Informational Differences Among Rival Firms: Measurement and Evidence from Hotels

Matthew Leisten*

Federal Trade Commission

mattleisten@gmail.com

This version: December 4, 2021

Click here for newest version

Abstract

Some firms are better informed about demand than others when setting prices. I establish identification of competing firms' information about demand in the presence of rich observed and unobserved heterogeneity in pricing. My strategy relies on the presence of market-wide demand shifters that are common knowledge among the firms. I use my results to quantify large differences in hotels' ability to forecast demand. Hotels affiliated with large chains surprisingly have worse information, on average, than small chains. Counterfactuals using my estimates show that these differences are explained by weaker incentives for large-chain hotel managers to exert information gathering effort.

*Website: mleisten.github.io. This article is based on the first chapter of my dissertation while at Northwestern University. Thanks to my dissertation committee: Robert Porter, Gaston Illanes, and Michael Mazzeo. I also benefited from discussions with Vivek Bhattacharya, Ronald Goettler, Daniel Hosken, Thomas Hubbard, Victoria Marone, Devesh Raval, Mar Reguant, Bill Rogerson, Nicholas Vreugdenhil, Gabriel Ziegler, and many others. Thanks to seminar participants at Northwestern University, Compass Lexecon, Charles River Associates, Cornerstone, University of Rochester Simon, DOJ, and FTC. Thanks to STR for providing the data and to the CSIO at Northwestern for funding. All errors are my own. Views and opinions expressed in this article are solely my own should not be interpreted as reflecting the views of the Federal Trade Commission, any of its individual Commissioners, or other members of the staff.

1 Introduction

The ability to predict demand before choosing prices is a key determinant of productivity. Firms that have better information about demand than their rivals can set more "correct" prices and will consistently earn higher profits given a fixed set of inputs. However, we know little about the extent to which firms differ in their ability to predict demand because it is difficult to measure information credibly. Empirical patterns consistent with "bad information" may exist for other reasons: Prices may appear to be random because firms are misinformed *or* because firms receive unobserved shocks. Prices may not respond to changes in demand because firms are misinformed *or* because firms' optimal prices do not depend strongly on this shift, even if fully informed. A firm may seem unresponsive to demand shifts because it does not know demand *or* because it knows demand but is mainly responding to rivals who do not.

In this article, I establish identification of firms' information in a manner robust to all of these challenges. I model information as a signal each firm receives about a market-wide demand shifter. Based on their signal, each firm rationally forms expectations over both demand and rivals' prices. These beliefs are then mapped to a price based on a flexible *price policy function*. Signals can be affiliated across firms. Price policy functions can be different for each firm and contain an idiosyncratic error that may represent a firm-specific pricing shock that is unobserved by the econometrician. I establish nonparametric identification of the signal structure and semiparametric identification of the price policy function under relatively mild restrictions. Thus, my framework allows for firm interaction, heterogeneity across firms, and unobserved pricing shocks, all while allowing for maximal flexibility in firms' information structure and policy functions.

Identification relies on observing a variable that shifts market-wide demand and that is *common knowledge* to all firms in that market when they set prices. This allows me to infer a firm's information quality from how responsive its prices are to demand shifts arising from this common-knowledge variable, compared to how its prices respond to other market-wide demand shifts *conditional on* this common-knowledge variable. I use this strategy to prove identification under two different informational schemes: one in which hotels have uncertainty over market-wide demand shocks but not over private demand shocks, and one in which hotels know the realization of a demand *index*, but cannot discern whether this index is driven my market-wide or hotel-specific shocks. The presence of market-wide, common-knowledge common to many settings. For instance, seasonality could be used as a common-knowledge demand shifter to identify information in many other markets.

I illustrate the usefulness of my results with data from the hotel industry. I measure differences across hotels in their ability to forecast demand and examine why some hotels are better-informed than others. My dataset contains a calendar year of hotel-specific nightly average prices and occupancies across eighteen different U.S. college towns. Using events such as home football games, commencement, and parents' weekend as common-knowledge demand shifters, I show that hotel prices respond more to demand shifts arising from these events than to similar-sized demand shifts not associated with these events. This descriptive strategy mirrors the intuition underlying my formal identification results, though it is insufficient on its own to pin down

To more precisely measure information, I estimate a parameterized version of my theoretical model using indirect inference. These estimates allow me to quantify differences in information quality across the hotels in my sample. I find substantial heterogeneity in information quality. The 90th percentile hotel's signal is roughly four times more informative than the 10th percentile hotel's signal in the following sense: a "high" signal generated from the 90th percentile hotel's signal distribution would induce Bayesian agents to increase their expectations of the demand shifter four times more than the same signal drawn from the 10th percentile hotel's distribution would. Counterfactual simulations of my model predict that hotels would increase operating profits by roughly 10% on average by moving from the 10th percentile of information quality across all hotels to the 90th percentile.

As an illustration, I use these results to examine two competing theories about the determinants of firms' information quality. According to one theory, larger firms have better information about demand because they have access to bigger data sets and more sophisticated analytical tools. Another theory posits that information quality depends primarily on managers' effort in gathering information. If managerial incentives are stronger at smaller firms, larger firms may be at an informational disadvantage. Large chain hotels, for instance, gather larger percent royalties from franchisees than small chains, thereby weakening franchisees' incentives to expend effort to learn demand. My estimates show that hotels affiliated with large chains tend to be worse-informed than small chain rivals.

As a final exercise, I quantify the effects of different managerial incentives on information gathering. I estimate hotels' costs of gathering information using my supply and demand estimates and a simple model of information-gathering effort. I then set the royalty rates the same and use this model, as well as my supply and demand estimates, to predict counterfactual information quality. In this scenario, large-chain hotels obtain better information than small-chain rivals. This difference, however, seems to be driven by large-chain hotels' greater revenue-generating ability (for instance, through their brand value) rather than having lower costs of obtaining information.

Related literature. This article measures information quality about a common demand shifter, so the closest methodological analogs in the common values auctions literature. Hendricks and Porter [1988] and Hendricks et al. [2003] show how additional information about the common value can help identify the model. My model goes further: if the analyst knows the common value *and* can assume that a component of this common value is common knowledge to the players, the model can be identified even if prices (or bids) contain idiosyncratic error. An even closer article is Compiani et al. [2020], who study common values auctions with unobserved heterogeneity. In their model, unobserved heterogeneity is auction-wide, and the object of interest is bidders' "pivotal values"– their expected value of winning in equilibrium, whereas in my setting unobserved heterogeneity is firm-specific and I am interested in recovering the information structure itself.

Some articles make similar identification arguments to study choice under incomplete information. Brown and Jeon [2020] study choice of insurance plan under limited information using an identification strategy similar to mine. They rely on differences in consumers' response to insurance premiums (assumed known) and expected out-of-pocket costs (not known). They argue that these should, under complete information, enter the choice problem the same way (i.e., through an index). Lu [2016] studies a single-agent discrete choice problem in which there are choice-specific tastes (analogous to the error in my model) and an unknown information structure. He shows that the agent's information structure is identified if the entire random choice rule is observed, which is often not the case empirically, including in my setting.

In my application, I consider information quality as a determinant of productivity.[1] The review by Syverson [2011] outlines several main explanations for productivity differences; my article relates to three in particular. The first is differences in managerial practices [Bloom and Van Reenen [2007], Bloom et al. [2016]]. The second is differences in IT adoption [Hubbard

[1] My mechanism is that productivity losses occur through *mispricing*, not through *lower prices* as is the case with some articles that use revenue as an econometric input for productivity, see Syverson [2004] on the shortcomings of that approach, referred to as "output price bias."

[2003], Bloom et al. [2012]]. The third is differences in vertical integration or firm size [Atalay et al., 2014].[2] The empirical analysis in this article is complementary to these papers. I take a narrower stance about the specific mechanism through which these factors affect productivity (by changing information about demand), allowing me to conduct counterfactual simulations that require a structural approach. For instance, I am able to examine information-gathering effort when the incentives for managers to gather information change.

Finally, this article relates to several recent articles on the industrial organization of hotels.[3] There are two key themes in these articles, both related to the effects of chain affiliation. The first are the consequences of demand uncertainty, as examined in various regards by Butters [2020], Cho et al. [2018], and Kalnins et al. [2017]. Huang [2021] studies how pricing frictions and the inability to process information about demand affect pricing by AirBnB hosts. The second set of articles is concerned with the importance of quality provision in hotels, with emphasis on vertical differentiation [Mazzeo, 2002], chain affiliation [Hollenbeck, 2017], and provision of amenities [Hubbard and Mazzeo, 2019]. Notably, Hollenbeck [2017] concludes that chains' main advantage is through higher demand, not lower costs.

2 Identification of Information

This section establishes the identification of firms' information and price policy functions using variation in prices and demand over time. The key to identification is that the econometrician observes a market-wide demand shifter that is common knowledge to all firms. By comparing the responsiveness of prices to this common-knowledge shifter and to an imperfectly known common demand state, I can recover information quality. The general idea is that, all else equal, the larger the wedge between firms' response to the unobserved shifter and the observed shifter, the less informed firms must be.

A simple example. To gain some intuition for this approach, consider a simple single-firm example. Let t index a series of markets (e.g., t indexes each night for a hotel). The demand state is $\theta_t = \theta_t^{known} + \theta_t^{unknown}$, where each of these components is an i.i.d. draw from a

[2] A few papers in the productivity literature have married some of these strands. For instance, the literature on CEO compensation [Bertrand and Mullainathan, 2003] examines how incentives induce best management practices. Bajari et al. [2019] study whether more data improves the ability to forecast demand using data from Amazon.

[3] There are also several articles examining the implications of the rise of the sharing economy in hospitality. See, e.g., Farronato and Fradkin [2018].

standard normal distribution on each date t. The firm knows θ_t^{known} but not $\theta_t^{unknown}$. The firm instead sees a signal $s_t = \theta_t^{unknown} + \varepsilon_t$ where ε_t is i.i.d. $\mathcal{N}(0, \rho)$. ρ governs the precision of this signal– that is, it governs information quality. The firm chooses prices according to the price policy function:

$$p_t = z\mathbb{E}(\theta_t | s_t, \theta_t^{known}) + \nu_t$$

where $\nu_t \sim \mathcal{N}(0, \sigma^2)$ is an i.i.d. pricing shock and $z > 0$ is a parameter. The econometrician views an arbitrarily long series of prices $p_t,\ t = 1, 2, \ldots$. Suppose first that the econometrician observes θ_t but not the individual components, the signals, or the pricing shocks. In this case, the parameters (z, ρ, σ) are not identified. The econometrician's "data" is the joint distribution of of θ and p, $F(\theta, p)$. Consider the distribution of prices conditional on θ, $F(p|\theta)$. The pricing equation can be rewritten

$$\begin{aligned} p_t =& \quad z\theta_t^{known} + z\mathbb{E}(\theta_t^{unknown} | s_t) + \nu_t \\ =& \quad z\theta_t^{known} + \frac{z}{1+\rho}(\theta_t^{unknown} + \varepsilon_t) + \nu_t \end{aligned}$$

where the second line uses the joint distribution of $\theta^{unknown}$ and the signal to compute firms' posterior beliefs. Then we can write

$$\begin{aligned} \theta =& \quad z(\theta - \theta^{unknown}) + \frac{z}{1+\rho}(\theta^{unknown} + \varepsilon) + \nu \\ =& \quad z\theta - \frac{z\rho}{1+\rho}\theta^{unknown} + \eta \end{aligned}$$

where the first line uses the definition of θ and the second line is just algebra, and $\eta \equiv \nu + \frac{z}{1+\rho}\varepsilon$. Conditional on θ, it is clear from the above that prices equal $z\theta$ plus two terms that are normally distributed and independent of each other. Using the properties of joint normality, it can be shown the second term above is normally distributed with mean $-\frac{z\rho}{1+\rho}\mathbb{E}(\theta^{unknown}|\theta) = -\frac{z\rho}{2(1+\rho)}\theta$ and variance $\frac{z^2\rho^2}{(1+\rho)^2}\text{Var}(\theta^{unknown}|\theta) = \frac{z^2\rho^2}{2(1+\rho)^2}$. The second term is normal with mean zero and variance $\text{Var}(\eta|\theta) = \sigma^2 + \frac{z^2\rho^2}{(1+\rho)^2}$. Thus, the distribution $F(p|\theta)$ is completely characterized by its mean $\frac{z(2-\rho)}{2(1+\rho)}\theta$ and variance $\frac{3z^2\rho^2}{2(1+\rho)^2} + \sigma^2$. This gives us two identifying moments with three unknowns: (ρ, z, σ^2). For instance, $F(p|\theta)$ would be the same if $(\rho, z, \sigma^2) = (1, 1, 1)$ and $(.5, .5, .875)$.

This example illuminates the key identification issue: higher z (firms care more about demand changes) implies higher responsiveness, but so does lower ρ (firms know more about demand changes). Essentially, "I don't know" is observationally the same as "I don't care." Decreasing both at the same time may leave measured responsiveness (the slope of $\mathbb{E}(p|\theta)$ with respect to θ) unchanged. Furthermore, without knowing σ^2, the conditional variance moment is also uninformative about z and ρ. If the econometrician knew σ^2, this would not be a problem, and the variance moment combined with the slope moment would identify the model. The presence of noisiness in pricing from a source *other than* differences in firms' beliefs is an important feature of my setting as well as many others common in economics. This is why my method is required.

In my setting, the econometrician observes prices, and both $\theta_t^{unknown}$ and θ_t^{known} for arbitrarily many t. Now the entire model is identified. Consider the distribution of prices conditional on both components of θ:

$$F(p|\theta^{known}, \theta^{unknown}) = F\left(z\theta^{known} + \frac{z}{1+\rho}\theta^{unknown} + \frac{z}{1+\rho}\varepsilon + \nu \middle| \theta^{known}, \theta^{unknown}\right).$$

This is normally distributed with mean $z\theta^{known} + \frac{z}{1+\rho}\theta^{unknown}$ and variance $\sigma^2 + \frac{z^2\rho^2}{(1+\rho)^2}$. Clearly, both z and ρ are identified from the conditional expectation of prices– specifically from the partial derivatives of this expectation with respect to θ^{known} and $\theta^{unknown}$– and σ^2 can be identified from the conditional variance of prices, given knowledge of z and ρ. Therefore, the model is identified if the econometrician observes the known demand shifter.

Model. I now establish identification in a more general model. Index firms who compete with each other (i.e., hotels in the same city) as $j = 1, 2, \ldots, J$ and the set of all of these firms as $\mathcal{J}$. Index a series of markets in which these firms compete by t. For instance, hotels in a given city compete for market share on that particular night. Demand in t is as follows:

$$\mathbf{q_t} = Q(\mathbf{p_t}, x_t, \lambda_t, \xi_\mathbf{t})$$

where $\mathbf{q_t}$ is a vector with generic entry q_{jt} being j's quantity in market t. $\mathbf{p_t}$ is the j-vector of prices. x_t and λ_t are market-wide demand shifts that increase demand regardless of price: $\frac{\partial Q_j}{\partial \lambda_t} > 0$ and $\frac{\partial Q_j}{\partial x_t} > 0, \forall j$. x_t is common knowledge, but λ_t is not. In the context of hotels, x_t may contain demand shifts arising from "it is a Monday," whereas λ_t contains variation in demand *conditional on* it being a Monday, such as a local arts festival on one particular Monday.

Assumption 1: Demand. I make the following assumptions about demand. λ_t is independent over time and independent of the other shifters. $\xi_\mathbf{t}$ is a j-vector of firm-specific demand shocks, with $\frac{\partial Q_j}{\partial \xi_{jt}} > 0, \forall j$. I assume ξ_{jt} are independent across firms and over time, and that they are independent of x_t and λ_t. In this sense, they are truly idiosyncratic demand shocks to each firm. Besides these restrictions, I place no assumptions on the demand function Q. The remainder of this section assumes demand is fully known. Estimation and identification of demand in my application are discussed in greater detail in Section 4. For now, I will treat x_t, λ_t and ξ_{jt} as data.

I now turn to supply. For firm j, there are three demand variables to keep track of: x_t, ξ_{jt}, and λ_t. x_t is common knowledge amongst all firms when they set prices; λ_t is possibly not common knowledge. ξ_{jt} is firm-specific. The supply model contains two components. The first is the mapping from firms' information sets to prices. Call this the "price policy function." The second component is each firm's information. The emphasis of this article is on the information structure, but to credibly estimate information, the policy functions must also be credibly estimated. There are four features of many markets, including hotels, that my supply model must incorporate in order to yield credible estimates. Below are four common features of markets where firms compete under uncertain demand that I must incorporate into my model.

Feature 1: Firms compete. Firms incorporate their beliefs over rivals' prices when setting their own price. In the hotel industry, this often comes in the form of monitoring a "comparison set" of rivals that are close geographically and in terms of quality tier. Formally, this feature means that the price policy function should include rivals' prices in some fashion.

Feature 2: Beliefs may be correlated. Broadly, if firm j thinks demand is high, rival firm j' may also tend to think demand is high. There may be two ways this happens. First, beliefs may be correlated because they are about a common demand shifter. I explicitly model firms' beliefs over the *market-wide* demand shifter λ_t. Second, beliefs may be correlated across firms, conditional on demand. That is, firms may tend to be "wrong" about demand in the same direction. This may be because they use similar data sources to gather information. In terms of my model, this means that signals firms receive about demand may be affiliated, even conditional on the actual demand state.

Feature 3: Black-box pricing. The mapping from a firm's information set to its prices may be difficult to deduce using economic principles alone. For instance, in revenue management industries like hotels and airlines, prices are based on approximate solutions to complicated dynamic stochastic programs. Given my data, I cannot fully estimate the parameters of a dynamic revenue management problem; in light of the inherent difficulty of solving these problems, it is unclear to what extent hotel managers are even able to do this. At the same time, a more traditional approach in which prices reflect a static (Bertrand) first order condition would also be misspecified. Instead, I work directly with the *price policy functions* while allowing for maximal *flexibility* in the price policy functions' functional form and *heterogeneity* across firms— every firm has a different price policy function that I estimate. All of this said, my framework can also accommodate pricing based on traditional Bertrand competition, so long as the best response functions satisfy the assumptions I outline later in this section.

Feature 4: Prices may be observed with error. The econometrician never sees everything that goes into a firm's pricing decision. The canonical example is that there are unobserved marginal cost shocks. Another example is measuerment error in the pricing data. My framework allows prices to be observed with some error. In the context of my application to hotels, I observe are average prices, known as "average daily rates," across many individual bookings. I do not see price dispersion arising from, for instance, whether someone booked a room at the last minute or whether someone used a loyalty program. The randomness involved in "who books what kind of room, and when" is reason enough to include an error term in the price policy function.

In light of these features, there are still assumptions needed to establish identification. I now outline the key assumptions of the supply side of the model. In general, the price policy function for firm j is

$$p_{jt} = z_{1j}\mathbb{E}(\bar{p}_{-j}|\mathcal{I}_{jt}) + \mathbb{E}(z_{2j}(\delta_{jt})|\mathcal{I}_{jt}) + \nu_{jt}. \tag{1}$$

where $\bar{p}_{-j}$ is the average price across all of j's rivals, $\delta_{jt} = \delta_j(x_t, \lambda_t, \xi_{jt})$ is hotel specific demand index that is based on j's demand shifters, $\mathcal{I}_{jt}$ is j's information set in market t, and ν_{jt} is an error term. One can think of δ_{jt} as being the mean utility for firm j in market t in a logit demand system, though it does not have to take this form. There are several assumptions embedded into this framework, which I now discuss.

Assumption 2: Information. All firms observe x_t and privately observe a signal s_{jt} about market demand λ_t, drawn from distribution $F(s_{jt}|s_{-jt}, \lambda_t)$, where s_{-jt} are the signals of all of j's rivals. The joint distribution of all signals and λ_t is $F(s_t, \lambda_t)$, where $s_t = [s_{jt}]_{j\in\mathcal{J}}$. Firms also either directly observe $\delta_{jt} = \delta_j(x_t, \lambda_t, \xi_{jt})$ but cannot distinguish between common λ_t and idiosyncratic ξ_{jt} (the assumption of *known draws*, 1-KD), or hotels know ξ_{jt} but do not know δ_{jt} because they do not know λ_t (the assumption of *unknown draws*, 1-UD). Concretely, this assumption is

$$\begin{aligned} \mathcal{I}_{jt} = & \ \{x_t, \xi_{jt}, s_{jt}\} \text{ (KD)} \\ \mathcal{I}_{jt} = & \ \{x_t, \delta_{jt}, s_{jt}\} \text{ (UD)} \end{aligned} \quad .$$

The signal structure F, the policy functions, and the marginal distribution of all of the demand variables, are common knowledge amongst all firms. That is, firms have a full understanding of the data generating process, even if they do not know the state of demand in a particular market.

Beliefs about demand are increasing in signals, $\frac{\partial \mathbb{E}(\lambda_t|s_{jt})}{\partial s_{jt})} > 0$, and signals are positively affiliated $\frac{\partial \mathbb{E}(s_{-jt}|s_{jt})}{\partial s_{jt})} > 0$. Signals are independent of the other demand variables, $\xi_t \equiv [\xi_{jt}]_{j\in\mathcal{J}}$ and x_t. That is, $F(s_t, \lambda_t|\xi_t, x_t) = F(s_t, \lambda_t)$. This precludes firms from making better (or worse) inferences about demand when x_t is large. In the instance of hotels, this might happen if a hotel does extra market research on weekends. I do not, however, preclude firms from making inferences about demand from observing δ_{jt}, because δ_{jt} depends on λ_t.

Assumption 3: Policy Functions. Equation 1 implies several restrictions on the policy function. First is an aggregation assumption: rivals' prices only enter firm j's price policy function through their average, $\bar{p}_{-j}$. This assumption could be relaxed so that $\bar{p}_{-j}$ is a weighted average of rivals' prices, which would be the case if the econometrician had reason to believe firms treated certain rivals as closer competitors than others. However, the weighting scheme must be known to the econometrician. This is to avoid a "too many parameters" problem where a weight would otherwise have to be estimated for every pair of firms in $\mathcal{J}$. Similarly, other statistics derived from rivals' prices, like the lowest rival's price or the spread of rivals' prices, cannot enter the policy function.

Next are monotonicity assumptions: $z_{2j}(.)$ is smooth, with $\frac{\partial z_{2j}}{\partial \delta} > 0$ and $z_{1j} > 0$. This ensures injectivity (i.e., to rule out multiple δ_{jt} that yield the same p_{jt}). It means that, all else equal, higher rivals' prices and higher demand yield higher prices for j. Together with the informational assumptions, it ensures that higher beliefs yield higher prices. I also make the assumption that the matrix

$$\begin{bmatrix} 1 & -\frac{z_{p1}}{J-1} & \dots & -\frac{z_{p1}}{J-1} \\ -\frac{z_{p2}}{J-1} & 1 & \dots & -\frac{z_{p2}}{J-1} \\ \vdots & \vdots & & \vdots \\ -\frac{z_{pJ}}{J-1} & -\frac{z_{pJ}}{J-1} & \dots & 1 \end{bmatrix}$$

has full rank. This is a sufficient condition to ensure that there exists a unique solution to the system of equations given by each firm's price policy function. The Bertrand pricing analogue is that best response curves do not slope upward too rapidly.

There is also an index restriction: the function $\delta(x_t, \lambda_t, \xi_{jt})$ is known to the econometrician. Though it does not have to be linear, $\delta(.)$ is assumed to be increasing and smooth in each of its arguments, for the same reasons $z_{2j}(.)$ must be increasing. It is clear why $\delta(.)$ must be known to the econometrician: because it is the only argument of $z_{2j}(.)$, it cannot be jointly identified with $z_{2j}(.)$. Later in this section, I discuss how the specification of $\delta(.)$ affects identification of $z_{2j}(.)$.

Finally, I make a separability assumption: Equation 1 also has three additively separable terms: one related to rivals' prices, one related to own demand shocks, and one related to unobserved pricing heterogeneity. This assumption is mainly for tractability and might be relaxed in future work. It rules out the possibility that, for instance, higher rivals' prices may change the marginal effect of a change in δ_{jt} on j's prices. This assumption is largely mitigated by two factors that are allowed by my framework: first, p_{jt} and p_{-jt} can be log prices (or any other transformation of prices), and $z_{2j}(.)$ can be nonlinear so long as it is smooth and monotonic. Because prices in the hotel industry tend to have longer right tails, i.e., particular high demand days have extremely high prices, I work with log prices in my application.

Assumption 4: Errors. Errors are i.i.d. across markets and firms, independent of demand shifters, and normally distributed with mean zero and unknown variance:

$$\nu_{jt} \sim \mathcal{N}(0, \sigma_j^2).$$

Notably, the variance of this error term may be different for different firms. As a result, a firm with "noisier" pricing need not necessarily be assigned noisier beliefs over demand. Independence is admittedly a strong assumption. In the context of hotels, errors in pricing may be correlated across hotels if, for instance, there were market-wide supply shocks on a particular night. Errors in pricing may be correlated over time if there were persistent supply shocks or optimization errors. Errors in pricing may be correlated with demand shocks if, for instance, hotels offered quantity discounts for large group bookings, or if high demand was also associated with a preponderance of loyalty customers. I cannot rule out any of these concerns; the stance of the present article is that i.i.d. error in pricing is better than no error. Future work may be able to account for correlated errors. For instance, if errors are correlated with demand shocks, instrumental variables that shift prices only through demand shocks may be useful. I discuss that extension in an online appendix.

Theorem 1. *Under **Assumptions 1-4 and KD**, the supply model, $\{z_{1j}, z_{2j}, \sigma_j\}_{j\in\mathcal{J}}$ and F, is identified from the joint distribution of prices $[p_{jt}]_{j\in\mathcal{J}}$ and demand shifters $x_t, \lambda_t, [\xi_{jt}]_{j\in\mathcal{J}}$.*

Proof: See Appendix A for the full proof; I oultline only the key steps here. The proof is broken into three lemmata. The first lemma shows that z_{1j} is identified from changes in x_t. As x_t increases, p_j and $\bar{p}_{-j}$ both increase. The ratio of the increase in p_j to the increase in $\bar{p}_{-j}$ pins down z_{1j}.

The second lemma shows that, given knowledge of z_{1j}, the joint distribution of p_j and p_{-j} jointly pins down the signal structure and σ_j. This lemma relies on a key observation: The joint distribution of prices *conditional on demand shifters* is a convolution of the $\mathcal{N}(0, \sigma_j^2)$ distribution of ν_j and the noise induced by the signal structure. Noisiness in prices must come from either the error term or the signal. I use a copula-based argument to show that, if the error term is larger, the signal must be more informative, which implies that the *marginal* distribution of prices must be more variable (this uses a comparative statics result from Vizcaíno and Mekonnen [2019]). Because prices are all observed, the marginal distribution of prices thus identifies the signal, and the conditional distribution identifies σ_j.

The final lemma shows that, given knowledge of the rest of the model, z_{2j} is identified from the response of prices to changes in δ, which is as good as observed because $\delta(.)$ is known. Thus, the model is identified. □

Assumption UD is more difficult to work with. To see, this, examine the price policy function under this assumption:

$$p_{jt} = z_{1j}\mathbb{E}\big(\bar{p}_{-jt}|x_{jt}, s_{jt}\big) + \mathbb{E}\big(z_{2j}(\delta_{jt})|x_t, \xi_{jt}, s_{jt}\big) + \nu_{jt}$$

as compared to the price policy function under Assumption KD:

$$p_{jt} = z_{1j}\mathbb{E}(\bar{p}_{-jt}|x_{jt}, \delta_{jt}, s_{jt}) + z_{2j}(\delta_{jt}) + \nu_{jt}.$$

Under Assumption KD, the second term is not an expectation, so j's response to changes in x_t, conditional on δ_{jt}, does not depend on the signal structure. This ensures that z_{1j} can be easily recovered from changes in x_t. Under Assumption UD, the second term is an expectation that depends on x_t, so changes in x_t, even conditional on some value of δ_{jt}, can change beliefs about δ_{jt} in a manner that depends on the signal structure. This renders it more difficult to separately identify z_{1j} and $z_{2j}(.)$. Nonetheless, the following theorem states that the model is *generically* identified under Assumption UD. The notion of genericity here corresponds to that of "prevalence" in Anderson and Zame [2001] and Andrews [2011], and roughly means that, out of the set of all models $\Gamma = \{[z_{1j}, z_{2j}(.), \sigma_j]_{j\in\mathcal{J}}, F\}$ that satisfy Assumptions 1-4 and KD, the subset of parameters for which there exists another set of parameters in Γ that generate the same data has measure zero.[4] That is, the model is identified for almost every set of parameters that satisfy the assumptions.

Theorem 2. *Under **Assumptions 1-4 and UD**, the supply model $\{z_{1j}, z_{2j}, \sigma_j\}_{j\in\mathcal{J}}$ and F, is generically identified from the joint distribution of prices $[p_{jt}]_{j\in\mathcal{J}}$ and demand shifters $x_t, \lambda_t, [\xi_{jt}]_{j\in\mathcal{J}}$*

Proof: See ppendix A]. The proof is similar to that of Theorem 1, except that establishing (generic) identification of z_{2j} is somewhat more involved because it lies inside of the expectation operator. $\square$

Though this genericity result is comforting, it does not provide sufficient conditions to *guarantee* the model is identified given the data. The following results provide some guidance on when the model is identified, under stronger assumptions. Suppose $z_{2j}(.)$ is an analytic function; that is, it can be represented as a polynomial of arbitrary degree $z_{2j}(\delta_{jt}) = \sum_{\tau=0}^{\tau^*} z_{2j}^{(\tau)}\delta_{jt}^{\tau}$, where $0 < \tau^* \leq \infty$. Furthermore, suppose all noncentral moments of the distribution of δ, conditional on ξ, exist, for each ξ. Define the family of functions $\mathcal{M}^{\tau^*}$ as follows:

$$\mathcal{M}^{\tau^*} = \left[\mathbb{E}(\delta^{\tau}|\xi_{jt})\right]_{\tau=1}^{\tau^*}.$$

That is, this is the set of noncentral moments of the distribution of δ_{jt}, conditional on ξ_{jt}. I express the τ-th element of $\mathcal{M}^{\tau^*}$ as the function $M^{(\tau)}(\xi_{jt}) = \mathbb{E}(\delta_{jt}^{\tau}|\xi_{jt})$. Define these moments as *complete* if the following condition holds: if

$$M^{(\bar{\tau})}(\xi) = \sum_{\tau=1}^{\tau^*} \alpha^{(\tau)} M^{(\tau)}(\xi)$$

[4]The concept of "measure zero" is problematic in infinite-dimensional spaces, such as function spaces, which forms the distinction between "prevalence" and genericity. See Anderson and Zame [2001].

for some series of constants $[\alpha^{(\tau)}]_{\tau=0}^{\tau^*}$, then $\alpha^{(\bar{\tau})} = 1$ and $\alpha^{(\tau')} = 0$ for all $\tau' \neq \bar{\tau}$. The completeness condition is equivalent to a rank condition if ξ were finite-valued. It roughly states that no moment of the conditional distribution of δ given ξ can be represented as a linear combination of other moments of the conditional distribution of δ given ξ, for all ξ, if these other moments are less than or equal to the τ^*-th moment. We have the following result:

Proposition 3. *If $z_{2j}(.)$ is a polynomial of order $\tau^* \in \mathbb{N}$, $0 < \tau^* \leq \infty$, for each j, the model is identified if and only if the class of functions $\mathcal{M}^{\tau^*}$ is* complete.

Proof: See Appendix. □

This proposition yields some useful insights regarding identification given the specification of $\delta_j(.)$. Informally speaking, the more the distribution of δ_{jt} is *distinct* for different values of ξ_{jt}, the more flexible $z_{2j}(.)$ can be. This means the proposition allows the econometrician to determine the appropriate choice of $\delta_j(.)$ to identify $z_{2j}(.)$ up to τ parameters. The following examples illustrate this.

Example 1. Suppose

$$\delta_{jt} = b_0 + \xi_{jt} + b_x x_t + b_\lambda \lambda_t,$$

with b_x and b_λ both greater than 0. The noncentral moments may be expressed as

$$\begin{aligned} M^{(\tau)}(\xi_{jt}) &= \mathbb{E}\big((b_0 + \xi_{jt} + b_x x_t + b_\lambda \lambda_t)^\tau | \xi_{jt}\big) \\ &= \mathbb{E}\Bigg(\sum_{k=0}^{J} \binom{\tau}{k} \xi_{jt}^k (b_0 + b_x x_t + b_\lambda \lambda_t)^{\tau-k} | \xi_{jt}\Bigg). \end{aligned}$$

Using the linearity of the expectation operator, we have

$$M^{(\tau)}(\xi_{jt}) = \sum_{k=0}^{\tau} \binom{\tau}{k} \xi_{jt}^k \; \mathbb{E}\big((b_0 + b_x x_t + b_\lambda \lambda_t)^{\tau-k} | \xi_{jt}\big)$$

which, because of independence, becomes

$$M^{(\tau)}(\xi_{jt}) = \sum_{k=0}^{\tau} \binom{\tau}{k} \xi_{jt}^k \; \mathbb{E}\big((b_0 + b_x x_t + b_\lambda \lambda_t)^{\tau-k}\big).$$

For some (finite) τ^*, we can then express $\mathcal{M}^{\tau^*}$ as a matrix:

$$\mathcal{M}^{\tau^*} = M \begin{bmatrix} \xi_{jt} \\ \xi_{jt}^2 \\ \vdots \\ \xi_{jt}^{\tau^*} \end{bmatrix}$$

where M is a $(\tau^* + 1)$-dimensional lower-triangular square matrix with generic entry $M_{\tau,k} = \binom{\tau}{k}\mathbb{E}\big((b_0 + b_x x_t + b_\lambda \lambda_t)^{\tau-k}\big)$. Completeness of $\mathcal{M}^{\tau^*}$ is equivalent to M being nonsingular. Because M is lower-triangular, it is nonsingular if and only if all of its diagonal entries are nonzero. This implies that $\mathbb{E}\big((b_0 + b_x x_t + b_\lambda \lambda_t)^0\big) \neq 0$, which is obviously true. This turns out to be true, by similar arguments, for any δ_{jt} that is additively separable in (x_t, λ_t) and ξ_{jt}, i.e. $\delta_{jt} = \delta_{1j}(x_t, \lambda_t) + \delta_{2j}(\xi_{jt})$.

Remark 1. *If $z_{2j}(.)$ is an analytic function, and δ_{jt} is an additive index of ξ_{jt}, x_t, and λ_t, then $z_{2j}(.)$ is identified.* △

Example 2. Suppose δ_{jt} followed a threshold rule:

$$\delta_{jt} = \begin{cases} 1 \text{ if } \lambda_t > \bar{\lambda}_j(x_t, \xi_{jt}) \\ 0 \text{ if } \lambda_t \leq \bar{\lambda}_j(x_t, \xi_{jt}) \end{cases}$$

where $\bar{\lambda}_j$ is decreasing in both of its arguments. Then the distribution of δ_{jt} conditional on ξ_t is Bernoulli with some probability that is increasing in ξ_{jt}. Note that the distribution of δ_{jt}^{τ} does not vary with τ if $\tau > 0$. As a result, $M^{(\tau)}(\xi_{jt})$ is the same for all $\tau \geq 1$. Thus, $z_{2j}(.)$ is only identified up to a constant and slope parameter (assuming λ_t is above the threshold with some probability strictly between 0 and 1).

Remark 2. *If δ_{jt} is binary, $z_{2j}(.)$ is only identified if it is an affine function.*

△

Discussion of policy function approach. In this analysis, I work directly with price policy functions instead of writing down a structural model including marginal costs and conduct. I do this for three reasons. First, it is most appropriate for my empirical setting of hotel pricing. Second, price policy functions are generally sufficient when the primary object of interest is the information structure. Third, so long as price policy functions satisfy Assumption 3, it is easy to microfound them without changing any results in the model. I discuss each reason in turn.

The policy function approach is most appropriate for hotel pricing. As I discuss in Section 3, because hotels face stochastic, dynamic demand and capacity constraints, hotel pricing is a complicated revenue management problem. In practice, pricing is often a black box. Any microfoundation of hotel pricing I could write down would likely be misspecified. Instead, my approach incorporates maximal flexibility in terms of functional form and cross-firm heterogeneity that the data will allow. Where I do impose stronger assumptions on functional form, they are drawn from the literature. For instance, the separability between rivals' prices and firms' demand is consistent with revenue management practices for hotels (see Cho et al. [2018], whose revenue management model yields a reduced form with this separability feature). Separability and monotonicity assumptions are commonly made in related settings such as auctions [Compiani et al., 2020].

Second, the questions I study in my application make the policy function approach most suitable. The main goals of my empirical exercise are to (1) measure differences in information quality and their determinants and (2) predict information gathering in a counterfactual where incentives to gather information are different, but hotels' objective functions remain the same. That last detail is crucial– because the only thing that changes is information quality, the price policy function remains unchanged. Therefore it is unnecessary to microfound the price policy function. In short, I am only as structural as I need to be.[5] This approach is taken in other empirical studies, such as Miller et al. [2019] and Benkard et al. [2019]. A shortcoming of my approach is that it is not appropriate for counterfactuals, such as mergers, where the price policy function would change.

The final reason is simple: any other model that microfounds price policy functions could still use my approach to identification so long as my assumptions are satisfied. Compiani et al. [2020] make a separability assumption on bidders' valuations, which in turn implies separability in their bid functions (the analog of policy functions). In this sense, it is easy to push the assumptions I make directly on the policy functions back one step to firms' objective functions.

[5] I thank an anonymous referee for this wonderfully pithy statement.

3 Hotel Industry and Data

Hotel Pricing. The majority of hotels in the United States are affiliated with a chain, or, equivalently, a "parent company." A parent company (e.g., Marriott International) maintains a portfolio of brands (e.g., Marriott owns Courtyard, Fairfield Inn, flagship Marriott hotels, and many other brands). Over the past twenty years, these parent companies have divested most of their properties [Roper, 2017]. As a result, chain-affiliated hotels are generally franchised and managed by either the franchisees themselves or third-party management companies. Generally the parent company provides branding, training, and information technology; in return, the franchisee pays the parent a fixed fee plus a percent royalty of top-line revenues. Royalty rates vary across brands but not within a brand. The franchisee is usually the residual claimant on hotel profits.[6]

Hotel chains delegate pricing to the franchisee. This may be because of the possibility that franchisees have some local knowledge about market conditions. As established in Aghion and Tirole [1997], delegating decision rights to the agent may provide stronger incentives for the agent to collect information, especially if the agent is best positioned to gather this information. Hotel franchisees typically have access to a proprietary revenue management system operated by the parent company through its centralized revenue management office. Through this system, the parent company *nonbinding* price recommendations to franchisees that may convey the chain's private information about market conditions. It is through this process that there may be scope for hotels affiliated with larger chains that have better data or more sophisticated IT capabilities to adjust prices to demand shocks more effectively. The algorithms parent companies use in generating price recommendations are largely a black box and may vary from hotel to hotel. Furthermore, franchisees may be able to collect information about market demand beyond whatever is conveyed to them by the parent company. There remains a notion that franchisees' local knowledge is important in the industry.

Data. My main data source is STR LLC., a hotel benchmarking company that collects price and occupancy data from its member hotels and offers members competitive reports on pricing and performance. The unit of observation is a hotel-night of stay (or hotel-night). For each hotel-night, I observe the average daily rate (or ADR, which I hereafter use interchangeably with "price") and occupancy rate across eighteen different U.S. markets for a calendar year,

[6] By law, franchisors who operate in Wisconsin must publicly post franchise disclosure documents online. These documents go through, in detail, the legal terms of the standard franchisee contract for a particular hotel brand. See https://www.wdfi.org/fi/securities/franchise/default.htm.

between July 2016 and June 2017. The ADR is the *average* price charged by that hotel on a particular night. As a result, I do not know how far in advance rooms were booked, through what channel they were booked, or the degree to which price dispersion exists for a specific hotel on a specific night. I also observe select hotel characteristics, such as anonymized franchisee, brand, management company, and parent company, as well as quality tier and capacity.

My eighteen geographic markets are all United States college towns with relatively large undergraduate populations, prominent Division IA football teams, and a degree of isolation from other metropolitan areas. Table 1 provides a sense of coverage in the eighteen geographic markets by comparing property counts in my data to those reported by the Census Bureau. I choose these markets because their boundaries are easily defined and it is easy to identify particular nights on which market-level demand for lodging will be high. These are the nights before and after undergraduate move-in, graduation, and home football games. I refer to these as "high demand days." All hotels in my data are within fifteen miles of the college's stadium, though most are within a five-mile radius of the stadium.

[Table 1 Here]

Finally, I match the hotels to the *2015/16 HVS Franchise Fee Guide*. The franchise fee guide is designed as a tool to help franchisees assess which brand is right for them. The fees are compiled by reviewing franchise disclosure documents and are reported as percentages of expected annual revenues for that brand. Though most brands levy royalties as percentages of revenues, some brands instead charge a flat or per-room fee. Brands that charge flat fees and independent hotels are assigned a 0% marginal royalty rate on revenues.

Because most of my analysis focuses on differences across parent companies, I report summary statistics of my data by anonymized parent company in Table 2. As is the case in the United States at large, five parent companies collectively dominate these markets, with a competitive fringe consisting of many smaller parents and independent hotels. I refer to the five parent companies as the "large chains" throughout this article. Of these parent companies, two ("A" and "B") exist primarily in higher-quality tiers, one ("C") is primarily midscale, and two ("D" and "E") are primarily lower-quality. Franchisees affiliated with the two large upscale parent companies in particular generate more revenue per available room, charge higher prices, and usually fill more occupancy than their rivals.

[Table 2 Here.]

I also report differences across parent companies in their marginal royalty rates in this table. Of the 768 hotels in my dataset, I match 729 of them to a marginal royalty rate. In general, the larger parent companies charge higher royalty fees than their rivals. Because franchisees are residual claimants on profits, this implies that franchisees likely have stronger incentives to gather information at smaller chains, all else equal.

[Figure 1 Here.]

Descriptive evidence. Overall, price and occupancy are positively correlated in my data. The binned scatterplot of ADR versus occupancy, which conditions on hotel fixed effects, in Figure 1 demonstrates this relationship. This figure is consistent with two features of the industry. First, demand is highly variable relative to supply for any particular hotel. This demand variation traces out a "reduced-form supply schedule" like that shown in the figure. Second, hotels increase prices as demand increases, possibly to capture surplus if they have market power and to account for the possibility that a sellout will occur if prices remain low. As demand increases, the probability of selling out increases, so prices increase steeply in order to ration capacity.

I now present descriptive evidence that hotels have incomplete information about demand. I use the identification results of the previous section as a rough guide: If a hotel is poorly informed about demand, its prices will vary more with common-knowledge demand shocks than possibly unknown demand shocks. For the descriptive section only, I assume that a hotel's occupancy is a reasonable proxy for shifts in its residual demand curve and that there is no supply-side variation at all; I relax this assumption in the more sophisticated analysis in the next section. If demand shifts result from common-knowledge demand shocks, prices should increase as occupancy increases, as in Figure 1. If demand shifts result from completely unknown demand shocks, prices will not respond, so the relationship between price and occupancy will be flat. Intermediate cases with incomplete information will result in an upward-sloping relationship between price and occupancy that is flatter than the common-knowledge case.

I construct empirical analogs of the relationships between price and occupancy driven by common-knowledge and possibly-unknown market-wide demand shocks as follows. First, I regress price and occupancy on variables that shift demand and are likely common-knowledge among all hotel managers. Specifically, I estimate

$$
\begin{aligned}
\log(ADR_{jt}) = & \quad \lambda_j^p + \beta^p X_{jt} + u_{jt}^p \\
Occupancy_{jt} = & \quad \lambda_j^q + \beta^q X_{jt} + u_{jt}^q.
\end{aligned}
\tag{2}
$$

for hotel j on night t, where X_{jt} contains the following common-knowledge demand shifters: dummies for a hotel's brand interacted with an indicator for being a weeknight, city interacted with day-of-week, and city interacted with a high demand day indicator. λ_j is a hotel fixed effect. The variation in prices and occupancy driven by common-knowledge demand shifts is given by the fitted values from these regressions after absorbing hotel fixed effects, e.g., $\hat{\beta}^p X_{jt}$ for prices. The effects of possibly-unknown variation in demand on prices and occupancy are reflected in the residuals, $\hat{u}_{jt}^p$ and $\hat{u}_{jt}^q$ respectively. I then plot the co-variation in prices and occupancy driven by common-knowledge demand shifts and possibly-unknown demand shifts in Figure 2. That is, I plot $\hat{\beta}^p x_{jt}$ against $\hat{\beta}^q x_{jt}$ and I plot $\hat{u}_{jt}^p$ against $\hat{u}_{jt}^q$. This figure is consistent with the notion that hotels have incomplete information about certain demand shocks: prices are more responsive to changes in occupancy when these changes are driven by common-knowledge demand shocks.

[Figure 2 Here.]

Figure 2 pools data across all hotels. I do the same analysis separately for each hotel and report summary statistics of each hotel's "responsiveness ratio"– the ratio of the slope of the solid line to the slope of the dotted line in Figure 2– in Table 3. Also report other descriptive measures of information quality in that table, broken down by the five big chains in Table 2. There are no obvious differences between large chain and small chain hotels. However, I cannot interpret any of these measures as purely reflective of information quality. For instance, large chain hotels seem responsive to demand shocks because they are responsive to rivals, who themselves are more responsive to demand shocks. To more precisely measure information, I require a more formal model of hotel supply and demand, which I develop in the next section and estimate using as guidance the identification results of the previous section.

4 Empirical Specification and Estimation

Demand. Demand is a nested logit. A market t is a city-night. I assume each market can be treated as separate. Let $w(t)$ indicate whether t is a weeknight (defined as Monday through Thursday nights). A hotel is indexed by j. The nest to which j belongs is $b(j) \in$

$\{downscale, midscale, upscale\}$, and the set of hotels belonging to nest b in market t is $\mathcal{J}_{bt}$. Hotel j is also affiliated with brand $r(j)$. Let $J = 0$ represent the outside option, which is its own nest. Consumer i chooses among $J+1$ total options. Utility is given by

$$u_{ijt} = -\alpha_{w(t)}\log(p_{jt}) + \lambda_{r(j),w(t)} + \beta_{w(t)}\tilde{x}_t + \lambda_t + \xi_{jt} + \varepsilon_{i,b(j),t}(\gamma_{w(t)}) + \varepsilon_{ijt}$$
$$u_{i0t} = \varepsilon_{i0t}$$

where p_{jt} is the hotel's average daily rate. $\tilde{x}_t$ includes a set of dummies for salient demand shifters; these are the shifters that I eventually assume are common knowledge. $\tilde{x}_t$ includes dummies for interactions between "high demand" status and city, interactions between month of year and city, and interactions between day of week and city. $\lambda_{r(j),w(t)}$ is a fixed effect for the brand $r(j)$ associated with hotel j on night-type $w(t)$. λ_t is a city-night fixed effect that captures that captures market-level demand variation that cannot be explained by $\tilde{x}_t$. These shocks are not considered common knowledge in my supply model. Finally, there is an unobserved hotel-level shock ξ_{jt}. ε_{ijt} and ε_{ij0t} are type-I extreme value draws that define consumers' idiosyncratic preferences for different hotels, and $\varepsilon_{i,b(j),t}$ represents a nest-wide preference shock that scales with parameter $\gamma \in [0,1]$. If $\gamma = 0$, $\varepsilon_{i,b(j),t}$ is degenerate at zero, and demand is a standard multinomial logit. If $\gamma = 1$, products are perfect substitutes within-nest.

I allow demand parameters to be different on weeknights versus weekends. This is a simple way to capture heterogeneity in tastes between business travelers and leisure travelers. This may, in turn, be associated with changes to hotels' supply strategies depending on whether the night is a weekend. Accordingly, in my supply specification later on, I will allow supply parameters to be different between weeknights and weekends. This allows differences in demand between weekdays and weeknights to involve both shifts and *rotations* of demand. Other nightly variation in demand is simply modeled as a shift, consistent with the model in Section 2.

Letting $\bar{u}_{jt} \equiv -\alpha_{w(t)}p_{jt} + \lambda_t + \beta\tilde{x}_t + \lambda_j + \xi_{jt}$ be the mean utility for hotel j on night t, this specification yields predicted quantities

$$q_{jt} = M_t \frac{\exp(\bar{u}_{jt}/(1-\gamma))}{\sum_{j' \in \mathcal{J}_b} \exp(\bar{u}_{j',t}/(1-\gamma))} \frac{\exp(V_{b(j)t})}{1 + \sum_b \exp(V_{bt})} \tag{3}$$

where $V_{bt} = \log\left((1-\gamma)\sum_{j\in\mathcal{J}_b} \exp(\bar{u}_{j',t}/(1-\gamma))\right)$ is the logit inclusive value of nest b. I set M_t, the market size, equal to double the total capacity of all hotels in a city. This allows for the possibility that a hotel will sell out for a sufficiently high draw of the unknown market-wide demand shifter, λ_t. As will be seen later in this section, this is key for the identification of the supply side of the model.

Due to the endogeneity of prices and within-nest shares (i.e., correlation of these terms with unobserved ξ_{jt}), I require instruments to identify the demand parameters. These instruments W must affect firm prices and within-nest shares while also satisfying the exclusion restriction

$$\mathbb{E}(W\xi|\lambda_{r(j)}, \lambda_t, \tilde{x}_t) = 0.$$

The most obvious candidates for instruments in such models are marginal cost shifters. However, marginal cost shifters are uncommon in the hotel industry; in part because there are hardly any marginal costs at all. Cho et al. [2019], for instance, cite a lack of cost shifters as a reason they rely on supply-side restrictions to identify hotel demand without using instruments at all. My data do contain variation in market structure across geographic markets, allowing me to use variation in rival characteristics as instruments that shift markups and shares, in the spirit of Berry et al. [1995]. These are selected based on my hypothesis is that larger parent companies price differently than small parent companies, due to differences in information quality, and that pricing strategy depends on the ownership and chain affiliation of the hotel as well as of its rivals. As a result, I use as instruments the number of within-nest rival rooms, the number of within-nest rival rooms at hotels affiliated with the same parent company, and the interaction of all of these terms with two market-wide exogenous demand shifters: a dummy for whether it is a weekend and a dummy for whether it is a high-demand day.

Supply. The identification strategy in Section 2 showed that parts of the model can be recovered nonparametrically. Because of the limited number of data points I have, in practice I impose more parametric restrictions on the model that I estimate. The supply model is based on features of pricing in the hotel industry that are corroborated by the evidence in other research that microfounds and tests for the approximate optimality of hotel pricing practices [Cho et al., 2018]. These features are:

1. Hotels that do not expect to sell out attempt to undercut the average price of all hotels in its “comp set.” These are hotels in the same geographic market and quality segment, i.e., hotels in the same “nest” per the demand specification.

2. Hotels increase their prices more rapidly as they increasingly expect to sell out at a given price.

3. Consumers arrive over many days leading up to the night of stay (the "booking window"). The earlier during the booking window the hotel learns that demand is high, the sooner it can increase its price, the higher average prices will be for that night. Thus, a hotel that learns demand *quickly* is observationally the same as a hotel that learns demand *precisely*. Both are interpreted as "better information."

Fix a geographic market and a weekday status $w(t) \in \{0, 1\}$. All supply parameters are allowed to be different between weekends and weekdays; keeping this in mind, I suppress w as an index on all supply parameters in this subsection. Consider a hotel j in nest $b(j)$. Let other nests in a particular market t be $-b$. Finally, let $x_t = \beta \tilde{x}_t$ be common-knowledge market-wide demand shifts.

I begin with the first part of pricing: each hotel's beliefs over rivals prices. j approximates each rival's price as log-linear in three demand shifters: x_t, ξ_{jt}, λ_t:

$$\log(p_{jt}) = \phi_{0j} + \phi_{xj} x_t + \phi_{\lambda j} \lambda_t + \phi_{\xi j} \xi_{jt} + u^p_{jt}. \tag{4}$$

Because ξ_{jt} and u^p_{jt} are, by construction, independent of the market-wide demand shifters and mean-zero, the average price in segment b, $\overline{\log(p_{bt})}$ is then approximated by

$$\overline{\log(p_{bt})} = \phi_{0b} + \phi_{xb} x_t + \phi_{\lambda b} \lambda_t$$

where $\phi_{0,b(j)} \equiv \frac{1}{J_b} \sum_{j:b(j)=b} \phi_{0j}$ and J_b is the number of hotels in segment b. j forms expectations given signal s_{jt} as follows:

$$\mathbb{E}(\overline{\log(p_{bt})} | x_t, s_{jt}) = \phi_{0b} + \phi_{xb} x_t + \phi_{\lambda b} \mathbb{E}(\lambda_t | s_{jt}). \tag{5}$$

The second factor in j's pricing is its expectation over selling out. Given demand shifters $\lambda_{r(j),w(t)}$, x_t, ξ_{jt}, and this approximation of rivals' pricing, a sellout boils down to whether or not λ_t is sufficiently large. I approximate the threshold λ_t above which j will sell out as follows: suppose all hotels, including j, had set prices according to Equation 4 for their respective nest, and all hotels but j had $\xi_{j't} = 0$. Given this, and the actual values of all remaining demand shifters, let $q_j(x_t, \xi_{jt}, \lambda_t)$ be j's quantity sold according to demand Equation 3. We then define function $\bar{\lambda}_j(x_t, \xi_{jt})$ as implicitly solving:

$$q_j(x_t, \xi_{jt}, \bar{\lambda}_j(x_t, \xi_{jt})) = \bar{q}_j \tag{6}$$

where $\bar{q}_j$ is j's capacity. $\bar{\lambda}$ can be interpreted as the threshold λ_t above which j would sell out of rooms, if rival hotels had zero idiosyncratic demand shock and all hotels including j priced according to Equation 4. Using the notation of Section 2, I specify δ_{jt} as follows:

$$\delta_{jt} = \begin{cases} 1 & if\ \lambda \geq \bar{\lambda}_j(x_t, \xi_{jt}) \\ 0 & if\ \lambda < \bar{\lambda}_j(x_t, \xi_{jt}). \end{cases}$$

That is, δ_{jt} is an indicator for a sellout. This means that $\mathbb{E}(\delta_{jt}|s_{jt}, x_t, \xi_{jt})$ is j's subjective probability it sells out, according to our approximations, if pricing follows Equation 4. Because δ_{jt} is binary as in Example 2, the function $z_{2j}(.)$, which measures j's responsiveness to the prospect of a sellout, is only identified up to a slope and intercept. Based on the features outlined in this section, I specify the price policy function as:

$$\log(p_{jt}) = z_{0j} + \underbrace{z_{1j}\mathbb{E}(\log(\bar{p}_{-jt})|s_{jt}, x_t)}_{\text{rivals' average price}} + (z_{2j} \underbrace{\mathbb{E}(\delta_{jt}|x_t s_{jt}, \xi_{jt})}_{\text{probability of sellout}} + \nu_{jt} \tag{7}$$

where ν_{jt} is an error distributed $\mathcal{N}(0, \sigma_j^2)$. I furthermore assume that x_t, λ_t, ξ_{jt} are normal and independent of each other with variances matching their empirical variance for that particular city.

Finally, I impose the following functional form on the information structure: the signal $s_{jt} = \lambda_t + \rho_j \sigma_\lambda \varepsilon_{jt}$, where σ_λ is the standard deviation of λ_t, ε_{jt} is standard normal, and $\rho_j \in \mathbb{R}^+$ is a parameter that governs the precision of j's signal. As ρ_j approaches zero, j obtains complete information; as ρ_j grows, j's information quality (or the speed at which j learns demand over the course of the booking window) worsens.

Under this assumption, using the properties of joint normality,

$$\mathbb{E}(z_{2j}\delta_{jt}|x_t, \xi_{jt}, s_{jt}) = z_{2j}\Phi\left(\frac{m_\lambda(\rho_j)s_{jt} - \bar{\lambda}(x_t, \xi_{jt})}{\hat{\sigma}_\lambda(\rho_j)}\right) \tag{8}$$

where $\Phi(.)$ is the standard normal cdf, $m_\lambda = \frac{1}{1+\rho}$ is the amount j's posterior expectation over λ changes when j's signal changes, and $\hat{\sigma}_\lambda = \sigma_\lambda\sqrt{\frac{\rho}{1+\rho}}$ is the standard deviation of j's posterior beliefs. Inserting Equations 5 and 8 into Equation 7, I obtain the supply model used for estimation:

$$\log(p_{jt}) = z_{0j} + z_{1j}\left[\phi_{0b(j)} + \phi_{xb(j)}x_t + \phi_{\lambda b(j)}m_\lambda(\rho_j)s_{jt}\right] + z_{2j}\Phi\left(\frac{m_\lambda(\rho_j)s_{jt} - \bar{\lambda}(x_t,\xi_{jt})}{\hat{\sigma}_\lambda(\rho_j)}\right) + \nu_{jt}. \quad (9)$$

Estimation. The supply parameters to be estimated are $z_{0j}^{w(t)}$, $z_{1j}^{w(t)}$, $z_{2j}^{w(t)}$, $\rho_j^{w(t)}$, and $\sigma_j^{w(t)}$, for each j and $w(t) \in \{weekend, weekday\}$. Estimation is by indirect inference. For each hotel in a city, I estimate the slope of log prices with respect to the demand shifters via the following linear regression:[7]

$$\log(p_{jt}) = \hat{\phi}_{0j}^{w(t)} + \hat{\phi}_{xj}^{w(t)}x_t + \hat{\phi}_{\lambda j}^{w(t)}\lambda_t + \hat{\phi}_{\xi j}^{w(t)}\xi_{jt} + \hat{u}_{jt}. \quad (10)$$

I then simulate the supply model for a guess of supply parameters $\rho_j^{w(t)}, z_{1j}^{w(t)}, z_{2j}^{w(t)}, z_{3j}^{w(t)}$ by drawing signals, forming hotels' beliefs, and using Equation 9 to predict prices. I then run a regression with the specification given by Equation 10 on the simulated data and search over the supply parameters until the regression coefficients from the simulation match the coefficients from the above regression. Details of the supply estimation are given in the online appendix.

Identification in practice. The constructive identification arguments in Section 2 provide guidance over which moments to use in estimation. I argue that the reduced-form coefficients I use do identify the model by mapping them to the underlying model parameters intuitively. Suppress the $w(t)$ superscript. $\hat{\phi}_{\xi j}$ serves to identify the responsiveness of hotel j to its own idiosyncratic demand shocks; because these shocks only affect j's probability of a sellout, and not rivals' prices, this serves to identify the z_{2j}. Given knowledge of z_{2j}, j's responsiveness to market-wide common-knowledge demand shocks ($\hat{\phi}_{xj}$), identifies z_{1j}. The difference between $\hat{\phi}_{xj}$ and $\hat{\phi}_{\lambda j}$ helps identify j's information quality: the closer these two are to each other, the better j's information, as it is equally responsive to both, known and unknown market-wide demand shocks.

As mentioned before, I posit a different set of demand and supply parameters on weekends versus weekdays. On the demand side, this means that differences in demand between weekends and weekdays, while common-knowledge and market-wide, involved both shifts and rotations

[7] I use ordinary least squares to estimate the regression. Feasible generalized least squares would yield, in this case, the same estimates. See Wooldridge [2010] Theorem 7.6.

of demand because of the different mix of travelers. Accordingly, variation in demand between weekends and weekdays may violate Assumption 1 in Section 2 and cannot be used to estimate supply. As a result, I estimate an entirely different supply-side for weekends versus weekdays as well. This adds realism: Hotel managers likely respond differently to demand shocks on weekends than on weekdays because the mix of travelers is different. I rely on the remaining common-knowledge sources of demand variation, such as high demand days and market interacted with the month of year and day of the week (conditional on being a weekend or weekday) for identification. Though I cannot rule out the same criticisms of these sources of variation that I just mentioned for weekends versus weekdays, I believe that pooling these days together is more plausible than pooling weekdays and weekends together.

A final caveat to identification of this model of supply: for z_{2j} to be identified, there must be variation in δ_{jt} across t. If a hotel has a sufficiently low draw of ξ_{jt} or $\lambda_{r(j),w(t)}$, it may not be possible that this hotel never expects to sell out, even if λ_t grows very large, in which case z_{2j} is not identified for this hotel (this follows from Proposition 3). I address this in part by allowing market size M_t to be quite large, but there is still no guarantee that a sufficiently unpopular hotel would ever expect a sellout. If δ_{jt} does not sufficiently vary for a hotel j over time, I set z_{2j} equal to zero by default.

5 Results

Demand. Demand parameter estimates from the nested logit are reported in Table 4. Because there is a city-night fixed effect λ_t, different specifications of market size do not substantially affect estimates of any parameters other than this fixed effect. The online appendix includes summary statistics of estimated own-price elasticities. These are roughly in line with other recent hotel demand estimates in the literature [Farronato and Fradkin, 2018]. A notable feature of my estimates is that consumers are more price-elastic on weekdays than on weekends. This is in contrast to the conventional wisdom that weekday business travelers are less price-sensitive than leisure travelers. My estimates suggest that this may be a unique feature of college towns, where weekend travelers are often families or alumni with high willingness to pay. To corroborate this, I run the same nested logit specification on seven Midwestern markets of similar size that are not considered college towns (e.g., Rochester, MN and Fort Wayne, IN). According to those estimates, consumers are considerably more price sensitive on weekends ($\alpha = -2.278$) than on weekdays ($\alpha = -1.306$).

[Table 4 Here.]

Table 4 also reports the variances of different demand shifters. Common-knowledge shifter $x_t + \lambda_j$ varies four times as much as λ_t across all observations; because these two shifters enter demand through the same index, this corresponds to common-knowledge shifters explaining three times as much of market-wide variation in demand over time as λ_t does. Conditioning on hotel reveals that much of the variation in $x_t + \lambda_j$ is *across* hotels and cities. The amount of demand variation for a given hotel over time explained by common-knowledge and unknown market-wide demand shifters are much more similar. Overall, the amount of variation in the demand shifters presented in Table 4 is consistent with common-knowledge and possibly-unknown demand shifters predicting similar amounts of variation in a hotel's prices over time, if the possibly-unknown demand shifters were indeed known. In the online appendix, I plot kernel densities of these shifters, conditional on hotel. These distributions appear approximately normal, consistent with my functional form assumptions.

Supply. I then estimate supply parameters according to the procedure described in the previous section for each of thirty-six submarkets: eighteen geographic markets, times two day of week categories $\{weekend, weekday\}$. There is considerable heterogeneity across hotels in information quality. I report these estimates in Table 5.

[Table 5 Here.]

As reported in Table 5, there is clearly substantial heterogeneity in information quality across hotels. The first measure of information quality I report is "signal responsiveness," $\frac{1}{1+\rho_j^{w(t)}}$. Given the assumptions made on joint normality of signals and λ_t, this translates to the amount by which a hotel's posterior expectation of λ_t changes if a signal changes:

$$\frac{1}{1+\rho_j^{w(t)}} = \frac{\partial \mathbb{E}(\lambda_t | s_{jt}, x_t)}{\partial s_{jt}}.$$

By construction, given the signal structure, this measure is between zero and one. It equals one if information is perfect and zero if the hotel has no information. The 90/10 ratio in this measure of responsiveness is 4.71. This means that the 90th percentile hotel in information quality, which has nearly perfect information, is over four times more responsive to changes in the signal it receives than the 10th percentile hotel. For the remainder of my analysis, I use Shannon information $H(\rho)$ as a measure of information quality. Shannon information, which is the expected change in entropy between the firm's prior and posterior from a signal, is computed as

$$H(\rho_{jw}) = \log \left(\frac{\sqrt{1+\rho_{jw}^2}}{\rho_{jw}} \right) \tag{11}$$

given the functional form assumptions placed on the signal structure. This measure of information quality has been microfounded and is widely used in the information economics literature, most notably the rational inattention literature [Caplin and Dean, 2015]. I also report dispersion in Shannon information in Table 5.

Predictors of information quality. I project my estimates of each hotel's information quality onto hotel characteristics. The results of these linear regressions are summarized in Table 6. In the first specification, corresponding to the first column of results in the table, I project information quality (in Shannons) on city fixed effects and an indicator for whether a hotel belongs to one of the five "big" chains, per the discussion in Section 2. Hotels associated with these large chains are significantly worse-informed than small-chain rivals. The point estimate of this difference is .191 Shannons, roughly equal to the difference between the 60th and 40th percentile estimate in Shannons across all hotel-1[weekdays]. As a robustness check, I also use the total number of hotel rooms in my dataset associated with each hotel chain as a continuous measure of hotel chain size and find similar results. These results run counter to the notion that large firms, equipped with bigger datasets and better analytical capabilities, are better at predicting demand.

[Table 6 Here.]

In specifications (3) and (4) of Table 6, I incorporate a dummy for whether a hotel is in the downscale quality segment. Two features of these regressions stand out. First, downscale hotels have better information quality than midscale hotels according to my point estimates.[8] Second, the difference between large and small-chain hotels' information vanishes. This is likely because several of the large hotel chains in my dataset— specifically C, D, and E per Table 2— operate many downscale hotels. Downscale hotels may have better information than midscale hotels because they are more fixated on price: whereas midscale hotels market themselves via non-price channels including brand loyalty and quality, downscale hotels have fewer levers to pull to ensure they attract customers without selling out of rooms at a price that is too low. This is corroborated by the additional specifications reported in the online

[8] Reminder that upscale hotels are excluded from my main analysis because they are less easily thought of as single product firms.

appendix: information quality is significantly higher for hotels on weekdays, which is also when consumers are more price-elastic per my demand estimates, and thus price is a more important variable for a hotel to control on weekdays.

Finally, in specification (5) of Table 6 I regress information quality on the *market-specific* share of rooms associated with the same parent company as each hotel. This allows me to differentiate between the overall size of a hotel chain and the magnitude of that chain's presence in a particular geographic market. I find that changes in market-specific chain share do not substantially predict a hotel's information quality. This suggests there may be no explicit relationship between access to richer market-specific data, *vis-a-vis* more data points for a specific market, and the ability to predict demand. Overall, these results descriptively suggest that larger chains do not have significantly better information than small-chain rivals, conditional on the quality segment in which a particular hotel competes.

Gains from information. I next investigate how hotel revenues change if the hotel obtains better or worse information. I obtain estimates of the gains from information as follows: using the policy function estimates I obtained, I simulate prices and occupancy for each day in my dataset. I then run the same simulation, but with ρ_j equal to the 10th percentile of ρ across all hotels in my dataset and again with ρ_j equal to the 90th percentile ($\rho = 2.05$). Because hotels interact with one another in my model, this may change rival hotels' pricing as well, which I account for in my simulation. Details of this simulation are in the online appendix. I use simulated prices and occupancies to generate annual revenues for each hotel in my data. The results are displayed in Figure 3.

[Figure 3 Here.]

Moving from the tenth percentile of information quality to (0.03 Shannons) to the 90th percentile (4.49 Shannons) increases gross annual revenues by over \$30,000 on average. To put this into context, I compute "calibrated operating profits" using the point estimates of operating costs per room for different-quality hotels in Hollenbeck [2017]. I subtract these costs from net revenues, defined as $(1 - Royalty_j) \times Revenues_j$ where $Royalty_j$ is hotel j's royalty fee per the franchise agreement specific to that particular brand of hotel. My estimates of these operating profits, in the right panel of Figure 3 range from \$255,000 to \$285,000 per year depending on information quality. This corresponds to hotels increasing profits by about 10% on average by moving from poor information to high-quality information. Finally, I note that the standard price quote for a 100-room hotel for Duetto, a common third-party revenue management tool, is \$12,000 per year. According to my estimates, a 100-room hotel with the median information quality (0.94 Shannons) would then be willing to buy a subscription to Duetto at this price if it increased its information quality to roughly the 90th percentile or better. Note that this analysis assumes that although *operating* costs of a hotel are nonzero, *marginal costs* are effectively zero, because changes to the number of rooms booked do not affect cost.

Incentives versus scale. As a final exercise, I compare the effects of incentives to those of scale in determining a hotel's ability to predict demand. I do this by using my estimates to (a) compute the marginal costs of gathering information for each hotel and then (b) solve for a hotel manager's optimal amount of information-gathering effort given a particular royalty fee. I employ a simple model information-gathering effort: A franchisee chooses information quality ρ_j to maximize profits at hotel j:

$$\max_{\rho_j}(1 - Royalty_j)\mathbb{E}\left[Revenues_j(\rho_j)\right] - c_j H(\rho_j) - F_j \tag{12}$$

where F_j is the operating cost of the hotel, $Royalty_j$ is the royalty rate on booking revenues, $Revenues_j(.)$ is the hotel's annual revenue, and c_j is a hotel-specific constant marginal cost of acquiring better information, measured in Shannons $H(\rho)$. The expectation is taken over the joint distribution of the hotel's signals and the demand shifters for each day of the year. The first order condition is

$$\begin{aligned} (1 - Royalty_j)\frac{\partial \mathbb{E}[Revenues_j(\rho_j)]}{\partial \rho_j} &= c_j \frac{\partial H(\rho_j)}{\partial \rho_j} \\ (1 - Royalty_j)\frac{\partial \mathbb{E}[Revenues_j(\rho_j)]}{\partial \rho_j} \Big/ \frac{\partial H(\rho_j)}{\partial \rho_j} &= c_j. \end{aligned}$$

Differentiating the formula for Shannon information, Equation 11, this equation becomes

$$-(1 - Royalty_j)\frac{\partial \mathbb{E}[Revenues_j(\rho_j)]}{\partial \rho_j}(\rho_j^3 + \rho_j) = c_j. \tag{13}$$

I then estimate the marginal revenues on the left hand side of the above equation by taking a linear approximation. Specifically, I simulate revenues at $0.8\rho_j$ and $1.2\rho_j$ per the same algorithm employed to generate Figure 3 and plug the difference in revenues, divided by $0.4\rho_j$, into Equation 13. I project these estimates of the marginal costs of information-gathering effort onto hotel characteristics, similarly to the regressions in Table 6. The results are in

Table 7.

[Table 7 Here.]

The results suggest that hotels belonging to large chains have higher marginal costs of information-gathering effort than small-chain and independent hotels, on average. This corresponds to large-chain hotels having less accurate demand predictions, as reported by Table 6, even though large chain hotels tend to generate higher revenues— and therefore higher gains from better demand predictions— and would thus be inclined to gather better information, all else equal. It does not suggest that large chains can more easily access information about demand simply because they are larger. For instance, there is no evidence of a "big data" advantage. Finally, because these estimates reflect the dilution in information-gathering incentives arising from royalty payments, they suggest that differences in royalty fees are not sufficient to explain the difference between large and small-chain hotels' information quality.

[Table 8 Here.]

I further illustrate how information gathering may respond to changes in incentives *vis à vis* the royalty fees by solving Equation 12 for a counterfactual scenario in which all firms face the same royalty of 12%. I use my estimated marginal costs of information-gathering effort, simulate revenues for each hotel over a grid of information quality, interpolate over this grid, and then deflate these predicted revenues according to the royalty fee before solving. I then re-estimate the specifications (1) and (3) from Table 6. The results are in Table 8: large chain hotels now have better information than small chain hotels, on average, though these differences are not statistically significant.

6 Conclusion

The results of my application suggest that any informational advantages large-chain hotels may have due to the effects of pure scale are smaller than the disadvantages of weaker managerial incentives. The estimates suggest that are part of large chains tend to have higher costs of obtaining information, consistent with the finding that these hotels tend to have worse predictions about demand than small-chain rivals. These differences are exacerbated by the weaker contractual incentives large-chain hotel managers have to gather information. At the same time, large-chain hotels tend to attract more demand, which gives their managers stronger incentives to gather information. My counterfactual analysis suggests that equalizing royalties across all hotels may lead to large-chain hotels having better-quality information than small-chain rivals *despite* facing higher costs of information gathering. However, this is driven by higher benefits of gathering information, not lower costs. Large-chain managers may have higher costs of gathering information because of the opportunity costs of their attention: these managers' time may be better spent ensuring high quality, for instance. This is pure speculation; the drivers of differences in information quality merit further investigation. There are, of course, limits to the external validity of this analysis: it is plausible that hotels in resort areas or large metropolitan areas may face a different informational scheme. Hotels may also be unique in that the production function for predicting demand is relatively reliant on the information-gathering effort of the downstream manager.

Other takeaways are more general. First, I show that it is possible to measure information credibly and precisely by leveraging common-knowledge demand shocks. The methods developed in this article can easily extend to other settings in which competing agents face uncertainty over some common parameter but may also face observable private shocks. This methodology can be applied, with minimal adjustments, to other settings where the analyst observes the unknown state *ex post* and can assume that some component of this state is common knowledge to the agents. A key assumption is that the pricing error term, ν_{jt}, is i.i.d. normal. Though formal extensions addressing this are not the primary objective of this article, I provide a discussion of addressing correlation between ν_{jt} and demand shifters in the online appendix. Theoretically, an instrumental variables approach would address this issue, but it is difficult to conceive of valid instruments. An alternative approach would be to impose more structure on the price policy functions, which may be sensible in settings such as auctions where pricing (bidding) is more straightforward. I note that common values auctions are a natural setting to apply my methods, so long as the econometrician observes a panel of auctions in which the same set of firms interact repeatedly, the econometrician can somehow measure the *ex post* realization of the common value, and the econometrician observes some variable that affects the value of the auctioned object that is common knowledge to bidders. Second, I demonstrate that informational differences across rivals are sizeable and that these informational differences yield sizeable differences in firm productivity and profitability. This point is likely to be true of many other industries. Further study should be devoted to the quantification of, attribution of, and consequences of informational differences across rival firms in other markets.

Table 1: Geographic markets in data

Market	Properties in Data	Census	Market	Properties in Data	Census
Ann Arbor, MI	75	44	Knoxville, TN	108	76
Auburn, AL	28	26	Madison, WI	79	83
Bloomington, IN	21	21	Oxford, MS	9	10
College Station, TX	43	43	South Bend, IN	59	44
Columbia, MO	29	34	Starkville, MS	11	13
Columbia, SC	97	77	State College, PA	28	26
East Lansing, MI	35	46	Tuscaloosa, AL	33	34
Gainesville, FL	33	38	Urbana-Champaign, IL	35	38
Iowa City, IA	26	29	West Lafayette, IN	20	27
Total					768

Source: STR LLC.. and U.S. Census Bureau *County Business Patterns* 2016. STR data include all hotels in data within fifteen miles of campus. Census data reported for the county in which the market is primarily located. Because these market definitions do not exactly coincide, either figure may be larger than the other. However, because most markets in my dataset are relatively isolated, these discrepancies are generally small. An exception is Ann Arbor; my dataset includes some properties to Ann Arbor's east as part of Wayne County in the Detroit metropolitan area.

Table 2: Data summary, by parent company

Parent	Number of Hotels				Mean		
	Upscale	Midscale	Downscale	Total	ADR ($)	Occ. (%)	Marginal Royalty (%)
A	9	97	0	106	122.92	73.37	13.65
B	12	94	0	106	118.45	70.67	12.97
C	1	79	15	95	104.57	68.41	13.13
D	0	58	86	144	77.60	58.08	11.24
E	0	5	119	124	68.42	52.66	12.54
All Others	1	53	111	165	78.26	64.81	9.13
Independent	16	7	5	28	122.11	59.88	0.0
Total	39	393	336	768	93.03	63.81	11.45

Data source: STR LLC. "Upscale," "Midscale," and "Downscale" classifications are based on a coarsening of STR classification of hotel segments. For instance, "Downscale" refers to STR "midscale" and "economy" segments, while "Midscale" refers to STR "upper midscale" and "midscale" STR segments. Mean ADR and Occupancy are calculated as the average across all hotel-nights in the dataset. For brands whose royalty fee is levied as a percentage of total revenues, the marginal royalty is equal to the percent royalty reported by HVS. For brands whose royalty fee is levied as a fixed rate (e.g., dollars per room per month), HVS reports royalties as an expected percentage of annual revenues; I record the *marginal* royalty as 0%.

Table 3: Big and small chain hotels' responsiveness

	Percentile		
Measure	25th	50th	75th
Rank Correlation of (Price, Occ)			
Big	0.40	0.52	0.61
Small	0.36	0.49	0.60
Responsiveness (Slope of Price w.r.t. Occ)			
Big	1.47	2.20	3.61
Small	1.76	2.63	3.93
Responsiveness Ratio (x100)			
Big	0.28	0.38	0.49
Small	0.28	0.39	0.52

Data source: STR LLC. Rank correlation is Spearman correlation between each hotel's price and occupancy. Responsiveness is β from a regression of the form $\text{ADR}_{jt} = \beta_j \text{Occ}_{jt} + \lambda_j + \epsilon_{jt}$. Responsiveness ratio is the ratio of the slope of the black line to the slope of the red line in Figure 2, multiplied by 100. Larger ratios imply better information. All measures calculated at the hotel level using variation over time; percentiles are taken over the sample of hotels.

Table 4: Demand estimates

Demand variable	Estimated coefficients (Std. Error)	
	Weekends	Weekdays
Price sensitivity, α	-1.332	-1.911
	(.069)	(.074)
Nesting parameter, γ	.353	.358
	(.009)	(.007)
	Variance	
	Across All Obs	Within Hotel
Common-knowledge shifters, $x_t + \lambda_j$	1.025	.484
Unknown common shifter, λ_t	.338	.321
Hotel-specific shifter, ξ_{jt}		.447

Data source: STR LLC. Estimates from nested logit. A market is a city-night; hotels classified based on a coarsening of the official STR classification. GMM standard errors reported. Within-hotel variances are reported as the average within-hotel variances across all hotels in my sample.

Table 5: Supply estimates

	Mean	Std. Dev.	Percentile 10th	Percentile 50th	Percentile 90th
Info quality, $\frac{1}{1+\rho}$	0.66	0.30	0.21	0.70	0.99
Shannon info, $H(\rho)$	1.70	1.68	0.03	0.94	4.49
Delta coefficient, z_{2j}	1.73	3.38	0.19	0.47	4.65
Price coefficient, z_{1j}	0.45	0.40	0.00	0.37	1.00

Data source: STR LLC. A different estimate of each parameter is calculated for each hotel-1[weekday]. Percentiles, mean and standard deviation are taken across all of these estimates. Shannon information calculated according to Equation 11.

Table 6: Information quality differences across parent companies

	(1) Shannons	(2) Shannons	(3) Shannons	(4) Shannons	(5) Shannons
Large Chain	−0.169		−0.057		−0.026
	(0.109)		(0.081)		(0.155)
Parent Tot. Rooms		−0.016		−0.005	
		(0.010)		(0.008)	
Downscale			0.436	0.439	0.437
			(0.244)	(0.246)	(0.243)
Parent Mkt. Share					−0.253
					(1.214)
Observations	1,444	1,444	1,444	1444	1,444
Adjusted R^2	0.044	0.044	0.063	0.063	0.062

Data source: STR LLC. An observation is a hotel*1[weekday]. Standard errors in parentheses are only for the final stage. They do not reflect the fact that information quality is itself estimated. The standard errors are clustered at the Market-Quality Segment-1[Weekday] level. All specifications include market-wide fixed effects. Parent total rooms are the sum of room counts across all hotels affiliated with a particular parent company in my dataset. This is measured in 1000s of rooms.

Table 7: Marginal cost regressions

	(1) Info Cost	(2) Info Cost	(3) Info Cost	(4) Info Cost
Large Chain	9731		7691	
	(7435)		(7104)	
Parent Tot. Rooms		1.302		1.112
		(0.887)		(0.877)
Downscale			-7935	-7807
			(10342)	(10409)
Observations	1,444	1,444	1,444	1,444
Adjusted R^2	0.154	0.154	0.154	0.154

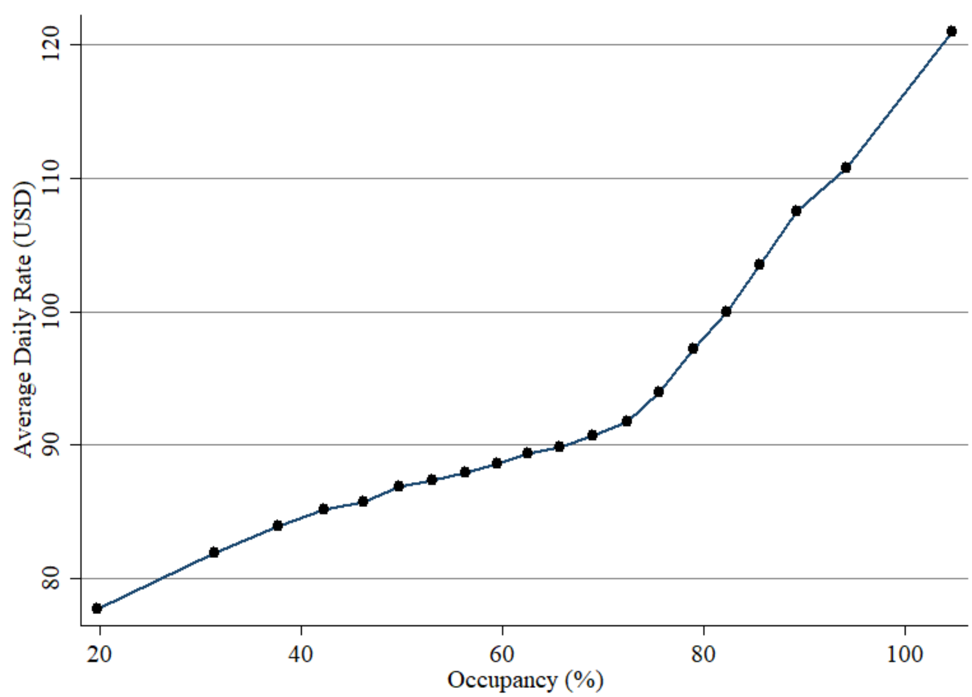
Data source: STR LLC. Standard errors reflect only error from final-stage regression and not that information quality, as well as other model parameters are estimated, and that counterfactual information quality relies on simulated revenues. Final-stage standard errors are clustered at the Market-Class-1[Weekday] level. Informational costs in dollars. All specifications include market-wide fixed effects. Parent total rooms are the sum of room counts across all hotels in my dataset affiliated with a particular parent company. This figure is reported in 1000s.

Table 8: Counterfactual information gathering

	(1) Shannons	(2) Shanons (CF)	(3) Shannons	(4) Shannons (CF)
Large Chain	-0.169	0.195	-0.057	0.196
	(0.109)	(0.163)	(0.081)	(0.153)
Downscale			0.436	0.004
			(0.244)	(0.234)
Observations	1,444	1,444	1,444	1,444
Adjusted R^2	0.044	0.053	0.063	0.052

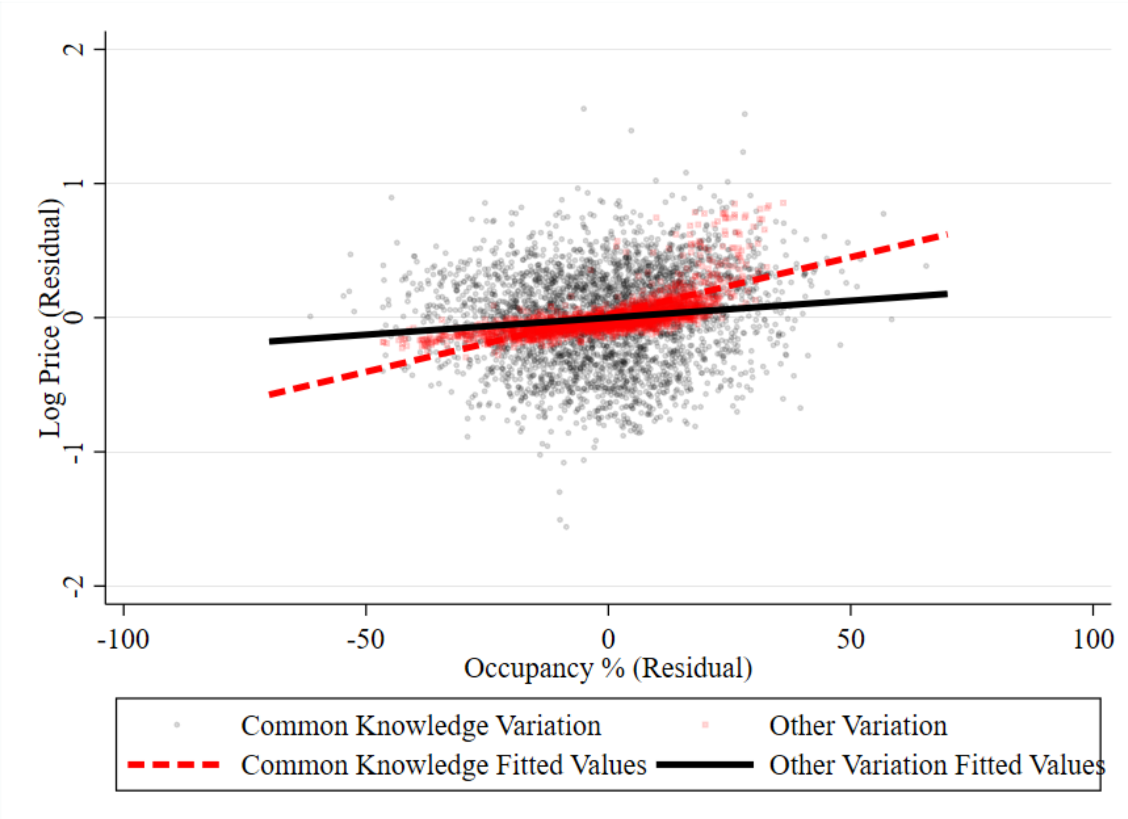
Data source: STR LLC. Standard errors reflect only error from final-stage regression and not that information quality, as well as other model parameters are estimated, and that counterfactual information quality relies on simulated revenues. All specification include market-wide fixed effects. Final-stage standard errors are clustered at the Market-Class-1[Weekday] level

Figure 1: Binned scatterplot, price vs. occupancy



Data source: STR LLC. Binscatter plot of average daily rate on occupancy, after absorbing hotel-level fixed effects.

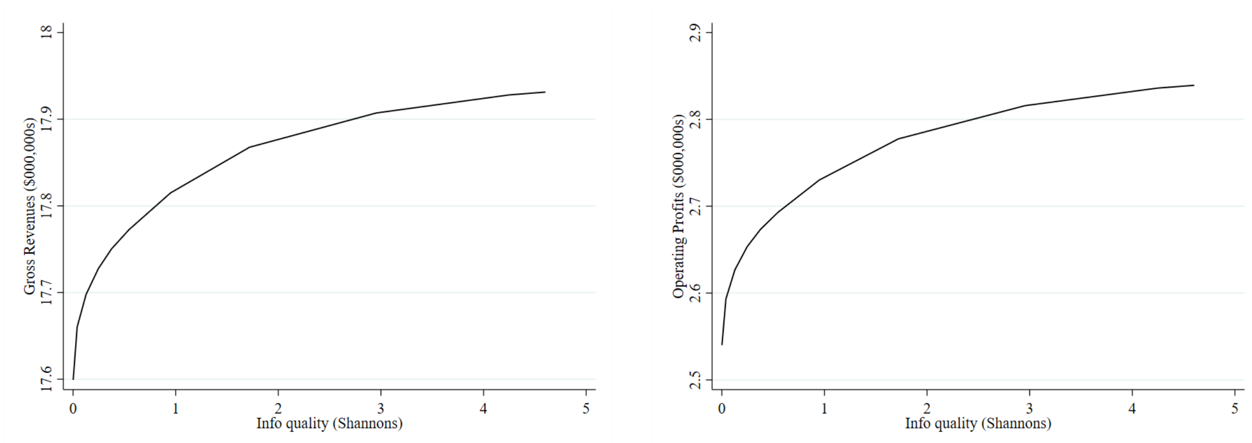
Figure 2: Descriptive evidence of imperfect information



Data source: STR LLC. Common knowledge variation is calculated from fitted values of regressions of price and average daily rate on common-knowledge demand shifters. Unpredictable variation is calculated from the residuals from these same regressions. Each red dot represents a different "bin" of these common-knowledge demand variables, as these variables are discrete. Because the residuals are not discrete, I plot only a random subsample of these residuals. The fitted lines for common knowledge variation are weighted to reflect the fact that some bins, and thus points on the graph, are affiliated with more observations and thus receive larger weight.

Figure 3: Gains from information quality

(a) Average annual revenues (b) Calibrated operating profits



Data source: STR LLC.

A Appendix: Proofs

I begin by proving Theorem 1, restated here:

Theorem 1. *Under **Assumptions 1-KD–7**, the supply model, $\{z_{1j}, z_{2j}, \sigma_j\}_{j\in\mathcal{J}}$ and F, is identified from the joint distribution of prices $[p_{jt}]_{j\in\mathcal{J}}$ and demand shifters $x_t, \lambda_t, [\xi_{jt}]_{j\in\mathcal{J}}$.*

Proof: First, I define some objects. Let $\Xi \equiv [p_t, \delta_t, x_t, \lambda_t]$ and define $G(p_t, \delta_t, x_t, \lambda_t)$ be the joint distribution of all of these variables. Because $\delta_j(x_t, \lambda_t, \xi_{jt})$ is known to the econometrician, all of the variables components of Ξ are as good as observed, so G is identified. Abusing notation slightly, $G(.|.)$ be the conditional distribution of some of these variables conditional on others. Finally, define $\hat{p}_{jt}(.) = \int p_{jt}\ dG(p_{jt}|.)$ be the expectation of j's prices conditional on some of the observed demand shifters. Note that, because G is identified, $\hat{p}_{jt}(.)$ is identified from the empirical conditional mean of j's prices.

The rest of the proof consists of three lemmata. First, I prove that z_{1j} is identified from variation in prices as x_t varies.

Lemma 1. *z_{1j} is identified, for each j, from the equations*

$$z_{1j} = (J-1)\frac{\frac{\partial \hat{p}_{jt}(x_t,\delta_{jt})}{\partial x_t}}{\sum_{j'\neq j}\frac{\partial \hat{p}_{j't}(x_t,\delta_{jt})}{\partial x_t}},\ \forall j.$$

Proof of Lemma 1: Examine

$$\hat{p}_{jt}(\delta_{jt}, x_t) = \int\int p_j(x_t, \delta_{jt}, s_{jt}, \nu_{jt})\ dF(s_t|x_t, \delta_t)\ dH(\nu_t).$$

Because ν_{jt} is i.i.d., additively separable, and mean zero, this becomes

$$\hat{p}_{jt}(\delta_{jt}, x_t) = \int p_j(x_t, \delta_{jt}, s_{jt}, 0)\ dF(s_t|x_t, \delta_t).$$

Using independence of signals from the demand shifters other than λ, this becomes

$$\hat{p}_{jt}(\delta_{jt}, x_t) = \int p_j(x_t, \delta_{jt}, s_{jt}, 0)\ dF(s_t).$$

Now examine changes in prices as x_t changes by differentiating with respect to x_t:

$$\frac{\partial \hat{p}_{jt}(\delta_{jt}, x_t)}{\partial x_t} = \int \frac{\partial p_j(x_t, \delta_{jt}, s_{jt}, 0)}{\partial x_t}\ dF(s_t)$$

Differentiating the policy function yields $\frac{\partial p_j}{\partial x_t}\Big|_{\delta,s} = z_{1j}\frac{\partial \mathbb{E}(\bar{p}_{-j}|x_t,s_{jt},\delta_{jt})}{\partial x_t}$. This means

$$\frac{\partial \hat{p}_{jt}(\delta_{jt},x_t)}{\partial x_t} = \int z_{1j}\frac{\partial \mathbb{E}(\bar{p}_{-j}|x_t,s_{jt},\delta_{jt})}{\partial x_t}\ dF(s_t).$$

Linearity of the expectation, integration, and differentiation operators allow us to rewrite the above as

$$\frac{\partial \hat{p}_{jt}(\delta_{jt},x_t)}{\partial x_t} = z_{1j}\int \mathbb{E}\Big(\frac{\partial \bar{p}_{-j}}{\partial x_t}|x_t,\delta_{jt},s_{jt}\Big)\ dF(s_{jt})$$

The term on the right is equal to $z_{1j}\mathbb{E}\bigg(\mathbb{E}\Big(\frac{\partial \bar{p}_{-j}}{\partial x_t}|x_t,\delta_{jt},s_{jt}\Big)|x_t,\delta_{jt}\bigg)$. By the law of total probability, this is equal to $z_{1j}\mathbb{E}\Big(\frac{\partial \bar{p}_{-j}}{\partial x_t}|x_t,\delta_{jt}\Big)$, yielding

$$\frac{\partial \hat{p}_{jt}(\delta_{jt},x_t)}{\partial x_t} = z_{1j}\mathbb{E}\Big(\frac{\partial \bar{p}_{-j}}{\partial x_t}|x_t,\delta_{jt}\Big)$$

Because the econometrician observes x_t and δ_{jt} for all j, the only objects on the right hand side of this equation that are unknown are z_{1j}. The left hand side is also known, given the data, for the same reason that x_t and δ_{jt} are observed. Because $\bar{p}_{-j} = \frac{1}{J-1}\sum_{j'\neq J} p_{j'}$, we can rewrite this equation as

$$\frac{\partial \hat{p}_{jt}(\delta_{jt},x_t}{\partial x_t} = z_{1j}\mathbb{E}\Big(\frac{1}{J-1}\sum_{j'\neq j}\frac{\partial p_{j't}}{\partial x_t}\Big|x_t,\delta_{jt}\Big)$$

This yields a solution:

$$z_{1j} = (J-1)\frac{\frac{\partial \hat{p}_{jt}(x_t,\delta_{jt})}{\partial x_t}}{\sum_{j'\neq j}\frac{\partial \hat{p}_{j't}(x_t,\delta_{jt})}{\partial x_t}},\ \forall j$$

All of the objects on the right hand side can be calculated (as changes in expectations of prices) from the data. □

The numerator of this equation is the change in j's prices, on average, when x_t changes but δ_{jt} does not change. This happens if x_t increases, but ξ_{jt} decreases such that δ_{jt} remains the same. The denominator is the change in rivals' prices, on average, when the same change occurs. By keeping δ_{jt} constant, I isolate the effects of increases in market-wide demand through the channel of rivals' prices only. I integrate over λ_t to form these conditions for a more subtle reason: if I were to write this equation conditional on λ_t instead, because firm j may have incorrect beliefs about λ_t, its response to changes in its beliefs over rivals' prices may diverge from its response to changes in rivals' actual prices.

Lemma 2. *If z_{1j} is known for all j, the signal structure F and σ_j are identified from the joint distribution of p_j and $\bar{p}_{-j}$.*

Proof of Lemma 2: Consider the joint distribution of p_j and the average price of j's rivals, $\bar{p}_{-jt} = \sum_{j' \neq j} p_{jt}$, conditional on x_t and δ_j: $G(p_{jt}, \bar{p}_{-jt}|x_t, \delta_{jt})$. Because all prices are observed, this distribution is identified. Furthermore, because we observe λ_t, the expectation of $\bar{p}_{-jt}$ conditional on λ_t, $\hat{p}_{-jt}(\lambda_t)$, is identified, so the joint distribution of p_{jt} and $\hat{p}_{-jt}(\lambda_t)$ is also identified. Call this $G(p_{jt}, \hat{p}_{-jt}(\lambda_t)|\delta_{jt}, x_t)$.

Suppose we guessed $\sigma_j = \hat{\sigma}_j$. Given knowledge of G, and this guess, we can solve the following functional equation to obtain the distribution of j's prices conditional on $\nu_{jt} = 0$, which we will call $\tilde{G}_{\hat{\sigma}_j}$:

$$G(p_{jt}, \hat{p}_{-jt}(\lambda_t)|\delta_{jt}, x_t) = \int \tilde{G}_{\hat{\sigma}_{jt}}(p_{jt} - \nu_{jt}, \hat{p}_{-jt}(\lambda_t)|\delta_{jt}, x_t)\ h_{\hat{\sigma}_j}(\nu_{jt})\ d\nu_{jt}.$$

This is a standard deconvolution of $h_{\hat{\sigma}_j}$, which is the pdf of a normal distribution with mean zero and variance $\hat{\sigma}_j$, from G. Now, for a joint distribution of two variables, $G(X, Y)$, define the following operator $\text{cop}(G, .) : \Delta\mathbb{R}^2 \times 2^{\{X,Y\}}$ that normalizes some dimensions of this joint distribution to be uniform on $[0, 1]$. For instance:

$$F(X, Y) = \text{cop}(G, X) \ \Leftrightarrow \ F(G(X), Y) = G(X, Y)$$

and

$$F(X, Y) = \text{cop}(G, \{X, Y\}) \ \Leftrightarrow \ F(G(X), G(Y)) = G(X, Y).$$

Note that if both random variables are included as arguments in $\text{cop}(G, .)$, we obtain the copula of G. Using this notation, the assumption that s_j is independent of δ_j and x, and the normalization that the marginal distribution of signals is uniform[9], we have

$$F(s_{jt}, \lambda_t) = \text{cop}(G(p_j, \hat{p}_{-jt}(\lambda_t)), p_j).$$

To see this, note that the monotonicity assumption implies that p_j is increasing in s_{jt} and $\hat{p}_{-jt}(\lambda_t)$ is increasing in λ_t. As a result, the probability that p_{jt} is below some $\hat{p}_{jt}$ and $\hat{p}_{-jt}(\lambda_t)$ is below some $\tilde{p}_{-jt}$ is the same as the probability that s_{jt} and λ_t are below some corresponding thresholds $\bar{s}_j(\hat{p}; x_t, \delta_j)$ and $\bar{\lambda}(\tilde{p}; x_t, \delta_j)$, respectively. This means that we can compute the joint distribution of

[9]This is by the integral probability transform.

j's signal and λ that is implied by a particular $\tilde{G}_{\hat{\sigma}_j}$:

$$\tilde{F}_{\hat{\sigma}}(s_{jt}, \lambda_t) = \text{cop}(\tilde{G}_{\hat{\sigma}_j}(p_{jt}, \hat{p}_{-jt}(\lambda_t), p_{jt}).$$

Now, note that the sum of two independent normally distributed random variables with mean zero is also a normal with mean zero. As a result, a deconvolution corresponding to $\hat{\sigma}_j$ from G to obtain $\tilde{G}_{\hat{\sigma}_j}$ can also be represented as the outcome of two successive deconvolutions of $\hat{\sigma}_j'$ and $\hat{\sigma}_j''$ noise from G, where $\hat{\sigma}_j^2 = (\hat{\sigma}_j')^2 + (\hat{\sigma}_j'')^2$. It follows that, if $\hat{\sigma}_j > \hat{\sigma}_j'$, $\tilde{G}_{\hat{\sigma}_j}$ dominates $\tilde{G}_{\hat{\sigma}_j'}$ in the Blackwell order.[10] Furthermore, because the Blackwell order is preserved under rescaling of the signal, $\tilde{F}_{\hat{\sigma}_j}$ dominates $\tilde{F}_{\hat{\sigma}_j'}$ in the Blackwell order.

Vizcaíno and Mekonnen [2019] show that, if signals improve for j in the Blackwell order, under the monotonicity conditions I impose, j's actions will become more "responsive." Their definition of responsiveness implies that the variance of p_{jt}, conditional on ν_{jt}, δ_{jt} and x_t, will increase. This is easy to see: conditional on these three variables, the only source of variation left in p_{jt} is variation in its beliefs about $\bar{p}_{-jt}$ that are induced by variation in s_{jt}. Specifically, using the policy function,

$$\text{var}(p_{jt}|\nu_{jt}, \delta_{jt}, x_t) = z_{1j}^2 \text{var}(\mathbb{E}(\hat{p}_{-jt}(\lambda_t, x_t)|s_{jt}, x_t, \delta_{jt})|\delta_{jt}, x_t). \tag{A.1}$$

The right hand side of this equation is z_{1j}^2 multiplied by the variance, across signals, of j's posterior expectation of rivals' prices. If j's signal structure $F(s_{jt}|\lambda_t)$ is completely uninformative, its posterior will be the same as its prior, and its posterior expectation over rivals' prices will not vary at all with the signal. If j's signal improves in the Blackwell order, j's beliefs will vary more with s_{jt}. This implies that the variance on the left hand side of this equation will increase, given knowledge of of z_{1j}. For instance, if s_{jt} is completely uninformative, beliefs will not vary at all with s_{jt}, so j's prices will not vary at all conditional on ν_{jt}, x_t, and δ_{jt}. For a guess of $\hat{\sigma}_j$, we can express $\text{var}_{\hat{\sigma_j}}(p_{jt}|\nu_{jt}, \delta_{jt}, x_t)$ using Equation A.1, given we also know $F_{\hat{\sigma}_j}$, z_{1j}, and $\mathbb{E}(\bar{p}_{-jt}|\lambda_t)$. Formally,

$$\begin{aligned}
\text{var}_{\hat{\sigma}_j}(p_{jt}|\nu_{jt}, \delta_{jt}, x_t) = \quad & z_{1j}^2 \text{var}_s(\mathbb{E}(\hat{p}_{-jt}(\lambda_t, x_t)|s_{jt}, x_t, \delta_{jt})|\delta_{jt}, x_t) \\
= \quad & z_{1j}^2 \text{var}_s\Big(\int \mathbb{E}(\hat{p}_{-jt}(x_t, \delta_{jt})|x_t, \delta_{jt}, \lambda_t)\ dF_{\hat{\sigma}_j}(\lambda|s_{jt})\Big)
\end{aligned}$$

where the subscript in var_s emphasizes that this variance is over signals s_{jt}. Because ν_{jt} is i.i.d., the variance of p_{jt} conditional on only δ_{jt} and x_t is will equal $\text{var}(p_{jt}|\nu_{jt}, \delta_{jt}, x_t) + \sigma_{jt}$. So, for our guess

[10]If a distribution $G_1(x|y)$ is a convolution of another distribution $G_2(x|y)$, i.e., G_1 is a convolution of G_2 with some non-degenerate distribution with mean zero, G_2 dominates G_1 in the Blackwell order.

of $\hat{\sigma}_j$, our "model-predicted" variance is

$$\overline{\text{var}}_{\hat{\sigma}_j}(p_{jt}|\delta_{jt}, x_t) = \text{var}_{\hat{\sigma}_j}(p_{jt}|\nu_{jt}, \delta_{jt}, x_t; \bar{p}_{-jt}) + \hat{\sigma}_j^2$$

This implies the following moment condition matching "model-predicted" variance to actual variance of prices for hotel j:

$$M_j(\hat{\sigma}_j) \equiv \text{var}_{\hat{\sigma}_j}(p_{jt}|\nu_{jt}, \delta_{jt}, x_t) + \hat{\sigma}_{jt} = \widehat{\text{var}}(p_{jt}|x_t, \delta_{jt}) \tag{A.2}$$

where $\widehat{\text{var}}(\hat{p}_{jt}(x_t, \delta_{jt}) = \int (p_{jt} - \mathbb{E}(p_{jt}|x_t, \delta_{jt}))^2 dG(p_{jt}|x_t, \delta_{jt})$ is the empirical variance of p_{jt} conditional on x_t and δ_{jt}. The left hand side of the above equation consists of two terms, both of which are increasing in $\hat{\sigma}_j$. We can formulate equivalent moment conditions for all of j's rivals, forming J moment conditions in total. Each condition M_j is strictly increasing in $\hat{\sigma}_j$ and invariant to changes in $\hat{\sigma}_{j'}$, $j' \neq j$. To see this, note that changes $\hat{\sigma}_j'$ change j's actions only through a strategic effect: the mapping from λ_t to $\bar{p}$ may change if j' becomes more informed about λ_t. However, because the left hand side of M_j is formed using the empirically observed $\hat{p}_{-jt}(x_t, \lambda_t)$, changes to j''s behavior have no bearing on the j's distribution over prices.

Because $M_j(\hat{\sigma}_j)$ is strictly increasing, it will only equal the right hand side of Equation A.2, which can be computed directly from the data and does not depend on any parameters, at the true $\hat{\sigma}_j$. Furthermore, because each moment M_j only depends on its own $\hat{\sigma}_j$, the Jacobian of the system of moments $\big[M_j\big]_{j \in \mathcal{J}}\big(\sigma_j\big)_{j \in \mathcal{J}}$ is diagonal, with every diagonal entry positive. This implies that there is a unique solution to this system of equations, equal to the true $\hat{\sigma}$. Thus, $\hat{\sigma}_j$ is identified for each j.

Given knowledge of $\hat{\sigma}$, the full signal structure $F(s_{1t}, s_{2t}, \ldots, s_{Jt}, \lambda_t)$ can be computed from $G(p_{1t}, p_{2t}, \ldots, p_{Jt}, \lambda_t)$ through the following two steps, which correspond to the steps taken to recover partial signals for each j in constructing M_j:

1. For each j, deconvolve $\nu_{jt} \sim \mathcal{N}(0, \sigma_j)$ from the jth dimension of G using the true σ_j. Call the resulting object $\tilde{G}(p_{1t}, p_{2t}, \ldots, p_{Jt}, \lambda_t)$.

2. Compute $F(s_{1t}, s_{2t}, \ldots, s_{Jt}, \lambda_t) = \text{cop}(\tilde{G}, \{s_{1t}, s_{2t}, \ldots, s_{Jt}\})$.

This recovers the signal structure F. □

Note that we have nonparametrically recovered a $J+1$-dimensional joint distribution F *and* J parameters $(\sigma_j)_{j \in \mathcal{J}}$ from a $J+1$-dimensional joint distribution of j prices and λ_t. How did we manage to do this? The key is the (wlog) normalization of signals so that the marginal distribution

$F(s_{jt})$ is uniform. This provides J restrictions on F that are not present in G, which is just enough information to recover the J parameters σ_j in addition to the signal structure.

The final lemma recovers z_{2j}.

Lemma 3. *$z_{2j}(.)$ is nonparametrically identified from the joint distribution of prices conditional on δ_j, given knowledge of z_{1j}.*

Proof of Lemma 3: Examine the expectation of j's prices, conditional on x_t and δ_{jt}:

$$\begin{aligned} \hat{p}(x_t, \delta_{jt}) = & \quad \mathbb{E}_{s_{jt}}(z_{1j}\bar{p}_{-jt}|s_{jt}, x_t, \delta_{jt}) + z_{2j}(\delta_{jt}) + \mathbb{E}(\nu_{jt}|x_t, \delta_{jt}) \\ = & \quad z_{1j}\mathbb{E}_{s_{jt}}(\mathbb{E}(\bar{p}_{-jt}|s_{jt}, x_t, \delta_{jt})|x_t, \delta_{jt}) + z_{2j}(\delta_{jt}). \end{aligned}$$

By the law of total probability,

$$\begin{aligned} \hat{p}(x_t, \delta_{jt}) = & \quad z_{1j}E(\bar{p}_{-j}|x_t, \delta_{jt}) + z_{2j}(\delta_{jt}) \\ = & \quad \hat{p}_{-jt}(x_t, \delta_{jt}) + z_{2j}(\delta_{jt}) \end{aligned}$$

which yields

$$z_{2j}(\delta_j) = \hat{p}_j(x_t, \delta_{jt}) - z_{1j}E(\bar{p}_{-jt}|x_t, \delta_{jt}).$$

Both terms on the right hand side of this equation can be computed from the data, given knowledge of z_{1j}. As a result, for any δ_{jt}, the left hand side is identified. This means that $z_{2j}(.)$ is identified. $\square$

Together, these three lemmata prove that the model is identified. $\square$

I restate Theorem 2 here:

Theorem 2. *Under* **Assumptions 1-UD–7**, *the supply model $\{z_{1j}, z_{2j}, \sigma_j\}_{j\in J_t}$ and F, is generically identified from the joint distribution of prices $[p_{jt}]_{j\in\mathcal{J}}$ and demand shifters $x_t, \lambda_t, [\xi_{jt}]_{j\in\mathcal{J}}$.*

The proof follows from the following four lemmata. The first Lemma shows that $z_{2j}(.)$ is generically identified up to an additive constant.

Lemma 4. *Let $z_{2j}(.) \equiv \hat{z}_{2j}(.) + \tilde{z}_{2j}$, where $\tilde{z}_{2j} = \mathbb{E}(z_{2j}(\delta_{jt})|\xi_{jt} = 0)$. Then $\tilde{z}_{2j}$ is generically nonparametrically identified, up to an additive constant $\tilde{z}_{2j}$, from covariation in p_j and ξ_j, for each $j \in \mathcal{J}$.*

Proof of Lemma 4: First note that, because ξ_{jt} is independent of the other demand shifters, rivals' average price, and j's beliefs over rivals' average price, do not depend on ξ_{jt}. The pricing equation is thus

$$p_{jt} = z_{1j}\mathbb{E}(\bar{p}_{-jt}|x_{jt}, s_{jt}) + \mathbb{E}(z_{2j}(\delta_j(x_t, \lambda_t, \xi_{jt}))|x_t, \xi_{jt}, s_{jt}) + \nu_{jt}.$$

Examine the expectation of j's prices, conditional on ξ_{jt}, and use the law of iterated expectations:

$$\begin{aligned}\hat{p}_{jt}(\xi_{jt}) = & \ \mathbb{E}_{x,s}\big[z_{1j}\mathbb{E}(\bar{p}_{-jt}|x_t, s_{jt}) + \mathbb{E}(z_{2j}(\delta_j(x_t, \lambda_t, \xi_{jt}))|x_t\xi_{jt}, s_{jt}\big) + \nu_{jt}|\xi_{jt}] \\ = & \ z_{1j}\mathbb{E}(\hat{p}_{-jt}) + \mathbb{E}\big(z_{2j}(\delta_j(x_t, \lambda_t, \xi_{jt})\big|\xi_{jt})\big)\end{aligned}$$

where $\hat{p}_{-jt}$ is the unconditional mean of j's rivals' prices, across all possible realizations of demand. Using our location normalization, $z_{2j}(.) \equiv \tilde{z}_{2j} + \hat{z}_{2j}(.)$, with $\mathbb{E}\big(\hat{z}_{2j}(\delta_j(x_t, \lambda_t, \xi_{jt})\big|\xi_{jt} = 0)\big) = 0$. We then have

$$\hat{p}_{jt}(\xi_{jt}) - \hat{p}_{jt}(0) = \mathbb{E}\big(\hat{z}_{2j}(\delta_j(x_t, \lambda_t, \xi_{jt})|\xi_{jt})\big). \tag{A.3}$$

Writing $\delta_{jt} = \delta_j(x_t, \lambda_t, \xi_{jt})$, this can be rewritten as

$$\begin{aligned}\hat{p}_{jt}(\xi_{jt}) - \hat{p}_{jt}(0) = & \ \mathbb{E}\big(\hat{z}_{2j}(\delta_{jt})|\xi_{jt}\big) \\ = & \int \hat{z}_{2j}(\delta_{jt})\ g(\delta_{jt}|\xi_{jt})\ d\delta_{jt}.\end{aligned}$$

The left hand side of this equation is identified by the conditional mean of j's prices. On the right hand side, because δ_{jt} is known $g(.)$ is identified from the empirical distribution of δ_{jt} conditional on ξ_{jt}. The only item that is unknown is the function $\hat{z}_{2j}(.)$. This equation is a Fredholm equation of the first kind, formally similar to those seen in the nonparametric IV literature [Newey and Powell, 2003]. Identification of $\hat{z}_{2j}(.)$ rests on whether the inverse of this equation exists and is unique, which is true if $g(\delta_{jt}|\xi_{jt})$ satisfies a *completeness* condition. Roughly speaking, this condition imposes that $g(\delta_{jt}|\xi_{jt})$ cannot be expressed as a weighted integral of other $g(\delta_{jt}|\xi'_{jt})$ and thus is an infinite-dimensional analogue to a rank condition for matrices. Under the assumption that $\mathbb{E}\big(\hat{z}_{2j}(\delta_{jt})^2\big|\xi_{jt})$ exists for all ξ_{jt}, Andrews [2011] establishes that completeness of $g(\delta_{jt}|\xi_{jt})$ is generic.[11]

This establishes that, under the assumption that $\mathbb{E}\big(z(\delta_{jt}|\xi_{jt} = 0)\big) = 0$, identification is generic. That is, $z_{2j}(.)$ is generically identified up to some location normalization, but this normaliation does not come without loss of generality. $\square$

Lemma 5. *If $\hat{z}_{2j}(.)$ is identified, z_{1j} is identified for each $j \in \mathcal{J}$.*

[11] Genericity is poorly defined for functions, because functional spaces are infinite-dimensional and there exists no natural extension of the Lebesgue measure to infinite-dimensional spaces. Andrews [2011] instead use *prevalence* as defined in Anderson and Zame [2001]. This is the precise notion of genericity I use here.

Proof of Lemma 5 The steps are largely similar to those of Lemma 1. Examine the expectation of p_{jt} conditional on x_t, again invoking the law of total probability to do so:

$$\hat{p}_{jt}(x_t) = z_{1j}\mathbb{E}(\bar{p}_{-jt}|x_t) + \mathbb{E}\big(\hat{z}_{2j}(\delta_{jt}) + \tilde{z}_{2j}|x_t\big).$$

Differentiate with respect to x_t to obtain

$$\begin{aligned}\frac{\partial \hat{p}_j(x_t)}{\partial x_t} &= z_{1j}\frac{\partial}{\partial x_t}\mathbb{E}(\bar{p}_{-jt}|x_t) + \frac{\partial}{\partial x_t}\mathbb{E}\big(\hat{z}_{2j}(\delta_{jt}) + \tilde{z}_{2j}|x_t\big) \\ &= z_{1j}\frac{\partial}{\partial x_t}\mathbb{E}(\bar{p}_{-jt}|x_t) + \frac{\partial}{\partial x_t}\mathbb{E}\big(\hat{z}_{2j}(\delta_{jt})|x_t\big).\end{aligned}$$

Rearrange and use the definition that $\bar{p}_{-jt} = \frac{1}{J-1}\sum_{j'\neq j} p_{j't}$ to obtain

$$z_{1j} = (J-1)\frac{\frac{\partial \hat{p}_j(x_t)}{\partial x_t} - \frac{\partial}{\partial x_t}\mathbb{E}\big(\hat{z}_{2j}(\delta_{jt})|x_t\big)}{\frac{\partial}{\partial x_t}\sum_{j'\neq j}\hat{p}_{j'}(x_t)}$$

where $\hat{p}_{j'}(x_t)$ is the expectation of a rival's prices conditional on x_t. All objects on the right hand side of this equation are identified given Lemma 4 and the joint distribution of prices and x_t. Thus, z_{1j} is identified. □

Lemma 6. *If z_{1j} and $\hat{z}_{2j}(.)$ are identified, $z_{2j}(.)$ is identified.*

Proof of Lemma 6: Examine the unconditional mean of j's prices $\hat{p}_j$:

$$\hat{p}_j = z_{1j}\mathbb{E}(\bar{p}_{-jt}) + \mathbb{E}\big(\hat{z}_{2j}(\delta_{jt}) + \tilde{z}_{2j}\big)$$

Rearrange to write

$$\tilde{z}_{2j} = \hat{p}_j - \mathbb{E}\big(z_{1j}\bar{p}_{-jt} - \hat{z}_{2j}(\delta_{jt})\big).$$

All objects on the right hand side are identified, so C_j is identified. $z_{2j}(.) = \hat{z}_{2j}(.) + \tilde{z}_{2j}$ is thus identified.

□

Lemma 7. *If z_{1j} and $z_{2j}(.)$ are identified, signal structure F and noise parameters σ_j, $j \in \mathcal{J}$ are jointly identified.*

Proof of Lemma 7: The proof is essentially the same as for Lemma 3: we match the variance of prices as predicted by the model and a guess of σ_j with the true variance of prices. The only difference is that we now condition on x_t and ξ_{jt}, and thus have to account for variation in $\mathbb{E}\big(z_{2j}(\delta_{jt})|x_t, \xi_{jt}, s_{jt}\big)$ across signals s_{jt} in computing model-predicted price variance. Regardless, after guessing a value for σ_j, we can compute the implied signal structure $F(s_{jt}|\lambda_t)$. Conditional on x_t and ξ_t, and given this guess and knowledge of z_{1j}, $z_{2j}(.)$, and how rivals' prices change with λ_t, we can model-predicted variance of prices for firm j. Because the monotonicity conditions still hold, the responsiveness result of Vizcaíno and Mekonnen [2019] still holds, and the J moments that correspond to Equation A.2, but now conditional on ξ_{jt}:

$$\tilde{\text{var}}_{\hat{\sigma}_j}(p_{jt}|\nu_{jt}, x_t, \xi_{jt}; \bar{p}_{jt}) + \hat{\sigma}_j = \widehat{\text{var}}(p_{jt}|x_t, \xi_{jt})$$

will still hold if and only if $\hat{\sigma}_j$ is correctly specified for all $j \in \mathcal{J}$. $\square$

I reprint the following proposition before proving it:

Proposition 1. *If $z_{2j}(.)$ is a polynomial of order $\tau^* \in \mathbb{N}$, $0 < \tau^* \leq \infty$, for each j, the model is identified if and only if the class of functions $\mathcal{M}^{\tau^*}$ is* complete. That is, if

$$M^{(\bar{\tau})}(\xi) = \sum_{\tau=1}^{\tau^*} \alpha^{(\tau)} M^{(\tau)}(\xi)$$

for some series of constants $\alpha^{(\tau)}$, then $\alpha^{(\bar{\tau})} = 1$ and $\alpha^{(\tau')} = 0$ for all $\tau' \neq \tau$.

Proof: Re-examine the identifying Equation A.3, reprinted below:

$$\hat{p}_{jt}(\xi_{jt}) - \hat{p}_{jt}(0) = \mathbb{E}\big(\hat{z}_{2j}(\delta_{jt})|\xi_{jt})\big).$$

Under this proposition's assumptions on $z_{2j}(.)$, we can rewrite this as

$$\hat{p}_{jt}(\xi_{jt}) - \hat{p}_{jt}(0) = \mathbb{E}\Big(\sum_{\tau=1}^{\tau^*} z_{2j}^{(\tau)} \delta^{\tau} \big| \xi_{jt})\Big)$$

where the superscript (τ) is an index, but the superscript τ without parentheses is an exponent. Using the linearity of the expectation operator, this is the same as

$$\hat{p}_{jt}(\xi_{jt}) - \hat{p}_{jt}(0) = \sum_{\tau=1}^{\tau^*} z_{2j}^{(\tau)} \mathbb{E}(\delta^\tau | \xi_{jt}).$$

Now, suppose that $z_{2j}(.)$ is not identified. In this case, either there exists another candidate for $\hat{z}_{2j}(.)$, $\tilde{z}_{2j}(.) \neq \hat{z}_{2j}(.)$, such that the above equation also holds. This means that:

$$0 = \sum_{\tau=1}^{\tau^*} (z_{2j}^{(\tau)} - \tilde{z}_{2j}^{(\tau)}) \mathbb{E}(\delta^\tau | \xi_{jt})$$

where $z_{2j}^{(\tau)} \neq \tilde{z}_{2j}^{(\tau)}$ for some τ. This can only be true if $\mathbb{E}(\delta^\tau | \xi_{jt}) = \sum_{\tau' \neq \tau} \alpha_{\tau'} \mathbb{E}(\delta_{\tau'} | \xi_{jt})$. This is precisely the (failure of) completeness condition for the first through τ^*-th uncentered moments.

□

References

Philippe Aghion and Jean Tirole. Formal and real authority in organizations. *Journal of Political Economy*, 105(1):1–29, 1997.

Robert M Anderson and William R Zame. Genericity with infinitely many parameters. *Advances in Theoretical Economics*, 1(1), 2001.

Donald WK Andrews. Examples of l2-complete and boundedly-complete distributions. *Cowles Foundation Discussion Paper*, (1801), 2011.

Enghin Atalay, Ali Hortaçsu, and Chad Syverson. Vertical integration and input flows. *American Economic Review*, 104(4):1120–48, 2014.

Patrick Bajari, Victor Chernozhukov, Ali Hortaçsu, and Junichi Suzuki. The impact of big data on firm performance: An empirical investigation. *AEA Papers and Proceedings*, 109:33–37, 2019.

C Lanier Benkard, Aaron Bodoh-Creed, and John Lazarev. Simulating the dynamic effects of horizontal mergers: U.s. airlines. *Mimeo*, 2019.

Steven Berry, James Levinsohn, and Ariel Pakes. Automobile prices in market equilibrium. *Econometrica: Journal of the Econometric Society*, pages 841–890, 1995.

Marianne Bertrand and Sendhil Mullainathan. Enjoying the quiet life? corporate governance and managerial preferences. *Journal of Political Economy*, 111(5):1043–1075, 2003.

Nicholas Bloom and John Van Reenen. Measuring and explaining management practices across firms and countries. *The Quarterly Journal of Economics*, 122(4):1351–1408, 2007.

Nicholas Bloom, Raffaella Sadun, and John Van Reenen. Americans do IT better: US multinationals and the productivity miracle. *American Economic Review*, 102(1):167–201, 2012.

Nicholas Bloom, Raffaella Sadun, and John Van Reenen. Management as a technology? *National Bureau of Economic Research Working Paper no. 22327*, 2016.

Zach Y Brown and Jihye Jeon. Endogenous information and simplifying insurance choice. Technical report, 2020.

R Andrew Butters. Demand volatility, adjustment costs, and productivity: An examination of capacity utilization in hotels and airlines. *American Economic Journal: Microeconomics*, 12(4): 1–44, 2020.

Andrew Caplin and Mark Dean. Revealed preference, rational inattention, and costly information acquisition. *American Economic Review*, 105(7):2183–2203, 2015.

Sungjin Cho, Gong Lee, John Rust, and Mengkai Yu. Optimal dynamic hotel pricing. *Working paper*, 2018.

Sungjin Cho, Gong Lee, John Rust, and Mengkai Yu. Semi-parametric instrument-free demand estimation: relaxing optimality and equilibrium assumptions. *Working paper*, 2019.

Giovanni Compiani, Philip Haile, and Marcelo Sant'Anna. Common values, unobserved heterogeneity, and endogenous entry in us offshore oil lease auctions. *Journal of Political Economy*, 128(10): 3872–3912, 2020.

Chiara Farronato and Andrey Fradkin. The welfare effects of peer entry in the accommodation market: The case of airbnb. *National Bureau of Economic Research Working Paper*, (24631), 2018.

Kenneth Hendricks and Robert H Porter. An empirical study of an auction with asymmetric information. *The American Economic Review*, pages 865–883, 1988.

Kenneth Hendricks, Joris Pinkse, and Robert H Porter. Empirical implications of equilibrium bidding in first-price, symmetric, common value auctions. *The Review of Economic Studies*, 70(1):115–145, 2003.

Brett Hollenbeck. The economic advantages of chain organization. *The RAND Journal of Economics*, 48(4):1103–1135, 2017.

Yufeng Huang. Seller-pricing frictions and platform remedies. *Mimeo*, 2021.

Thomas N Hubbard. Information, decisions, and productivity: On-board computers and capacity utilization in trucking. *American Economic Review*, 93(4):1328–1353, 2003.

Thomas N Hubbard and Michael J Mazzeo. When demand increases cause shakeouts. *American Economic Journal: Microeconomics*, 2019.

Arturs Kalnins, Luke Froeb, and Steven Tschantz. Can mergers increase output? evidence from the lodging industry. *The RAND Journal of Economics*, 48(1):178–202, 2017.

Jay Lu. Random choice and private information. *Econometrica*, 84(6):1983–2027, 2016.

Michael J Mazzeo. Product choice and oligopoly market structure. *RAND Journal of Economics*, pages 221–242, 2002.

Keaton S Miller, Amil Petrin, Robert Town, and Michael Chernew. Optimal managed competition subsidies. *National Bureau of Economic Research Working Paper*, (25616), 2019.

Whitney K Newey and James L Powell. Instrumental variable estimation of nonparametric models. *Econometrica*, 71(5):1565–1578, 2003.

Angela Roper. *Vertical Disintegration in the Corporate Hotel Industry: The End of Business as Usual*. Routledge, 2017.

Chad Syverson. Market structure and productivity: A concrete example. *Journal of Political Economy*, 112(6):1181–1222, 2004.

Chad Syverson. What determines productivity? *Journal of Economic Literature*, 49(2):326–65, 2011.

René Leal Vizcaíno and Teddy Mekonnen. Bayesian comparative statics. 2019.

Jeffrey M Wooldridge. *Econometric analysis of cross section and panel data*. MIT press, 2010.