$\frac{\partial\mathbb{W}_{(1)}}{\partial x} = \frac{\partial\mathbb{W}_{(1)}}{\partial y}\frac{\partial y}{\partial x} \tag{80}$$

$$= -2\phi\left(\overline{\boldsymbol{p}}-\overline{\boldsymbol{c}}\right)^2 x\frac{\phi\rho\widetilde{\Sigma}\widetilde{N}\left(5y\right)+\left(\phi\rho\widetilde{\Sigma}\right)^2\left(8y+3\phi\rho\widetilde{\Sigma}\widetilde{N}\right)}{\left[3y+\left(\phi\rho\widetilde{\Sigma}\right)\left(4\widetilde{N}+\phi\rho\widetilde{\Sigma}\right)\right]^3} \tag{81}$$

$$< 0 \tag{82}$$

Therefore, the first part is decreasing over the data asymmetry

The second term:

$$\mathbb{W}_{(2)} = \phi\frac{\left\{\left[5\widetilde{N}^2+\phi\rho\widetilde{\Sigma}\left(4\widetilde{N}+\phi\rho\widetilde{\Sigma}\right)\right]\left(2\widetilde{N}+\phi\rho\widetilde{\Sigma}\right)+\widetilde{N}\left(\widetilde{N}+\phi\rho\widetilde{\Sigma}\right)^2\right\}\left(\widetilde{N}-\widetilde{\Sigma}\right)+x^2\left[11x^2-14\widetilde{N}^2+7\left(1-\phi\rho\right)\right.}{\left[3\left(\widetilde{N}^2-x^2\right)+\left(\phi\rho\widetilde{\Sigma}\right)\left(4\widetilde{N}+\phi\rho\widetilde{\Sigma}\right)\right]^2} \tag{83}$$

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 7: (Aggregation effect) Firm-level markup wedge increases in data.

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

$$M_i^f = \frac{\mathbf{E}[\tilde{\boldsymbol{q}}_i' \tilde{\boldsymbol{p}}_i]}{\mathbf{E}[\tilde{\boldsymbol{q}}_i' \tilde{\boldsymbol{c}}_i]} = \frac{\mathbf{E}[\tilde{\boldsymbol{q}}_i]' \mathbf{E}[\tilde{\boldsymbol{p}}_i] + \mathbf{trCov}\left(\tilde{\boldsymbol{p}}_i, \tilde{\boldsymbol{q}}_i\right)}{\mathbf{E}[\tilde{\boldsymbol{q}}_i' \tilde{\boldsymbol{c}}_i]} = \frac{\sum_{l=1}^N \mathbf{E}[\tilde{\boldsymbol{q}}_{i,l}] \mathbf{E}[\tilde{\boldsymbol{p}}_{i,l}] + \sum_{l=1}^N \mathbf{Cov}_{i,l}}{\sum_{l=1}^N \mathbf{E}[\tilde{\boldsymbol{q}}_{i,l}] \tilde{\boldsymbol{c}}_{i,l}} \tag{84}$$

where $\mathbf{Cov}_{i,l}$ as the l^{th} diagonal values of covariance matrix (**??**) is

$$\mathbf{Cov}_{i,l} = \frac{\hat{\boldsymbol{H}}_{i,l}}{\left(1 + \sum_{s=1}^{n_F} \frac{1}{\phi} \hat{\boldsymbol{H}}_{s,l}\right)^2} \left[\sum_{s=1, s\neq i}^{n_F} \left(\frac{\hat{\boldsymbol{H}}_{s,l}}{\phi}\right)^2 \boldsymbol{K}_{s,l} + \boldsymbol{K}_{i,l} \left(1 + \sum_{s=1, s\neq i}^{n_F} \frac{\hat{\boldsymbol{H}}_{s,l}}{\phi}\right)^2 \right] \tag{85}$$

We already know that

$$\begin{aligned}
\frac{\partial \boldsymbol{K}_{k,l}}{\partial \Sigma_{\epsilon_i,j}^{-1}} &= \delta_{ik} \delta_{jl} \Sigma_{b_i,j}^2 \\
\frac{\partial \hat{\boldsymbol{H}}_{k,l}}{\partial \Sigma_{\epsilon_i,j}^{-1}} &= \delta_{ik} \delta_{jl} \rho_i \hat{\boldsymbol{H}}_{i,j}^2 \Sigma_{b_i,j}^2 \\
\frac{\partial \mathbf{E} \tilde{\boldsymbol{p}}_{k,l}}{\partial \Sigma_{\epsilon_i,j}^{-1}} &= \frac{\partial \boldsymbol{D}_l}{\partial \Sigma_{\epsilon_i,j}^{-1}} = \delta_{jl} \frac{1}{\phi} \frac{\rho_i \hat{\boldsymbol{H}}_{i,j}^2 \Sigma_{b_i,j}^2}{1 + \frac{1}{\phi} \sum_{s=1}^{n_F} \hat{\boldsymbol{H}}_{s,j}} \left(\tilde{\boldsymbol{c}}_{i,j} - \boldsymbol{D}_j\right) \\
\frac{\partial \mathbf{E} \tilde{\boldsymbol{q}}_{k,l}}{\partial \Sigma_{\epsilon_i,j}^{-1}} &= \delta_{jl} \rho_i \hat{\boldsymbol{H}}_{i,j}^2 \Sigma_{b_i,j}^2 \left(\boldsymbol{D}_j - \tilde{\boldsymbol{c}}_{i,j}\right) \left(\delta_{ik} - \frac{\frac{1}{\phi} \hat{\boldsymbol{H}}_{k,j}}{1 + \frac{1}{\phi} \sum_{s=1}^{n_F} \hat{\boldsymbol{H}}_{s,j}}\right)
\end{aligned} \tag{86}$$

Thus the derivative of numerator is

$$\begin{aligned}
\frac{\partial \mathbf{E} \tilde{\boldsymbol{q}}_{i,l} \mathbf{E} \tilde{\boldsymbol{p}}_{i,l}}{\partial \Sigma_{\epsilon_i,j}^{-1}} &= \delta_{jl} \rho_i \hat{\boldsymbol{H}}_{i,j}^2 \Sigma_{b_i,j}^2 \left(\boldsymbol{D}_j - \tilde{\boldsymbol{c}}_{i,j}\right) \left(\frac{1 - \frac{1}{\phi} \hat{\boldsymbol{H}}_{i,j} + \frac{1}{\phi} \sum_{s=1, s\neq i}^{n_F} \hat{\boldsymbol{H}}_{s,j}}{1 + \frac{1}{\phi} \sum_{s=1}^{n_F} \hat{\boldsymbol{H}}_{s,j}} \boldsymbol{D}_j + \frac{\frac{1}{\phi} \hat{\boldsymbol{H}}_{i,j} \tilde{\boldsymbol{c}}_{i,j}}{1 + \frac{1}{\phi} \sum_{s=1}^{n_F} \hat{\boldsymbol{H}}_{s,j}} \right) \geq 0 \\
\frac{\partial \mathbf{Cov}_{i,l}}{\partial \Sigma_{\epsilon_i,j}^{-1}} &= \delta_{jl} \hat{\boldsymbol{H}}_{i,j} \Sigma_{b_i,j}^2 \left(\frac{\rho_i \left(1 - \frac{1}{\phi} \hat{\boldsymbol{H}}_{i,j} + \frac{1}{\phi} \sum_{s=1, s\neq i}^{n_F} \hat{\boldsymbol{H}}_{s,j}\right)}{\left(1 + \sum_{s=1}^{n_F} \frac{1}{\phi} \hat{\boldsymbol{H}}_{s,j}\right)} \mathbf{Cov}_{i,j} + \frac{\left(1 + \sum_{s=1, s\neq i}^{n_F} \frac{1}{\phi} \hat{\boldsymbol{H}}_{s,j}\right)^2}{\left(1 + \sum_{s=1}^{n_F} \frac{1}{\phi} \hat{\boldsymbol{H}}_{s,j}\right)^2} \right) \geq 0
\end{aligned} \tag{87}$$

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

$$\frac{\partial M_i^f}{\partial \Sigma_{\epsilon_i,j}^{-1}} = \frac{\left(\frac{\partial \mathbf{E} \tilde{\boldsymbol{q}}_{i,j} \mathbf{E} \tilde{\boldsymbol{p}}_{i,j}}{\partial \Sigma_{\epsilon_i,j}^{-1}} + \frac{\partial \mathbf{Cov}_{i,j}}{\partial \Sigma_{\epsilon_i,j}^{-1}} \right)}{\sum_{l=1}^N \mathbf{E} \tilde{\boldsymbol{q}}_{i,l} \tilde{\boldsymbol{c}}_{i,l}} - \frac{\left(\sum_{l=1}^N \mathbf{E}[\tilde{\boldsymbol{q}}_{i,l} \tilde{\boldsymbol{p}}_{i,l}] \right) \frac{\partial \mathbf{E} \tilde{\boldsymbol{q}}_{i,j}}{\partial \Sigma_{\epsilon_i,j}^{-1}} \tilde{\boldsymbol{c}}_{i,j}}{\left(\sum_{l=1}^N \mathbf{E} \tilde{\boldsymbol{q}}_{i,l} \tilde{\boldsymbol{c}}_{i,l} \right)^2} \quad \text{and} \quad \lim_{\rho_i \to 0} \frac{\partial M_i^f}{\partial \Sigma_{\epsilon_i,j}^{-1}} = \frac{\hat{\boldsymbol{H}}_{i,j} \Sigma_{b_i,j}^2 \frac{\left(1 + \sum_{s=1, s\neq i}^{n_F} \frac{1}{\phi} \hat{\boldsymbol{H}}_{s,j}\right)^2}{\left(1 + \sum_{s=1}^{n_F} \frac{1}{\phi} \hat{\boldsymbol{H}}_{s,j}\right)^2}}{\sum_{l=1}^N \mathbf{E} \tilde{\boldsymbol{q}}_{i,l} \tilde{\boldsymbol{c}}_{i,l}} > 0 \tag{88}$$

□

Proof of Proposition 8: Industry-level markup wedge increases in data.

Proof. The cost weight for firm i is

$$w_i^{cost} = \frac{\mathbf{E}\left[\tilde{\boldsymbol{q}}_i'\tilde{\boldsymbol{c}}_i\right]}{\sum_{k=1}^{n_F}\mathbf{E}\left[\tilde{\boldsymbol{q}}_k'\tilde{\boldsymbol{c}}_k\right]} = \frac{\sum_{l=1}^{N}\mathbf{E}[\tilde{\boldsymbol{q}}_{i,l}]\tilde{\boldsymbol{c}}_{i,l}}{\sum_{k=1}^{n_F}\sum_{l=1}^{N}\mathbf{E}[\tilde{\boldsymbol{q}}_{k,l}]\tilde{\boldsymbol{c}}_{k,l}} \tag{89}$$

This weight is increasing in data for the firm i since

$$\begin{aligned}\frac{\partial w_i^{cost}}{\partial\Sigma_{\epsilon_i,j}^{-1}} &= \frac{\frac{\partial\mathbf{E}[\tilde{\boldsymbol{q}}_{i,j}]}{\partial\Sigma_{\epsilon_i,j}^{-1}}\tilde{\boldsymbol{c}}_{i,j}\left(\sum_{k=1,k\neq i}^{n_F}\mathbf{E}\left[\tilde{\boldsymbol{q}}_k'\tilde{\boldsymbol{c}}_k\right]\right) - \mathbf{E}\left[\tilde{\boldsymbol{q}}_i'\tilde{\boldsymbol{c}}_i\right]\sum_{k=1,k\neq i}^{n_F}\tilde{\boldsymbol{c}}_{k,j}\frac{\partial\mathbf{E}[\tilde{\boldsymbol{q}}_{k,j}]}{\partial\Sigma_{\epsilon_i,j}^{-1}}}{\left(\sum_{k=1}^{n_F}\mathbf{E}\left[\tilde{\boldsymbol{q}}_k'\tilde{\boldsymbol{c}}_k\right]\right)^2} \\ &= \rho_i\hat{\boldsymbol{H}}_{i,j}^2\Sigma_{b_i,j}^2\left(\boldsymbol{D}_j - \tilde{\boldsymbol{c}}_{i,j}\right)\left[\frac{\tilde{\boldsymbol{c}}_{i,j}\left(\sum_{k=1,k\neq i}^{n_F}\mathbf{E}\left[\tilde{\boldsymbol{q}}_k'\tilde{\boldsymbol{c}}_k\right]\right)}{\left(\sum_{k=1}^{n_F}\mathbf{E}\left[\tilde{\boldsymbol{q}}_k'\tilde{\boldsymbol{c}}_k\right]\right)^2}\frac{1+\sum_{s\neq i,s=1}^{n_F}\frac{\hat{\boldsymbol{H}}_{s,j}}{\phi}}{1+\sum_{s=1}^{n_F}\frac{\hat{\boldsymbol{H}}_{s,j}}{\phi}} + \frac{\mathbf{E}\left[\tilde{\boldsymbol{q}}_i'\tilde{\boldsymbol{c}}_i\right]\sum_{k=1,k\neq i}^{n_F}\frac{\frac{1}{\phi}\hat{\boldsymbol{H}}_{k,j}\tilde{\boldsymbol{c}}_{k,j}}{1+\frac{1}{\phi}\sum_{s=1}^{n_F}\hat{\boldsymbol{H}}_{s,j}}}{\left(\sum_{k=1}^{n_F}\mathbf{E}\left[\tilde{\boldsymbol{q}}_k'\tilde{\boldsymbol{c}}_k\right]\right)^2}\right] \geq 0\end{aligned} \tag{90}$$

This result indicates that high-data firms produce more on average and have larger impacts on cost-weighted industry markup, thus increasing the industry-level markup wedge.

□

B.1 Cyclical Markups, proposition 9

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

$$\begin{aligned}&\hat{\boldsymbol{H}}_i = \left(\frac{\boldsymbol{I}_N}{\phi} + \rho_i\mathbf{Var}(\tilde{\boldsymbol{p}}_i|\mathcal{I}_i)\right)^{-1} \quad \text{and} \quad \mathbf{Var}(\tilde{\boldsymbol{p}}_i|\mathcal{I}_i) = \left(\Sigma_b^{-1} + \Sigma_{\epsilon_i}^{-1}\right)^{-1} \\ &\Rightarrow \lim_{\Sigma_b\to\infty}\mathbf{Var}(\tilde{\boldsymbol{p}}_i|\mathcal{I}_i) = \Sigma_{\epsilon_i}, \quad \tilde{\hat{\boldsymbol{H}}}_i := \lim_{\Sigma_b\to\infty}\hat{\boldsymbol{H}}_i = \left(\frac{\boldsymbol{I}_N}{\phi} + \rho_i\Sigma_{\epsilon_i}\right)^{-1}\end{aligned} \tag{91}$$

The equilibrium price is given by

$$\mathbf{E}\left[\tilde{\boldsymbol{p}}_i\right] = \boldsymbol{D} = \left(\boldsymbol{I}_N + \frac{1}{\phi}\sum_{i=1}^{n_F}\hat{\boldsymbol{H}}_i\right)^{-1}\left(\bar{\boldsymbol{p}} + \frac{1}{\phi}\sum_{i=1}^{n_F}\hat{\boldsymbol{H}}_i\boldsymbol{c}_i\right) \tag{92}$$

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

$$\begin{aligned}\tilde{\boldsymbol{p}} := \lim_{\Sigma_b\to\infty}\mathbf{E}\left[\tilde{\boldsymbol{p}}_i\right] &= \left(\boldsymbol{I}_N + \frac{1}{\phi}\sum_{i=1}^{n_F}\lim_{\Sigma_b\to\infty}\hat{\boldsymbol{H}}_i\right)^{-1}\left(\bar{\boldsymbol{p}} + \frac{1}{\phi}\sum_{i=1}^{n_F}\lim_{\Sigma_b\to\infty}\hat{\boldsymbol{H}}_i\boldsymbol{c}_i\right) \\ &= \left[\boldsymbol{I}_N + \sum_{i=1}^{n_F}\left(\boldsymbol{I}_N + \phi\rho_i\Sigma_{\epsilon_i}\right)^{-1}\right]^{-1}\left[\bar{\boldsymbol{p}} + \sum_{i=1}^{n_F}\boldsymbol{c}_i\left(\boldsymbol{I}_N + \phi\rho_i\Sigma_{\epsilon_i}\right)^{-1}\right]\end{aligned} \tag{93}$$

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.

$$\boldsymbol{q}_i = \boldsymbol{A}\tilde{\boldsymbol{q}}_i \quad \text{and} \quad \boldsymbol{p}_i = \boldsymbol{A}\tilde{\boldsymbol{p}}_i \Rightarrow \overline{M}^p = \frac{1}{N}\frac{1}{n_F}\sum_{i=1}^{n_F}\sum_{j=1}^{N}\frac{(\boldsymbol{A}\mathbf{E}[\tilde{\boldsymbol{p}}_i])_j}{(\boldsymbol{A}\boldsymbol{c}_i)_j} \quad \text{converges.} \tag{94}$$

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

$$\frac{\partial \mathbf{E}[\tilde{\boldsymbol{p}}_{i,j}]}{\partial \Sigma_{b,j}} = \frac{(\bar{\boldsymbol{c}}_j - \bar{\boldsymbol{p}}_j)\frac{1}{\phi}\sum_{i=1}^{n_F}\frac{\partial \hat{\boldsymbol{H}}_{i,j}}{\partial \Sigma_{b,j}}}{\left(1 + \frac{1}{\phi}\sum_{i=1}^{n_F}\hat{\boldsymbol{H}}_{i,j}\right)^2} \quad \text{and} \quad \frac{\partial \hat{\boldsymbol{H}}_{i,j}}{\partial \Sigma_{b,j}} = -\frac{\hat{\boldsymbol{H}}_{i,j}^2 \rho_i \Sigma_{\epsilon_i,j}^2}{\left(\Sigma_{b.j} + \Sigma_{\epsilon_i,j}\right)^2} \leq 0 \tag{95}$$

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.

$$\begin{aligned}\mathbf{Cov}(\tilde{\boldsymbol{p}}_i, \tilde{\boldsymbol{q}}_i) = &\left(\boldsymbol{I}_N + \sum_{j=1}^{n_F}\frac{\hat{\boldsymbol{H}}_j}{\phi}\right)^{-1}\sum_{j=1}^{n_F}\hat{\boldsymbol{H}}_j\mathbf{Var}(\boldsymbol{K}_j\boldsymbol{s}_j)\hat{\boldsymbol{H}}_j\left(\boldsymbol{I}_N + \sum_{j=1}^{n_F}\frac{\hat{\boldsymbol{H}}_j}{\phi}\right)^{-1}\frac{\hat{\boldsymbol{H}}_i}{\phi^2} + \mathbf{Var}(\boldsymbol{K}_i\boldsymbol{s}_i)\hat{\boldsymbol{H}}_i \\ &- \left(\boldsymbol{I}_N + \sum_{j=1}^{n_F}\frac{\hat{\boldsymbol{H}}_j}{\phi}\right)^{-1}\hat{\boldsymbol{H}}_i\mathbf{Var}(\boldsymbol{K}_i\boldsymbol{s}_i)\frac{\hat{\boldsymbol{H}}_i}{\phi} - \mathbf{Var}(\boldsymbol{K}_i\boldsymbol{s}_i)\hat{\boldsymbol{H}}_i\left(\boldsymbol{I}_N + \sum_{j=1}^{n_F}\frac{\hat{\boldsymbol{H}}_j}{\phi}\right)^{-1}\frac{\hat{\boldsymbol{H}}_i}{\phi}\end{aligned} \tag{96}$$

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{97}$$

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

$$\mathbf{Cov}\left(\tilde{\boldsymbol{p}}_i, \tilde{\boldsymbol{q}}_i\right) = \phi\left[\boldsymbol{Y}^{-1}\sum_{j=1}^{n_F}\boldsymbol{x}_i Z_j \boldsymbol{x}_i \boldsymbol{Y}^{-1}\boldsymbol{x}_i + Z_i\boldsymbol{x}_i - \boldsymbol{Y}^{-1}\boldsymbol{x}_i Z_i \boldsymbol{x}_i - Z_i \boldsymbol{x}_i \boldsymbol{Y}^{-1}\boldsymbol{x}_i\right] \tag{98}$$

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.

$$\begin{aligned}\mathbf{Cov}_{i,k} &:= \mathbf{Cov}\left(\boldsymbol{p}_{i,k}, \tilde{\boldsymbol{q}}_{i,k}\right)\\ &= \phi\left[\boldsymbol{Y}_k^{-1}\sum_{j=1}^{n_F}\boldsymbol{x}_{j,k}Z_{j,k}\boldsymbol{x}_{j,k}\boldsymbol{Y}_k^{-1}\boldsymbol{x}_{i,k} + Z_i\boldsymbol{x}_{i,k} - \boldsymbol{Y}_k^{-1}\boldsymbol{x}_{i,k}Z_{i,k}\boldsymbol{x}_{i,k} - Z_{i,k}\boldsymbol{x}_{i,k}\boldsymbol{Y}_k^{-1}\boldsymbol{x}_{i,k}\right]\\ &= \phi\frac{\boldsymbol{x}_{i,k}}{\boldsymbol{Y}_k^2}\left[\sum_{j\neq i, j=1}^{n_F}\boldsymbol{x}_{j,k}^2 Z_{j,k} + Z_{i,k}\left(\boldsymbol{x}_{i,k} - \boldsymbol{Y}_k\right)^2\right]\end{aligned} \tag{99}$$

The limiting behavioral for all variables are

$$\begin{aligned}\lim_{\Sigma_{b,k}\to\infty}\boldsymbol{x}_{i,k} &= \left(1 + \phi\rho_i\Sigma_{\epsilon_i,k}\right)^{-1}\\ \lim_{\Sigma_{b,k}\to\infty}\boldsymbol{Y}_k &= 1 + \sum_{j=1}^{n_F}\lim_{\Sigma_{b,k}\to\infty}\boldsymbol{x}_{j,k} = 1 + \sum_{j=1}^{n_F}\left(1 + \phi\rho_j\Sigma_{\epsilon_j,k}\right)^{-1}\\ \lim_{\Sigma_{b,k}\to\infty}\frac{Z_{i,k}}{\Sigma_{b,k}} &= \lim_{\Sigma_{b,k}\to\infty}\frac{\frac{\Sigma_{b,k}^2}{\Sigma_{b,k}+\Sigma_{i,k}}}{\Sigma_{b,k}} = 1\end{aligned} \tag{100}$$

The ratio of covariance to shock variance converges as

$$\lim_{\Sigma_{b,k}\to\infty}\frac{\mathbf{Cov}_{i,k}}{\Sigma_{b,k}} = \frac{\phi\left(1+\phi\rho_i\Sigma_{\epsilon_i,k}\right)^{-1}\left[\sum_{j\neq i,j=1}^{n_F}\left(1+\phi\rho_j\Sigma_{\epsilon_j,k}\right)^{-2} + \left(1 + \sum_{j=1,j\neq i}^{n_F}\left(1+\phi\rho_j\Sigma_{\epsilon_j,k}\right)^{-1}\right)^2\right]}{\left(1+\sum_{j=1}^{n_F}\left(1+\phi\rho_j\Sigma_{\epsilon_j,k}\right)^{-1}\right)^2} \tag{101}$$

we have

$$\begin{aligned}M_i^f &= \frac{\mathbf{E}[\tilde{\boldsymbol{q}}_i'\tilde{\boldsymbol{p}}_i]}{\mathbf{E}[\tilde{\boldsymbol{q}}_i'\boldsymbol{c}_i]} = \frac{\mathbf{E}[\tilde{\boldsymbol{q}}_i]'\mathbf{E}[\boldsymbol{p}] + \mathbf{tr}\left[\mathbf{Cov}(\tilde{\boldsymbol{p}}_i, \tilde{\boldsymbol{q}}_i)\right]}{\mathbf{E}[\tilde{\boldsymbol{q}}_i'\boldsymbol{c}_i]}\\ &= \frac{\sum_{j=1}^{N}(\mathbf{E}(\tilde{\boldsymbol{p}}_{i,j}) - \boldsymbol{c}_{i,j})\mathbf{E}(\tilde{\boldsymbol{p}}_{i,j})\hat{\boldsymbol{H}}_{i,j} + \sum_{j=1}^{N}\mathbf{Cov}_{i,j}}{\sum_{j=1}^{N}(\mathbf{E}(\tilde{\boldsymbol{p}}_{i,j}) - \boldsymbol{c}_{i,j})\boldsymbol{c}_{i,j}\hat{\boldsymbol{H}}_{i,j}}\end{aligned} \tag{102}$$

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

$$\begin{aligned}\alpha_i &:= \lim_{\Sigma_b\to\infty}\frac{M_i^f}{\Sigma_b} = \frac{\sum_{j=1}^{N}\lim_{\Sigma_b\to\infty}\frac{\mathbf{Cov}_{i,j}}{\Sigma_{b,k}}}{\sum_{j=1}^{N}(\tilde{\boldsymbol{p}}_j - \boldsymbol{c}_{i,j})\boldsymbol{c}_{i,j}\tilde{\hat{\boldsymbol{H}}}_{i,j}} > 0\\ \gamma_i &:= \lim_{\Sigma_b\to\infty}\left(M_i^f - \alpha_i\Sigma_b\right) = \frac{\sum_{j=1}^{N}(\tilde{\boldsymbol{p}}_j - \boldsymbol{c}_{i,j})\tilde{\boldsymbol{p}}_j\tilde{\hat{\boldsymbol{H}}}_{i,j} + \widetilde{\mathbf{Cov}}_{i,j}}{\sum_{j=1}^{N}(\tilde{\boldsymbol{p}}_j - \boldsymbol{c}_{i,j})\boldsymbol{c}_{i,j}\tilde{\hat{\boldsymbol{H}}}_{i,j}}\end{aligned} \tag{103}$$

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

$$\widetilde{\mathbf{Cov}}_i := \lim_{\Sigma_b\to\infty}\left(\mathbf{Cov}_{i,k} - \left(\lim_{\Sigma_b\to\infty}\frac{\mathbf{Cov}_{i,k}}{\Sigma_b}\right)\Sigma_b\right)$$
$$= -\frac{\phi\left(1+\phi\rho_i\Sigma_{\epsilon_i,k}\right)^{-1}\left[\sum_{j\neq i,j=1}^{n_F}\left(1+\phi\rho_j\Sigma_{\epsilon_j,k}\right)^{-2}\Sigma_{\epsilon_j,k} + \left(1+\sum_{j=1,j\neq i}^{n_F}\left(1+\phi\rho_j\Sigma_{\epsilon_j,k}\right)^{-1}\right)^2\Sigma_{\epsilon_i,k}\right]}{\left(1+\sum_{j=1}^{n_F}\left(1+\phi\rho_j\Sigma_{\epsilon_j,k}\right)^{-1}\right)^2} \tag{104}$$

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.

$$w_i := \lim_{\Sigma_b\to\infty} w^{\boldsymbol{H}_i} = \frac{\sum_{j=1}^{N}(\tilde{\boldsymbol{p}}_j - \boldsymbol{c}_{i,j})\boldsymbol{c}_{i,j}\tilde{\tilde{\boldsymbol{H}}}_{i,j}}{\sum_{i=1}^{n_F}\sum_{j=1}^{N}(\tilde{\boldsymbol{p}}_j - \boldsymbol{c}_{i,j})\boldsymbol{c}_{i,j}\tilde{\tilde{\boldsymbol{H}}}_{i,j}} \Rightarrow M^m \text{ approaches } \sum_{i=1}^{n_F} w_i\alpha_i\Sigma_b + \sum_{i=1}^{n_F} w_i\gamma_i \tag{105}$$

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{106}$$

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

$$\frac{\partial\mathbf{Cov}_{i,k}}{\partial\Sigma_{b,k}} = \phi\frac{\boldsymbol{x}_{i,k}}{\boldsymbol{Y}_k^2}\left[\boldsymbol{G}_1+\boldsymbol{G}_2\right] \tag{107}$$

where

$$
\begin{aligned}
\boldsymbol{G}_1 &:= Z_{i,k}\left(\boldsymbol{x}_{i,k}-\boldsymbol{Y}_k\right)^2 \frac{\Sigma_{b,k}+(1+\boldsymbol{x}_{i,k})\Sigma_{\epsilon_i,k}}{\Sigma_{b,k}+\Sigma_{\epsilon_i,k}} - 2Z_{i,k}\left(\boldsymbol{Y}_k-\boldsymbol{x}_{i,k}\right)\frac{\boldsymbol{x}_{i,k}}{\boldsymbol{Y}_k}\sum_{j\neq i,j=1}^{n_F}\frac{\boldsymbol{x}_{j,k}(1-\boldsymbol{x}_{j,k})\Sigma_{\epsilon_j,k}}{\Sigma_{b,k}+\Sigma_{\epsilon_j,k}} \\
\boldsymbol{G}_2 &:= \frac{\Sigma_{b,k}+\boldsymbol{x}_{i,k}\Sigma_{\epsilon_i,k}}{\Sigma_{b,k}+\Sigma_{\epsilon_i,k}}\sum_{j\neq i,j=1}^{n_F}\boldsymbol{x}_{j,k}^2 Z_{j,k} + \frac{2}{\boldsymbol{Y}_k}\sum_{j=1}^{n_F}\frac{\boldsymbol{x}_{j,k}(1-\boldsymbol{x}_{j,k})\Sigma_{\epsilon_j,k}}{\Sigma_{b,k}+\Sigma_{\epsilon_j,k}}\sum_{j\neq i,j=1}^{n_F}\boldsymbol{x}_{j,k}^2 Z_{j,k} \\
&\qquad + \sum_{j\neq i,j=1}^{n_F}\boldsymbol{x}_{j,k}^2 Z_{j,k}\frac{\Sigma_{\epsilon_j,k}}{\Sigma_{b,k}+\Sigma_{\epsilon_j,k}}\left[1-2(1-\boldsymbol{x}_{j,k})\right]
\end{aligned} \tag{108}
$$

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

$$
\boldsymbol{G}_1 \geq 0 \Leftrightarrow Z_{i,k}\left(\boldsymbol{x}_{i,k}-\boldsymbol{Y}_k\right)^2 \frac{\Sigma_{b,k}+(1+\boldsymbol{x}_{i,k})\Sigma_{\epsilon_i,k}}{\Sigma_{b,k}+\Sigma_{\epsilon_i,k}} \geq 2Z_{i,k}\left(\boldsymbol{Y}_k-\boldsymbol{x}_{i,k}\right)\frac{\boldsymbol{x}_{i,k}}{\boldsymbol{Y}_k}\sum_{j\neq i,j=1}^{n_F}\frac{\boldsymbol{x}_{j,k}(1-\boldsymbol{x}_{j,k})\Sigma_{\epsilon_j,k}}{\Sigma_{b,k}+\Sigma_{\epsilon_j,k}} \tag{109}
$$

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

$$
\begin{aligned}
Z_{i,k}\left(\boldsymbol{x}_{i,k}-\boldsymbol{Y}_k\right)^2 \frac{\Sigma_{b,k}+(1+\boldsymbol{x}_{i,k})\Sigma_{\epsilon_i,k}}{\Sigma_{b,k}+\Sigma_{\epsilon_i,k}} &\geq Z_{i,k}\left(\boldsymbol{Y}_k-\boldsymbol{x}_{i,k}\right)^2 \\
&\geq Z_{i,k}\left(\boldsymbol{Y}_k-\boldsymbol{x}_{i,k}\right)\frac{2\boldsymbol{x}_{i,k}}{\boldsymbol{Y}_k}\left(1+\sum_{j\neq i,j=1}^{n_F}\boldsymbol{x}_{j,k}\right) \\
&\geq Z_{i,k}\left(\boldsymbol{Y}_k-\boldsymbol{x}_{i,k}\right)\frac{2\boldsymbol{x}_{i,k}}{\boldsymbol{Y}_k}\sum_{j\neq i,j=1}^{n_F}\frac{\boldsymbol{x}_{j,k}(1-\boldsymbol{x}_{j,k})\Sigma_{\epsilon_j,k}}{\Sigma_{b,k}+\Sigma_{\epsilon_j,k}}
\end{aligned} \tag{110}
$$

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

$$
\begin{aligned}
&\Sigma_{b,k} \geq \Sigma_{\epsilon_j,k} \Rightarrow \frac{\Sigma_{b,k}}{\Sigma_{b,k}+\Sigma_{\epsilon_i,k}} \geq \frac{\Sigma_{\epsilon_j,k}}{\Sigma_{b,k}+\Sigma_{\epsilon_j,k}} \\
&\Rightarrow \frac{\Sigma_{b,k}+\boldsymbol{x}_{i,k}\Sigma_{\epsilon_i,k}}{\Sigma_{b,k}+\Sigma_{\epsilon_i,k}}\sum_{j\neq i,j=1}^{n_F}\boldsymbol{x}_{j,k}^2 Z_{j,k} \geq \sum_{j\neq i,j=1}^{n_F}\boldsymbol{x}_{j,k}^2 Z_{j,k}\frac{\Sigma_{\epsilon_j,k}}{\Sigma_{b,k}+\Sigma_{\epsilon_j,k}} \\
&\Rightarrow \boldsymbol{G}_2 \geq \sum_{j\neq i,j=1}^{n_F}\boldsymbol{x}_{j,k}^2 Z_{j,k}\frac{\Sigma_{\epsilon_j,k}}{\Sigma_{b,k}+\Sigma_{\epsilon_j,k}}\left[2-2(1-\boldsymbol{x}_{j,k})\right] \geq 0
\end{aligned} \tag{111}
$$

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

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.

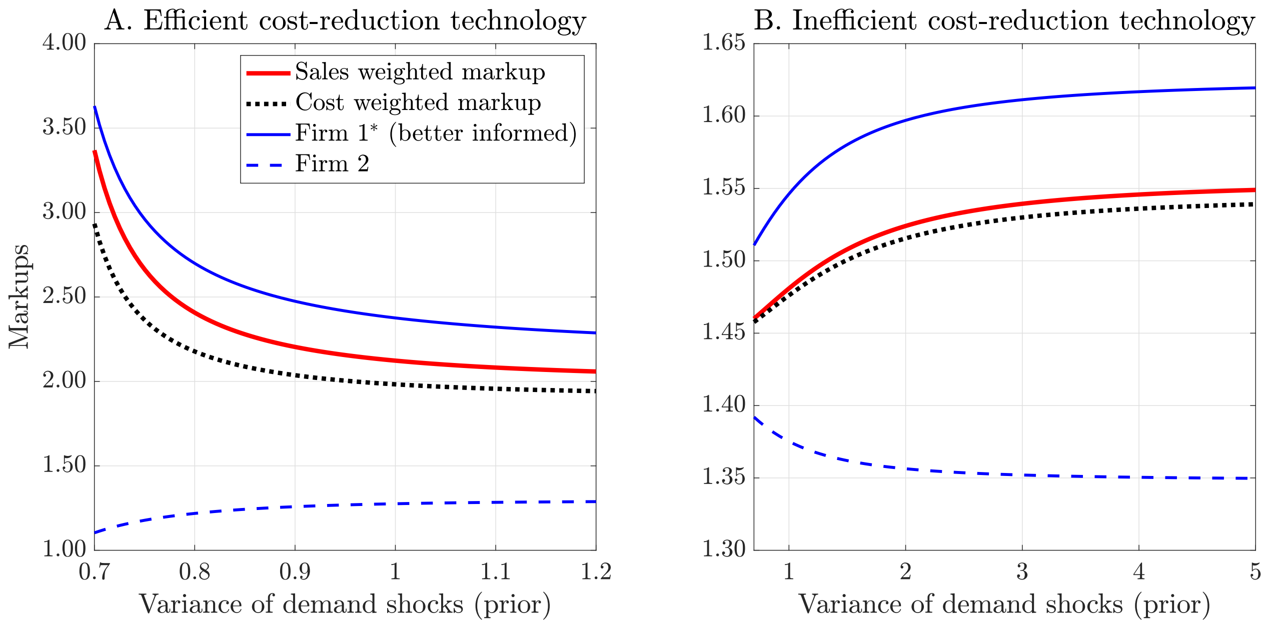

Figure 8: Comparative static: markups and economic cycles, $\rho = 1$

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, $\rho_1 = \rho_2 = 1$.

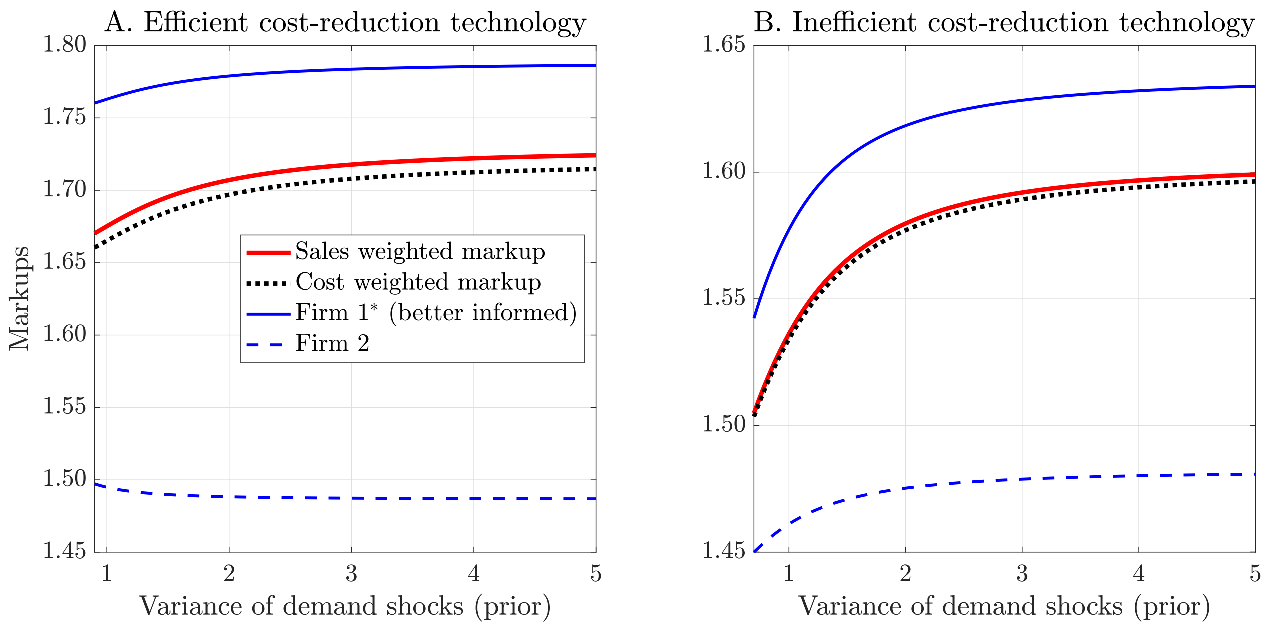

Figure 9: Comparative static: markups and economic cycles, $\rho = 4$

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.

C.2 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 $\bar{\boldsymbol{p}}$

$$\frac{1}{\phi}\sum_{i=1}^{n_F}\tilde{\boldsymbol{q}}_i = \bar{\boldsymbol{p}} + \boldsymbol{b} - \boldsymbol{p} \tag{112}$$

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.

C.3 Solution

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

$$\tilde{\boldsymbol{q}}_i = \left(\rho_i \mathbf{Var}\left[\boldsymbol{p}_i|\mathcal{I}_i\right] - \frac{\partial \mathbf{E}\left[\boldsymbol{p}_i|\mathcal{I}_i\right]}{\partial \tilde{\boldsymbol{q}}_i}\right)^{-1}\left(\mathbf{E}\left[\boldsymbol{p}_i|\mathcal{I}_i\right] - \boldsymbol{c}_i\right) \tag{113}$$

From differentiating the pricing function (112), we find that the price impact of one additional unit of attribute output is

$$\frac{\partial \mathbf{E}\left[\boldsymbol{p}_i|\mathcal{I}_i\right]}{\partial \tilde{\boldsymbol{q}}_i} = -\frac{1}{\phi}\boldsymbol{I}_N \Rightarrow \hat{\boldsymbol{H}}_i \equiv \left(\rho_i \mathbf{Var}\left[\boldsymbol{p}_i|\mathcal{I}_i\right] + \frac{\boldsymbol{I}_N}{\phi}\right)^{-1} \tag{114}$$

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 $\{\boldsymbol{h}_i\}_{i=1,\dots,n_F}$.

$$\boldsymbol{p} = \boldsymbol{D} + \boldsymbol{F}\boldsymbol{b} + \sum_{i=1}^{n_F}\boldsymbol{h}_i\boldsymbol{\varepsilon}_i \tag{115}$$

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

$$\mathbf{E}[\boldsymbol{p}|\boldsymbol{s}_i] = \boldsymbol{D} + \boldsymbol{\beta}_i\boldsymbol{s}_i \text{ where } \boldsymbol{\beta}_i = \mathbf{Cov}(\boldsymbol{p},\boldsymbol{s}_i)\mathbf{Var}(\boldsymbol{s}_i)^{-1} \tag{116}$$

The variance of price forecast error is

$$\mathbf{Var}[\boldsymbol{p}|\boldsymbol{s}_i] = \mathbf{Var}(\boldsymbol{p}) - \mathbf{Cov}(\boldsymbol{p},\boldsymbol{s}_i)\mathbf{Var}(\boldsymbol{s}_i)^{-1}\mathbf{Cov}(\boldsymbol{p},\boldsymbol{s}_i)' \tag{117}$$

The optimal production is

$$\begin{aligned} \tilde{\boldsymbol{q}}_i &= \hat{\boldsymbol{H}}_i(\boldsymbol{D} - \boldsymbol{c}_i + \boldsymbol{\beta}_i \boldsymbol{s}_i) \\ \mathbf{E}[\tilde{\boldsymbol{q}}_i] &= \hat{\boldsymbol{H}}_i(\boldsymbol{D} - \boldsymbol{c}_i) \end{aligned} \tag{118}$$

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

$$\begin{aligned} & \frac{1}{\phi}\sum_{i=1}^{n_F} \tilde{\boldsymbol{q}}_i = \bar{\boldsymbol{p}} + \boldsymbol{b} - \boldsymbol{p} \\ \Rightarrow & \frac{1}{\phi}\sum_{i=1}^{n_F} \hat{\boldsymbol{H}}_i(\boldsymbol{D} - \boldsymbol{c}_i + \boldsymbol{\beta}_i \boldsymbol{s}_i) = \bar{\boldsymbol{p}} + \boldsymbol{b} - \left(\boldsymbol{D} + \boldsymbol{F}\boldsymbol{b} + \sum_{i=1}^{n_F} \boldsymbol{h}_i \boldsymbol{\varepsilon}_i\right) \\ \Rightarrow & \left(\boldsymbol{F} - \boldsymbol{I}_N + \frac{1}{\phi}\sum_{i=1}^{n_F} \hat{\boldsymbol{H}}_i \boldsymbol{\beta}_i\right)\boldsymbol{b} + \sum_{i=1}^{n_F}\left(\boldsymbol{h}_i + \frac{1}{\phi}\hat{\boldsymbol{H}}_i \boldsymbol{\beta}_i\right)\boldsymbol{\varepsilon}_i + \left(\boldsymbol{I}_N + \frac{1}{\phi}\sum_{i=1}^{n_F} \hat{\boldsymbol{H}}_i\right)\boldsymbol{D} - \bar{\boldsymbol{p}} - \frac{1}{\phi}\sum_{i=1}^{n_F} \hat{\boldsymbol{H}}_i \boldsymbol{c}_i = 0 \end{aligned} \tag{119}$$

So the coefficients must satisfy

$$\begin{aligned} \boldsymbol{F} &= \boldsymbol{I}_N - \frac{1}{\phi}\sum_{i=1}^{n_F} \hat{\boldsymbol{H}}_i \boldsymbol{\beta}_i \\ \boldsymbol{h}_i &= -\frac{1}{\phi}\hat{\boldsymbol{H}}_i \boldsymbol{\beta}_i, \ \ \forall\, i = 1, \ldots, n_F \\ \boldsymbol{D} &= \left(\boldsymbol{I}_N + \frac{1}{\phi}\sum_{i=1}^{n_F} \hat{\boldsymbol{H}}_i\right)^{-1}\left(\bar{\boldsymbol{p}} + \frac{1}{\phi}\sum_{i=1}^{n_F} \hat{\boldsymbol{H}}_i \boldsymbol{c}_i\right) \end{aligned} \tag{120}$$

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

$$\begin{aligned} \boldsymbol{\beta}_i &= (\boldsymbol{F}\Sigma_b + \boldsymbol{h}_i \Sigma_{\epsilon_i})(\Sigma_b + \Sigma_{\epsilon_i})^{-1} \\ \hat{\boldsymbol{H}}_i &= \left[\rho_i\left(\boldsymbol{F}\Sigma_b \boldsymbol{F}' + \sum_{i=1}^{n_F} \boldsymbol{h}_i^2 \Sigma_{\epsilon_i} - (\boldsymbol{F}\Sigma_b + \boldsymbol{h}_i \Sigma_{\epsilon_i})(\Sigma_b + \Sigma_{\epsilon_i})^{-1}(\boldsymbol{F}\Sigma_b + \boldsymbol{h}_i \Sigma_{\epsilon_i})'\right) + \frac{\boldsymbol{I}_N}{\phi}\right]^{-1} \end{aligned} \tag{121}$$

C.4 Markups

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

$$\overline{M}^p = \frac{1}{N}\frac{1}{n_F}\sum_{i=1}^{n_F}\sum_{j=1}^{N} M_{ij}^p = \frac{1}{N}\frac{1}{n_F}\sum_{i=1}^{n_F}\sum_{j=1}^{N}\frac{\boldsymbol{D}(j)}{\boldsymbol{c}_i(j)} \tag{122}$$

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

$$\begin{aligned} M_i^f = \frac{\mathbf{E}[\tilde{\boldsymbol{q}}_i' \boldsymbol{p}]}{\mathbf{E}[\tilde{\boldsymbol{q}}_i' \boldsymbol{c}_i]} &= \frac{\mathbf{E}[\tilde{\boldsymbol{q}}_i]'\mathbf{E}[\boldsymbol{p}] + \mathbf{tr}\left[\mathbf{Cov}(\boldsymbol{p}_i, \tilde{\boldsymbol{q}}_i)\right]}{\mathbf{E}[\tilde{\boldsymbol{q}}_i' \boldsymbol{c}_i]} \\ &= \frac{(\boldsymbol{D} - \boldsymbol{c}_i)'\hat{\boldsymbol{H}}_i \boldsymbol{D} + \mathbf{tr}\left(\hat{\boldsymbol{H}}_i \boldsymbol{\beta}_i \mathbf{Var}(\boldsymbol{s}_i)\boldsymbol{\beta}_i'\right)}{(\boldsymbol{D} - \boldsymbol{c}_i)'\hat{\boldsymbol{H}}_i \boldsymbol{c}_i} > \frac{(\boldsymbol{D} - \boldsymbol{c}_i)'\hat{\boldsymbol{H}}_i \boldsymbol{D}}{(\boldsymbol{D} - \boldsymbol{c}_i)'\hat{\boldsymbol{H}}_i \boldsymbol{c}_i} \end{aligned} \tag{123}$$

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

Economy-level markup The industry markup is

$$M^m := \frac{\mathbf{E}\left[\sum_{i=1}^{n_F} \tilde{\boldsymbol{q}}_i' \boldsymbol{p}_i\right]}{\mathbf{E}\left[\sum_{i=1}^{n_F} \tilde{\boldsymbol{q}}_i' \boldsymbol{c}_i\right]} = \frac{\sum_{i=1}^{n_F} \mathbf{E}\left[\tilde{\boldsymbol{q}}_i' \boldsymbol{p}_i\right]}{\sum_{i=1}^{n_F} \mathbf{E}\left[\tilde{\boldsymbol{q}}_i' \boldsymbol{c}_i\right]} = \sum_{i=1}^{n_F} w^{\boldsymbol{H}_i} M_i^f \quad \text{where} \quad w^{\boldsymbol{H}_i} = \frac{\mathbf{E}\left[\tilde{\boldsymbol{q}}_i' \boldsymbol{c}_i\right]}{\sum_{i=1}^{n_F} \mathbf{E}\left[\tilde{\boldsymbol{q}}_i' \boldsymbol{c}_i\right]}. \tag{124}$$

C.5 Aggregate Demand Model: Cyclical Markup Fluctuations

Proposition 10. 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.

$$\begin{aligned} \boldsymbol{\beta}_i &= \frac{\Sigma_b M_i}{I_N + \sum_{j=1}^{n_F} \Sigma_b M_j \frac{\hat{\boldsymbol{H}}_j}{\phi}} \\ \boldsymbol{h}_i &= -\boldsymbol{\beta}_i \frac{\hat{\boldsymbol{H}}_i}{\phi} = -\frac{\Sigma_b M_i \frac{\hat{\boldsymbol{H}}_i}{\phi}}{I_N + \sum_{j=1}^{n_F} \Sigma_b M_j \frac{\hat{\boldsymbol{H}}_j}{\phi}} \\ \boldsymbol{F} &= I_N + \sum_{j=1}^{n_F} \boldsymbol{h}_i = \frac{1}{1 + \sum_{j=1}^{n_F} \Sigma_b M_j \frac{\hat{\boldsymbol{H}}_j}{\phi}} \end{aligned} \tag{125}$$

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

$$\left(\frac{\hat{\boldsymbol{H}}_i}{\phi}\right)^{-1} = I_N + \rho_i \phi \left(\boldsymbol{F}^2 \Sigma_b + \sum_{i=1}^{n_F} \boldsymbol{h}_i^2 \Sigma_{\epsilon_i} - \boldsymbol{\beta}_i^2 (\Sigma_b + \Sigma_{\epsilon_i})\right) \tag{126}$$

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

$$\begin{aligned} \left(\frac{\hat{\boldsymbol{H}}_{i,k}}{\phi}\right)^{-1} &= 1 + \rho_i \phi \frac{\left(\Sigma_{\epsilon_{i,k}} + \Sigma_{b,k} + \Sigma_{\epsilon_{i,k}} \frac{\hat{\boldsymbol{H}}_{i,k}}{\phi}\right)^2 \Sigma_{b,k} + n_F \frac{\hat{\boldsymbol{H}}_{i,k}^2}{\phi^2} \Sigma_{b,k}^2 \Sigma_{\epsilon_{i,k}} - \Sigma_{b,k}^2 (\Sigma_{b,k} + \Sigma_{\epsilon_{i,k}})}{\left(\Sigma_{b,k} + \Sigma_{\epsilon_{i,k}} + (\Sigma_{\epsilon_{i,k}} + n_F \Sigma_{b,k}) \frac{\hat{\boldsymbol{H}}_{i,k}}{\phi}\right)^2} \\ &= 1 + \rho_i \phi \frac{\left(1 + \frac{\hat{\boldsymbol{H}}_{i,k}}{\phi}\right)\left(2 + \left(1 + \frac{\hat{\boldsymbol{H}}_{i,k}}{\phi}\right) \frac{\Sigma_{\epsilon_{i,k}}}{\Sigma_{b,k}}\right) + n_F \frac{\hat{\boldsymbol{H}}_{i,k}^2}{\phi^2} \Sigma_{\epsilon_{i,k}} - \Sigma_{\epsilon_{i,k}}}{\left(1 + \frac{\Sigma_{\epsilon_{i,k}}}{\Sigma_{b,k}} + (\frac{\Sigma_{\epsilon_{i,k}}}{\Sigma_{b,k}} + n_F) \frac{\hat{\boldsymbol{H}}_{i,k}}{\phi}\right)^2} \end{aligned} \tag{127}$$

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

$$(\boldsymbol{F}^2 - \boldsymbol{\beta}_i^2)\Sigma_b = \left(1 + \sum_{j=1}^{n_F} \Sigma_b M_j \frac{\hat{\boldsymbol{H}}_j}{\phi}\right)^{-2} \frac{2\Sigma_{\epsilon_i}\left(1 + \frac{\hat{\boldsymbol{H}}_i}{\phi}\right) + \frac{\left(\Sigma_{\epsilon_i}\left(1 + \frac{\hat{\boldsymbol{H}}_i}{\phi}\right)\right)^2}{\Sigma_b}}{\left(1 + \frac{\Sigma_{\epsilon_i}}{\Sigma_b}\left(1 + \frac{\hat{\boldsymbol{H}}_i}{\phi}\right)\right)^2} \tag{128}$$

So the RHS of equation (126) is bounded and $\hat{\boldsymbol{H}}_i$ is positive in the limit. The product-level markup is clearly convergent because $\lim_{\Sigma_b\to\infty} \mathbf{E}[\boldsymbol{p}_i] = \lim_{\Sigma_b\to\infty} \left(\boldsymbol{I}_N + \frac{1}{\phi}\sum_{i=1}^{n_F}\hat{\boldsymbol{H}}_i\right)^{-1}\left(\bar{\boldsymbol{p}} + \frac{1}{\phi}\sum_{i=1}^{n_F}\hat{\boldsymbol{H}}_i\boldsymbol{c}_i\right)$ exists.

□

Proposition 11. The firm-level and economy-level markups are strictly increasing if the shock variance is large enough, and approach their 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

$$\begin{aligned} M_i^f &= \frac{\mathbf{E}[\tilde{\boldsymbol{q}}_i'\boldsymbol{p}_i]}{\mathbf{E}[\tilde{\boldsymbol{q}}_i'\boldsymbol{c}_i]} = \frac{\mathbf{E}[\tilde{\boldsymbol{q}}_i]'\mathbf{E}[\boldsymbol{p}] + \mathbf{tr}\left[\mathbf{Cov}(\boldsymbol{p}_i, \tilde{\boldsymbol{q}}_i)\right]}{\mathbf{E}[\tilde{\boldsymbol{q}}_i'\boldsymbol{c}_i]} \\ &= \frac{\sum_{j=1}^{N}(\mathbf{E}(\boldsymbol{p}_{i,j}) - \boldsymbol{c}_{i,j})\mathbf{E}(\boldsymbol{p}_{i,j})\hat{\boldsymbol{H}}_{i,j} + \sum_{j=1}^{N}\hat{\boldsymbol{H}}_{i,j}\frac{\boldsymbol{\beta}_{i,j}^2\Sigma_{b,j}^2}{\Sigma_{b,j}+\Sigma_{\epsilon_i,j}}}{\sum_{j=1}^{N}(\mathbf{E}(\boldsymbol{p}_{i,j}) - \boldsymbol{c}_{i,j})\boldsymbol{c}_{i,j}\hat{\boldsymbol{H}}_{i,j}} \end{aligned} \tag{129}$$

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

$$\alpha_i := \lim_{\Sigma_b\to\infty}\frac{M_i^f}{\Sigma_b} = \frac{\sum_{j=1}^{N}\tilde{\hat{\boldsymbol{H}}}_{i,j}\tilde{\boldsymbol{\beta}}_j^2}{\sum_{j=1}^{N}(\tilde{\boldsymbol{p}}_j - \boldsymbol{c}_{i,j})\boldsymbol{c}_{i,j}\tilde{\hat{\boldsymbol{H}}}_{i,j}}, \quad \gamma_i := \lim_{\Sigma_b\to\infty}\left(M_i^f - \alpha_i\Sigma_b\right) = \frac{\sum_{j=1}^{N}\left((\tilde{\boldsymbol{p}}_j - \boldsymbol{c}_{i,j})\tilde{\boldsymbol{p}}_j - \tilde{\boldsymbol{\beta}}_j^2\Sigma_{\epsilon_i,j}\right)\tilde{\hat{\boldsymbol{H}}}_{i,j}}{\sum_{j=1}^{N}(\tilde{\boldsymbol{p}}_j - \boldsymbol{c}_{i,j})\boldsymbol{c}_{i,j}\tilde{\hat{\boldsymbol{H}}}_{i,j}} \tag{130}$$

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.

$$w_i := \lim_{\Sigma_b\to\infty} w^{\boldsymbol{H}_i} = \frac{\sum_{j=1}^{N}(\tilde{\boldsymbol{p}}_j - \boldsymbol{c}_{i,j})\boldsymbol{c}_{i,j}\tilde{\hat{\boldsymbol{H}}}_{i,j}}{\sum_{i=1}^{n_F}\sum_{j=1}^{N}(\tilde{\boldsymbol{p}}_j - \boldsymbol{c}_{i,j})\boldsymbol{c}_{i,j}\tilde{\hat{\boldsymbol{H}}}_{i,j}} \Rightarrow M^m \text{ approaches } \sum_{i=1}^{n_F}w_i\alpha_i\Sigma_b + \sum_{i=1}^{n_F}w_i\gamma_i \tag{131}$$

□

C.6 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

$$\begin{aligned}
&\boldsymbol{p}_i = \bar{\boldsymbol{p}} + \boldsymbol{b}_i - \frac{1}{\phi}\sum_{j=1}^{n_F} \tilde{\boldsymbol{q}}_j \\
\Rightarrow &\mathbf{E}\left[\boldsymbol{p}_i|\mathcal{I}_i\right] = \bar{\boldsymbol{p}} + \mathbf{E}\left[\boldsymbol{b}_i|\mathcal{I}_i\right] - \frac{1}{\phi}\sum_{j=1}^{n_F} \mathbf{E}\left[\tilde{\boldsymbol{q}}_j|\mathcal{I}_i\right] \\
\Rightarrow &\mathbf{E}_i\boldsymbol{p}_i = \bar{\boldsymbol{p}} + \mathbf{E}_i\boldsymbol{b}_i - \frac{1}{\phi}\sum_{j=1}^{n_F} \mathbf{E}_i\tilde{\boldsymbol{q}}_j
\end{aligned} \tag{132}$$

So the optimal output in the incomplete information setup is

$$\begin{aligned}
&\tilde{\boldsymbol{q}}_i = \left(\rho_i \mathbf{Var}\left[\boldsymbol{p}_i|\mathcal{I}_i\right] - \frac{\partial \mathbf{E}\left[\boldsymbol{p}_i|\mathcal{I}_i\right]}{\partial \tilde{\boldsymbol{q}}_i}\right)^{-1}\left(\mathbf{E}\left[\boldsymbol{p}_i|\mathcal{I}_i\right] - \boldsymbol{c}_i\right) \\
\Rightarrow &\left(\rho_i \mathbf{Var}\left[\boldsymbol{p}_i|\mathcal{I}_i\right] - \frac{\partial \mathbf{E}\left[\boldsymbol{p}_i|\mathcal{I}_i\right]}{\partial \tilde{\boldsymbol{q}}_i}\right)\tilde{\boldsymbol{q}}_i = \mathbf{E}\left[\boldsymbol{p}_i|\mathcal{I}_i\right] - \boldsymbol{c}_i \\
\Rightarrow &\left(\rho_i \mathbf{Var}\left[\boldsymbol{p}_i|\mathcal{I}_i\right] + \frac{1}{\phi}\boldsymbol{I}_N\right)\tilde{\boldsymbol{q}}_i = \bar{\boldsymbol{p}} + \mathbf{E}_i\boldsymbol{b}_i - \frac{1}{\phi}\sum_{j=1}^{n_F}\mathbf{E}_i\tilde{\boldsymbol{q}}_j - \boldsymbol{c}_i \\
\Rightarrow &\left(\rho_i \mathbf{Var}\left[\boldsymbol{p}_i|\mathcal{I}_i\right] + \frac{2}{\phi}\boldsymbol{I}_N\right)\tilde{\boldsymbol{q}}_i = \bar{\boldsymbol{p}} + \mathbf{E}_i\boldsymbol{b}_i - \frac{1}{\phi}\sum_{j=1,j\neq i}^{n_F}\mathbf{E}_i\tilde{\boldsymbol{q}}_j - \boldsymbol{c}_i \\
\Rightarrow &\tilde{\boldsymbol{q}}_i = \boldsymbol{H}_i\left(\bar{\boldsymbol{p}} + \mathbf{E}_i\boldsymbol{b}_i - \boldsymbol{c}_i - \frac{1}{\phi}\sum_{j=1,j\neq i}^{n_F}\mathbf{E}_i\tilde{\boldsymbol{q}}_j\right),\ \forall\, i = 1,\ldots,n_F
\end{aligned} \tag{133}$$

C.7 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

$$\tilde{\boldsymbol{q}}_i = \boldsymbol{\alpha}_i + \gamma_i\mathbf{E}_i\boldsymbol{b}_i = \boldsymbol{\alpha}_i + \gamma_i\boldsymbol{K}_i\boldsymbol{s}_i \tag{134}$$

Then the optimal action function (133) across all firms is

$$\begin{aligned}
&\tilde{\boldsymbol{q}}_i = \boldsymbol{H}_i\left(\bar{\boldsymbol{p}} + \mathbf{E}_i\boldsymbol{b}_i - \boldsymbol{c}_i - \frac{1}{\phi}\sum_{j=1,j\neq i}^{n_F}\mathbf{E}_i\tilde{\boldsymbol{q}}_j\right) \\
\Rightarrow &\boldsymbol{\alpha}_i + \gamma_i\mathbf{E}_i\boldsymbol{b}_i = \boldsymbol{H}_i\left(\bar{\boldsymbol{p}} + \mathbf{E}_i\boldsymbol{b}_i - \boldsymbol{c}_i - \frac{1}{\phi}\sum_{j=1,j\neq i}^{n_F}\mathbf{E}_i\left(\boldsymbol{\alpha}_j + \gamma_j\mathbf{E}_j\boldsymbol{b}_j\right)\right) \\
\Rightarrow &\boldsymbol{\alpha}_i + \gamma_i\mathbf{E}_i\boldsymbol{b}_i = \boldsymbol{H}_i\left(\bar{\boldsymbol{p}} + \mathbf{E}_i\boldsymbol{b}_i - \boldsymbol{c}_i - \frac{1}{\phi}\sum_{j=1,j\neq i}^{n_F}\boldsymbol{\alpha}_j\right) \\
\Rightarrow &\boldsymbol{\alpha}_i - \boldsymbol{H}_i\left(\bar{\boldsymbol{p}} - \boldsymbol{c}_i - \frac{1}{\phi}\sum_{j=1,j\neq i}^{n_F}\boldsymbol{\alpha}_j\right) + \left(\gamma_i - \boldsymbol{H}_i\right)\mathbf{E}_i\boldsymbol{b}_i = 0,\ \forall\, i = 1,\ldots,n_F
\end{aligned} \tag{135}$$

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

$$
\begin{aligned}
\boldsymbol{\alpha}_i &= \boldsymbol{H}_i \left(\bar{\boldsymbol{p}} - \boldsymbol{c}_i - \frac{1}{\phi} \sum_{j=1, j\neq i}^{n_F} \boldsymbol{\alpha}_j \right) \\
\gamma_i &= \boldsymbol{H}_i = \hat{\boldsymbol{H}}_i \left(\boldsymbol{I}_N + \frac{1}{\phi} \hat{\boldsymbol{H}}_i \right)^{-1}
\end{aligned}
\tag{136}
$$

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

$$
\begin{aligned}
\boldsymbol{\alpha}_i &= \left(\boldsymbol{H}_i^{-1} - \frac{\boldsymbol{I}_N}{\phi} \right)^{-1} \left[\left(\boldsymbol{I}_N + \frac{1}{\phi} \sum_{j=1}^{n_F} \left(\boldsymbol{H}_j^{-1} - \frac{\boldsymbol{I}_N}{\phi} \right)^{-1} \right)^{-1} \left(\bar{\boldsymbol{p}} + \frac{1}{\phi} \sum_{j=1}^{n_F} \left(\boldsymbol{H}_j^{-1} - \frac{\boldsymbol{I}_N}{\phi} \right)^{-1} \boldsymbol{c}_j \right) - \boldsymbol{c}_i \right] \\
&= \hat{\boldsymbol{H}}_i \left[\left(\boldsymbol{I}_N + \frac{1}{\phi} \sum_{j=1}^{n_F} \hat{\boldsymbol{H}}_j \right)^{-1} \left(\bar{\boldsymbol{p}} + \frac{1}{\phi} \sum_{j=1}^{n_F} \hat{\boldsymbol{H}}_j \boldsymbol{c}_j \right) - \boldsymbol{c}_i \right]
\end{aligned}
\tag{137}
$$

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

$$
\begin{aligned}
\tilde{\boldsymbol{q}}_i &= \hat{\boldsymbol{H}}_i \left[\left(\boldsymbol{I}_N + \frac{1}{\phi} \sum_{j=1}^{n_F} \hat{\boldsymbol{H}}_j \right)^{-1} \left(\bar{\boldsymbol{p}} + \frac{1}{\phi} \sum_{j=1}^{n_F} \hat{\boldsymbol{H}}_j \boldsymbol{c}_j \right) - \boldsymbol{c}_i \right] + \boldsymbol{H}_i \mathbf{E}_i \boldsymbol{b}_i \\
&= \hat{\boldsymbol{H}}_i \left(\boldsymbol{D} - \boldsymbol{c}_i \right) + \hat{\boldsymbol{H}}_i \left(\boldsymbol{I}_N + \frac{1}{\phi} \hat{\boldsymbol{H}}_i \right)^{-1} \boldsymbol{K}_i \boldsymbol{s}_i
\end{aligned}
\tag{138}
$$

The equilibrium price and output are

$$
\begin{aligned}
\mathbf{E}(\tilde{\boldsymbol{q}}_i) &= \hat{\boldsymbol{H}}_i \left(\boldsymbol{D} - \boldsymbol{c}_i \right) \\
\boldsymbol{p}_i &= \boldsymbol{D} + \boldsymbol{b}_i - \frac{1}{\phi} \sum_{j=1}^{n_F} \hat{\boldsymbol{H}}_j \left(\boldsymbol{I}_N + \frac{1}{\phi} \hat{\boldsymbol{H}}_j \right)^{-1} \boldsymbol{K}_j \boldsymbol{s}_j \Rightarrow \mathbf{E}(\boldsymbol{p}_i) = \boldsymbol{D}
\end{aligned}
\tag{139}
$$

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

$$
\overline{M}^p = \frac{1}{N} \frac{1}{n_F} \sum_{i=1}^{n_F} \sum_{j=1}^{N} M_{ij}^p = \frac{1}{N} \frac{1}{n_F} \sum_{i=1}^{n_F} \sum_{j=1}^{N} \frac{\boldsymbol{D}(j)}{\boldsymbol{c}_i(j)}
\tag{140}
$$

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

$$
\begin{aligned}
M_i^f &= \frac{\mathbf{E}[\tilde{\boldsymbol{q}}_i' \boldsymbol{p}]}{\mathbf{E}[\tilde{\boldsymbol{q}}_i' \boldsymbol{c}_i]} = \frac{\mathbf{E}[\tilde{\boldsymbol{q}}_i]' \mathbf{E}[\boldsymbol{p}] + \mathbf{tr}\left[\mathbf{Cov}(\boldsymbol{p}_i, \tilde{\boldsymbol{q}}_i) \right]}{\mathbf{E}[\tilde{\boldsymbol{q}}_i' \boldsymbol{c}_i]} \\
&= \frac{(\boldsymbol{D} - \boldsymbol{c}_i)' \hat{\boldsymbol{H}}_i \boldsymbol{D} + \mathbf{tr}\left(\left(\boldsymbol{I}_N - \frac{\boldsymbol{H}_i}{\phi} \right) \boldsymbol{K}_i \mathbf{Var}(\boldsymbol{s}_i) \boldsymbol{K}_i' \boldsymbol{H}_i \right)}{(\boldsymbol{D} - \boldsymbol{c}_i)' \hat{\boldsymbol{H}}_i \boldsymbol{c}_i} > \frac{(\boldsymbol{D} - \boldsymbol{c}_i)' \hat{\boldsymbol{H}}_i \boldsymbol{D}}{(\boldsymbol{D} - \boldsymbol{c}_i)' \hat{\boldsymbol{H}}_i \boldsymbol{c}_i}
\end{aligned}
\tag{141}
$$

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

$$M^m := \frac{\mathbf{E}\left[\sum_{i=1}^{n_F}\tilde{\boldsymbol{q}}_i'\boldsymbol{p}_i\right]}{\mathbf{E}\left[\sum_{i=1}^{n_F}\tilde{\boldsymbol{q}}_i'\boldsymbol{c}_i\right]} = \frac{\sum_{i=1}^{n_F}\mathbf{E}\left[\tilde{\boldsymbol{q}}_i'\boldsymbol{p}_i\right]}{\sum_{i=1}^{n_F}\mathbf{E}\left[\tilde{\boldsymbol{q}}_i'\boldsymbol{c}_i\right]} = \sum_{i=1}^{n_F} w^{\boldsymbol{H}_i} M_i^f \quad \text{where} \quad w^{\boldsymbol{H}_i} = \frac{\mathbf{E}\left[\tilde{\boldsymbol{q}}_i'\boldsymbol{c}_i\right]}{\sum_{i=1}^{n_F}\mathbf{E}\left[\tilde{\boldsymbol{q}}_i'\boldsymbol{c}_i\right]}. \tag{142}$$

C.8 Private Information Model: Cyclical Markup Behavior

Proposition 12. 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

$$\hat{\boldsymbol{H}}_i = \left[\rho_i \mathbf{Var}(\boldsymbol{b}_i|\mathcal{I}_i) + \frac{\boldsymbol{I}_N}{\phi} + \rho_i(n_F-1)\frac{\hat{\boldsymbol{H}}_i}{\phi}\left(\boldsymbol{I}_N + \frac{1}{\phi}\hat{\boldsymbol{H}}_i\right)^{-1}\mathbf{Var}(\boldsymbol{K}_i\boldsymbol{s}_i)\left(\boldsymbol{I}_N + \frac{1}{\phi}\hat{\boldsymbol{H}}_i\right)^{-1}\frac{\hat{\boldsymbol{H}}_i}{\phi}\right]^{-1}$$

$$\Rightarrow \hat{\boldsymbol{H}}_{i,k}^{-1} = \rho_i\frac{\Sigma_{b,k}\Sigma_{\epsilon_i,k}}{\Sigma_{b,k}+\Sigma_{\epsilon_i,k}} + \frac{1}{\phi} + \rho_i(n_F-1)\left(\frac{\hat{\boldsymbol{H}}_{i,k}}{\phi+\hat{\boldsymbol{H}}_{i,k}}\right)^2\frac{\Sigma_{b,k}^2}{\Sigma_{b,k}+\Sigma_{\epsilon_i,k}},\ \forall k = 1,\dots,N \tag{143}$$

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

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

$$-\left(\hat{\boldsymbol{H}}_{i,k}^{-2} + \rho_i(n_F-1)\frac{\Sigma_{b,k}^2}{\Sigma_{b,k}+\Sigma_{\epsilon_i,k}}\frac{2\phi\hat{\boldsymbol{H}}_{i,k}}{(\hat{\boldsymbol{H}}_{i,k}+\phi)^3}\right)\frac{\partial\hat{\boldsymbol{H}}_{i,k}}{\partial\Sigma_{b,k}} = \frac{\rho_i\Sigma_{\epsilon_i,k}^2}{\Sigma_{b,k}+\Sigma_{\epsilon_i,k}} + \rho_i(n_F-1)\left(\frac{\hat{\boldsymbol{H}}_{i,k}}{\phi+\hat{\boldsymbol{H}}_{i,k}}\right)^2\frac{\Sigma_{b,k}(\Sigma_{b,k}+2\Sigma_{\epsilon_i,k})}{\Sigma_{b,k}+\Sigma_{\epsilon_i,k}} \tag{144}$$

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 (143) is unbounded while the LHS is bounded. The product-level markup $\overline{M}^p$ is convergent:

$$\overline{M}^p = \frac{1}{N}\frac{1}{n_F}\sum_{i=1}^{n_F}\sum_{j=1}^{N} M_{ij}^p = \frac{1}{N}\frac{1}{n_F}\sum_{i=1}^{n_F}\sum_{j=1}^{N}\frac{\mathbf{E}(\boldsymbol{p}_{i,j})}{\boldsymbol{c}_{i,j}} \quad \text{and} \quad \lim_{\Sigma_b\to\infty}\overline{M}^p = \frac{1}{N}\frac{1}{n_F}\sum_{i=1}^{n_F}\sum_{j=1}^{N}\frac{\bar{\boldsymbol{p}}_j}{\boldsymbol{c}_{i,j}} \tag{145}$$

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

□

Proposition 13. 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:

$$M_i^f = \frac{\mathbf{E}[\tilde{\boldsymbol{q}}_i' \boldsymbol{p}_i]}{\mathbf{E}[\tilde{\boldsymbol{q}}_i' \boldsymbol{c}_i]} = \frac{\mathbf{E}[\tilde{\boldsymbol{q}}_i]' \mathbf{E}[\boldsymbol{p}] + \mathbf{tr}\left[\mathbf{Cov}(\boldsymbol{p}_i, \tilde{\boldsymbol{q}}_i)\right]}{\mathbf{E}[\tilde{\boldsymbol{q}}_i' \boldsymbol{c}_i]}$$
$$= \frac{(\boldsymbol{D} - \boldsymbol{c}_i)' \hat{\boldsymbol{H}}_i \boldsymbol{D} + \mathbf{tr}\left(\left(\boldsymbol{I}_N + \frac{1}{\phi}\hat{\boldsymbol{H}}_i\right)^{-1} \mathbf{Var}(\boldsymbol{K}_i \boldsymbol{s}_i) \left(\boldsymbol{I}_N + \frac{1}{\phi}\hat{\boldsymbol{H}}_i\right)^{-1} \hat{\boldsymbol{H}}_i\right)}{(\boldsymbol{D} - \boldsymbol{c}_i)' \hat{\boldsymbol{H}}_i \boldsymbol{c}_i} \tag{146}$$

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

$$M_i^f = \frac{\sum_{j=1}^N (\boldsymbol{D}_j - \boldsymbol{c}_{i,j}) \boldsymbol{D}_j + \sum_{j=1}^N \left(1 + \frac{1}{\phi}\hat{\boldsymbol{H}}_{i,j}\right)^{-2} \frac{\Sigma_{b,j}^2}{\Sigma_{b,j} + \Sigma_{\epsilon_i,j}}}{\sum_{j=1}^N (\boldsymbol{D}_j - \boldsymbol{c}_{i,j}) \boldsymbol{c}_{i,j}} \tag{147}$$

M_i^f also admits an asymptote as

$$\alpha_i := \lim_{\Sigma_b \to \infty} \frac{M_i^f}{\Sigma_b} = \frac{N}{\sum_{j=1}^N (\bar{\boldsymbol{p}}_j - \boldsymbol{c}_{i,j}) \boldsymbol{c}_{i,j}} \quad \text{and} \quad \gamma_i := \lim_{\Sigma_b \to \infty} \left(M_i^f - \alpha_i \Sigma_b\right) = \frac{\sum_{j=1}^N \left((\bar{\boldsymbol{p}}_j - \boldsymbol{c}_{i,j}) \bar{\boldsymbol{p}}_j - \Sigma_{\epsilon_i,j}\right)}{\sum_{j=1}^N (\bar{\boldsymbol{p}}_j - \boldsymbol{c}_{i,j}) \boldsymbol{c}_{i,j}} \tag{148}$$

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 $w^{\boldsymbol{H}_i} = \frac{\mathbf{E}[\tilde{\boldsymbol{q}}_i' \boldsymbol{c}_i]}{\sum_{i=1}^{n_F} \mathbf{E}[\tilde{\boldsymbol{q}}_i' \boldsymbol{c}_i]}$. The weight $w^{\boldsymbol{H}_i}$ converges to w_i as shock variance goes to infinity.

$$w_i := \lim_{\Sigma_b \to \infty} w^{\boldsymbol{H}_i} = \frac{\sum_{j=1}^N (\bar{\boldsymbol{p}}_j - \boldsymbol{c}_{i,j}) \boldsymbol{c}_{i,j}}{\sum_{i=1}^{n_F} \sum_{j=1}^N (\bar{\boldsymbol{p}}_j - \boldsymbol{c}_{i,j}) \boldsymbol{c}_{i,j}} \Rightarrow M^m \text{ approaches } \sum_{i=1}^{n_F} w_i \alpha_i \Sigma_b + \sum_{i=1}^{n_F} w_i \gamma_i \tag{149}$$

□

C.9 Choosing A Location in Product Space

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

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

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

$$max_{a_i, q_i} \mathbf{E}\left[q_i \boldsymbol{a}_i' (\tilde{\mathbf{p}} - \boldsymbol{c}_i) | \mathcal{I}_i\right] - \frac{\rho_i}{2} \mathbf{Var}\left[q_i \boldsymbol{a}_i' (\tilde{\mathbf{p}} - \boldsymbol{c}_i) | \mathcal{I}_i\right] - g(\chi_c, \boldsymbol{c}_i), \tag{150}$$

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 1.1

$$max_{c_i}\mathbf{E}\left[\ \mathbf{E}\left[q_i\boldsymbol{a}_i'\left(\tilde{\mathbf{p}}-\boldsymbol{c}_i\right)|\mathcal{I}_i\right]-\frac{\rho_i}{2}\mathbf{Var}\left[q_i\boldsymbol{a}_i'\left(\tilde{\mathbf{p}}-\boldsymbol{c}_i\right)|\mathcal{I}_i\right]\right]-g(\chi_c,\boldsymbol{c}_i). \tag{151}$$

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.

$$q_i\boldsymbol{a}_i=\left(\rho_i\mathbf{Var}\left[\boldsymbol{p}_i|\mathcal{I}_i\right]+\frac{\partial\mathbf{E}\left[\boldsymbol{p}_j|\mathcal{I}_j\right]}{\partial\boldsymbol{q}_i}\right)^{-1}\left(\mathbf{E}\left[\boldsymbol{p}_i|\mathcal{I}_i\right]-\boldsymbol{c}_i\right) \tag{152}$$

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

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.