

Costs of Managerial Attention and Activity as a Source of Sticky Prices: Structural Estimates from an Online Market

Sara Fisher Ellison
M.I.T. and CESifo

Christopher M. Snyder
Dartmouth College and N.B.E.R.

Hongkai Zhang
M.I.T.

March 2015

Abstract: We study the pricing decisions of a set of rival firms selling memory modules over the internet. We have fine-grained data—hourly data on their prices for about a year—which will allow us to estimate a dynamic structural model of competition and back out parameters, such as managerial attention cost, that inform the micro-foundations of the price inertia we observe. We couple this analysis with manager interviews and some supporting statistical evidence, shedding light on the daily process of price-setting that they implement. We estimate substantial managerial costs associated with price-setting and substantial heterogeneity in costs across firm types. Our estimation is in the tradition of structural IO, but our results complement the large and growing empirical macroeconomic literature on price-setting.

JEL Codes: L11, C73, D21, L81

Contact Information: Ellison: Department of Economics, M.I.T., 50 Memorial Drive, Cambridge, MA 02142; email sellison@mit.edu. Snyder: Department of Economics, Dartmouth College, 301 Rockefeller Hall, Hanover, NH 03755; email chris.snyder@dartmouth.edu. Zhang: Department of Economics, M.I.T., 50 Memorial Drive, Cambridge, MA 02142; email zhanghk@mit.edu.

Acknowledgments: The authors are grateful to Michael Baye, Glenn Ellison, Whitney Newey, and Peter Klenow for their insightful advice and to seminar participants at M.I.T., Xiamen University, and University of Bologna for helpful comments.

1. Introduction

One of the most important decisions firms make is how to price their products. It is not a surprise, then, that many fields of economics have studied firms' pricing decisions intensively. Dating back to Bertrand and Cournot, game theorists have formulated models of optimal pricing in static games, and, more recently, dynamic and repeated games. Building on the theoretical literature, industrial organization economists have structurally estimated dynamic games in a number of market settings. In contrast, organizational economists have been more interested in the processes internal to the firm that result in pricing decisions, one example being the classic study of pricing in a department store by Cyert and March (1963). Macroeconomists, too, have a keen interest in microeconomic price-setting behavior. In many monetary models, price inertia plays a key role in the monetary transmission mechanism, so macroeconomists have tried to identify the practical concerns that firms face in setting prices, such as menu costs, the costs of monitoring rivals' prices, managerial inattention, and the implicit contracts that firms have with their customers to maintain prices.

Our goal is to contribute, in a small way, to many of these literatures, as well as starting to bridge the gap between them. We study the pricing decisions of a set of rival firms selling memory modules over the internet. We have extremely fine-grained data—hourly data on their prices for about a year—which will allow us to estimate a dynamic model of competition and back out parameters, such as managerial attention cost, that inform the micro-foundations of the price inertia we observe. This estimation is in the tradition of structural work in IO. We couple this analysis with manager interviews and some supporting statistical evidence, shedding light on the daily process of price-setting that they implement. Although our methods tend to be more familiar to IO economists, we think our results nicely complement the large and growing empirical macroeconomic literature on price-setting.

Macroeconomists studying price stickiness tend to use large datasets containing a sample of retail prices covering much of the economy, such as the datasets underlying the computation of the CPI. Being able to make statements about large sectors of the economy is useful, but there are trade-offs when using such datasets. Those data are typically collected monthly, so price movements within a month are not recorded. We have hourly data, and we do, in fact, observe many price changes within a month. We also have prices on a (close to) complete set of the rivals operating in a particular market, allowing us to estimate how firms react to rivals' price movements. CPI data would not allow such an analysis since it only contains a sample of prices. We follow firms long enough to observe dozens of price changes for many of them. This time series allows us to make detailed inferences into what is driving price changes at the firm level. We also have unusual and

useful auxiliary data on our market. We have a daily series of wholesale cost of memory modules, which should be common to all market participants, as well as knowledge of other elements of their cost structure. We can, therefore, estimate the importance of cost shocks to price changes. We have quantity data from one market participant, which allows us to estimate demand, providing useful parameters for our estimation. Finally, we have access to a manager who has provided qualitative information about the price-setting process.

One fortunate feature of our market is the fact that wholesale prices of memory modules were quite volatile over the period we study. As a result, we see a large number of both retail price increases and decreases. While this feature might be unusual, it is quite useful to the researcher interested in studying price movements and rigidity. We simply have many observations to study.

A limit of our approach is that our conclusions will be based on one small and specialized market, not a set of markets that cover a large fraction of the U.S. economy. But even then, we would argue that more general lessons can be drawn from our findings here, and the detail of the findings more than outweighs the disadvantage of specificity. Among useful general lessons we can draw are how pricing decisions and price stickiness are influenced by internet technologies. Much of the empirical evidence on price stickiness so far has come from traditional brick and mortar markets.¹ There are reasons to believe that price-setting in online markets could be quite different, both because of increased price transparency and decreased menu costs. As online retail continues to grow, understanding price-setting online will be increasingly important. Also, we would argue that simply understanding the price-setting mechanisms in one market very well, as we can with our unusually detailed data, can help us predict which factors are likely to be important in many other markets. Finally, using a structural approach, as we do, allows us to simulate counterfactual cost structures or market settings. We include results from a counterfactual simulation where one firm receives shocks to its managerial costs.

Our main empirical output is the structural estimation of competition and price-setting behavior in this market. In order to implement our structural estimation, there are various elements that we need to specify. The first is a reduced-form policy function, describing firms' equilibrium pricing behavior—in particular what triggers price changes and what determines their size. The second is a one-period profit function. We will draw on our knowledge of the firms' cost structures and business strategies, as well as estimates of demand, to specify the profit function. This profit function will include parameters for costs of managerial attention and activity, which we will estimate. The third piece, the value function, will be an aggregation forward of the one-period profit functions derived by simulation based on firm play dictated by the estimated policy functions.

¹Exceptions include Lünne-man and Wintr (2006), Chakrabarti and Scholnick (2005), and Arbatskaya and Baye (2004). The ambitious “Billion Prices Project” (Cavallo and Rigobon 2012) will surely generate additional evidence.

Using this structure, we will be able to infer the costs associated with managerial attention and activity. Those costs will be the ones that rationalize the price-setting behavior we observe in the data given our specifications of the policy and profit functions.

For some of these elements, we rely on previous research in this market. For instance, two of the authors here performed a reduced-form analysis of the factors that drive price changes in the online market for generic memory modules (see Ellison and Snyder (2013)). They find that strategic considerations are important—What rank are you in a price-sorted list of competitors? How close in price space are your rivals adjacent to you in rank?—but evidence of managerial inattention or costs associated with monitoring and changing prices was also strong. We will use a similar reduced-form analysis as the policy function. Another related paper, Ellison and Ellison (2009a), provides us with information on costs, business strategy, and demand, which we can use to specify the profit function.

Our method for estimating our structural model is based on Bajari, Benkard, and Levin (2007) (BBL) in which a two-step method is proposed to estimate payoff-relevant parameters in a dynamic game. Their first step involves estimating reduced-form strategies that constitute a Markov perfect equilibrium. Their second step is to find the set of parameters of the profit function that rationalize the estimated strategies such that no player can improve his payoff by switching to a different strategy.

In our case, we use the rule of thumb strategies similar to those estimated in Ellison and Snyder (2013) for the first step and then choose parameters in the firms' profit functions which rationalize their observed behavior within the set of similar rule of thumb strategies. This is a modification of BBL in that firms formulate their policies based on just a subset consisting of the most salient of the hundreds of possible state variables in our market. We make an additional modification to the BBL framework: while it allows heterogeneity of firms up to a distribution, we explicitly postulate three types of firms. We have found that allowing for three types is a parsimonious and tractable but still quite flexible way to allow for the heterogeneity we observe.

The next section reviews the related literature. Section 3 provides details on the Pricewatch market we study motivating the structural model of managerial costs we go on to estimate. Section 4 describes the data. Section 5 specifies the three components of our structural model, the value function, profit function, and policy function. Section 6 provides further details on the estimation of the policy function and the estimation results themselves. Section 7 provides further details on the estimation of structural parameters and the estimation results themselves. Section 8 conducts counterfactual simulations in which we show how a firm's pricing behavior changes when we posit different managerial costs for it than those estimated. The last section concludes.

2. Literature Review

Our paper builds on several previous studies of the same Pricewatch market using the same data: Ellison and Ellison (2009a, 2009b) and Ellison and Snyder (2013). Ellison and Ellison (2009a) estimates demand in the market, but does not consider dynamic pricing patterns or firm interactions, which is the focus of the current study. Instead, we will use the demand estimates from that paper as an input into our analysis. Ellison and Ellison (2009b) also estimates aspects of demand to see how sensitive consumers are to sales-tax savings from purchasing online and whether consumers exhibit home-state “biases” in their purchasing. Ellison and Snyder (2013), also using the same data collected from an online market, poses questions similar to the ones we pose here but uses different methodologies to investigate them. In particular, one could think of Ellison and Snyder as a reduced-form investigation of the factors which drive price changes and this paper as a structural estimation of the firms’ price-setting behavior, using estimated policy functions from Ellison and Snyder as an input.

We also build on a literature in organizational economics, anchored by the classic book “A Behavioral Theory of the Firm” by Cyert and March (1963). In it, the authors document the algorithm, complete with detailed decision diagrams, used by an unnamed department store for pricing its merchandise. Considerations included the wholesale acquisition cost, the price of similar items, what competitors were charging, the season, how long the item had been in inventory, how well the item had been selling, the final digit of the price, whether the desired price was within 2% of an “alliterative figure,” and so forth. It discussed when managerial latitude was allowed and when exceptions to the algorithm were made. Not only does the decision diagram suggest a list of important considerations and triggers in price-setting behavior, but it also highlights the potential importance of agency within the firm in such decisions. Having only firm-level data, agency issues are not ones that we cannot address directly, but keeping them in mind could be helpful. Early pioneers of theory of the firm performed complementary theoretical research, most notably Simon (1955, 1962). In fact, he foreshadowed our use of a reduced state space for structural estimation of the parameters of price-setting behavior: “Price setting involves an enormous burden of information gathering and computation that precludes the use of any but simple rules of thumb as guiding principles (Simon, 1962).”²

Substantively, our analysis is closely related to a large literature in macroeconomics on sticky prices. The recent wave of this literature is nicely summarized in Klenow and Malin’s (2011) Handbook chapter “Microeconomic Evidence on Price-Setting.”³ They discuss in detail the types

²We first came across this quotation in Artinger and Gigerenzer (2012).

³Other important papers include Blinder et al. (1998), Bils and Klenow (2004), Nakamura and Steinsson (2008),

of evidence that macroeconomists have brought to bear on the question of price stickiness, ranging from manager surveys to supermarket scanner data to prices collected by the BLS (and other national statistical agencies) to compute inflation indexes. These studies are typically reduced-form analyses documenting facts about frequency and size of price changes as well as attempting to explain the significant heterogeneity across sectors in price-setting behavior. For instance, prices tend to be more rigid in markets for nondurable goods and services, those for “processed” goods, and those for noncyclical goods. Also, less competitive markets (as measured by concentration indexes) tend to be more rigid as well. This literature has also noted that hazard rates for price changing tend to be flat, suggesting that prices are more state-dependent than time-dependent. Finally, this literature notes interesting facts from manager surveys: firms review prices much more often than they change them, menu costs and the costs associated with gathering information are not important determinants of price-setting, and considerations such as implicit and explicit contracts with buyers, wholesale costs, and “coordination failure” are more important. (“Coordination failure” refers to a phenomenon that we would more likely label dynamic competition.)

Artinger and Gigerenzer (2012) is an interesting paper on used car prices in Germany, which combines these two traditions, using techniques from both organizational economics and macroeconomics. They conduct extensive manager surveys on what factors cause price changes, following on the work of Blinder et al. (1998). They also gather data online from 748 car dealers to test Simon’s (1955) model of aspirational pricing. Their goals are closely aligned with ours—to shed light on what triggers price changes—even if their techniques differ. They do not, for instance, attempt to estimate parameters in firms’s profit functions that may shed light on these questions, as we do.

As our introduction suggested, interest in pricing and, in fact, price rigidity, has a long and distinguished history in industrial economics as well. Gardiner Means, a Harvard industrial economist, testified to Congress about inflexibility of prices during the Depression (Means, 1935) and was later interviewed by Studs Terkel on the topic for his oral history of the Depression (Terkel, 1970). More recently, Stigler and Kindahl (1970) and Carlton (1986) established a number of stylized facts about price rigidity using transactions data. This work was, in part, a critique of previous research on price rigidity performed using BLS price indices. Most of this work focused on industrial, or wholesale, prices.

This particular interest in rigidity complements the long tradition in industrial economics of focus on strategic aspects of pricing. This literature is vast and rich, and we will not attempt to discuss it here, but we will mention one natural empirical outgrowth, the estimation of dynamic

and Eichenbaum, Jaimovich, and Rebelo (2011).

models of pricing. So far this research has focused on the competitive aspects of pricing decisions, abstracting away from menu costs, implicit contracts with customers, costs of managerial attention, and other practical issues which have important impacts on pricing decisions. Much of the research has been set in retail and wholesale gasoline markets, perhaps because prices change often and both rise and fall, so those settings seem like fruitful places to study pricing behaviors. Also, price data in those markets have been relatively easy to come by. Some papers in this literature include Castanias and Johnson (1993), Noel (2007a, 2007b), Lewis (2008), and Wang (2009). A somewhat more detailed review of that literature is in Ellison and Snyder (2013).

Finally, we draw very heavily on the methodological literature on the structural estimation of dynamic games. The most prominent papers are Aguirregabiria and Mira (2007), Bajari, Benkard, and Levin (2007), and Pakes, Porter, Ho, and Ishii (2015). Our estimation strategy is based on BBL.

Our paper applies BBL's techniques, but with a novel feature: instead of assuming that firms are fully rational and play Markov perfect Nash equilibria, say, conditional on all observable variables, we instead allow firms to be boundedly rational, in other words following simple rule of thumb strategies based on a subset of the most salient state variables. One reason that we might be interested in this bounded rationality approach could be that it eases our computational burden and allows us to estimate parameters of interest. More to the point, though, we think it is difficult to justify a fully-rational approach in our empirical setting. There are dozens of firms operating in this market and hundreds of state variables. It is unimaginable that firms could actually be computing best responses every period in such a rich and complicated empirical setting, especially one where the cost to a manager of even monitoring rivals' prices is a significant deterrent to managerial action, let alone performing some wildly complicated calculation. In other words, we argue that the bounded rationality approach makes more sense in this empirical setting and would be preferable to a full rationality approach even if we had infinite computational resources (because managers do not). Furthermore, we provide evidence in the previous paper that the rule of thumb pricing strategies that we estimate not only fit the data well but provide strikingly similar price paths to the actual ones when used to guide simulated firm interactions. And finally, we present results from a goodness of fit test to allow readers to judge how accurate the simulations using the simplified policy functions are in generating profit-relevant statistics.

3. Empirical Setting

Our empirical setting is the online marketplace for computer components mediated by Pricewatch, previously studied by Ellison and Ellison (2009a, 2009b) and Ellison and Snyder (2013). See those

papers for additional details on the empirical setting. During the 2000–01 period during which our data were collected, the Pricewatch universe was characterized by a large number of small, undifferentiated e-retailers selling memory upgrades, CPUs, and other computer parts. These retailers tended to do little or no advertising, have rudimentary websites, receive no venture capital, and run efficient, profit-maximizing operations. They also tended to receive a large fraction of their customers through Pricewatch. The retailers paid monthly fees to list products, had to abide by Pricewatch rules about transparency of prices, and paid no click-through fees. Potential customers could use Pricewatch to locate a product in one of two ways. They could either type a technical product description, such as “Kingston PC2100 512MB,” into a search box, or they could run through a multilayered menu to select one of a number of predefined product categories, e.g., clicking on “System Memory” and then on “PC100 128MB SDRAM DIMM.” In that case, they would receive back a list of products sorted from cheapest to most expensive in a format with twelve listings per page. These pre-defined categories may contain as many as 350 listings from 100 different websites. Figure 1 contains the first page of a typical list for PC100 128MB memory modules downloaded during our sample period. October 12, 2000.

Given our focus on the dynamic aspects of firms’ pricing strategies, it is worth noting that the Pricewatch ranking exhibits substantial turnover from day to day and even from hour to hour. On average, five of the twenty-four retailers on the first two pages of the above-mentioned list will change their prices on a given day. Each price change typically moves several other retailers up or down one place on the list. Some retailers pursued a strategy of remaining near the top of the Pricewatch list; others spent most of the time lower down. We will explore heterogeneity in strategies in the next section when we introduce the dataset. Whichever part of the ranking a given retailer might target, retailer did not maintain rigid ranks. For example, only rarely would the same retailer remain in the first position for as long as a week.

Based on information from a detailed interview of a manager of one of the retailers participating on Pricewatch, we can identify several possible reasons for this churn in the rankings. First, wholesale prices could be quite volatile for some products. Retailers would receive wholesale-price quotes via daily emails. Wholesale prices often fluctuated from day to day. Short-term fluctuations could be up or down, but as these are electronic components, the long-term trend was downward. The daily price quotes were relevant to the manager’s operations as they typically carried little or no inventory, ordering enough to cover just the sales since the previous day’s order. A second reason for turnover in the rankings was that managers did not continuously monitor each individual product’s rank on the Pricewatch website, making the instantaneous changes needed to maintain a rank. Retailers typically offered scores of products in different categories on Pricewatch; it would

be impractical for a manager to continuously monitor all of them even he or she attended to no other managerial tasks. Certain high-volume products, such as the specific memory modules we study, could merit more attention, but this might involve checking at most one or two times a day. During our study period, managers had to enter price changes manually into Pricewatch's database, as automated price setting was not introduced until 2002. Each adjustment would thus involve a fixed cost in both determining and entering the appropriate price.

Retailers could infer important information from sudden changes in sales (which could be monitored in real time as orders flowed in) without checking Pricewatch. For example, a sudden decrease could indicate that the retailer has been bumped to a higher-price rank. The first panel of Figure 2, derived from demand estimates from Ellison and Ellison (2009a), indicates how a retailer's daily sales vary with rank. The bulk of sales go to the two or three lowest-priced retailers in this market, but positive sales still accrue to many additional retailers on the list. For later reference in our calculation of firm profits, we will label the curve the $Q(Rank_{it})$ function.

A significant source of profit in this market comes from the "upselling" strategy documented by Ellison and Ellison (2009a). A firm can attract potential customers with low prices for the "base" memory module, but then try to induce the consumer to upgrade to a more expensive memory module. The second panel of Figure 2 shows the estimate of the hourly profit from this upselling strategy as a function of the firm's rank. The shape results from two opposing forces: at higher ranks the firm has fewer potential customers, but the customers it does attract are advantageously selected to be more likely to upgrade. For later reference in our calculation of profits, which will incorporate returns from the upselling strategy, we will label the curve the $U(Rank_{it})$ function.

The background information on the Pricewatch market provided so far provides motivation for our subsequent structural model incorporating managerial costs of monitoring the market and further costs of changing prices, allowing us to derive structural estimates of those cost parameters. More formal motivation is provided by findings in Ellison and Snyder (2013). In a simple probit regression, price changes were significantly less likely the longer the time elapsed since the last change, consistent with spells of inactivity rather than price changing at systematic intervals. Price changes were also significantly less likely at night and on weekends, when managerial attention and activity is likely to be particularly costly.

Figure 3, reproduced from Ellison and Snyder (2013), provides finer evidence on managerial inattention. The figure plots the residual probability of a price change (after partialling out other covariates as controls) during each hour measured in Eastern Time, estimated separately for retailers operating on the East and West Coasts. Retailers supply a national market via Pricewatch, so if there were no managerial costs, presumably the timing of activity would be similar on the

two coasts, responding to the same national demand factors. In fact the probability functions peak at different points, at 11 a.m. for East Coast and 8 p.m. for West Coast retailers. Interestingly, the peak for the West Coast retailers is not simply shifted by the three-hour difference between Eastern Time and the local time for West Coast retailers, as might be expected if the retailers on the two coasts served separate markets but faced the same pattern of managerial costs during the day. Instead, the peak on the West Coast is shifted by ten hours, to 8 p.m. Eastern Time, which is 5 p.m. in their local time. One explanation, supported by our interview subject, is that the market has already been operating for a few hours by the time the West Coast managers arrive at work, so they are well-advised to set prices in the evening before they leave. The evening might also be a less busy time for them since orders might have started falling off at least from customers in the East. In contrast, the East Coast manager would arrive to a very slow order flow at 8 a.m. Eastern Time and have the leisure to adjust prices at that point before orders picked up for the day. In the absence of managerial costs, it is difficult to rationalize why retailers on the two coasts would pick the opposite ends of the day to do most of their price changes.

Ellison and Snyder (2013) documented another feature of the empirical setting that we will incorporate into the structural model: substantial heterogeneity in retailers' strategies, including how often they change price, the price they settle on conditional on changing, the sensitivity to rivals' actions, and so forth. This heterogeneity is illustrated in Figure 4, reproduced from that paper, showing movements in price and rank for three firms over a particular month in the sample. The firm in the first column actively changed price and occupied relatively low ranks. The firm in the last column also actively changed price but occupied much higher ranks. The firm in the middle column shows much less price activity. It starts the month at rank 6, but due to its inactivity is bumped up to rank 14 by the end of the month.

Ellison and Snyder (2013) balanced flexibility against parsimony by classifying retailers as one of three types and allowing estimates to vary freely across the types. See that paper for a complete description of the method used to classify firms, called cluster analysis. This is a standard grouping algorithm used in various social sciences, described in the textbook treatment by Romesburg (2004). Here, we will simply adopt Ellison and Snyder's (2013) classification of firms into three types, denoted $\tau \in \{1, 2, 3\}$. Our preferred specification will allow estimated structural parameters to vary freely across τ .

Ellison and Snyder (2013) are agnostic about what drives this heterogeneity in firm behavior. Differences in firm costs, such as the cost incurred by a manager to monitor market conditions or to compute and enter a price change, could explain some of the heterogeneity. We will later estimate managerial costs along with other structural parameters. Allowing the estimates to differ across

firm types will allow us to determine if differences in managerial costs are sufficient to explain this observed strategy heterogeneity.

4. Data

Our data are the same as in Ellison and Snyder (2013). Their primary data source was the first two pages, including the 24 lowest-price offerings, from the category of 128MB PC100 memory modules. These pages, consisting of the lowest 24 price offerings, were scraped for the price, name of the retailer, and detailed product description associated with each listing. The first two pages were scraped each hour from May 2000 to May 2001. Products in this category are, physically at least, fairly homogeneous. We do keep track of product offerings, registering a product change when a retailer changes the name of its product.

We also use supplementary data from a retailer participating on Pricewatch. This retailer provided data on its sales and wholesale acquisition costs. We take this cost to be common across retailers justified by the fact that the typical retailer carried little inventory and had access to a similar set of wholesalers.

A large number of firms made brief appearances on the Pricewatch lists. Since we are interested in the dynamics of firms' pricing patterns, we study firms that were present for at least 1,000 hours during the year (approximately one-eighth of our sample period) and changed price while staying on the list at least once. For a small number of firms who had multiple products on the first two pages of Pricewatch simultaneously during some periods, we excluded their observations during those periods. We were left with 43 firms that appear at some point during the year, with at most 24 present on the first two pages of Pricewatch at any particular moment. Although the excluded retailers do not constitute observations, we do use them to compute relevant state variables for rivals including rank, density of neighboring firms, etc.

Based on these data, we created a number of variables to describe factors that might be important to firms' decisions about timing and magnitude of price changes. Figure 2 demonstrated the importance of the rank of that firm on the Pricewatch list for firm outcomes. We also included margin, length of time since its last price change, number of times a firm has been "bumped" (i.e., had its rank changed involuntarily) since its last price change, and so forth. Table 1 provides a description of these variables and summary statistics.

Most of the variables can be understood from the definitions in the table, but a few require additional explanation. *Placement* measures where a firm is between the next lower- and next higher-priced firms in price space. For example, if three consecutive firms were charging \$85, \$86, and \$88, the value for *Placement* for the middle firm would be $0.33 = (86 - 85) / (88 - 85)$.

Density is a measure of the crowding of firms in the price space around a particular firm. It is defined as the difference between the price of the next higher-priced firm minus the price of the firm three spaces below divided by 4. For example, if five consecutive firms charged \$84, \$84, \$85, \$86, and \$88, the value for *Density* for the firm charging \$86 would be $1 = (88 - 84)/4$. *QuantityBump* reflects relative changes in a retailer's order flow caused by being bumped from its rank. It is calculated using the $Q(Rank_{it})$ function in Figure 2, in particular proportional to $\ln(Q(Rank_{it})/Q(Rank_{it'}))$, where t is the current period and t' is the period in which the retailer last changed price. *CostTrend* and *CostVol* are computed by regressing the previous two weeks of costs on a time trend and using the estimated coefficient as a measure of the trend and the square root of the estimated error variance as a measure of the volatility. The definitions of the remaining variables are self-explanatory.

Turning to the descriptive statistics, note first that the average price charged by our firms for a PC100 128MB memory module during this period was about \$69 with a very large range, \$21 up to \$131. Most of this variation is over time, with prices typically above \$100 at the beginning of the period and down in the \$20s by the end, mirroring a large decline in the wholesale cost of these modules. Price changes are on average around five days (117.55 hours) apart. Contrast this with findings in the macro literature using BLS data that the median length of a price before it is changed is 4.3 months (Bils and Klenow, 2004). Finally note that wholesale cost is falling quickly on average, \$0.17 a day, but is quite volatile.

Table 2 provides a selection of descriptive statistics broken down by the three types of firm. The means for *Margin* show that type-1 firms earn the lowest margins (at least captured by this measure), followed by type 2, followed by type 3. The means of *Rank* show the same pattern, with type 1 occupying the lowest ranks, followed by type 2, followed by type 3, consistent with Figure 4. The means of *SinceChange* show that types 1 and 3 change price more than twice as often as type 2. The higher standard deviation on *Rank* for type 2 reflects the fact that its inactivity in price leads to greater variation in rank.

5. Model

This section presents the model that will be the basis for our structural estimation. We begin in the next subsection by discussing the key object in dynamic structural estimation, the firm's value function. The subsections after that discuss the components that go into computation of the value function.

5.1. Value Function

Firm i participates in the market each period t from the current one, $t = 0$, until it exits at time T_i . Its objective function is the present discounted value of the stream of profits given by the value function

$$V_i(s_t; \sigma, \theta) = E \left[\sum_{k=t}^{T_i} \delta^k \pi_{ik} \right], \quad (1)$$

where s_t represents the state of the game in period t , σ is the vector of firms' strategies, δ is the discount factor,

$$\pi_{it} = \pi(s_t, \sigma(s_t), \theta) \quad (2)$$

is the static profit function, reflecting the payoff from one period (one hour in our empirical setting) of play, and θ is a vector of parameters. The expectation is taken over the distribution of all possible game play and evolution of private shocks starting from s_t .

The focus of this study is on obtaining structural estimates of θ , which will include measures of managerial costs, upselling profits, and other variables of central interest. Our methods of estimating θ follow Bajari, Benkard, and Levin (2007) (BBL) closely. In essence, we will compare i 's value function when all players play according to their equilibrium strategies $\sigma_i(s_i)$ to i 's value when it deviates to some other strategy $\tilde{\sigma}_i(s_i)$, where s_i denotes that part of the state vector s observable to i . The estimated θ will minimize according to some criterion function violations of the optimality of the equilibrium strategies. Performing this comparison requires us to compute the value function in and out of equilibrium, which in turn requires three components: (a) specification of the profit function π_{it} ; (b) an estimate $\hat{\sigma}(s)$, which following BBL we will use in place of the equilibrium strategy profile $\sigma(s)$; and (c) treatment of the expectations operator. The remainder of this section will be devoted to specifying the profit function and a model which will be used to estimate firm i 's strategy, also called its policy function. We will compute expectations by averaging the stream of profits from many simulated runs of the market. Details on the simulation methodology are deferred to Section 7 on the structural estimation.

5.2. Profit Function

Our detailed specification of the profit function π_{it} draws on our rich information about the business strategies of the firms operating in this market and the costs they face, as well as institutional details and estimates from Ellison and Ellison (2009a):

$$\pi_{it} = Base_{it} + Upsell_{it} - \mu_{\tau(i)} Monitor_{it} - \chi_{\tau(i)} Change_{it} \neq 0\}, \quad (3)$$

where

$$Base_{it} = Q(Rank_{it})(Price_{it} - Cost_t) \quad (4)$$

$$Upsell_{it} = U(Rank_{it}). \quad (5)$$

The first term $Base_{it}$ accounts for the profits from the sale of the base version of the memory modules, equaling the quantity sold times the markup per unit. The quantity sold is the function $Q(Rank_{it})$ from Panel A of Figure 2.

The second term $Upsell_{it}$ accounts for the upselling strategy discussed in Section 3, whereby the firm attracts potential customers with low prices for the base product but then induces some of them to upgrade to more expensive versions of the memory module. It is given by the function $U(Rank_{it})$ from Panel B of Figure 2⁴.

The final two terms in (3) reflect the two types of managerial costs that we would like to estimate, both of which vary by firm type. The coefficient $\mu_{\tau(i)}$ on the indicator $Monitor_{it}$ for whether firm i monitors in period t will provide an estimate of the cost of monitoring. The last term involves $Change_{it}$, which is our label for the indicator $1\{\Delta_{it} \neq 0\}$ for whether i changes price in period t , where we define $\Delta_{it} = Price_{it} - Price_{i,t-1}$. The coefficient $\chi_{\tau(i)}$ on this indicator will provide an estimate of the cost of price change over and above any monitoring cost. Monitoring behavior is not observed, but in the simulations we will use to compute expectations of the value function, just as we will generate simulated values for $Rank_{it}$, $Price_{it}$, and other variables, we can simulate $Monitor_{it}$ using the model of monitoring behavior from our policy function, which we turn to next.

5.3. Policy Function

The next piece needed for structural estimation is a policy function, i.e., a model of firms' strategies. An estimate of this model provides the $\hat{\sigma}(s_t)$ that will be substituted for equilibrium strategies $\sigma(s_t)$ to compute the value function (1) in the simulations. The model reflects the reality that price changes in this market are infrequent relative to our data frequency. To a first order, the best prediction of the firm's price next hour is its current price. We will thus focus on modeling the timing and size of price change episodes, with the firm maintaining the price outside of these episodes. Consistent with our belief that firms are not engaging in complicated calculations of optimal policy based on hundreds of state variables each hour, we want the model to be simple, streamlined, and reflective of just the variables that firms are likely to be able to monitor and process. Also, we want

⁴We have subtracted \$4.7, which is our approximation of a firm's hourly profit at rank 25, from the numbers in Panel B for normalization purpose. The structural estimates are robust to the choice of this number.

the model to be a good empirical description of what firms actually do.

We thus model price changes as coming from a two-step process. The manager knows some components of the market state vector at all times, information he receives essentially “for free.” So first, the manager must decide, based on these state variables, whether to attend to the market to gather information needed for a pricing decision. We will call this behavior “monitoring” and denote the decision to do so with the indicator function $Monitor_{it}$. In our setting, we think of monitoring as visiting the Pricewatch website and making price decisions, involving an opportunity cost of cognition and time.⁵ Through monitoring, the manager gains additional information on state variables, including current rank and the distribution of competitors’ prices, and computes the new desired price. If the new desired price is different from the current price, to carry out the price change, he then enter it in the Pricewatch form, and change it on one’s own website, again involving costs in terms of cognition and time.

The two-step process can account for periods of excess inertia, during which the manager keeps price constant even though market conditions would warrant a price change. Inertia can come from three sources. First, the manager may not be aware of the changed market conditions because he or she did not monitor. Second, the benefit from making a desired price change, especially a small price change, may not justify the managerial cost of entering it. Third, if the desired price change is smaller than a whole dollar unit in which Pricewatch prices are denominated, price may stay constant. The two-step process is also consistent with anecdotal evidence from our interview subject and broader survey evidence (see Blinder et al., 1998) that managers often monitor market conditions including rival prices without changing their own price.

We specify the manager’s latent desire to monitor, $Monitor_{it}^*$, as

$$Monitor_{it}^* = X_{it}\alpha_{\tau(i)} + e_{it}, \quad (6)$$

where X_{it} is a vector of explanatory variables, $\alpha_{\tau(i)}$ is a vector of coefficients to be estimated, which are allowed to differ across firm types $\tau(i)$, and e_{it} is an error term. If $Monitor_{it}^* \geq 0$, then the firm monitors; i.e., $Monitor_{it} = 1$. Otherwise, if $Monitor_{it}^* < 0$, then the firm does not monitor; i.e., $Monitor_{it} = 0$.

Practically tautological but still important to emphasize, the explanatory variables can only include state variables known by the manager before monitoring. In addition, the variables must be important shifters of either the cost or benefit of monitoring. We specify the following parsimo-

⁵Monitoring may include performing some calculations, consistent with the idea that additional effort is necessary and the results of those calculations are not gained “for free.”

nious list:

$$X_{it} = \left(Night_t, Weekend_t, CostVol_t, CostTrend_t^+, |CostTrend_t^-|, \right. \\ \left. QuantityBump_{it}^+, |QuantityBump_{it}^-|, \ln SinceChange_{it}, (\ln SinceChange_{it})^2 \right) \quad (7)$$

We assume the manager is automatically aware of the time and day. The variables $Night_{it}$ and $Weekend_{it}$ are included to reflect the high cost of monitoring instead of sleeping or engaging in leisure activities as well as the low benefit when few sales are made. The manager is also assumed to be aware of the day's wholesale cost—recall that he receives emails every day from the wholesalers—and can glean volatility and trends from the pattern of costs over the past several weeks. Presumably the gains to monitoring are greater the more conditions including costs are fluctuating. We include $CostVol_{it}$ to capture unpredictable fluctuations and $CostTrend_{it}$ predictable fluctuations. Both rapidly rising and rapidly falling costs would lead the manager to monitor more. To allow asymmetry⁶ in the concern for rising or falling costs, a rising cost trend, $CostTrend_t^+ = CostTrend_t \times 1\{CostTrend_t > 0\}$, enters (6) separately from a falling cost trend, $CostTrend_t^- = |CostTrend_t| \times 1\{CostTrend_t < 0\}$. The latter variable appears in absolute value so that in theory the two variables' coefficients should have the same sign if not magnitude.

We assume that the manager is roughly aware of changes in his or her order flow resulting from being bumped in the ranks and will be more likely to monitor if there has been a large change, whether an increase or decrease. Thus (6) includes $QuantityBump_{it}$. Recall that this variable is computed by translating the current rank and rank at the previous price change into a quantity change using the function in Panel A of Figure 2. This predicted change in order flow is a proxy for the quantity signal the manager observes. The proxy diverges from the signal because the firm's actual sales depend on random market fluctuations on top of any predictable effect of a rank change. The proxy also diverges from the signal because the manager may only be vaguely aware of actual sales in any given hour. Again, to allow for asymmetries, increases in order flow, $QuantityBump_{it}^+$, enter separately from decreases, $QuantityBump_{it}^-$.

The last set of variables, functions of $SinceChange_{it}$, enter in a flexible, nonlinear way, allowing for various patterns of managerial attention, including monitoring the market at regular intervals as well as periods of intense monitoring, in which several price changes may follow in succession, followed by periods of inattention during which price stays constant independent of market condi-

⁶Some evidence suggests that prices could be stickier in one direction versus the other. For instance, Borenstein, Cameron, and Gilbert (1997) identify an asymmetry in the response of wholesale gasoline prices to cost increases versus decreases.

tions. Including this variable can help us tease out time dependence from state dependence in price monitoring behavior.

Conditional on monitoring, the manager may decide to change price based on the information acquired. The latent size of the price change, $\Delta_{it}^* = Price_{it}^* - Price_{i,t-1}$, is given by

$$\Delta_{it}^* = \begin{cases} Z_{it}\beta_{\tau(i)} + u_{it} & \text{if } Monitor_{it} = 1, \\ 0 & \text{if } Monitor_{it} = 0, \end{cases} \quad (8)$$

where Z_{it} is a vector of explanatory variables, $\beta_{\tau(i)}$ is a vector of coefficients to be estimated, which again are allowed to differ across firm types $\tau(i)$, and u_{it} is an error term. Because prices were denominated in whole dollars on Pricewatch, the latent price change, which is continuous, is not realized. The integer price change $\Delta_{it} = Price_{it} - Price_{i,t-1}$, derived from the latent price change Δ_{it}^* , is realized instead. Let

$$\dots < C_{-5} < C_{-4} < C_{-3} < C_{-2} < C_{-1} < C_0 < C_1 < C_2 < C_3 < C_4 < C_5 < \dots \quad (9)$$

be a set of cut points on the real line. Then

$$\Delta_{it} = \begin{cases} k & \text{if } \Delta_{it}^* \in (C_k, C_{k+1}), \\ -k & \text{if } \Delta_{it}^* \in (C_{-(k+1)}, C_{-k}). \end{cases} \quad (10)$$

Thus, for example, an observed price increase of \$1 corresponds to a latent price change satisfying $\Delta_{it}^* \in (C_1, C_2)$. If the manager monitored but did not change price, then $\Delta_{it}^* \in (C_{-1}, C_1)$. If we knew that the manager rounded the latent price change to the nearest dollar to compute the actual price change and we knew the distribution of the error term, we could take the cut points to be the corresponding integers, $C_k = k$. In the absence of such knowledge, we will let C_k be free parameters to be estimated as in an ordered probit.

The explanatory variables can include state variables the manager learns as a result of monitoring in addition to those known before. We specify the following parsimonious set:

$$Z_{it} = \left(CostTrend_t, CostChange_{it}, Margin_{it}, NumBump_{it}, \right. \\ \left. Density \times NumBump, Placement_{it}, Rank_{it}, RankOne_{it} \right). \quad (11)$$

The higher are forecasted costs as proxied by $CostTrend_t$ and the more costs have risen since price

was last changed as measured by $CostChange_t$, the higher the firm's desired price. $Margin_{it}$ may also factor into price changes, a firm with low margins being more likely to increase price and one whose margins are already high less likely to further increase price.

The remaining variables in Z_{it} are the sort of state variables revealed by monitoring. After visiting the Pricewatch website, the firm learns its current rank and thus the number of ranks it was bumped since the last price change. The firm can use this information to return itself to its desired rank, so decreasing price if it was bumped up in the ranks (gauged by a positive $NumBump_{it}$) and increasing price if it was bumped down. A low-ranked firm may have less of an incentive to cut price to increase its sales; this effect may be particularly strong for a firm occupying rank 1: further price reductions will only result in a small increase in sales because most of the demand elasticity is with respect to rank, which the firm cannot improve beyond 1. We, therefore, include an indicator, $RankOne_{it}$. We include $Density_{it}$ because the presence of a thicket of close competitors may affect pricing incentives. For example, in a dense price space, firms may have less incentive to increase price because this will result in a more severe rank change. Specifically, we include $Density_{it}$ interacted with $NumBump_{it}$ because density will likely matter more for firms that have a need to change price as proxied by a bump from the previous preferred ranking. A firm's $Placement_{it}$ between its nearest rivals will also affect its desired price in potentially complex ways.

The list of explanatory variables in Z_{it} deliberately excludes some of the variables that appear in the monitoring equation. For example, $Night_t$ is included in the monitoring equation to reflect the fact that checking the Pricewatch website at 2 a.m. would typically be more costly than during the workday. However $Night_t$ should have little effect on the desired price change conditional on monitoring because, conditional on already being on the website, changing price is no more difficult at 2 a.m. than 10 a.m. The same logic applies to $Weekend_t$. While the time since price was last changed may affect the desire to monitor—one possibility is that with more time for market conditions to change, more information is to be gained—conditional on the information gained through monitoring such as $Rank_{it}$ and $NumBump_{it}$, $SinceChange_{it}$ has no obvious role in price setting and thus is excluded from Z_{it} .

6. Estimation of the Policy Function

6.1. Identification

One might wonder how the coefficients $\alpha_{\tau(i)}$ in the monitoring equation can be estimated when the dependent variable $Monitor_{it}$ is not observed in the data. The answer is that the dependent variable in the price-change equation is observable, and $Monitor_{it}$ enters that equation through (8). In cases

in which equations (8) and (10) predict large-magnitude price changes, either positive or negative, but these do not materialize in the data, this discrepancy will either be attributed to a large error u_{it} , which carries a likelihood penalty, or will be picked up by the coefficients on correlated variables in the monitoring equation. In this way the monitoring equation provides explanatory power.

The power to separately identify determinants of monitoring separately from determinants of desire to change price comes from variables included in the price-change equation that are excluded from the monitoring equation and vice versa. There are variables that cannot affect the monitoring decision because they are unknown until monitoring has been done but are important in driving pricing decisions. In principle, that exclusion restriction is all that is necessary for identification, but we also have variables that enter the monitoring equation but not the price change equation, as discussed earlier. These are variables, such as $Night_t$ and $Weekend_t$, that were excluded from the pricing equation based on the argument that, conditional on monitoring, they should have not incremental effect on the price change decision.

Another way to think about identifying parameters in the monitoring equation is to consider two different periods where variable values in the price change equation would predict a price change. If there is a price change in the first and not the second, our model posits that one possibility is that monitoring occurred in the first and not the second, so price could not be changed in the second even if conditions warranted. It is precisely that variation that allows us to identify what factors drive the monitoring decision.

While it is easy to justify the inclusion of the explanatory variables in the respective functions, it may be harder to assert that this strategy model is sufficient to describe firms' behavior. In reality, firms can execute much more complex strategies, from ones with nonlinear terms to ones with additional variables such as their neighbors' characteristics. Furthermore, one might be concerned that formulating such an empirical model is, by its nature, arbitrary and undisciplined by theory. However, we feel that there are several reasons that our approach is appealing in this particular setting. First, it appears as if managers in our market face monitoring and decision costs, for acts as simple as logging onto a website, that are significant enough to alter behavior in important ways. In such a situation, we find it implausible to imagine that complicated calculations of expected value functions based on dozens of state variables each period are being carried out. Far more plausible is a type of simplified rule of thumb decision-making. Second, Ellison and Snyder (2013) have established that these rule of thumb functions do a good job of describing firm behavior in standard statistical fit metrics and also generate histories very similar to the actual price paths in simulation studies. Third, these estimated strategies give the same expected behavior conditional on the set of variables in this model as the firms' true strategies. Fourth, there are significant

computational advantages to our approach, relative to the fully-rational model. Finally, as we will detail in Section 6.4, we have encouraging results from a goodness of fit test to determine how accurately the estimated strategies can generate the exact profit-relevant statistics we use as input. In our opinion, these facts bolster the case for using these policy functions in the first stage of estimation.

6.2. Estimation Details

The policy function consists of equations (6)–(11). We estimate these equations jointly using maximum likelihood taking the errors e_{it} and u_{it} to be independent standard normal random variables. Without the monitoring stage, the price-change stage of the model would be equivalent to an ordered probit, where various intervals would correspond to various discrete price changes. The monitoring stage adds a factor to incorporate into the likelihood function. The likelihood of observations in which firm i does not change price at time t is

$$L(\Delta_{it} = 0) = \underbrace{1 - \Phi(X_{it}\alpha_{\tau(i)})}_{\Pr(\text{Monitor}_{it}=0)} + \underbrace{\Phi(x_{it}\alpha_{\tau(i)})}_{\Pr(\text{Monitor}_{it}=1)} \underbrace{[\Phi(C_1 - Z_{it}\beta_{\tau(i)}) - \Phi(C_{-1} - Z_{it}\beta_{\tau(i)})]}_{\Pr(\Delta_{it}^* \in (C_{-1}, C_1))}, \quad (12)$$

where Φ is the standard normal distribution function. The likelihood of, for example, a k dollar price increase is

$$L(\Delta_{it} = k) = \underbrace{\Phi(x_{it}\alpha_{\tau(i)})}_{\Pr(\text{Monitor}_{it}=1)} \underbrace{[\Phi(C_{k+1} - Z_{it}\beta_{\tau(i)}) - \Phi(C_k - Z_{it}\beta_{\tau(i)})]}_{\Pr(\Delta_{it}^* \in (C_k, C_{k+1}))}. \quad (13)$$

Adding the monitoring stage scales up the probability of no price change and scales down the probability of any given sized price change. Harris and Zhao (2007) describes this type of two-stage model as the zero-inflated ordered probit model (ZIOP) and uses Monte Carlo results to show the maximum likelihood estimator of this model has good finite sample performance. We group the few price reductions of \$5 or more together and similarly group the few price increases of \$5 or more so that we only need to estimate thresholds down to C_{-5} and up to C_5 .

6.3. Results

Table 3 presents estimates of $\alpha_{\tau(i)}$ and $\beta_{\tau(i)}$. The first two columns present estimates on the combined sample of firms and the next three sets estimates from separate estimation for each of the three firm types. Focusing first on the combined estimates, we see that *Night* and *Weekend* are very important determinants in whether a firm monitors. We do not find this result surprising, but we

think it is a telling indicator of the importance of managerial attention: managers take the evenings and weekends off and do not fool around monitoring their firms products during these times. Managers are also more likely to monitor if wholesale cost is volatile, if there has been a significant trend in wholesale cost recently, if the order flow has changed a lot, or if it has been a long time since the last price change. These estimates consistent with our intuition, and note that they are typically quite precisely estimated.

Conditional on monitoring, a price increase is associated with a wholesale cost increase, a firm being bumped down, rank being low, rank being 1, and there being few firms close by in price space. Again, these results are consistent with our intuition of the factors that drive price changes as well as our conversations with firm managers.

We estimate the policy function separately by firm type and we can reject policy function homogeneity, but most of the differences across firm types are in terms of degree, as opposed to there being large qualitative differences. Our framework is flexible enough to accommodate the difference in activeness between type 2 firms and the rest. In the monitor equation, the coefficients for $\ln \text{SinceChange}$ and $(\ln \text{SinceChange})^2$ of type 2 firms constitute a quadratic function of $\ln \text{SinceChange}$ that is decreasing when $\text{SinceChange} < 586$. So a type 2 firm would only monitor when it has accumulated enough change in QuantityBump or it has been a very long time since the last price change. Conditional on monitoring, a type 2 firm requires a more significant latent size of price change to trigger actual price changes because the size of the interval of no price change, $C_1 - C_{-1}$ is larger for type 2 than that for the rest.

6.4. Goodness of Fit

We will use a series of figures to assess the goodness of fit of the estimated policy function $\hat{\sigma}$, in particular whether it fits well enough to be suitable for the later structural estimation. As discussed in Section 7.2, given the structural parameters, firm i 's value function $\widehat{EV}_i(s; \hat{\sigma}, \theta)$ depends linearly on just a few present discounted values: in particular, the presented discounted value of the hours i spends at each rank and the present discounted value of the number of i 's price changes. For our purposes, these present discounted values can be taken to be simple sums. This is because we truncate the value function after 720 hours in the structural estimation to lessen simulation error and computational burden; moreover, the assumed annual discount factor $\delta = 0.95$ is sufficiently high that it entails essentially no discounting over a short 720-hour period.

Thus we start our assessment of goodness of fit in Figure 5 by comparing the time firms spend at each rank when they behave according to the estimated policy function over a 720-hour forward simulation to that during the same time period in the actual sample. The graphs break the results

down by the three firm types. To perform the simulations, we start with each it observation for a firm of that type in the data and construct 20 simulated forward histories lasting 720 hours. The simulations use the actual market data for state variables where possible (i.e., for cost histories), simulating firm behavior by substituting current state variables as well as random draws for error terms e_{it} and u_{it} into equations (6) and (8) of the policy function. The average from the simulations is compared to the average in the actual data of 720-hour forward histories starting from each candidate it observation.

Panel A of Figure 5 compares simulated time spent at each rank from the estimated policy function (the solid black curves) to the averages in the actual data (the dashed grey curves) for the three firm types. The actual curves have quite different levels and shapes across firm types, yet the simulated curves are able to fit each quite well. Panel B compares the simulated number of price changes from the estimated policy function (the black squares) to the averages in the actual data (the grey circles). Again, the fit is quite close, with the markers essentially overlapping and moving together across the types of firm: moderate for type 1, low for type 2, and high for type 3. The close fit is not surprising given the policy function is a reduced form that was estimated in part to maximize the likelihood of the observed frequency of price changes. However, the close fit was not guaranteed. The maximum likelihood estimation targeted size as well as number of price changes; we see that the joint estimation does not harm the fit for number alone. More importantly, the policy functions were estimated to fit individual behavior; letting their interaction on the market play out over a length of time could generate feedbacks causing behavior to diverge from actual outcomes. We see in Panel A that this is not the case. Taken together, the results from Figure 5 suggest that the estimated policy function will provide good estimates of the sums that are the essential inputs into the value functions in the structural estimation.

Figure 6 provides a more refined assessment of goodness of fit. The policy function should not only be able to fit the forward history on average across states but also fit the forward history in any state s the firm finds itself. The figure compares simulated to actual profit conditional on various states including various initial ranks in Panel A and various initial markups in Panel B. (To save space, we just show the results for type-1 firms; the fit for types 2 and 3 is similar.) We focus on profit in this figure as opposed to time spent at each rank in the previous figure because profit is a convenient summary statistic for the distribution of times spent at each rank, allowing us to reduce the dimensionality of the graph. While we do not yet have all the components of profit π_{it} from equation (3)—the managerial-cost components are of course yet to be structurally estimated—we can estimate the first two components $Base_{it}$ and $Upsell_{it}$ using equations (4) and (5). Because monetary profits are not observed even in the actual data, they have to be estimated in

both cases—using simulated prices in the former case and actual prices in the latter. Because the figure displays the results down separately for each initial state s , each average is now taken only over initial observations it that qualify for state s .

Panel A compares the simulated to actual monetary profit going forward for 720 hours conditional on rank in the initial period. The solid black curve for profit based on simulated prices closely matches the dashed grey one for profit based on actual prices over the whole range of the horizontal axis. Of course the closeness of the graphs could be a symptom, not of good fit, but of the stability of the environment, with initial rank correlating highly with profits at least over an horizon as short as 720 hours. To investigate this possibility, we have added a curve (lighter with dot markers) representing the naïve forecast that the firm earns the same profit in each of the 720 hours as it does in the first. This naïve forecast ends up overestimating profit conditional on top ranks, because it has not taken into account the other firms’ reaction to firm i ’s top rank, which is to undercut firm i . For similar reasons it ends up underestimating profit conditional on bottom ranks. Our estimated policy function fits dramatically better for both high and low ranks, suggesting that our policy function likely captures this dynamic correctly.

Panel B similarly compares simulated, actual, and naïve estimates of monetary profit conditional on a type-1 firm’s markup (price minus wholesale cost in levels) in the initial state s . Again our estimated policy simulation fares well, while the naïve prediction underestimates profit conditional on negative markups and overestimates profit conditional on positive markups because does not properly incorporate the firm’s future price adjustments and the cascade of rival responses. This comparison suggest that our policy function has likely captured this dynamic correctly.

7. Estimation of Structural Parameters

7.1. Identification

We follow the broad outlines of the approach to estimating a dynamic structural model in Bajari, Benkard, and Levin (2007) (BBL). BBL assume the econometrician observes the game play of a Markov perfect Nash equilibrium in which the strategy $\sigma_i(s)$, also referred to a policy function, for each player i is a function only of the public state of the game s plus a random shock. So first, the econometrician can use the observed data to obtain consistent estimates of the policy function and the distribution of the random shock. Second, given a specification of the per-period profit function, the econometrician can use simulation to calculate the value function (1) for each player i . Third, the econometrician chooses structural parameters θ such that, when the value function is evaluated at these parameters, i cannot increase its value function through deviating from the

estimated strategy. The implied identification condition is

$$V_i(s; \hat{\sigma}, \theta) \geq V_i(s; \sigma_i, \hat{\sigma}_{-i}, \theta) \quad \text{for all } i, s, \sigma_i. \quad (14)$$

Operationally, θ is estimated by considering deviations from the policy functions estimated in the first stage (for example, perturbing the estimated policy-function parameters), simulating the value functions from those deviations, and then choosing the $\hat{\theta}$ such that the value functions associated with the deviations are no greater than the equilibrium value functions.

Our approach to identification differs in some details from BBL's to suit our application. Because our policy function is a simplified rule of thumb based on a few salient state variables, it may be heroic to assume it will dominate any possible deviation strategies for all possible, potentially quite complex, states. Instead we will only require the estimated policy function to dominate deviations in the class of strategies, denoted $PF(\alpha_{\tau(i)}, \beta_{\tau(i)}, C_k)$, formed by choosing alternative values for the coefficients estimated for the policy function. Further, we will only require the estimated policy function to dominate in expectation over states in the firms' consideration sets, defined as that part of the state vector that factors into firms' policy functions. Assuming firms have consistent beliefs about game play—the premise of the BBL method—then a natural approximation of the distribution of states in firms' consideration sets is the empirical distribution observed in the data, denoted \hat{S} . This approximation is also consistent with our policy-function estimation, which matches firms' behavior to the same distribution of states. Our modification of the identification condition (14) is thus

$$E_{s \in \hat{S}}[V_i(s; \hat{\sigma}, \theta)] \geq E_{s \in \hat{S}}[V_i(s; \tilde{\sigma}_i, \hat{\sigma}_{-i}, \theta)] \quad \text{for all } \tilde{\sigma}_i \in PF(\alpha_{\tau(i)}, \beta_{\tau(i)}, C_k). \quad (15)$$

We estimate θ by finding values that satisfy (15) for a large number of deviations $\tilde{\sigma}_i(s_i)$. The details of our specific estimator and our choice of deviations are provided in the next subsection.

To gain some intuition for how different deviations can be used to identify the main structural parameters—the managerial costs of monitoring and price changing—consider an artificial market in which two identical firms compete for the top rank by undercutting each other. The firms occupy the top rank stochastically, depending on who monitored most recently. Suppose a firm deviates in a way that increases the frequency of monitoring or changing price. This firm would end up occupying the top rank more than half the time, earning lower average margins on greater average sales. For the equilibrium strategy to dominate this deviation, it must be that the gain in profit, if any, is swamped by increased managerial costs.

It is possible to separately identify the cost of monitoring from the cost of changing prices

because there exist deviations that change the correlation between monitoring and price changes. For example, a spread in the thresholds C_k away from 0 would reduce the frequency of price changes but leave the monitoring rate unchanged. On the other hand, an increase in the coefficient on the constant term in the monitoring equation (6) would increase the rate of monitoring and price changing. Changing both together in certain proportions could increase the monitoring rate leaving the rate of price change constant. Not all deviations can give us useful information. For example, a deviation that leads the firm to earn negative margins a large part of the time would be dominated by the equilibrium strategy for any managerial-cost parameters. We will choose deviations that have a realistic chance of being profitable to maximize the power of our identification assumption (15). The restriction to deviations $\tilde{\sigma}_i(s_i) \in PF(\alpha_{\tau(i)}, \beta_{\tau(i)}, C_k)$ allows for the possibility that firms follow behavioral pricing rules short of fully rational strategies. Fully rational strategies dominate all conceivable deviations, generating higher expected present value of profits than, for example, a one-time price increase of \$1 in any given hour. This deviation is not in $PF(\alpha_{\tau(i)}, \beta_{\tau(i)}, C_k)$, because given the difficulty in even solving for optimal rational strategies in our setting, we are reluctant to examine deviations that only “work” (i.e., generate inequalities in the correct direction) if firms are fully rational. The restricted set of deviations we consider reflects the idea that firms are able to experiment among pricing rules in a simple class to discover the most profitable of them. By restricting the set of deviations, our identification assumption is weaker than the standard assumption in BBL, allowing for estimation of structural parameters that is robust to certain forms of behavioral pricing.

7.2. Estimation Details

To transform our identification condition (15) into an estimator of θ requires empirical analogs to the expectations over value functions appearing there, which we will compute via simulation. Our empirical analogue to the first expectation, $E_{s \in \hat{S}}[V_i(s; \hat{\sigma}, \theta)]$, is

$$\widehat{EV}_i(s; \hat{\sigma}, \theta) = \frac{1}{MN} \sum_{m=1}^M \sum_{n=1}^N \sum_{t_{mn}=0}^{\min\{T, 720\}} \delta^{t_{mn}} \pi(s_{t_{mn}}, \hat{\sigma}(s_{t_{mn}}), \theta). \quad (16)$$

There are three summations in (16). The first sum simulates the expectation over initial states represented by the $E_{s \in \hat{S}}$ operator. We do this by taking M draws from the set of state vectors observed in the data, \hat{S} , and averaging the result (hence the division by M). The second sum simulates the expectation implicit in the value function V_i ; this expectation is over the distribution of all possible histories of game play and private shocks starting from the given initial state. We compute this expectation by simulating N histories for each of the M initial states and averaging

the resulting value functions (hence the division by N). The third sum adds up the profit stream implicit in the value function.

The other expectation, $E_{s \in \hat{S}}[V_i(s; \hat{\sigma}_i, \hat{\sigma}_{-i}, \theta)]$ from the identification condition (15) is similarly transformed into its empirical analogue $\widehat{EV}_i(s; \hat{\sigma}_i, \hat{\sigma}_{-i}, \theta)$. Instead of all firms behaving according to the estimated policy function $\hat{\sigma}$, firm i deviates to the policy function $\tilde{\sigma}_i \in PF(\alpha_{\tau(i)}, \beta_{\tau(i)}, C_k)$ in the simulation, resulting in profit $\pi(s_{t_{mn}}, \tilde{\sigma}_i(s_{t_{mn}}), \hat{\sigma}_{-i}(s_{t_{mn}}), \theta)$ as the new summand in (16).

The upper limit in the third sum is modified from the value function as originally appears in (1). Instead of calculating the value function over an indeterminate number of periods ending with firm i 's exit at time T_i , we just add up the profit during the first month—720 hours to be precise. We do this to reduce the accumulation of simulation errors as the period becomes longer. Given the nature of deviations we are considering, with the firm deviating to a whole new policy function for the entire game, profits are stationary in all simulations. Hence the average per-period profit over 720 hours is an unbiased estimate of the average over the whole game. This is not true of some excluded deviations, for example, a one-time increase in price of \$1; such a deviation could generate a complicated impulse response, which would lead the average profit over a truncated period to diverge that over the full game. Because we are restricting attention to an horizon of about a month, discounting is fairly inconsequential, so we simplify the simulation by taking $\delta = 1$.

Calculating \widehat{EV}_i is further expedited following BBL's insight that when the profit function is linear in the structural parameters, this linearity is inherited by \widehat{EV}_i because it is essentially an average over these linear profits. Thus, for example, we can write $\widehat{EV}_i(s; \hat{\sigma}, \theta)$ as the dot product of two vectors

$$\widehat{EV}_i(s; \hat{\sigma}, \theta) = \overrightarrow{EV}_i(s; \hat{\sigma}) \cdot \vec{\theta}_{\tau(i)}. \quad (17)$$

Here $\overrightarrow{EV}_i(s; \hat{\sigma})$ is a vector with four components, corresponding to the four terms in the profit function in (3). The first component is the sum of the stream of i 's base profits in a simulation, $\sum_{t_{mn}=0}^{\min\{T, 720\}} Base_{it_{mn}}$, averaged over the MN simulations, where firms behave according to their estimated policy functions in each simulation. Similarly, the remaining three components are the averages over the MN simulations of, respectively, the total upselling profit over a simulation, $\sum_{t_{mn}=0}^{\min\{T, 720\}} Upsell_{it_{mn}}$; the number of times i monitored market conditions during a simulation, $\sum_{t_{mn}=0}^{\min\{T, 720\}} Monitor_{it_{mn}}$; and the number of times i changed price during a simulation, $\sum_{t_{mn}=0}^{\min\{T, 720\}} 1\{\Delta_{it_{mn}} \neq 0\}$. The second vector in (17) includes the structural parameters for a firm of type $\tau(i)$: i.e., $\vec{\theta}_{\tau(i)} = (1, v_{\tau(i)}, \mu_{\tau(i)}, \chi_{\tau(i)})$. The linearity in (17) means that the summary statistics $\overrightarrow{EV}_i(s; \hat{\sigma})$ are all that need to be saved from the MN simulations to later compute $\widehat{EV}_i(s; \hat{\sigma}, \theta)$ for any given θ because one just needs to take the dot product of the summary statistics and sub-vectors of the given θ . Hence the simulation and estimation steps can essentially be conducted

independently.

The other expectation estimate can be expressed similarly as

$$\widehat{EV}_i(s; \tilde{\sigma}_i, \hat{\sigma}_{-i}, \theta) = \overline{EV}_i(s; \tilde{\sigma}_i, \hat{\sigma}_{-i}) \cdot \vec{\theta}_{\tau(i)}, \quad (18)$$

in this case conducting the simulations with firm i using its deviation strategy $\tilde{\sigma}_i$ in the simulation while other firms use their estimated policy functions.

With these estimates of the expectations in identification condition (15) in hand, we can proceed to estimation of the structural parameters θ . Following BBL, define Q to be the change in the value function caused by a deviation in strategy:

$$Q(\hat{\sigma}, \tilde{\sigma}_i, \theta_{\tau(i)}) = \left[\widehat{EV}_i(s; \hat{\sigma}, \theta) - \widehat{EV}_i(s; \tilde{\sigma}_i, \hat{\sigma}_{-i}, \theta) \right] \quad (19)$$

$$= \left[\overline{EV}_i(s; \hat{\sigma}) - \overline{EV}_i(s; \tilde{\sigma}_i, \hat{\sigma}_{-i}) \right] \cdot \vec{\theta}_{\tau(i)}, \quad (20)$$

where (20) follows from (17) and (18). The force of identification assumption (15) here is that Q will be non-negative for sufficiently accurate estimates of the expectations and for θ sufficiently close to the true structural parameters. We will estimate θ by assessing a penalty for violations of non-negativity and choosing the value that minimizes the sum of squared penalties over the deviations considered:

$$\hat{\theta} = \underset{\{\vec{\theta}_{\tau(i)} | \tau(i)=1,2,3\}}{\operatorname{argmin}} \left\{ \sum_{\{\tilde{\sigma}_i | \tau(i)=1,2,3\}} \left[\min\{Q(\hat{\sigma}, \tilde{\sigma}_i, \theta_{\tau(i)}), 0\} \right]^2 \right\}. \quad (21)$$

For each firm type, we considered a set of 1,800 deviations σ_i from the set of policy functions $PF(\alpha_{\tau(i)}, \beta_{\tau(i)}, C_k)$. A third of these were generated by multiplying the coefficients $\hat{\alpha}_{\tau(i)}$, $\hat{\beta}_{\tau(i)}$ and the cut points \hat{C}_k in the estimated policy function by log-uniform noise terms. A third were generated by, in addition to multiplying the log-uniform noise terms, adding correlated perturbations to the constant term in $\hat{\alpha}_{\tau(i)}$ and the cut points \hat{C}_k , such that the deviations amount to an experiment with changing the frequency of monitoring and frequency of price change conditional on monitoring in opposite ways. The last third were generating by, in addition to multiplying the log-uniform noise terms, adding correlated perturbations to the cut points \hat{C}_k , such that the deviations amount to experiments with trade off between small and large price changes. For each deviation $\tilde{\sigma}_i$, we calculate the value function statistics $\overline{EV}_i(s; \tilde{\sigma}_i, \hat{\sigma}_{-i})$ with $M = 10,000$ random draws of initial states and $N = 1$ simulation for each initial state. For the estimated policy, we calculate the value function statistics $\overline{EV}_i(s; \hat{\sigma})$ using the actual sample of observed states for each type for all initial states and

$N = 20$ simulations for each state.

We bootstrap the standard error of $\hat{\theta}$ by subsampling from the 1800 random deviations. Because we have chosen a large number MN of simulations for each deviation, the expression $\overrightarrow{EV}_i(s; \hat{\sigma}) - \overrightarrow{EV}_i(s; \tilde{\sigma}_i, \hat{\sigma}_{-i})$ is estimated with very small standard error (around 5%) for a given deviation. The randomness in $\hat{\theta}$ is therefore mainly caused by the choice of deviations, as some deviations are more informative or cause greater violation of optimality than others. We randomly select a bootstrapped sample of 900 deviations with replacement from our original set and estimate θ using this bootstrapped sample. We perform this procedure 1,000 times to obtain the bootstrapped distribution of each estimated structural parameter.⁷ Because the random errors in our estimate of the policy function and the estimates of $Q(Rank_{it})$ and $U(Rank_{it})$ we cited from Ellison and Ellison (2009a) could contribute additional variance to the structural estimates, we note that our calculation of the standard error is an optimistic approximation. We feel these additional sources of errors do not affect the structural estimation as large as randomness in choice of deviations, and given the complexity of the model and the computational burden, it would be very difficult to incorporate those errors rigorously.

7.3. Results

Table 4 presents the estimates of the structural parameters from the profit equation (3). Recall that we rely on estimates of $Base_{it}$ and $Upsell_{it}$ from previous research, leaving only managerial costs μ_τ and ξ_τ as the only structural parameters to be estimated. We measure μ_τ and ξ_τ in dollars. Their confidence intervals are obtained by taking quantiles from the subsample estimates.

First consider the estimates of the monitoring cost μ_τ . Type 1 and 3 firms are estimated to have similar monitoring costs, both around \$60 per episode. For type 1 firms, their monitoring costs are about three times the maximum profit they can earn in an hour. Type 2's monitoring costs are significantly lower, around \$10. The magnitudes of these estimates may seem surprisingly large—could it really cost a manager \$60 to log onto a website and do a few calculations?—but a consideration of opportunity cost suggests that these magnitudes are at least plausible. These managers could have dozens of other products to monitor, disputes to resolve, personnel matters to address, and so forth. Even a few moments of their time could be quite valuable.

There is, however, a nonintuitive pattern to the relative magnitudes. Type 2 firms have the lowest estimated monitoring costs but they are the least active of the firms. For this reason, it is important to consider these costs together with the other managerial cost, χ_τ , the cost of changing

⁷In total we execute some 55 million simulations. Each simulation takes about 0.2 seconds, so in total this would take a couple months of computing on a single core. Actual computing time was reduced significantly by running simulations in parallel on a multi-core server.

price. Again, types 1 and 3 have similar estimated costs, around \$10 or less. Note, however, that we cannot reject that either is equal to zero. The managerial cost structures of types 1 and 3 are consistent with the intuition that the internet technology has almost eliminated the physical menu costs, while the cognitive costs occurring in decision making persist. Type 2's cost of price change is much higher, around \$160, meaning that their total managerial costs⁸ are substantially higher, \$168.78 versus \$76.64 and \$59.53. This total is consistent with our intuition, even if the cost of monitoring taken on its own is not.

It is interesting to speculate on what could be driving this very different split of managerial costs across firm types. The explanation could be a simple econometric one: we have many fewer observations of type 2 firms and those firms are also less active, so perhaps we simply do not have enough of the type of variation we need to robustly identify these two types of costs separately for type 2. An explanation more grounded in economics could be that type 2 firms really do behave quite differently because, say, their websites are much more difficult to alter when making a price change. (This could happen if the price were hardwired into the programming of the website and needed to be changed in various locations in the code.) We still find it puzzling, though, that type 2 firms' cost of monitoring would be less than the other two types.

It is also worth noting that the total managerial costs are quite large, comparable to, perhaps, an hour or two (or more) of wages for a typical manager. One implication of this observation is that a particular firm can gain a lot if adjusting price becomes easier for the firm. In the years following the time period of our data, firms did exactly that by adopting automated pricing software.

8. Counterfactuals

To assess the importance of managerial costs for the firm's pricing behavior, in this section we present counterfactual exercises in which we shock a single firm i 's managerial costs and see how its pricing behavior and profit change, leaving other firms the same as before. In effect, this section performs the inverse exercise from the structural estimation in the previous section. In the structural estimation, we estimated a policy function from actual pricing behavior and used this to infer firms' managerial costs. Here, we posit a vector of managerial costs for i and search over policy functions for the one maximizing its simulated profits. For each candidate policy function, we want to simulate i 's new pricing behavior and profit assuming that all other firms continue with their originally estimated managerial costs, equivalent to assuming that their pricing

⁸The total managerial cost here sums up the cost of one monitoring and the cost of one price change for simplicity. In fact, according to our estimated policy, type 1, 2 and 3 firms on average monitor 1.4, 2.3 and 1.5 times per price change, respectively. At these ratios the comparison still holds.

behavior is given by the originally estimated policy functions. This is exactly the simulation we did for the 1800 deviations during the structural estimation, so we can treat these deviations as candidate policy functions, and use the recorded simulations to recalculate simulated profits with the new managerial costs for these deviations, and pick the one with the highest profits as the firm's optimal policy at the new managerial costs. In fact, for this purpose, we did a much finer search by simulating for additional several thousands candidate policy functions that are drawn sequentially from the proximity of the existing promising ones. To save space we will focus on the case in which i is a type-1 firm. The results are shown in Figure 7.

Panel A shows the monthly number of price changes for the firm as the costs of monitoring and price change vary from 0 to 200 dollars. With a large yet finite pool of high quality candidate policy functions, we are able to identify 49 policy functions as optimal for some given vectors of managerial costs. The surface is therefore marked by 49 plateaus, which indicate regions of costs for which the same pricing strategy is optimal. As managerial costs increase, the plateaus cover broader regions, reflecting the fact that a given increase in the cost level represents a smaller percentage increase. Even as costs of monitoring and price changing go to 0, the frequency of price change does not grow without bound. In other words, in the complete absence of frictions associated with price change, it is not optimal for a firm to continuously tweak its price. There are three reasons for this. First, other firms have retained their positive costs, so they do not respond continuously. This results in stretches when state variables do not change during which i 's optimal strategy is to keep price constant. Second, changing price, especially downwards, may trigger other firms' reaction and intensity future competition, so there is a dynamic incentive to wait for a while between price changes. Third, prices in this market are posted in whole dollar amounts, so even if a firm continuously monitored its optimal continuous price, it still would not want to change the price until the optimal price exceeded the threshold necessary to move the price a whole dollar up or down. When i has no costs of monitoring or price changing, it ends up changing price around once a day.

Panel B explores the same counterfactual exercise but now focuses on a different outcome variable: i 's monthly profits. These are the net profits from equation (3), which subtract off the new managerial costs with which we are shocking firm i , accumulated over the month. While the surface in the previous panel had discrete jumps reflecting the discrete changes to firm i 's pricing policy, the continuous changes in i 's costs smooth out its profit function in this panel. We highlight several important features of the graph. First, it shows that the division between the two types of managerial costs matters for profit. Fixing the total managerial cost of a joint episode of monitoring and price changing at say \$200, its monthly profits would vary from about \$1,800 to

\$2,400, depending on how the \$200 was divided between the two costs. Second, the height of the surface indicates the potentially large gain to adopting technologies that decrease managerial costs, potentially thousands of dollars a month just for this one product. In particular, based on numbers in this graph, the monthly net profits would increase \$1620 if the managerial costs decrease from the estimated levels to zero. One would under-estimate the profits improvement assuming the firm simply saves the managerial cost, which is about a half of \$1620, without changing the policy. Reoptimizing the policy function to a more proactive one contributes the other half of the profits improvement. The gain would presumably be multiplied if the technology could be used to reduce managerial costs for the scores of other products the retailers marketed on Pricewatch. In fact the retailers did move to automated pricing soon after the time period of our data, consistent with our estimates of potentially large gains.

9. Conclusion

We have provided a framework for estimating the costs of managerial attention and activity in the particular case of price changes. We examine a setting where managers can change prices at will and face rapidly changing game states but still seem to exhibit a large degree of inertia in pricing. We build a dynamic model of firm decision-making which features boundedly rational firms who only consider a subset of all state variables each period. The model also allows for a limited degree of firm heterogeneity. Using this model along with assumptions about the (bounded) optimality of firms' decisions, we can use their observed actions to back out estimates of managerial costs. These cost estimates vary by firm type but are quite substantial, ranging from \$60 to \$169. For the majority of the firms (type 1 and 3), the managerial costs primarily occur in the monitor stage, which involves information collection and decision making, and the cost of carrying out the price change online is not significant, perhaps due to the technology advantage of e-commerce.

This paper fills a number of existing gaps in the economics literature. First, it incorporates a recognition that firms are often limited in the information they can process and the calculations they can perform. As such, it is the first paper to our knowledge which uses structural techniques to estimate quantities of interest but does not assume that firms calculate optimal responses based on a full set of state variables. We feel that, in some settings, these assumptions on firm behavior are excessive and unrealistic, and we wanted to offer an alternative where firms used simple, robust, rule of thumb decision rules. Second, we address some of the same questions posed in an emerging empirical macroeconomics literature on why retail prices are so sticky. Like that literature, we suggest that the costs of managerial attention and activity may be important. Our novel contribution is a carefully specified dynamic model of pricing behavior that can generate actual estimates of

these costs. Finally, we contribute to a broader literature in behavioral economics in two ways: we allow firms to behave in a boundedly rational way, which we think is consistent with both formal and casual empiricism, and, since we provide an estimate of the managerial cost of changing price, one could compare our estimates with “engineering estimates,” giving us insight into the psychological barriers of taking an affirmative action such as changing a price.

References

- Aguirregabiria, Victor and Pedro Mira. (2007) “Sequential Estimation of Dynamic Discrete Games,” *Econometrica* 75: 1–53.
- Arbatskaya, Maria and Michael Baye. (2004) “Are Prices ‘Sticky’ Online? Market Structure Effects and Asymmetric Responses to Cost Shocks in Online Mortgage Markets,” *International Journal of Industrial Organization* 22: 1443–1462.
- Artinger, Florian and Gerd Gigerenzer. (2012) “Aspiration-adaption, Price Setting, and the Used Car Market,” mimeo.
- Bajari, Patrick C., Lanier Benkard, and Jonathan Levin. (2007) “Estimating Dynamic Models of Imperfect Competition,” *Econometrica* 75: 1331–1370.
- Bils, Mark and Peter Klenow. (2004) “Some Evidence on the Importance of Sticky Prices,” *Journal of Political Economy* 112: 947–985.
- Blinder, Alan, Eile Canetti, David Lebow, and Jeremy Rudd. (1998) *Asking About Prices—A New Approach to Understanding Price Stickiness*. New York: Russell Sage Foundation.
- Borenstein, Severin, A. Colin Cameron, and Richard Gilbert. (1997) “Do Gasoline Prices Respond Asymmetrically to Crude Oil Price Changes?” *Quarterly Journal of Economics* 112: 305–339.
- Carlton, Dennis. (1986) “The Rigidity of Prices,” *American Economic Review* 76: 637–658.
- Castanias, Rick and Herb Johnson. (1993) “Gas Wars: Retail Gasoline Price Fluctuations,” *Review of Economics and Statistics* 75: 171–174.
- Cavallo, Alberto and Roberto Rigobon. (2012) “The Distribution of the Size of Price Changes,” National Bureau of Economic Research Working Paper no. w16760.
- Chakrabarti, Rajesh and Barry Scholnick. (2005) “Nominal Rigidities without Literal Menu Costs: Evidence from E-Commerce,” *Economics Letters* 86: 187–191.
- Cyert, Richard and James March. (1963) *Behavioral Theory of the Firm*. Oxford: Blackwell Publishers.
- Eichenbaum, Martin, Nir Jaimovich, and Sergio Rebelo. (2011) “Reference Prices, Costs, and Nominal Rigidities,” *American Economic Review* 101: 234–262.
- Ellison, Glenn and Sara Fisher Ellison. (2009a) “Search, Obfuscation, and Price Elasticities on the Internet,” *Econometrica* 77: 427–452.
- Ellison, Glenn and Sara Fisher Ellison. (2009b) “Tax Sensitivity and Home State Preferences in Internet Purchasing,” *American Economic Journal: Economic Policy* 1: 53–71.
- Ellison, Sara Fisher and Christopher M. Snyder. (2013) “An Empirical Study of Pricing Strategies in an Online Market with High-Frequency Price Information,” mimeo.

- Harris, Mark N. and Xueyan Zhao. (2007) “A zero-inflated ordered probit model, with an application to modelling tobacco consumption,” *Journal of Econometrics* 141: 1073-1099.
- Klenow, Peter J. and Benjamin A. Malin. (2011) “Microeconomic Evidence on Price-Setting,” in Benjamin Friedman and Michael Woodford, eds., *Handbook of Monetary Economics*, vol. 3. Amsterdam: North-Holland, 231–284.
- Lewis, Matthew. (2008) “Price Dispersion and Competition with Differentiated Sellers,” *Journal of Industrial Economics* 56: 654–678.
- Lünneman, Patrick and Ladislav Wintr. (2006) “Are Internet Prices Sticky?,” ECB Working Paper no. 645.
- Means, Gardiner. (1935) “Industrial Prices and Their Relative Inflexibility,” U.S. Senate Document 13, 74th Congress, 1st Session.
- Nakamura, Emi and Jón Steinsson. (2008) “Five Facts About Prices: A Reevaluation of Menu Cost Models,” *Quarterly Journal of Economics* 123: 1415–1464.
- Noel, Michael. (2007a) “Edgeworth Price Cycles, Cost-Based Pricing, and Sticky Pricing in Retail Gasoline Markets,” *Review of Economics and Statistics* 89: 324–334.
- Noel, Michael. (2007b) “Edgeworth Price Cycles: Evidence from the Toronto Retail Gasoline Market,” *Journal of Industrial Economics* 55: 69–92.
- Pakes, Ariel, Jack Porter, Katherine Ho, and Joy Ishii. (2015) “Moment Inequalities and Their Application,” *Econometrica* 83: 315–334.
- Romesburg, H. Charles. (2004) *Cluster Analysis for Researchers*. North Carolina: Lulu Press.
- Simon, Herbert. (1955) “A Behavioral Model of Rational Choice,” *Quarterly Journal of Economics* 69: 99–118.
- Simon, Herbert. (1962) “New Developments in the Theory of the Firm,” *American Economic Review* 52: 1–15.
- Stigler, George J. and James K. Kindahl. (1970) *The Behavior of Administered Prices*. New York: Columbia University Press.
- Turkel, Studs. (1970) *Hard Times: An Oral History of the Great Depression*. New York: Pantheon Books.
- Wang, Zhongmin. (2009) “(Mixed) Strategy in Oligopoly Pricing: Evidence from Gasoline Price Cycles Before and Under a Timing Regulation,” *Journal of Political Economy* 117: 987–1030.

Table 1: Descriptive Statistics for Dataset

Variable	Definition	Mean	Std. dev.	Min.	Max.	Obs.
Firm-level variables						
<i>Density</i>	Measure of density in price space of firms with nearby ranks	0.60	0.40	0	3	111,276
<i>Margin</i>	Percentage markup over wholesale cost, $100(\text{Price} - \text{Cost})/\text{Cost}$	1.01	5.65	-20.50	20.38	111,276
<i>NumBump</i>	Net number of ranks bumped since last price change	1.30	3.40	-22	21	111,276
<i>Placement</i>	Placement between adjacent firms in price space	0.58	0.42	0	1	111,276
<i>Price</i>	Current listed price in dollars	69.0	35.1	24.2	132.6	111,276
<i>QuantityBump</i>	Relative change in hourly sales resulting from rank bump	-0.16	0.36	-2.08	2.48	111,276
<i>Rank</i>	Rank of listing in price-sorted order	10.75	6.77	1	24	111,276
<i>RankOne</i>	Indicates whether firm is at rank 1	0.06	0.24	0	1	111,276
<i>SinceChange</i>	Hours since firm last changed price	117.55	146.45	1	1,113	111,276
Market-level variables						
<i>Cost</i>	Wholesale cost	70.44	36.35	24	129	7,018
<i>CostTrend</i>	Trend in <i>Cost</i> over previous two weeks	-0.17	0.74	-2.06	1.53	7,108
<i>CostVol</i>	Volatility of <i>Cost</i> over previous two weeks	1.73	1.10	0.00	4.36	7,018
<i>Night</i>	Indicates hour from midnight to 8 a.m. EST	0.33	0.47	0	1	7,018
<i>Weekend</i>	Indicates Saturday or Sunday	0.28	0.45	0	1	7,108

Table 2: Selected Descriptive Statistics by Firm Type

Variable	Type 1 (22 firms)		Type 2 (8 firms)		Type 3 (13 firms)	
	Mean	Std. dev.	Mean	Std. dev.	Mean	Std. dev.
<i>Margin</i>	-0.72	5.02	1.24	5.54	6.20	4.18
<i>Rank</i>	7.78	5.35	13.37	6.20	18.06	4.34
<i>SinceChange</i>	99.88	130.12	225.08	200.43	93.33	107.80
<i>NumBump</i>	1.00	3.08	2.12	4.46	1.64	3.32

Table 3: Maximum Likelihood Estimates of Policy Function

Variable	Firms combined		Type-1 firms		Type-2 firms		Type-3 firms	
	Coefficient	Std. err.	Coefficient	Std. err.	Coefficient	Std. err.	Coefficient	Std. err.
Monitoring estimates α								
Constant	-2.37***	(0.14)	-2.25***	(0.10)	-1.71***	(0.54)	-2.85***	(0.25)
Night	-0.73***	(0.13)	-0.68***	(0.05)	-0.84***	(0.23)	-1.30***	(0.21)
Weekend	-0.49***	(0.07)	-0.47***	(0.04)	-0.61***	(0.20)	-0.71***	(0.11)
CostVol	0.03***	(0.01)	0.04***	(0.01)	0.04	(0.06)	0.05*	(0.03)
CostTrend ⁺	0.14***	(0.03)	0.13***	(0.04)	0.33**	(0.15)	0.18	(0.12)
CostTrend ⁻	0.12***	(0.04)	0.09***	(0.03)	0.02	(0.21)	0.24***	(0.07)
QuantityBump ⁺	0.44***	(0.08)	0.39***	(0.08)	0.69***	(0.26)	0.13	(0.35)
QuantityBump ⁻	0.39***	(0.08)	0.36***	(0.05)	0.15	(0.26)	0.20	(0.19)
InSinceChange	0.27***	(0.08)	0.25***	(0.05)	-0.23	(0.17)	0.47***	(0.12)
(InSinceChange) ²	-0.05***	(0.01)	-0.05***	(0.01)	0.02	(0.02)	-0.07***	(0.02)
Price change estimates β								
CostTrend	0.05	(0.05)	0.11**	(0.05)	0.25	(0.22)	-0.31**	(0.13)
CostChange	0.08***	(0.01)	0.06***	(0.01)	0.06	(0.04)	0.09***	(0.03)
Margin	-0.02	(0.01)	-0.02**	(0.01)	0.04	(0.04)	-0.11***	(0.03)
NumBump	-0.05***	(0.02)	-0.05***	(0.02)	0.07	(0.05)	-0.05	(0.04)
Density \times NumBump	-0.10***	(0.03)	-0.10***	(0.04)	-0.28***	(0.08)	-0.06	(0.06)
Placement	0.18*	(0.10)	0.30***	(0.08)	0.19	(0.32)	0.07	(0.17)
Rank	-0.02**	(0.01)	-0.04***	(0.01)	-0.09**	(0.04)	-0.05**	(0.02)
RankOne	0.58***	(0.13)	0.62***	(0.11)	<i>a</i>	<i>a</i>	<i>a</i>	<i>a</i>
Cutoff C ₋₁	-0.09	(0.13)	-0.15	(0.14)	-1.80	(0.48)	-1.62	(0.37)
Cutoff C ₁	0.64***	(0.12)	0.76***	(0.14)	0.55	(0.87)	-0.44	(0.43)
Log likelihood	-8,025.7		-6,174.2		-383.0		-1,337.9	
Observations	111,276		71,460		16,904		22,912	

Notes: Maximum likelihood estimates of coefficients from equations (6) and (10). Heteroskedasticity-robust standard errors clustered by firm reported in parentheses. *Because most of the observations with *RankOne* = 1 are in group 1, equations estimated for groups 2 and 3 constrain *RankOne* coefficient to be the same as estimated for group 1, 0.62. Model includes cutoffs C_k for $k \in \{-5, -4, -3, -2, -1, 1, 2, 3, 4, 5\}$; for space considerations we only report C_{-1} and C_1 . Statistically significant in a two-tailed test at the * 10% level, ** 5% level, *** 1% level.

Table 4: Estimates of the Structural Parameters

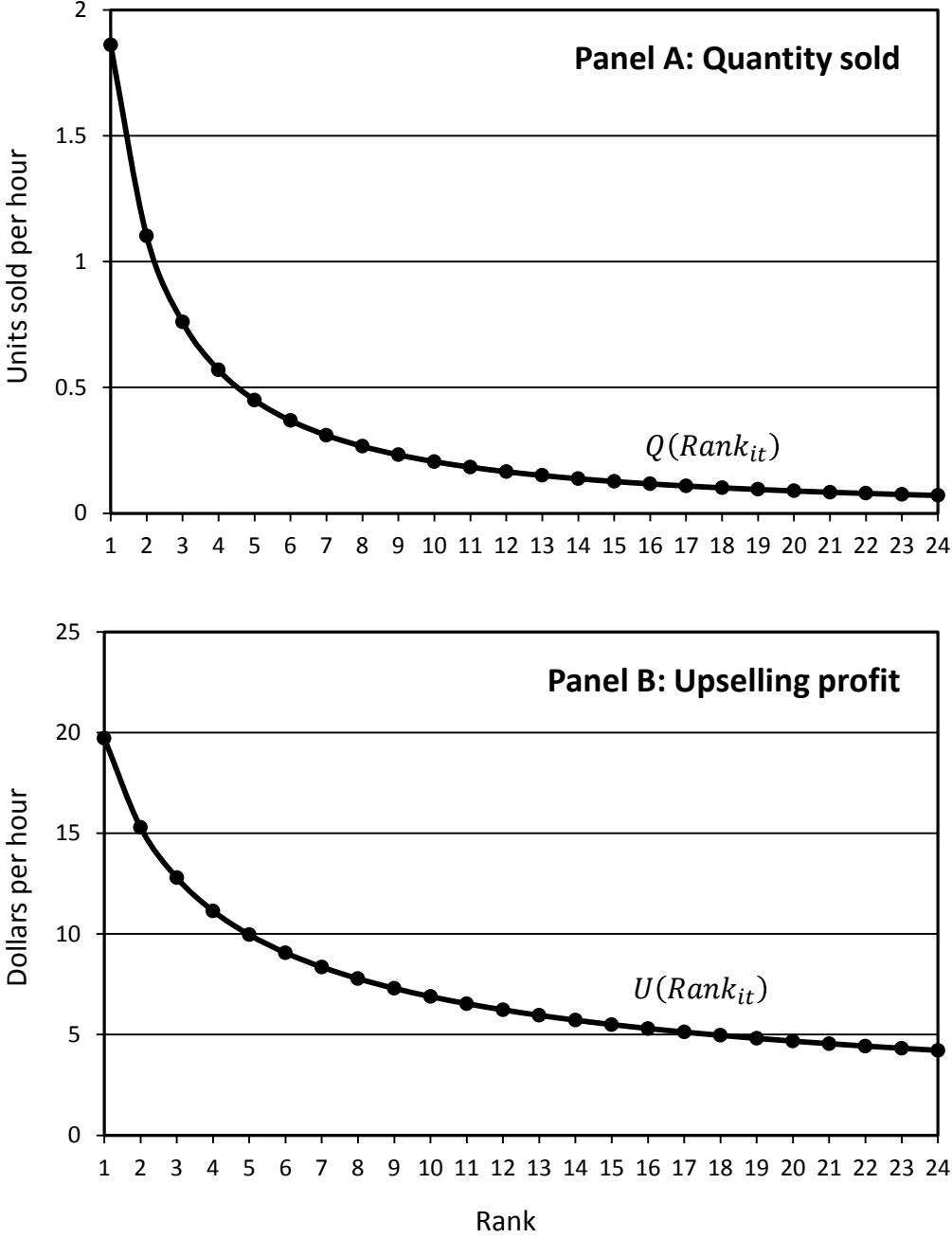
Parameter	Type-1 firms	Type-2 firms	Type-3 firms
Cost of monitoring, μ	66.35	10.38	56.78
95% confidence interval	[55.11, 84.07]	[2.22, 18.98]	[47.76, 69.71]
Cost of changing price, χ	10.29	158.40	2.75
95% confidence interval	[-14.09, 26.54]	[134.30, 184.70]	[-15.02, 17.76]

Figure 1: Example Pricewatch webpage

BRAND	PRODUCT	DESCRIPTION	PRICE	SHIP	DATE/HR	DEALER/PHONE	ST	PART#
<u>Generic</u>	PRICE FOR ONLINE ORDERS ONLY - 128MB PC100 SDRAM DIMM - 8ns Gold Leads	- * LIMIT ONE - Easy installation - in stock	\$ 68	9.69 INSURED	10/12/00 12:40:05 AM CST	Computer Craft Inc. 800-487-4910 727-327-7559 Online Ordering	FL	MEM-128-100PCT
<u>Generic</u>	ONLINE ORDERS ONLY - 128MB SDRAM PC100 16x64 168pin	- * LIMIT ONE	\$ 69	INSURED\$9.95	10/11/00 10:59:56 PM CST	Connect Computers 888-277-6287 949-367-0703 Online Ordering	CA	-
<u>Generic</u>	PRICE FOR ONLINE ORDER - 128MB PC100 SDRAM DIMM	- * LIMIT ONE - - InStock, 16x64-Gold Leads	\$ 70	10.75	10/11/00 2:11:00 PM CST	1st Choice Memory 949-888-3810 -- P.O.'s accepted Online Ordering	CA	-
<u>Generic</u>	PRICE FOR ONLINE ORDER - 128mb True PC100 SDRAM EEPROM DIMM16x64 168pin 6ns/7ns/8ns Gold Leads	- * LIMIT ONE - in stock - with Lifetime Warranty	\$ 72	9.85	10/10/00 11:30:39 AM CST	pcboost.com 800-382-6678 -- P.O.'s accepted Online Ordering	CA	-
<u>Generic</u>	IN STOCK, 128MB PC100 3.3volt unbuffered SDRAM Gold Lead 168 Pin, 7/8ns - with Lifetime warranty	- * LIMIT ONE Not compatible with E Machine	\$ 74	10.95- UPS INSURED	10/11/00 12:44:00 PM CST	Memplus.com 877-918-6767 626-918-6767	CA	- 880060
<u>Generic</u>	PRICE FOR ONLINE ORDERS ONLY - 128MB True PC100 SDRAM DIMM - 8ns Gold - warranty	- * LIMIT ONE	\$ 74	10.25	10/9/00 6:53:25 PM CST	Portatech 800-487-1327 .	CA	-
<u>House Brand</u>	128MB PC100 3.3volt SDRAM 168 Pin, 7/8ns - with LIFETIME WARRANTY	- * LIMIT ONE	\$ 74	10.50 FedEx	10/11/00 10:20:23 AM CST	1st Compu Choice 800-345-8880 800-345-8880	OH	-
<u>Generic</u>	128MB 168Pin TRUE PC100 SDRAM - OEM 16X64	DIMM16x64 168pin 6ns/7ns/8ns Gold Leads	\$ 75	\$10	10/11/00 2:37:00 PM CST	Sunset Marketing, Inc. 800-397-5050 410-626-0211 -- P.O.'s accepted	MD	-
<u>Generic</u>	128MB 16x64 PC100 8ns SDRAM.	- * LIMIT ONE	\$ 77	10.90	10/12/00 9:37:45 AM CST	PC COST 800-877-9442 847-690-0103 Online Ordering	IL	-
<u>Generic</u>	IN STOCK, PC100, 128MB, 168pins DIMM NonECC, - with Lifetime warranty	- * LIMIT 5	\$ 77	\$10.95 & UP For UPS Ground	10/9/00 5:11:10 PM CST	Augustus Technology, Inc 877-468-5181 909-468-1883 Online Ordering	CA	-
<u>Generic</u>	128MB PC100 8NS 16x64 SDRAM - one year warranty	- * LIMIT ONE	\$ 78	Ups Ground \$10.62	10/11/00 5:16:36 PM CST	Computer Super Sale 800-305-4930 847-640-9710 Online Ordering	IL	-
<u>Generic</u>	PRICE FOR ONLINE ORDERS ONLY - PC100 128MB NonBuffered, NonECC 16x64 SDRAM DIMM 3.3V 8ns mem module	- * LIMIT ONE - with lifetime warranty	\$ 78	10.95	10/5/00 6:29:59 PM CST	Jazz Technology USA, LLC 888-485-8872 909-869-8859	CA	ME-GBP100128

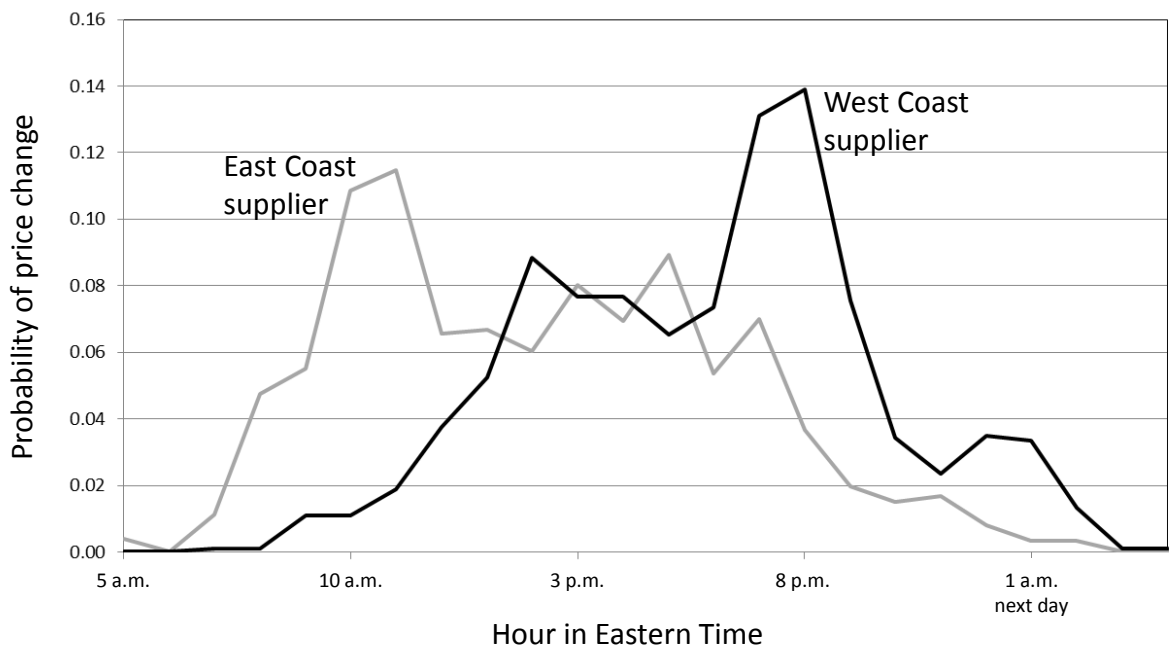
Notes: Page downloaded October 12, 2000.

Figure 2: Effect of rank on various retailer outcomes



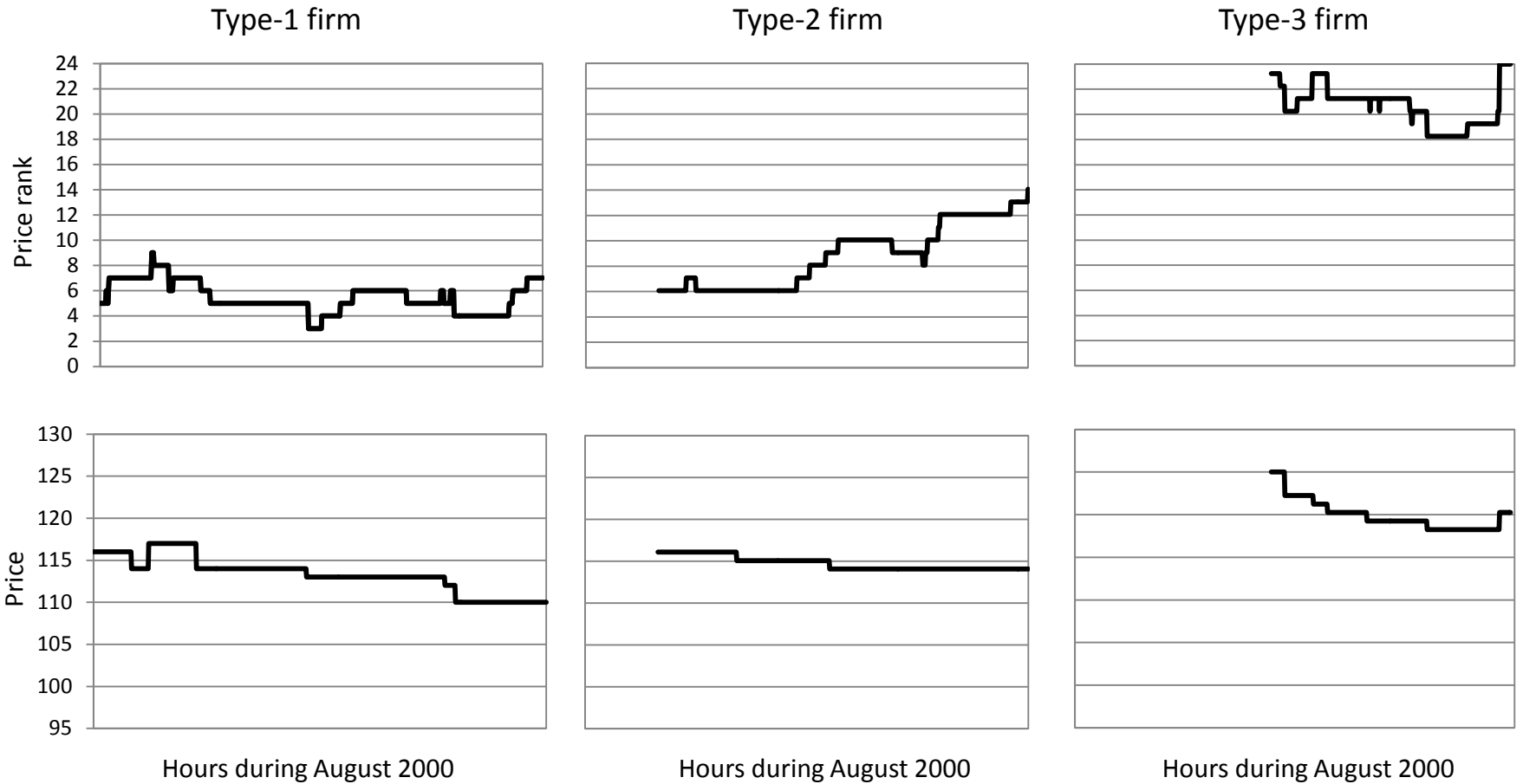
Notes: Panel A is derived from Ellison and Ellison’s (2009a) demand estimates, based on their regression of the natural log of quantity on linear rank. Exponentiating yields the equation graphed, $Q(Rank_{it}) = 4.56(1 + Rank_{it})^{-1.29}$. Panel B is derived from Ellison and Ellison’s (2009a) estimates of upselling profit. They estimate that a retailer sells an additional $0.97(1 + Rank_{it})^{-0.77}$ units of a medium-quality product with an average markup of \$15.69 and an additional $0.49(1 + Rank_{it})^{-0.51}$ units of a high-quality product with an average markup of \$31.45, for total upselling profit of $U(Rank_{it}) = 15.18(1 + Rank_{it})^{-0.77} + 15.48(1 + Rank_{it})^{-0.51}$, the equation graphed.

Figure 3: Price-changing activity during the day for retailers on different coasts



Notes: Reproduced from Figure 6B of Ellison and Snyder (2013), showing the residual probability of price change each hour estimated after partialling out other covariates. So that the probability functions integrate to 1, they were converted into conditional probabilities, i.e., conditional on a price change occurring during the day.

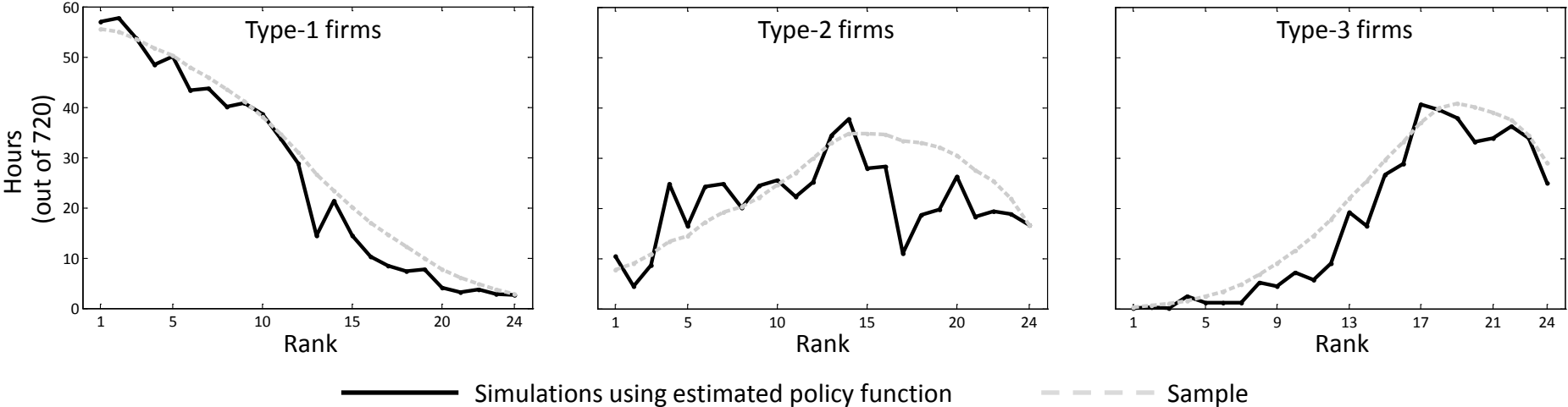
Figure 4: Price and rank series for representative firms of each type



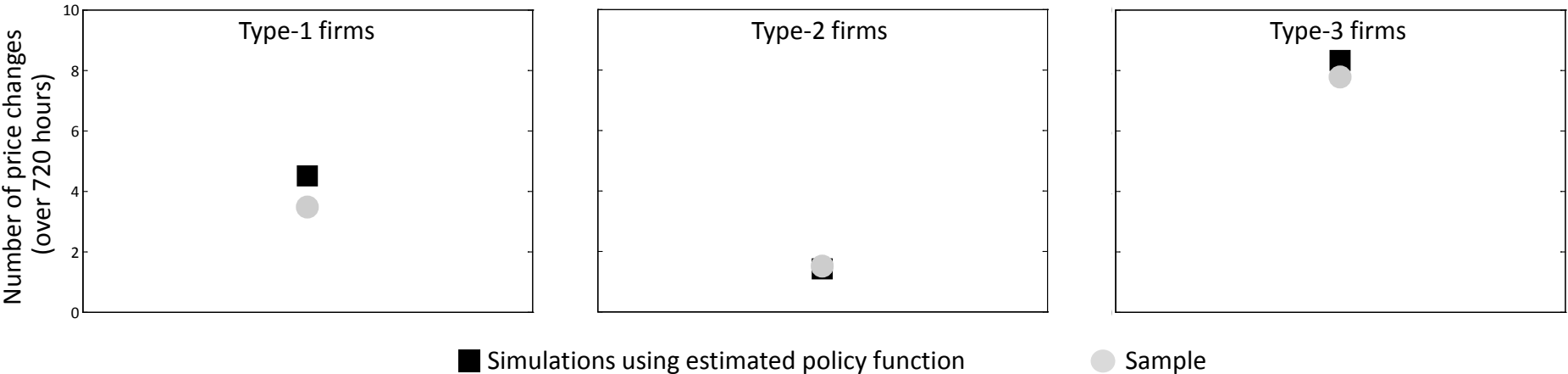
Notes: Reproduced from Figure 7 of Ellison and Snyder (2013), showing price and rank series for representative firm of each type.

Figure 5: Goodness of fit of estimated policy function by firm type

Panel A: Fit to time spent at each rank

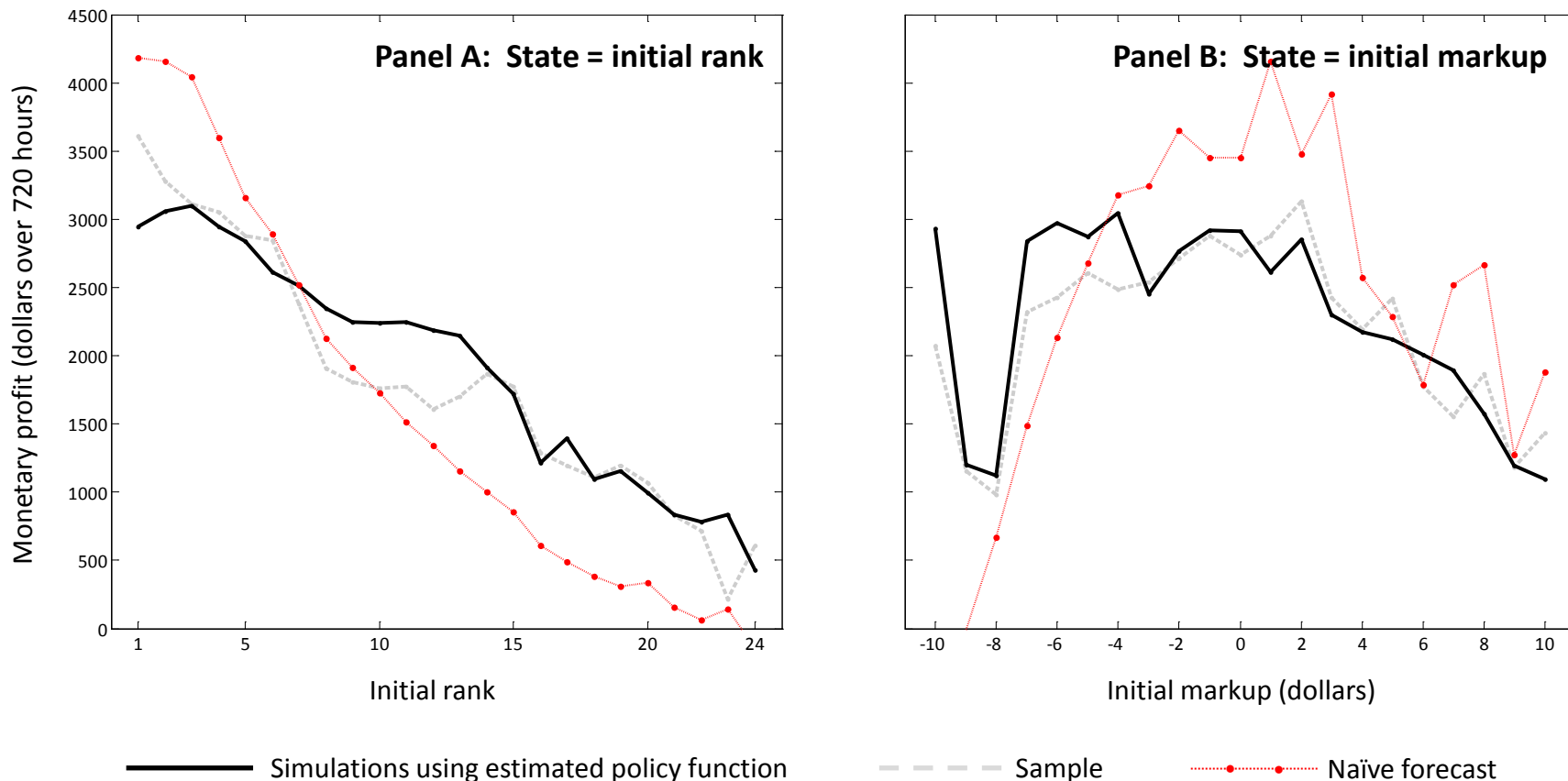


Panel B: Fit to number of price changes



Notes: Variables on vertical axis are discounted using same 0.95 annual factor used to compute value functions in the structural estimation. Over the short 720-hour horizon, discounting has a negligible effect on the graphs.

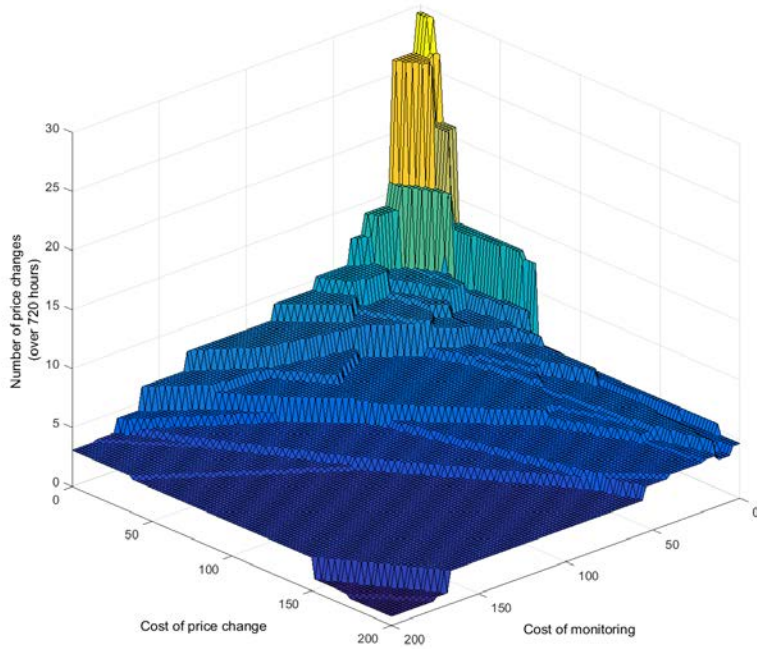
Figure 6: Comparing fit of estimated policy function to deviation across various initial states



Notes: Monetary profit involves just the first two components of π_{it} , $Base_{it} + Upsell_{it}$, not the managerial costs structurally estimated later. Monetary profit for the actual sample is also estimated but estimated based on firms' actual prices as opposed to the simulations, which use prices from the estimated policy function. Profits discounted using same 0.95 annual factor used to compute value functions in the structural estimation. Over the short 720-hour horizon, discounting has a negligible effect on the graphs.

Figure 7: Counterfactual scenarios with different managerial costs

Panel A: Number of price changes



Panel B: Profit net of managerial costs

