

A Dynamic Model of Prices and Margins in the LCD TV Industry

Christopher T. Conlon

November 2, 2012

Abstract

Inter-temporal tradeoffs are an important part of the consumer decision making process. These tradeoffs are especially important in markets for high-tech consumer goods where prices and costs fall rapidly over time. These tradeoffs are important not only for understanding consumer behavior, but for understanding firm pricing behavior as well. Prices in markets for durable goods may be low for one of several reasons. For example, prices may be low because costs are low, prices may be low because high-value consumers have already made purchases, or prices may be low because consumers anticipate lower prices in the future. This paper estimates a dynamic model of demand and supply in order to measure the relative magnitudes and importance of these effects in the market for LCD Televisions. This paper contributes to the existing empirical literature on dynamic durable goods in three ways. It improves the estimation by employing an empirical likelihood estimator instead of a generalized method of moments estimator. It improves the computation by recasting the problem in the language of constrained optimization, which makes the dynamic problem only slightly more difficult to solve than the static problem. Finally, it makes use of reliable marginal cost data from the LCD television industry, which makes it possible to re-compute markups under counterfactual scenarios without worrying about recovering marginal costs from the model.

1 Introduction

An important aspect of consumer choice is that for many products, consumers face inter-temporal tradeoffs. That is consumers may buy a good today and enjoy the consumption value for several periods or they may wait and make a purchase later when prices are likely to be lower. The option to wait is important in understanding consumer choices for high-technology products, where markets are often characterized by rapidly increasing product quality and rapidly decreasing prices. Consumers decide not only which product to purchase, but also when to make a purchase.

It is important to understand the relationship between dynamic consumer behavior and firm pricing decisions for two reasons. The first is that the ability of firms to extract rents is an important driver of innovation and research and development in high-technology markets. The second is the Coasian argument that when firms compete with their own products over time, it reduces the effects of market power. The ability to decompose dynamic consumer behavior can help us understand why prices are low (or high) in durable goods market. For example, prices may be low because costs are low; prices may be low because consumers have a low cost of waiting; or prices may be low because firms have already sold to high-value consumers. These three possibilities can have very different implications for policy makers and firms.

This paper focuses on how dynamic consumer behavior influences the prices firms charge (or would charge) in equilibrium by decomposing consumer dynamics into two major parts. The first aspect is that as consumers make purchases, the distribution of remaining consumers changes over time. This gives firms an incentive to engage in inter-temporal price discrimination. Firms set initial prices high, and sell only to consumers with the highest valuations. Over time, firms reduce prices and sell to the highest value consumers in the remaining population. This leads to a decline in prices over time. The second aspect is that when prices are decreasing, consumers have an option value associated with waiting. If prices are too high, consumers can wait and make a purchase at a later time. In addition to responding to these two effects, firms also take into account how prices affect both the future distribution of consumers and the perceived value option value associated with waiting.

This paper examines these dynamics through the lens of the LCD television industry. This is a large industry that has been characterized by substantial process improvements and rapid declines in price. One nice feature of the industry that this paper is able to exploit is that the primary input into LCD Televisions are LCD panels; which provide information about how the costs of LCD televisions change over time. A second feature of the industry is that prior to 2006, there were no LCD televisions, thus partially alleviating the initial conditions problem faced in many dynamic settings. In the period from 2002-2006, the LCD panel makers (providing displays for phones and laptops) were widely known to have engaged in one of the largest price fixing conspiracies of all time. This paper looks at the first two generations of the LCD-TV industry that followed the end of the price-fixing regime from 2006-2009. During this period, it was widely reported that the industry invested on the order of \$50 Billion, by 2008 margins of LCD televisions were effectively zero, and nearly the entire industry suffered major losses. This paper provides a framework for understanding whether

or not firms simply overestimated the value generated by LCD televisions or whether some combination of competition and forward looking behavior made that value impossible to extract.

This paper improves upon the existing empirical literature on dynamic consumer behavior in three ways. First, it improves the computation by recasting the problem in the language of the constrained optimization. In this formulation, the dynamic problem is not appreciably harder to solve than the static problem. Additionally, this framework accommodates an estimator based on Empirical Likelihood (EL) rather than Generalized Method of Moments (GMM), which is higher-order efficient and may be less biased in finite samples (Conlon 2012). Finally, it makes use of a new dataset on LCD televisions. One advantage of working with this industry is that consumers rarely discard or replace their LCD TV's during the sample period, which simplifies the model of consumer behavior. Also, this dataset includes not only sales and prices, but marginal costs as well. By observing marginal costs instead of inferring them from pricing decisions, it is possible to consider price setting behavior by firms.

There is a large theoretical literature on the durable goods problem, including the Coase Conjecture Coase (1972). This literature studies how firms compete with their own products over time, and how this competition limits the rents extracted by firms. In general, the theoretical literature Stokey (1982), Bulow (1982), has focused on establishing closed-form results for tractable models (monopoly, single-product, linear demand, etc.), but few results exist for the more complicated multi-firm, multi-product, multi-characteristic settings that are common in the empirical literature on differentiated products markets.

There is also a small and growing empirical literature which aims to expand static oligopoly models of differentiated products demand (Berry, Levinsohn, and Pakes 1995) to markets with dynamic consumer behavior; this literature has generally taken two directions. The first relies primarily on scanner data, and examines how consumers stockpile inventories of products when prices are low. Examples include Erdem, Imai, and Keane (2003), Hendel and Nevo (2007b), and Hendel and Nevo (2007a). The other direction focuses on adoption of high-tech products such as digital cameras and video-games; examples include: Melnikov (2001), Gowrisankaran and Rysman (2009), Carranza (2007), Nair (2007), Zhao (2008), and Lee (2009). With a few exceptions, previous studies have focused on the dynamic decisions of consumers but not dynamic pricing decisions of firms. One exception is Nair (2007), who considers the problem of a video game seller as a single product durable-goods monopolist with constant marginal costs. Another exception is Zhao (2008), who uses a dynamic Euler equation approach motivated by Berry and Pakes (2000) in order to recover marginal costs from a dynamic model of supply and demand.

One reason for the paucity of empirical work on models of dynamic supply and demand is that the task of estimating dynamic models of demand is quite challenging by itself. Most approaches require solving an optimal stopping problem akin to Rust (1987) inside of the fixed point of static demand model like Berry, Levinsohn, and Pakes (1995). Such an approach presents both numerical and computational challenges; and often requires highly specialized algorithms and weeks of computer time. This paper takes a different approach,

and considers a model similar to that of Gowrisankaran and Rysman (2009), but employs the MPEC (Mathematical Programming with Equilibrium Constraints) method of Su and Judd (2008) to simplify estimation. One way to understand this approach, is that instead of solving the optimal stopping problem at each step of the demand problem, it is possible to define additional variables and consider a single constrained problem which treats the two problems simultaneously. This makes it possible to use state-of-the-art commercial solvers for estimation, instead of relying on custom algorithms. It is possible to use this single framework, for estimation, inference and testing, Conlon (2012) develops a method for estimating moment condition models via the MPEC approach using empirical likelihood (EL) instead of generalized method of moments (GMM). The EL estimator admits a simple test statistic that can be inverted to obtain confidence intervals in the MPEC framework. Evidence indicates that these confidence intervals have less bias, and better coverage properties than confidence intervals implied by the asymptotic distribution of GMM estimators (Newey and Smith 2004). This EL framework also makes it possible to directly test different modeling assumptions and conduct inference on counterfactual predictions.

In order to understand the effect durable goods have on prices, it is necessary to consider the dynamic pricing problem firms face. In a dynamic context, prices affect both the sales of other products today, and residual demand in the future. Both firms and consumers may have beliefs (and strategies) regarding the future path of prices, which may depend on the full history of previous actions. The resulting pricing strategies can be complex, and not necessarily unique. As in much of the previous literature, this paper focuses on a specific set of Markov Perfect Equilibria (MPE) in order to avoid these complications. Specifically, I consider a Markov Perfect Equilibrium where firms only take into account of the consumer types remaining in the market, and each type's reservation value when setting prices. In several cases where demand is deterministic, the resulting equilibrium is subgame perfect and can be solved by backward induction.

In this simplified framework, it is possible to conduct counterfactual experiments to decompose the effect that inter-temporal price discrimination, the value of waiting, and the indirect effects have on prices. For example, to measure the direct effect of price discrimination, one constructs a counterfactual equilibrium where the distribution of consumer types is fixed over time. The other effect of price discrimination is that prices in one period may affect demand in other periods. However, it is possible to construct counterfactual experiments where firms set prices without internalizing the effects on other periods. It is possible to construct similar experiments for the option value of waiting. By comparing predicted prices in both of these cases to a baseline case, it is possible to separately measure the impact that changes in the consumer population and the option value of waiting have on prices.

The principal empirical finding from these counterfactual experiments is that the distribution of consumer types (the price discrimination motive) appears to have the most substantial impact on prices (about 50%). The option value of waiting has a smaller, but still significant impact on prices (about 20%). Meanwhile, the indirect effects have relatively small effects on prices, even for dominant firms. These results highlight important differences in durable goods markets between monopoly and oligopoly cases. The distribution of

consumer types describes how much surplus remains in the market, while the option value of waiting limits firms' ability to extract surplus from consumers. In the oligopoly framework, competition also limits a firm's ability to extract surplus from consumers. Similarly, competition reduces the extent to which firms internalize the effects of prices today on tomorrow's state, since there is no guarantee that a marginal consumer today will be the same firm's consumer tomorrow.

The rest of this paper is organized as follows. Section 2 provides additional details regarding the LCD television industry and the dataset used in the empirical exercise. Section 3 describes a dynamic model of demand for durable goods similar to Gowrisankaran and Rysman (2009) and makes some modifications specific to the LCD TV industry. When multiple purchases are ruled out, the problem is substantially simplified, and the value of waiting can be recovered without further restrictions. Section 4 discusses the MPEC estimation procedure that makes it possible to estimate the model, as well as presenting the empirical likelihood estimator which offers an alternative to the traditional GMM approach. Section 5 provides an in-depth examination of firms' Markov-Perfect pricing strategies, and describes how counterfactual experiments separate the effects of different aspects of consumer behavior on prices. Section 6 presents the empirical results, and the Appendix provides additional details on the EL/MPEC approach including inference.

2 Description of Data and Industry

This paper makes use of a new dataset provided by NPD-DisplaySearch which tracks the prices, costs, and sales of LCD televisions over 13 quarters from 2006-2009. The market for LCD Televisions is a good example of a durable goods market, where consumers have strong incentives to strategically time their purchases. Over the sample, the average LCD television declined over 60% in price, with some price declines in excess of 80%. Although the dataset tracks LCD televisions as small as 10 inches, this paper focuses only on High Definition LCD TV's that are 27" or larger. For this sample, the observed time period roughly corresponds to the entire universe of sales. This makes it ideal for studying durable goods markets, because consumers do not begin the sample with an existing inventory of LCD televisions. Moreover, survey data indicates that repeated purchases in the 27"+ category are rare, so most consumers only purchase an LCD television once during the sample period. An additional benefit of studying LCD TV's, is that the panel itself is the major cost driver in the manufacturing process. Panels are typically produced by separate OEMs and panel prices are observable. In conjunction with other engineering tear-down analysis from NPD-DisplaySearch this makes it possible to construct an accurate measure for marginal costs.

The LCD TV industry is important and interesting in its own right, with annual sales in excess of \$25 billion per year in the United States, and \$80 billion worldwide. By 2008, nearly 90% of overall television sales were flat-panel LCD televisions.¹ Televisions are an important

¹The remaining television sales were older CRT televisions (at the lower end of the size-price spectrum), and plasma display panel (PDP) televisions. PDP TV's face higher materials costs that make it difficult for small panels to be priced competitively, and sell mostly larger, high-end televisions. In Q1 2009, the market

aspect of consumer spending, and are often considered a bellwether for consumer sentiment. Substantial declines in prices over short periods of time help explain why televisions (along with personal computers) are widely known to be a major challenge in the construction of price indices (see Pakes (2003) and Erickson and Pakes (2008)).

The dataset is constructed by matching shipments and revenues of LCD televisions from manufacturers to average selling prices (ASP) and cost estimates for each panel. The ASP data roughly correspond to a sales-weighted average of prices paid at the retail point-of-sale (POS). For each panel, the key characteristics are the size and the resolution. The resolution is an important characteristic because it determines which inputs the television can display natively or at full quality. The convention is to describe resolution by the number of horizontal lines. LCD Televisions generally come in one of two resolutions: HD (720p), and Full HD (1080p).²

Both price and shipment data are reported quarterly at the manufacturer-panel level. The data are recorded as *Samsung 46" 1080p 120Hz Q4 2008* rather than *Samsung LN46B750* or *Samsung LN46A650*. Some information is lost at this level of aggregation. For example, the BLS tracks the number of video inputs, whether or not the TV supports picture-in-picture (PiP), and a number of other features. Tracking at this level of detail is problematic because it dramatically increases the number of models. Major manufacturers such as Samsung and Sony, often offer models that are specific to major retail customers, like Best-Buy and Wal-Mart. This kind of behavior makes models difficult to tell apart, since much of the differentiation comes in the number of ports, or the software and menus on the television.

In some cases I aggregate data over manufacturers. There are two reasons that necessitate this. The first reason is that aggregation avoids numerical issues when market shares are very small. The second is that Akerberg and Rysman (2005) show adding more brands mechanically increases the overall quality of the market. In order to mitigate these two potential problems, manufacturers with overall sales of less than twenty-thousand units are aggregated into a catch-all category. Additionally, manufacturers who sell less than 100 units of a particular (size,resolution,quarter)-tuple are also aggregated. This leaves 1406 "model"-quarter observations. The number of observations still varies from period to period, which can be explained by two factors. Over time, more FHD (1080p) TV's and fewer HD TV's are offered; these trends are displayed in Table 1. Also, more manufacturers enter the larger size segment over time, which is demonstrated in Table 2.

LCD manufacturers typically assemble televisions from parts purchased from original equipment manufacturers (OEM's). These components include: the panel itself, the power supply, the tuner, and the bezel. A combination of OEM prices and engineering tear-down

share of PDP TV's was insignificant in all size categories except for 50"+ televisions, while most sales of LCD TV's (and most of the overall market) are between 30 – 46 inches.

²As an example, standard definition TV is broadcast in 480i which means 480 lines of resolution. DVD players usually output 480p or 720p, High-Definition TV broadcasts are either in 720p (ABC, ESPN) or 1080i (FOX, NBC, CBS), while Blu-Ray disc players (BDP) and some video game consoles output 1080p. The i and the p denote whether an input is scanned progressively or in an interlaced format. It is believed that progressive scanning produces a better quality image, particularly for fast motion, but at the cost of additional bandwidth.

analysis makes it possible to estimate the manufacturing costs at the panel-quarter level, for example *46" 1080p 120Hz in Q4 2007*. An example of this cost breakdown is presented in Table 3. The most important input, both as a fraction of the cost and as a source of cost variation is the panel. Panels are produced by upstream panel manufacturers that are independent of the TV manufacturers.³ In the sample, the panel represents 67% of the cost for the average television, and represents nearly 80% of the cost variation. The panel makes up a larger portion of the overall costs for larger televisions (≥ 40 "), at 72%, than it does for smaller televisions (< 40 "), at only 61%. The share of panel price over time is plotted in Figure 3. The share of panel price decreases as panels become less expensive while other input prices remain largely the same (plastic, glass, power supply, etc.). The panel production process is similar to that of microprocessors (CPUs), in which panels are produced in batches rather than sequentially and each batch has a failure rate that engineers improve over time. The source of the decline in panel prices comes from improvements in the yield of the panel manufacturing process, which is plausibly exogenous.⁴

An important feature of the market is that panel prices fall sharply during the period of observation. Steeply falling prices (and costs) are important for understanding the dynamic tradeoffs faced by both consumers and manufacturers. From 2006 to 2009, consumers paid on average 11% less per year for televisions, and prices of similar televisions fell 17% – 28% per year. Table 4 reports the results of a hedonic regression of prices on functions of size and resolution. To summarize the results: product characteristics alone explain about 66% of the price variation in the market, and the addition of a linear time trend increases the R^2 to 0.87, while inclusion of manufacturer dummies increases the R^2 to 0.92. The regression results imply that introduction of a new product characteristic (Full HD/1080p resolution) commands a price premium equivalent to approximately three quarters of price declines. Table 5 reports similar results for marginal costs, where product characteristics and a time trend explain 97% of cost variation. It is also important to note that costs fall about 4% per quarter while prices fall almost twice as fast at 8% per quarter (and on a larger basis) according to the hedonic estimates. This motivates the need for a dynamic model which explains declining markups and addresses the inter-temporal nature of the consumer's problem.

3 Demand Model

There is a growing and recent literature on extending models of differentiated products demand (such as Berry, Levinsohn, and Pakes (1995)) to incorporate consumer dynamics and durable goods. Many of these approaches exploit a unique property of logit-type models that allows the expected utility of making a purchase to be written in a closed-form that does not

³A notable exception is that Sony and Samsung buy most of their panels through S-LCD, a joint venture owned 50-50 by the two manufacturers.

⁴It is important to note that the LCD panel manufacturing industry settled one of the largest price-fixing cases of all time in 2006. However, this price fixing behavior was alleged to have taken place back in 2003 and pertained to smaller panels used in laptop computers, cell phones, and portable music players.

depend on which product is purchased. This literature begins with Melnikov (2001) and has been employed in a number of studies of the digital camera market including: Chintagunta and Song (2003), Gowrisankaran and Rysman (2009), Carranza (2007), and Zhao (2008). This section presents the model of Gowrisankaran and Rysman (2009) that has been modified to fit the LCD TV industry.

In this model a consumer i with tastes α_i decides whether to purchase good j in period t . The utility of the purchase depends on observable characteristics x_{jt} (size, resolution, brand, and price) and unobservable quality ξ_{jt} . Consumers may also elect not to make a purchase in period t and enjoy their existing television stock v_{i0t} . Each consumer has an idiosyncratic taste ε_{ijt} that is IID (across time, products, and individuals) Type I extreme value:

$$u_{ijt} = v_{ijt} + \varepsilon_{ijt} = x_{jt}\alpha_i^x + \xi_{jt} - \alpha_i^p p_{jt} + \varepsilon_{ijt} \quad \text{or} \quad u_{i0t} = v_{i0t} + \varepsilon_{i0t}$$

The consumer decision problem is dynamic, as consumers decide whether to wait until next period, or make a purchase today. In this context Ω_t is a state variable that contains all of the relevant information for the consumer's decision. This dynamic optimization problem is described by the following Bellman Equation:

$$V_i(\varepsilon_{it}, v_{i0t}, \Omega_t) = \max \left\{ u_{i0t} + \beta E[V_i(\varepsilon_{i,t+1}, v_{i0t}, \Omega_{t+1}) | \Omega_t], \max_j u_{ijt} + \beta E[V_i(\varepsilon_{ij,t+1}, v_{ijt}, \Omega_{t+1}) | \Omega_t] \right\}$$

For studying the LCD Television market, it is helpful to assume that consumers exit the market after making a purchase, or that consumers do not “upgrade” their televisions. This is reasonable in the for the LCD television market because consumers rarely discard or replace LCD televisions during the sample period. This may seem strange in light of the fact that many US households hold in excess of four televisions. Households that do purchase multiple televisions will locate them in separate rooms, and it is helpful to think of them as separate consumers, rather than consumers making replacement decisions.⁵

When repeat purchases are ruled out, it is without loss of generality to assume that all utility is earned at the time of purchase (since utility now represents the present discount value of flow utilities $\frac{1}{1-\beta}v_{ijt}$). It is also possible to normalize the expected utility of the outside good $E[u_{i0t}] = 0$.

Assumption 1. *No Upgrades:* $v_{i0t} = 0$ and $V_i(\varepsilon_{ij,t+1}, v_{ijt}, \Omega_{t+1}) = 0$ for all $j \neq 0$.

When we rule out upgrades, we do not need to keep track of a consumer's existing stock of televisions v_{i0t} , since consumers exit the market after making a purchase. This simplifies

⁵In a world with “upgrades” consumers respond to utility differences between a 32” HD TV and a 55” FHD TV, however, after purchasing a 55” FHD TV the 32” HD TV is lost forever. In a world without upgrades we might think of the same household as two consumers a bedroom consumer might be price sensitive and adopt the 32” HD TV early, while the living room consumer might have a higher taste for size and resolution but adopt the 55” FHD TV later. If utilities are additively separable in televisions, we can consider these as separate individuals. This is in contrast with the model of Gowrisankaran and Rysman (2009). In the market for digital cameras, it is more reasonable to assume that consumers purchase a 3MP digital camera, and later upgrade to a 10MP digital camera. After the “upgrade”, the 3MP camera is likely to sit in a drawer and not see much use, unlike the second best television.

the state space of the dynamic problem.

$$\begin{aligned} u_{ijt} &= x_{jt}\alpha_i^x + \xi_{jt} - \alpha_i^p p_{jt} + \varepsilon_{ijt} \quad \text{or} \quad u_{i0t} = \varepsilon_{i0t} \\ V_i(\varepsilon_{it}, \Omega_t) &= \max \left\{ \varepsilon_{i0t} + \beta E[E_\varepsilon V_i(\varepsilon_{it}, \Omega_{t+1}) | \Omega_t], \max_j u_{ijt} \right\} \end{aligned} \quad (1)$$

An important feature of the logit error is that it is possible to consider the expected utility consumer i earns from making a purchase in period t without conditioning on which product is purchased, this is known as the logit inclusive value:

$$\delta_{it} = \delta_{it}(\Omega_t) = E_\varepsilon[\max_j u_{ijt}] = \log \left(\sum_j \exp[x_{jt}\alpha_i^x + \xi_{jt} - \alpha_i^p p_{jt}] \right) \quad (2)$$

Following Rust (1987), it is possible to integrate the logit error ε_{ijt} out of the value function, and we simplify the expression through the standard abuse of notation:

$$\begin{aligned} V_i(\Omega_t) &= \int V(\varepsilon_{it}, \Omega_t) f(\varepsilon) \\ V_i(\Omega_t) &= \max \left\{ \varepsilon_{i0t} + \beta E[V_i(\Omega_{t+1}) | \Omega_t], \max_j u_{ijt} \right\} \\ &= \ln(\exp(\delta_{it}) + \exp(\beta E[V_i(\Omega_{t+1}) | \Omega_t])) + \eta \end{aligned} \quad (3)$$

In the above expression, η is merely the constant of integration. It is important to note that $V(\Omega_t)$ only depends on the current and future inclusive values δ_{it} and not directly on prices, product characteristics, or other features of the market. One challenge of dynamic models is that when the state space is large, they become difficult to solve. In order to avoid this challenge, and in light of the expression in (3), Melnikov (2001) and others assume that the continuation value is fully described by the current period inclusive value δ_{it} . Such an assumption would imply not only that $V_i(\Omega_t) = V_i(\delta_{it})$, but also that the evolution of the state space $P_i(\Omega_{t+1} | \Omega_t) = P_i(\delta_{t+1} | \delta_t)$. One must also specify consumer beliefs in order to calculate $E[V_i(v_{ijt}, \delta_{i,t+1}) | \delta_{it}]$. The literature considers different functional forms for $E[\delta_{t+1} | \delta_t]$, the most common is of which is an $AR(1)$: $\delta_{i,t+1} = \gamma_{1i} + \gamma_{2i}\delta_{it} + \nu_{it}$. Functional form assumptions on beliefs can be a bit controversial when they do not represent an implication of an economic restriction on the model. While the demand model may indicate that the inclusive value is sufficient for understanding a consumer's two period decision, nothing indicates that the inclusive value today is sufficient for predicting tomorrow's inclusive value. Also, when considering a model of supply as well as demand, firms may try to game the functional form of consumer beliefs.

Assumption 2. *Perfect Foresight:* $E[V_{i,t+1}(\Omega_{t+1}) | \Omega_t] = v_{i,t+1}$.

This paper considers an alternative where consumers have perfect foresight regarding $E[V_i(\Omega_{t+1}) | \Omega_t]$. In this case, the relevant state variable is now just the period t . It is helpful to define a new variable v_{it} which represents the continuation value of consumer i in period

t , so that $v_{it} = V_i(\Omega_t)$. It is also important to note that when we consider a finite number of consumer types i and periods t , the resulting state space (and the dimension of v_{it}) is finite.

$$\begin{aligned} v_{it} &= \ln(\exp(\delta_{it}) + \exp(\beta E[v_{i,t+1}|\Omega_t])) + \eta \\ &= \ln(\exp(\delta_{it}) + \exp(\beta v_{i,t+1})) + \eta \end{aligned} \quad (4)$$

This assumption may not be quite so strong in light of the existing functional form assumptions. The logit inclusive value implies that consumers only need to predict the expected utility of tomorrow's state, not necessarily every price and product characteristic. Each period consumers only need to decide if any products exceed their reservation utility $v_{i,t}$; the content of this assumption is that v_{it} is correct up to the logit error. Over short periods of time, the LCD television market is reasonably predictable, as the hedonic regressions in Table 4 indicate.

Under this assumption, the probability that consumer i buys good j in period t can be expressed as the product of the probability that the consumer buys any good in period t and the conditional probability of choosing good j given some purchase in t .

$$s_{ijt} = \frac{e^{\delta_{it}}}{e^{v_{it}}} \cdot \frac{e^{x_{jt}\alpha_i^x + \xi_{jt} - \alpha_i^p p_{jt}}}{e^{\delta_{it}}} = \exp[x_{jt}\alpha_i^x + \xi_{jt} - \alpha_i^p p_{jt} - v_{it}] \quad (5)$$

We consider a finite distribution of consumer types i , so that each type has a corresponding weight w_{it} . Consumers leave the market after making a purchase, and the weights adjust deterministically. It is important to note that while the initial weights sum to unity $\sum_i w_{i0} = 1$, but subsequent weights do not. The aggregate market shares are just the sum of the type-specific market shares:

$$w_{i,t+1} = w_{it}(1 - s_{i0t}) \quad (6)$$

$$s_{jt} = \sum_j w_{it}s_{ijt} \quad (7)$$

The two dynamic relationships evolve in separate ways, as is indicated in Figure 4. The distribution of consumer types w_t evolves forwards over time; future values of the types distribution depend only the outside good share and the current distribution of consumer types. The reservation value depends the current inclusive value and the next period reservation value. The fact that these processes evolve separately, and in different directions helps to simplify both the estimation and the firms' pricing problem in the following sections.⁶

3.1 Other Details

In order to estimate this model it is necessary to choose an initial distribution w_{i0} for consumer tastes α_i . Once the initial distribution is pinned down, the future distribution is

⁶The model can be extended to include multiple purchases by expanding the space of consumer types. This can be accomplished by discretizing consumer inventories v_{i0t} , and the new type space is the product of the existing type space and the space of consumer inventories. Rather than leave the market after making a purchase, consumers transition to the "type" of consumer with the same tastes, but a different inventory holding. This is omitted because it is not an important feature of the market for LCD televisions.

determined by the purchase decisions of consumers. Like in much of the literature, we assume that tastes for individual product characteristics are independent and normally distributed. Over time, as consumers make purchases, the resulting distribution may no longer be normally distributed.

When we assume that α_i are distributed as independent normal random variables, tastes can be broken up between the population average $\bar{\alpha}_k$ and the individual deviation v_{ik} :

$$\alpha_i^k = \bar{\alpha}_k + \sigma_k v_{ik}$$

There are several ways to choose the random tastes and the corresponding population weights (v_i, w_{i0}). The Monte-Carlo approach is to set $w_{i0} = \frac{1}{ns}$ and then randomly sample v_{ik} from the standard normal distribution. A better method is to choose the draws and weights in accordance with some quadrature rule. This is particularly effective when the dimension of integration (number of random coefficients) is low. Theoretically, this approach can be thought about by approximating the function of interest with a high-order polynomial and then integrating the polynomial exactly. In practice, it simply provides a set of points and a set of weights (w_{i0}, v_i) in a non-random way. This approach is especially effective for this type of problem because the integrand s_{ijt} is an analytic function, meaning it is well approximated by a Taylor series, and it is also bounded between $[0, 1]$, which guarantees that the tails die off sufficiently fast.

This paper chooses the initial distribution of consumer types w_{i0} according to the Sparse-Grids approach of Smolyak (1963). Heiss and Winschel (2008) determine an efficient way of nesting quadrature rules for ML estimation of multinomial logit models. The authors find that for three-dimensional logit-normal integrals, 87 quadrature points are more accurate than 2500 quasi-random draws. By using the quadrature points provided on the authors' website (<http://www.sparse-grids.de>), it is possible to accurately approximate a normal distribution of consumer tastes with a relatively small number of consumer types.

4 Estimation

The economic model imposes the following constraints:

$$s_{ijt} = \exp[x_{jt}\alpha_i^x - \alpha_i^p p_{jt} + \xi_{jt} - v_{it} + \eta] \quad (8)$$

$$\delta_{it} = \log\left(\sum_j \exp[x_{jt}\alpha_i^x - \alpha_i^p p_{jt} + \xi_{jt}]\right) \quad (9)$$

$$v_{it} = \log(\exp(\delta_{it}) + \exp(\beta v_{i,t+1})) \quad (10)$$

$$w_{i,t+1} = w_{i,t} \left(1 - \sum_j s_{ijt}\right) \quad (11)$$

$$s_{jt} = \sum_i w_{i,t} s_{ijt} \quad (12)$$

Estimation adds two constraints.⁷ The first constraint is that the observed shares S_{jt} match the predicted shares. The second constraint is that the demand shock ξ_{jt} is orthogonal to some function of the observable variables x_{jt} and the excluded instruments z_{jt} , so that $Z_{jt} = [z_{jt}, x_{jt}]$.

$$S_{jt} = s_{jt} \tag{13}$$

$$E[\xi_{jt}Z_{jt}] = 0 \tag{14}$$

In the presence of over-identifying restrictions, it is not usually possible to solve (14) exactly. Instead, estimation usually proceeds by choosing an appropriate weight matrix for the moment restrictions and minimizing a quadratic objective function via Generalized Method of Moments (Hansen 1982).

The empirical likelihood (EL) estimator of Owen (1990) provides an alternative to GMM methods. One way to think about EL is as an extension of Nonparametric Maximum Likelihood Estimation (NPMLE) to moment condition models. The empirical likelihood estimator re-interprets the moment condition by attaching a probability weight to each observation. These weights, ρ_{jt} , are constructed so that equation (14) holds exactly. If each observation in the dataset were distributed as an independent multinomial, then the corresponding likelihood function would be $l(\rho, \theta) = \sum_{\forall j,t} \log \rho_{jt}$. The resulting optimization problem is to find a set of weights that maximizes the likelihood subject to the following additional constraints:

$$\sum_{\forall j,t} \rho_{jt} \xi_{jt} Z_{jt} = 0 \tag{15}$$

$$\sum_{\forall j,t} \rho_{jt} = 1 \tag{16}$$

Empirical likelihood estimators provide a number of properties that are desirable to applied researchers. Many of these properties are related to the fact that EL avoids estimating the weighting matrix. The estimates derived from EL estimators have the same asymptotic distribution as GMM estimators, but are higher order efficient (Newey and Smith 2004). The appendix provides more details on the construction of the EL estimator. Conlon (2012) provides an examination of the computational properties of empirical likelihood for static demand models, and Kitamura (2006) provides a general survey of the EL literature.

In much of the literature, such as Gowrisankaran and Rysman (2009), Carranza (2007), and Lee (2009), estimation follows a multi-step procedure that involves iterating over three loops. The innermost loop involves solving a dynamic optimal stopping problem similar to Rust (1987) for the value function. The middle loop consists of a modified version of the contraction mapping in Berry, Levinsohn, and Pakes (1995) to find ξ_{jt} , and constructing consumer expectations about the evolution of δ_{it} . The outer loop involves a nonlinear search over parameters of a nonlinear (GMM) objective function formed from the moment conditions (14). Such a method is often quite difficult to implement, and may take several days to achieve convergence.

⁷This formulation of the problem nests the static problem of Berry, Levinsohn, and Pakes (1995). In that problem, $w_{i,t}$ is fixed, and $v_{it} = 1 + \exp[\delta_{it}]$.

This paper takes a different approach to estimation, and solves (8) - (16) directly using the MPEC method of Su and Judd (2008). The MPEC method was later adapted to static (and a simple dynamic) demand estimation problem by Dube, Fox, and Su (2009). The key intuition of the MPEC approach is that instead of solving for equilibrium outcomes at each iteration, these equilibrium outcomes can be viewed as constraints that only need to hold at the optimum. For example, instead of iteratively solving for the value function for each guess of the parameters, it is sufficient to ensure that (10) is satisfied at the final estimate $\hat{\theta}$. The MPEC method works by using constrained optimization routines to directly solve the system of nonlinear equations implied by the model. This is markedly different from other approaches in the literature such as: Rust (1987), Hotz and Miller (1993), Berry, Levinsohn, and Pakes (1995), Aguirregabiria and Mira (2007) which manipulate equations and solve for parameters as implicit functions of other parameters in order to reduce the number of parameters and eliminate the constraints.

It is well established in the literature on optimization (Nocedal and Wright (2006) and Boyd and Vandenberghe (2008)) that sparse and convex problems are easier to solve than dense and non-convex problems. Fortunately, (8)-(14) are mostly convex equations⁸ and reasonably sparse. Sparsity refers to the resulting Hessian matrix of an optimization problem. It is easy to see that a problem is sparse when many variables only enter one or two equations (like w_{it}) or enter the model linearly. The “trick” in many MPEC problems is how to re-write the problem in a way which makes it more sparse. In this case, the “trick” is to replace the value function with an extra variable, v_{it} . The difficulty of the MPEC method (and constrained optimization in general) depends more on convexity and sparsity than the number of unknown parameters. As written, the model implies an extremely large number of parameters ($v_{it}, w_{it}, s_{ijt}, \delta_{it}, \xi_{jt}, \alpha_i$), which would make it nearly impossible to solve using traditional nested fixed-point methods.

The advantage of the MPEC formulation for the dynamic demand problem is twofold. The first advantage is that it is substantially easier from a computational perspective. In fact, the dynamic demand problem is not appreciably harder than the static demand problem and can be solved in about an hour. One way to see this relationship is to recognize that the static problem is just the special case of the dynamic problem where $v_{i,t+1} = 0$ in (10). The second advantage of the MPEC method is that it allows us to define the dynamic quantities directly as extra variables. Rather than implicitly solving the value function, the dynamic behavior is just a set of constraints on the (v_{it}, w_{it}) variables. Moreover, these constraints help exploit the sparsity, since v_{it} depends only on the one period ahead value $v_{i,t+1}$ and $w_{i,t}$ depends only on the one period lagged value $w_{i,t}$. It should be pointed out that, this system of equations represents an exact solution to the value function, and does not require

⁸Note: For an optimization problem to be convex, it must have only affine equalities. This problem has several nonlinear equality constraints, although they are themselves convex. For example, the logit inclusive value: $\log \sum \exp(\cdot)$ is convex (Boyd and Vandenberghe 2008). The share equation is the ratio of two exponentials $\frac{e^{v_{ijt}}}{e^{v_{it}}}$, which are convex, but the ratio is not. This is not enough to guarantee that the overall optimization is globally convex, and hence has a single unique optimum that can be found in constant time. However, these equations are well behaved enough that it is usually possible to find the optimum using off the shelf nonlinear solvers.

approximation on a grid or via polynomials. Additionally, it makes it possible to estimate the model without additional functional form assumptions on beliefs or the evolution of the state space, and without assuming that the inclusive value encodes all of the information about a particular state.

The overall optimization problem is:

$$\begin{aligned}
& \max_{(\rho_{jt}, v_{it}, w_{it}, s_{ijt}, \delta_{it}, \xi_{jt}, \alpha_i)} \sum_{j,t} \log \rho_{j,t} && \text{s.t.} \\
& s_{ijt} = \frac{e^{x_{jt}\alpha_i^x - \alpha_i^p p_{jt} + \xi_{jt}}}{e^{v_{it} + \gamma_{it}}} \\
& s_{jt} = \sum_i w_{i,t} s_{ijt} \\
& w_{i,t+1} = w_{i,t} (1 - \sum_j s_{ijt}) \\
& \exp[\delta_{it}] = \sum_j \exp[x_{jt}\alpha_i^x - \alpha_i^p p_{jt} + \xi_{jt}] \\
& v_{it} = \log(\exp(\delta_{it}) + \exp(\beta v_{i,t+1})) + \gamma_{it} \\
& S_{jt} = s_{jt} \\
& \sum_{\forall j,t} \rho_{jt} \xi_{jt} Z_{jt} = 0 \\
& \sum_{\forall j,t} \rho_{jt} = 1
\end{aligned}$$

5 Pricing

An important aspect of oligopoly models is that prices are not exogenous, but rather the result of the profit-maximizing behavior of firms. There is a large literature on static oligopoly models, where inferred equilibrium behavior of firms is used to recover markups and costs. The idea behind this is that when observed prices represent equilibrium outcomes of a static differentiated products Bertrand pricing game, there is a one to one relationship between prices and costs that can be inverted. This relationship is described in detail in Berry (1994) and Nevo (2000). This paper takes a different approach to the supply-side decisions of firms. An important feature of the dataset is that marginal cost data are observable, so that costs do not need to be recovered from the pricing equilibrium.⁹

Instead, observed costs are used as an input into several counterfactual pricing equilibria. The goal is to construct counterfactual prices from scenarios where firms consider some aspects of dynamic consumer behavior but not others. By isolating different aspects of consumer behavior, it is possible to understand why prices are low (or what would make

⁹Zhao (2008) develops a novel procedure based on Berry and Pakes (2000) that exploits the orthogonality of the dynamic pricing Euler equation with some instruments to recover costs without explicitly solving the pricing problem.

them higher). We re-compute markups when firms account for changes in the distribution of types when setting prices, but not consumers' value of waiting (and vice versa). Furthermore, it is possible to ask what prices would be in a world where consumers had no option to wait, or one in which the distribution of consumer types was fixed over time. By comparing prices in these scenarios to the observed prices (and to the other scenarios) it is possible to measure the extent to which different types of dynamic behavior influence equilibrium prices.

The next subsection reviews the price setting behavior of a static oligopolist. This provides an important building block and intuition for some of the counterfactual experiments. The section after that examines pricing setting behavior in a dynamic oligopoly setting. Under some simplifying assumptions, and by exploiting the structure of the demand model which lets us separate the value of waiting v from the distribution of consumer types w , the resulting equilibrium is deterministic and straightforward. The final subsection applies the static and dynamic equilibrium pricing strategies to different scenarios for consumer behavior.

5.1 Static Pricing

It is helpful to define the $J \times J$ matrix of same-brand price elasticities, where J_g represents the set of products owned by multi-product firm g .

$$A_{jk} = \frac{\partial s_{kt}}{\partial p_{j\tau}} \text{ when } (j, k) \in J_g, t = \tau \quad 0 \quad \text{o.w.}$$

Assume there is a fixed population of consumers M , then firms choose a set of prices p_{jt} for products they own J_g in order to maximize profits by examining the FOC:

$$\begin{aligned} \max_{p_{jt} \in J_g} \pi_{gt} &= \max_{p_j \in J_g} \sum_{j \in A_g} M(p_{jt} - c_{jt}) s_{jt}(\mathbf{p}_t, \theta) \\ \Rightarrow s_{jt} &= \sum_{k \in J_g} (p_{kt} - c_{kt}) \frac{\partial s_{kt}}{\partial p_{jt}} = A(\mathbf{p} - \mathbf{c}) \\ \Rightarrow \mathbf{p} &= \mathbf{c} + A^{-1} s(\mathbf{p}, \theta) \end{aligned}$$

The extension of this approach to multiple periods is not trivial. At the minimum, oligopolist firms play a repeated game over several periods, and static Nash-Bertrand is only one possible outcome. Firms can condition their pricing strategy to depend on actions of other firms. For example, firms may collude in some periods and engage in price wars in other periods. The dynamic nature of the consumer's problem makes this more challenging, since sales in period t depend not only on $s(\mathbf{p}_t, \theta, v_t)$ but also on beliefs about the future v_t .

5.2 Dynamic Pricing

Firms, subscripted by g , now solve a more challenging problem where $\tilde{\sigma}$ is a state variable that contains information about the past history of prices and beliefs by firms and consumers about which strategy is being played. The additional challenge is that firms are now

maximizing a discounted stream of future profits V_g with discount factor β_m .

$$V_g(\tilde{\sigma}_t) = \max_{p_{jt} \in A_g} E\pi_{gt}(\tilde{\sigma}_t, \mathbf{p}_t) + \beta_m \int V_g(\tilde{\sigma}_{t+1}) Pr(\tilde{\sigma}_{t+1} | \tilde{\sigma}_t, \mathbf{p}_t) \quad (17)$$

The demand model is fully described by a set of $(\alpha_i, \xi_{jt}, w_{it}, v_{it})$. If we know the set of consumer tastes and product quality $\theta = (\xi, \alpha_i)$, then demand is described in each period by (v_t, w_t) , because v_{it} is a sufficient statistic at time t for a consumer's beliefs about the future. Moreover, both v_t and w_t evolve in a simple deterministic fashion. We assume that firms set prices based only on the demand state, and no other factors. More formally,

Assumption 3. *The state variable $\tilde{\sigma} = (v_t, w_t, c_t)$ where v_t and w_t are the set of all v_{it} and $w_{it} \forall i$ respectively and c_t is the cost vector for all products j at time t .*

Assumption 4. *Firms know $c_{jt}, \xi_{jt} \forall j, t$ at time $t = 0$, as well as α_i .*

Assumption 4 makes it possible to suppress c_t from the state space and instead include it in the t subscript on the value function. In the LCD TV industry, it is probably reasonable to assume that firms are able to accurately forecast future costs. This is because costs are driven primarily by panel prices, and panel prices decline in a predictable way through the engineering process. The hedonic marginal cost regression provides empirical support for this assumption in Table 5.

Assumption 3 is a stronger assumption, which prevents firms from conditioning their actions on the actions of other firms, or on the full history of the demand state. This makes it possible to compute pricing equilibria in a straightforward way, although at the risk of ruling some potentially interesting behavior. Firms essentially respond to the demand state, and don't worry about inter-temporal effects of their decisions or their competitors decisions, except as they influence the demand state.

These assumptions make it possible to simplify the dynamic problem that firms solve in (17) so that:

$$V_{gt}(v_t, w_t) = \max_{p_{jt} \in A_g} \sum_{j \in A_g} (p_{jt} - c_{jt}) \cdot s_{jt}(\mathbf{p}_t, v_t, w_t) + \beta_m V_{g,t+1}(v_{t+1}, w_{t+1}) \quad \forall g, t \quad (18)$$

$$w_{i,t+1} = w_{i,t} s_{i0t} = w_{i,t} \frac{e^{\delta_{it}(\mathbf{p}_t)}}{e^{v_{it}}} \quad (19)$$

$$v_{i,t} = \log(\exp(\delta_{it}(\mathbf{p}_t)) + \exp(\beta v_{i,t+1})) \quad (20)$$

The key aspect of this formulation of the problem is that firms are required only to choose prices in each period. Once prices are chosen, all other variables respond in a deterministic way. Moreover, the state variables depend on prices only through the exponentiated inclusive value, which has a very simple price derivative, and guarantees that the states are smooth in prices. Moreover

$$\frac{\partial \exp[\delta_{it}]}{\partial p_{jt}} = -\alpha_i^p \quad \frac{\partial \exp[\delta_{it}]}{\partial p_{j\tau}} = 0 \text{ when } \tau \neq t$$

It is possible to construct a sequence of prices, reservation utilities, and a distribution of consumer types (p_t, v_t, w_t) that satisfy (18) -(20) and define a Markov Perfect Equilibrium.

There are some further simplifications, that can also be helpful to consider. If v_{it} was not affected by the prices firms set in later periods, the distribution of consumer types would be the only state variable (in addition to costs) and the price setting game in each period $t + 1$ would be a strict subgame of the period t game, and the resulting equilibrium could be solved by backward induction. Likewise if the distribution of consumer types were fixed over time, so that v_{it} were the only state variable, then the game in each period $t - 1$ would be a strict subgame of the period t game and an equilibrium could be solved by forward induction. The intuition behind this is suggested in Figure 4.

5.3 Constructing Counterfactual Pricing Equilibria

The goal of the counterfactual experiments is to understand how prices are affected by different aspects of dynamic consumer behavior. The advantage of the setup in (18)-(20) is that the value of waiting is captured by v_{it} and the distribution of consumer types is captured separately by w_{it} . This enables us to consider the effect of both of these factors separately when trying to understand Coasian price dynamics. That is, we could consider a world where consumers correctly incorporate the value of waiting when making a decision, but where the distribution of consumer types is fixed. Likewise, we can consider a world where firms engage in “skimming” against naive consumers. The differences in prices between these two cases and the actual prices, gives us an idea what the overall pricing impact is of these two aspects of Coasian dynamics.

Mechanically, these quantities affect prices in different ways and have different economic interpretations. distribution of consumer types w_t determines which consumers are left in the market, which enters demand linearly $s_{jt} = \sum_i w_{it}s_{ijt}$. The other effect of w_{it} is that firms take into account the effect that prices today have on the distribution of consumer types tomorrow $w_{i,t+1}$ through (19). Likewise v_{it} directly affects the outside good share today $s_{i0t} = \frac{e^{\delta_{it}(\mathbf{p}_t)}}{e^{v_{it}}}$ (but not relative marketshares of products), and future prices determine earlier values of v_{it} through (20).

These relationships suggest some counterfactual experiments. The goal of the counterfactual experiments is to understand how much higher (or lower) prices would be if firms did not account for these aspects of consumer behavior. In order to make this more clear, these aspects are broken out into three parts. The first part is the dynamic or “indirect” effect. This looks at deviations between a firm that statically set prices against residual demand each period, and a firm who fully incorporated the effect that prices had on other periods. Consider the effects of a price increase in period t , it increases the value of the next period $t + 1$, since more consumers remain in the market, and it increases the value of the previous period $t - 1$ since it makes waiting until period t look less attractive. This leads us to expect that firms which do not account for these dynamic effects will price lower than firms which do. This is addressed in experiment 1.

The second aspect is that the option value of waiting influences the prices that firms

charge. A high value of waiting means that firms have less ability to extract surplus from consumers. We know from (20) that a high value of waiting might be due to lower prices or higher quality in the future (through the inclusive value). We expect that the option to wait should be most valuable when there are substantial changes in the inclusive value, that is when $\Delta_{\delta_{it}} = \exp[\delta_{it}] - \beta \exp[\delta_{i,t+1}]$ is large. As an alternative, we can consider consumers who cannot wait: $v_{i,t+1} = 0$ (or are extremely impatient: $\beta = 0$) so that $v_{it} = \delta_{it}$. We can then compare the prices firms would charge facing these consumers to the observed prices in order to understand how the option to wait affects prices. This is explored in experiment 2.

The third aspect is that the distribution of consumer types influences the prices that firms charge. As high value consumers make purchases, they leave the market and the residual demand curve becomes more elastic over time. This means that over time, the amount of surplus left in the market should be decreasing.¹⁰ As the amount of potential surplus declines over time, this should lead firms to set lower prices. By comparing the observed prices to those constructed from a world where the distribution of consumer types does not change over time (that is consumers who make purchases do not leave the market, but are “reborn” next period), it is possible to understand the effect that changes in the distribution of consumer types have on prices. This is explored in experiment 3.

It is also important to mention that we don’t need to make specific assumptions on the beliefs of consumers in addition to the model of dynamic demand that has already been specified. This is because the v_{it} are the beliefs of consumers. It is also important to understand that our markov assumption means we do not worry about the beliefs of firms, rather firms wake up each period and respond to the demand state (v_{it}, w_{it}) .

Experiment 1: Firms engage in static pricing.

This experiment considers the pricing decision where firms correctly observe the state each period (v_{it}, w_{it}) but do not take into account how prices affect demand in other periods. However, in this experiment, only the initial distribution of types w_0 is determined by the demand model. Now, the reservation values v_{it} depend on the observed price sequence, and (v_{it}, w_{it}) are allowed to adjust according to the demand model. However, firms do not account for their effects on this adjustment. For a given guess of v_t markups can be computed the same way they are in the static model. The pricing equilibrium is defined by the system of equations system of equations below, where A represents the proper ownership-elasticity matrix for (v_t, w_t) :

$$\begin{aligned} \mathbf{p}_t &= \mathbf{c}_t + A_t(v_t, w_t)^{-1} s_t(\mathbf{p}_t, v_t, w_t, \theta) \\ w_{i,t+1} &= w_{i,t} s_{i0t} \\ v_{i,t} &= \ln(\exp(\delta_{it}(\mathbf{p}_t)) + \exp(\beta v_{i,t+1})) \end{aligned}$$

Experiment 2: Consumers are myopic.

Consumers are myopic and do not realize that they can make a purchase in a later period so that $v_{i,t+1} = 0$ everywhere. The initial distribution of consumers w_{i0} is used as a starting point and firms set prices in response to the distribution of consumer types each period.

¹⁰There is a potential source of new “surplus” for firms because marginal costs are also declining over time, but not for consumers.

Additionally, when setting prices firms consider the distribution of consumer types in the future. The resulting pricing equilibrium is subgame perfect and can be solved by induction.

$$\begin{aligned} v_{it} &= 0 \quad \forall i, t \\ w_{i,t+1} &= w_{i,t} s_{i0t} \end{aligned}$$

The goal of experiments 2 is to measure how much higher prices would be if consumer were myopic. This experiment simulates a sequence of prices where the rest of the model dynamics take place, but where consumers make myopic purchase decisions, and leave the market after making a purchase.

Experiment 3: Distribution of Consumer Types is Fixed. Firms account for inter-temporal effects.

In this experiment the distribution of consumer types is fixed at w_{i0} , but consumers have beliefs about the future v_{it} that are consistent with the prices set by firms. Firms set prices each period in response to the (v_{it}, w_{i0}) and incorporate how prices affect reservation values v_{it} . Prices are the result of a subgame perfect equilibrium defined by the following system of equations, and can be solved by induction, although once again we must numerically interpolate the firm’s value function $V_{gt}(v_t, w_t)$.

$$\begin{aligned} w_{it} &= w_{i0} \quad \forall t \\ v_{i,t} &= \log(\exp(\delta_{it}(\mathbf{p}_t)) + \exp(\beta v_{i,t+1})) \end{aligned}$$

6 Empirical Results

The demand model is estimated via the empirical likelihood method described in a previous section. Three specifications are reported: a static model similar to Berry, Levinsohn, and Pakes (1995) and the dynamic model with perfect foresight γ_{it} .¹¹ Both specifications are estimated from the same moment condition $E[\xi_{jt} z_{jt}] = 0$, and the same set of instruments. The instruments are: cost shifters from the marginal cost data, BLP-style instruments (average characteristics of other brand products in the same market), and non-price x_{jt} explanatory variables. All models are estimated using the MPEC method and the KNITRO solver, a modern interior-point solver that handles both constrained and unconstrained problems (Byrd, Nocedal, and Waltz 2006). As reported in Table 6, the static model implies dramatically different price sensitivities from the dynamic model, and consumers appear to be far more price sensitive. This bias is well documented in the empirical literature on dynamic durable goods models, and is due to the fact that the static model does not account for consumers who do not purchase the good in the current period in order to purchase the good later. In a comparison of empirical likelihood values (or by constructing the ELR test statistic) we fail to reject the model without upgrades when compared to the model with upgrades. Thus the model without upgrades is the preferred specification.

¹¹The model without perfect foresight is estimated as a robustness check and gives extremely similar results, and small values for γ_{it}

The implied price elasticities and substitution probabilities are reported in Table 7. The first two rows are the 95% confidence intervals for average own-price elasticities. The average own price elasticity is $E[\frac{\partial s_{jt}}{\partial p_{jt}}]$, where the expectation is computed over the empirical likelihood probability weights. The 95% confidence interval is constructed by inverting the χ^2 test statistic. Also in Table 7, the following experiment is conducted, the price of a 32" HD Sony TV in 2008Q1 is increased by 10% from \$890 to \$980. All expectations about the future are held constant. That is, v_{it} is held fixed for all other periods. This represents a one time shock akin to a pricing mistake where all retailers mislabel the price of the unit. The table reports how many consumers substitute to another Sony TV today, how many substitute to the same size and resolution TV today, how many buy the same Sony unit in the next period, and how many buy some other product (either now or in the future). The same experiment is repeated for the 32" HD Vizio TV in the same period where the price is increased from \$596 to \$655. For both TV's, few consumers substitute to the same brand television in a different period, at least when compared with an arbitrary consumer. However, overall substitution to later periods is quite large. There are two ways to interpret this finding. One interpretation is that because the market is fairly competitive firms do not compete closely with their own products over time. The other interpretation is that in the absence of brand specific tastes, and in the presence of a logit error, we shouldn't expect strong inter-temporal correlation among tastes for brand. Although the Vizio TV sells more units than the Sony 240,000 to 170,000, Vizio faces more price sensitive consumers and earns less of a brand margin, whereas Sony consumers are more likely to stay with the brand in the event of a price increase.

6.1 Counterfactual Experiments

The results of the experiments (in terms of sales-weighted average prices) are reported in Table 9 and in Figure 4.

The figure shows pricing under several different scenarios. The first scenario is that of a constant markup. This would be akin to costs changing, but no characteristics of demand or the competitive environment changing. Not shown on the figure is the case of static pricing. Static pricing is a good approximation in later periods when margins are lower, but does not fit the data particularly well in the early periods. In the later periods, static pricing would lead to prices that are only about 3% – 5% higher than the ones we see in the data in 2009. This gives some indications (at least in the later periods when margins are low) that firms are not considering the effect that prices today have on the demand state. In part, this is a feature of the logit property of the model, since consumers have no “brand loyalty” from period to period.

Experiment 1 seeks to answer the question, how much higher/lower than the observed prices would prices be if firms did not account for the inter-temporal effects of pricing decisions. Theory suggests that when firms do not account for the fact that higher prices improve the profitability both of later periods, because the number of consumers in the future population increases, and the profitability of earlier periods since higher prices in the future

reduce value of waiting. In the case of experiment 1, the prices are quite similar, but not always lower than the observed prices. One potential explanation is that the pricing model assumes the costs are perfectly measured, and there may be unobserved cost variation. Also, the pricing model assumes away not only the dynamic effects on consumers, but also the dynamic effects on other firms.

Experiment 2 is similar to experiment 1 except that consumers are also myopic, that is $v_{i,t+1} = 0$. This produces prices that are initially substantially higher (around 30%), but the difference between the predicted prices and the observed prices declines over time. This makes sense if the consumer's value of waiting declines over time. For consumers who prefer small to medium sized televisions, price declines towards the end of the sample are small, and the value of waiting is declining. Thus the gap between a consumer with no value of waiting and one who has a declining value of waiting narrows over time.

Experiment 3 is similar to experiment 1 except that while consumers are able to anticipate the future, the distribution of consumer types is fixed over time. Intuition suggests that firms will sell to high-value consumers over and over. This is a bit more complicated than the traditional static setting, because costs fall in each period. In this case, prices fall over time, but fall more quickly early on when costs are also falling fast. In the later periods when costs are falling more slowly (and there is less incentive to wait) prices also fall more slowly. Also, unlike in the observed data, prices fall more slowly than costs overall.

When compared to theoretical predictions for the monopoly case, there are some important differences in oligopoly markets. The distribution of consumer types determines the amount of surplus remaining in the market, which is not reduced by competition. Counterfactual experiments demonstrate that were it not for changes in the distribution of consumers over time (the price discrimination motive) prices might be around 50% higher than they are today. Similarly, the option value to wait limits the amount of surplus firms are able to extract, but so do competitors. In the absence of the option value prices would be only around 15-20% higher. Both competition and the logit error may reduce the extent to which firms engage in inter-temporal competition with their own products. Empirically, firms do not seem to take into account the effect that prices have on other periods, or these effects are small relative to the other effects, and unlike in the monopoly case, static pricing does not appear to be bad approximation.

7 Conclusion

This paper adapts the existing methodology for dynamic models of differentiated products to the MPEC framework. In this framework, dynamic consumer behavior places only a few simple constraints on the static demand model. This makes it possible to estimate the model directly without making any additional functional form assumptions or the inclusive value sufficiency assumption. Empirical likelihood estimators can be adapted quite naturally to the MPEC framework. The resulting estimator is likely to be more efficient, and easier to compute than standard approaches based on GMM estimators and fixed point algorithms.

In addition to improving the statistical and computational properties, this approach

simplifies the economics as well. It is possible to estimate the model directly without making additional functional form assumptions or the inclusive value sufficiency assumption. In industries where repeat purchases are not an important aspect of consumer demand, it is possible to consider the value of waiting as an additional variable. When this is the case, it is possible to identify the changes in the distribution of consumer types separately from the value of waiting. Separating these quantities simplifies the dynamic pricing problem firms face, and makes it possible to quantify how the value of waiting and changes in the distribution of consumer types differently effect the prices we observe in equilibrium.

This paper provides a simple framework for beginning to understand the dynamics of supply and demand in differentiated products oligopoly settings. However, much remains to be done. For example, this approach does not consider the dynamic effects that firms and prices have on each other (such as collusion, price-fixing, etc.) though these are often an interesting and important aspect of markets with high-technology products and fast price declines. Also, the pricing problem uses the simplifying assumption that consumers are able to exactly predict their expected utility of future states. Likewise, both supply and demand exploit the assumption that there are no repeat purchases. These assumptions are perhaps reasonable over a short period of time in the LCD TV industry, but may be more problematic when adapted to other industries.

References

- ABOWD, J. M., AND D. CARD (1989): “On the Covariance Structure of Earnings and Hours Changes,” *Econometrica*, 57(2), 411–445.
- ACKERBERG, D., AND M. RYSMAN (2005): “Unobserved product differentiation in discrete-choice models: estimating price elasticities and welfare effects,” *RAND Journal of Economics*, 36(4).
- AGUIRREGABIRIA, V., AND P. MIRA (2007): “Sequential Estimation of Dynamic Discrete Games,” *Econometrica*, 75(1), 1–53.
- ALTONJI, J. G., AND L. M. SEGAL (1996): “Small-Sample Bias in GMM Estimation of Covariance Structures,” *Journal of Business and Economic Statistics*, 14(3), 353–366.
- BERRY, S. (1994): “Estimating discrete-choice models of product differentiation,” *RAND Journal of Economics*, 25(2), 242–261.
- BERRY, S., J. LEVINSOHN, AND A. PAKES (1995): “Automobile Prices in Market Equilibrium,” *Econometrica*, 63(4), 841–890.
- BERRY, S., AND A. PAKES (2000): “Estimation from The Optimality Conditions of Dynamic Controls,” Working Paper.
- BOYD, S., AND L. VANDENBERGHE (2008): *Convex Optimization*. Cambridge University Press, 6th edition edn.
- BULOW, J. I. (1982): “Durable-Goods Monopolists,” *Journal of Political Economy*, 90(2), 314.
- BYRD, R. H., J. NOCEDAL, AND R. A. WALTZ (2006): *KNITRO: An Integrated Package for Nonlinear Optimization* pp. 35–59. Springer-Verlag.
- CARRANZA, J. (2007): “Demand for durable goods and the dynamics of quality,” Unpublished Manuscript. University of Wisconsin.
- CHINTAGUNTA, P., AND I. SONG (2003): “A Micromodel of New Product Adoption with Heterogeneous and Forward Looking Consumers: Application to the Digital Camera Category,” *Quantitative Marketing and Economics*, 1(4), 371–407.
- COASE, R. (1972): “Durability and Monopoly,” *Journal of Law and Economics*, 15(1), 143–149.
- CONLON, C. (2012): “The Empirical Likelihood MPEC to Demand Estimation,” Unpublished Manuscript. Columbia University.
- DUBE, J., J. T. FOX, AND C.-L. SU (2009): “Improving the numerical performance of BLP static and dynamic demand estimation,” Working Paper.

- ERDEM, T., S. IMAI, AND M. KEANE (2003): “A Model of Consumer Brand and Quantity Choice Dynamics under Price Uncertainty,” *Quantitative Marketing and Economics*, 1(1), 5–64.
- ERICKSON, T., AND A. PAKES (2008): “An Experimental Component Index for the CPI: From Annual Computer Data to Monthly Data on Other Goods,” Working Paper.
- GOWRISANKARAN, G., AND M. RYSMAN (2009): “Dynamics of Consumer Demand for New Durable Goods,” Working Paper.
- HANSEN, L. P. (1982): “Large Sample Properties of Generalized Method of Moments Estimators,” *Econometrica*, 50, 1029–1054.
- HEISS, F., AND V. WINSCHHEL (2008): “Likelihood approximation by numerical integration on sparse grids,” *Journal of Econometrics*, 144(1), 62 – 80.
- HENDEL, I., AND A. NEVO (2007a): “Measuring Implications of Sales and Consumer Inventory Behavior,” *Econometrica*.
- HENDEL, I., AND A. NEVO (2007b): “Sales and Consumer Inventory,” *The RAND Journal of Economics*.
- HOTZ, V. J., AND R. A. MILLER (1993): “Conditional Choice Probabilities and the Estimation of Dynamic Models,” *Review of Economic Studies*, 60(3), 497–529.
- KITAMURA, Y. (2006): “Empirical Likelihood Methods in Econometrics: Theory and Practice,” Cowles Foundation Discussion Paper 1569.
- LEE, R. S. (2009): “Vertical Integration and Exclusivity in Platform and Two-Sided Markets,” Working Paper.
- MELNIKOV, O. (2001): “Demand for Differentiated Products. The case of the US computer printer market.,” Working Paper. Cornell University.
- NAIR, H. (2007): “Intertemporal Price Discrimination with Forward-looking Consumers: Application to the US Market for Console Video-Games,” *Quantitative Marketing and Economics*, 5(3), 239–292.
- NEVO, A. (2000): “A Practitioner’s Guide to Estimation of Random Coefficients Logit Models of Demand (including Appendix),” *Journal of Economics and Management Strategy*, 9(4), 513–548.
- NEWKEY, W. K., AND R. J. SMITH (2004): “Higher Order Properties of GMM and Generalized Empirical Likelihood Estimators,” *Econometrica*, 72(1), 219–255.
- NOCEDAL, J., AND S. WRIGHT (2006): *Numerical Optimization*, Springer Series in Operations Research and Financial Engineering. Springer, 2nd edn.

- OWEN, A. (1990): “Empirical Likelihood Ratio Confidence Regions,” *The Annals of Statistics*, 18(1), 1725–1747.
- PAKES, A. (2003): “A Reconsideration of Hedonic Price Indices with an Application to PC’s,” *American Economic Review*, 93(5), 1578–1596.
- RUST, J. (1987): “Optimal Replacement of GMC Bus Engines: An Empirical Model of Harold Zurcher,” *Econometrica*, 55(5), 999–1033.
- SMOLYAK, S. (1963): “Quadrature and interpolation formulas for tensor products of certain classes of function,” *Soviet Mathematics Doklady*, 4, 240–243.
- STOKEY, N. L. (1982): “Rational expectations and durable goods pricing,” *Bell Journal of Economics*, 12(1), 112–128.
- SU, C.-L., AND K. L. JUDD (2008): “Constrained optimization approaches to estimation of structural models,” Working Paper.
- ZHAO, E. (2008): “Why Are Prices Falling Fast? An Empirical Study of the US Digital Camera Market,” Unpublished Manuscript. Queens College.

Quarter	# Prod (HD)	# Prod (FHD)	Total
2006q1	59	8	67
2006q2	61	14	75
2006q3	66	22	88
2006q4	70	28	98
2007q1	77	26	103
2007q2	85	39	124
2007q3	82	49	131
2007q4	78	50	128
2008q1	72	52	124
2008q2	65	63	128
2008q3	66	64	130
2008q4	49	60	109
2009q1	45	56	101
Total	875	531	1406

Table 1: Number of Products By Period

Quarter	26	32	37	40	42	46	47	52	55	57	65
2006q1	11	16	13	6	5	1	1	0	0	1	1
2006q2	13	17	13	9	5	4	2	0	0	2	1
2006q3	12	18	15	9	7	9	3	2	0	2	1
2006q4	12	18	19	10	10	9	6	3	0	2	1
2007q1	15	20	18	9	15	9	5	3	0	2	1
2007q2	16	20	23	10	22	11	9	5	0	2	1
2007q3	18	20	22	10	24	11	9	9	0	3	2
2007q4	17	20	21	10	23	11	10	9	0	3	2
2008q1	17	21	21	10	22	10	9	9	0	3	1
2008q2	15	24	21	10	26	9	9	10	0	2	1
2008q3	14	26	25	9	24	9	7	8	2	1	1
2008q4	15	23	19	7	18	7	7	9	2	0	1
2009q1	14	22	17	5	16	8	5	9	3	0	1

Table 2: Number of Products By Size for Selected Segments

Δp	(1)	(2)	(3)	(4)
Δc	1.962*** (0.241)	1.498*** (0.268)	1.808*** (0.226)	1.483*** (0.264)
Constant	-20.71** (9.820)	2.686 (10.32)	5.672 (13.92)	13.80 (16.09)
Observations	1,218	1,218	1,218	1,218
R-squared	0.183	0.225	0.219	0.243
Panel FE	NO	YES	NO	YES
Manuf FE	NO	NO	YES	YES

Table 3: Pass-through Regressions

Input	Q4 2007	Q1 2008
LCD Module Price in Previous Quarter	810.31	789.91
Inverter	0.00	0.00
NTSC Tuner	5.16	0.00
ATSC Tuner Demod	21.85	5.40
Image Processing	21.38	21.32
Audio Processing	10.30	8.60
Power	20.02	25.00
Other Electronics	29.83	28.48
PCB Mechanical	5.76	5.00
Other Mechanical	92.24	85.30
Packaging&Accessories	16.09	16.04
Royalties	10.00	10.00
Labor Overhead	62.58	59.70
Warranty for 12-18 Months	31.29	29.85
USA Import Duty	0.00	0.00
Freight to USA	6.52	6.46
Insurance	4.57	4.36
Handling & Surface	8.04	7.67
Ex-Hub	1155.92	1103.09
Brand Margin (%)	21%	24%
Reseller Margin (%)	25%	27%
Average Street Price	1954.16	1976.00

Table 4: Cost Breakdown Example

$\log(\text{Price})$	(1)	(2)	(3)	(4)
<i>size</i>	0.0519*** (0.0015)	0.0498*** (0.00094)	0.0487*** (0.00074)	0.0496*** (0.00071)
<i>size</i> ²	0.000978*** (0.00010)	0.000867*** (0.000063)	0.000749*** (0.000051)	0.000744*** (0.000048)
<i>FHD</i>	0.0861*** (0.025)	0.262*** (0.016)	0.231*** (0.013)	0.223*** (0.012)
<i>trend</i>		-0.0828*** (0.0018)	-0.0838*** (0.0014)	-0.0890*** (0.0022)
Constant	4.936*** (0.051)	5.577*** (0.035)	5.519*** (0.039)	5.585*** (0.041)
Fixed Effects	x	x	Manuf	Manuf + Time
Observations	1406	1406	1406	1406
R^2	0.66	0.87	0.92	0.93

Table 5: Hedonic Regression for Prices

$\log(\text{Cost})$	(1)	(2)	(3)
<i>size</i>	0.0678*** (0.0080)	0.0573*** (0.0028)	0.0551*** (0.0018)
<i>size</i> ²	-0.000205** (0.000094)	-0.0000959*** (0.000033)	-0.0000714*** (0.000021)
<i>FHD</i>	0.0427 (0.044)	0.0946*** (0.015)	0.104*** (0.0097)
<i>trend</i>		-0.0379*** (0.00064)	-0.0413*** (0.0010)
<i>Constant</i>	3.980*** (0.16)	4.841*** (0.058)	5.138*** (0.045)
Fixed Effects	x	x	Time
Observations	506	506	506
R^2	0.74	0.97	0.99

Table 6: Hedonic Regression for Marginal Costs

	Static	Dynamic
Price	-.0034	-0.0291
FHD	2.2423	1.8412
Size	-0.0251	0.0184
$Size^2$	-.0006	-0.002
σ_p	0.0009	0.0059
σ_s	0.0113	0.00162
σ_{FHD}	0.0233	0.0621
FE	Manuf + Q4	Manuf + Q4
EL	11438	10941

Table 7: Demand Estimates

	Sony 32in HD	Vizio 32in HD
Own Elasticity (Lower)	-0.83	-0.94
Own Elasticity (Upper)	-0.61	-0.72
Same Model	48%	14%
Same Brand-Period	11%	10%
Same Period-Other Brand	27%	39%
Same Brand-Other Period	4%	6%
Other-Other	10%	31%

Table 8: Elasticities and Product Substitution

	25th percentile	75th percentile
δ_{it} only	.62	.89
3 Term Chebyshev	.81	.97

Table 9: R^2 for Regressions of R_{it} on $f(\delta_{it})$

	Q1'06	Q2'06	Q3'06	Q4'06	Q1'07	Q2'07	Q3'07	Q4'07	Q1'08	Q2'08	Q3'08	Q4'08
Cost	843	810	702	601	568	523	517	529	505	502	484	428
Price	1437	1345	1172	899	868	795	727	673	698	662	622	518
Experiment 1	1393	1266	1170	897	848	769	708	678	647	644	621	522
Experiment 3	1453	1373	1191	985	916	830	808	814	765	749	691	611
Experiment 5	1505	1446	1254	1073	1014	934	923	945	902	896	864	764

Table 10: Counterfactual Experiments for 32" HD TV

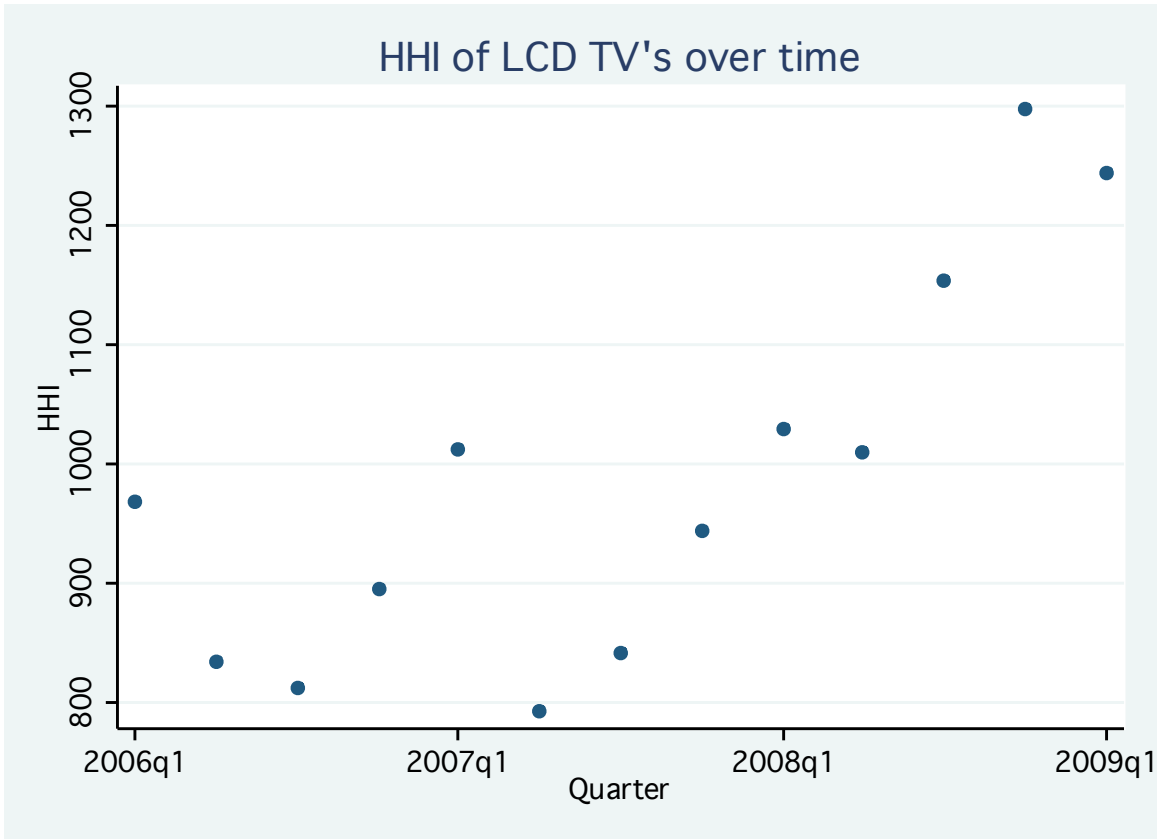


Figure 1: HHI of LCD TV Industry over Time

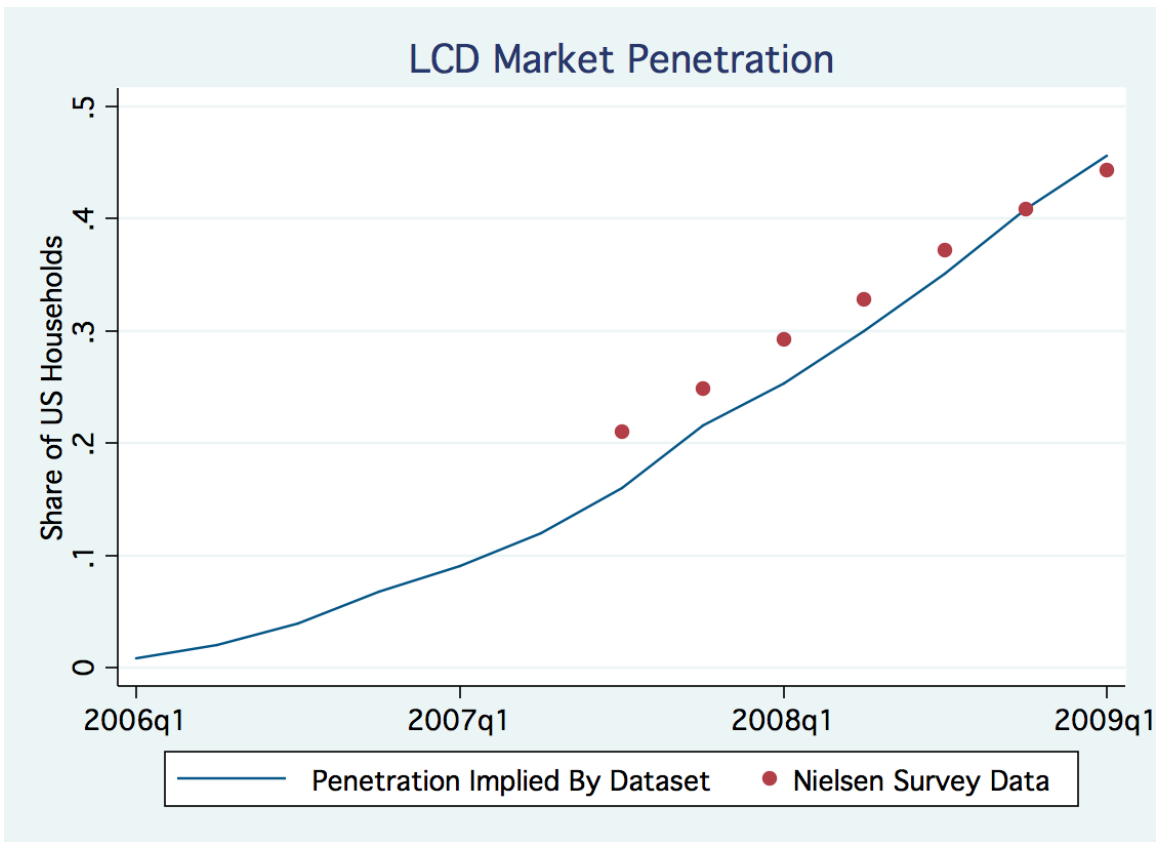


Figure 2: Comparison of Penetration Implied by Dataset (assuming no upgrades) with Household Survey Data

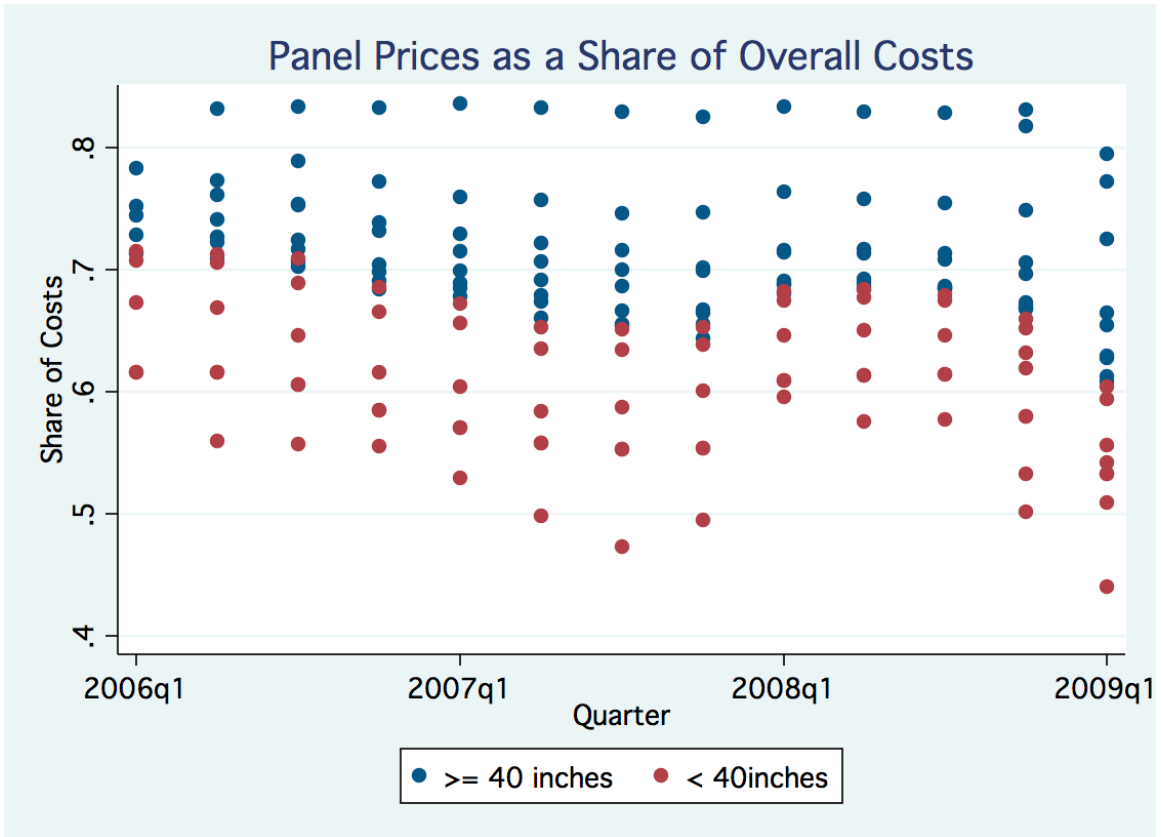


Figure 3: Panel Price as a Share of Overall Costs

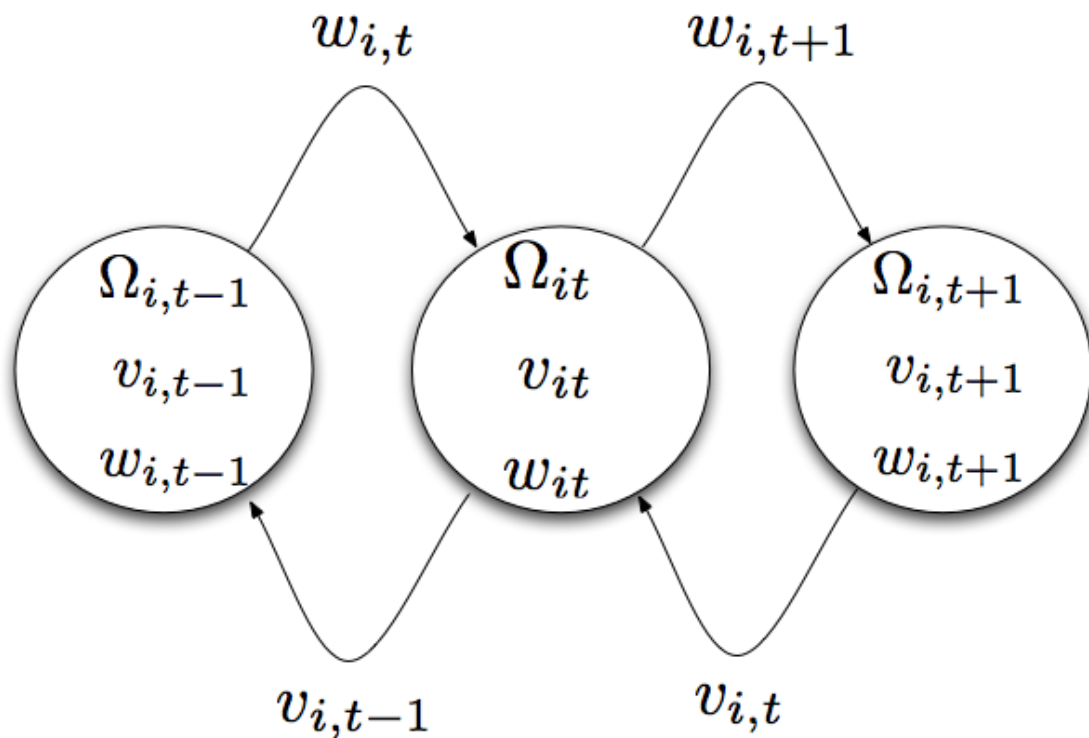


Figure 4: Timing of Dynamic Problem

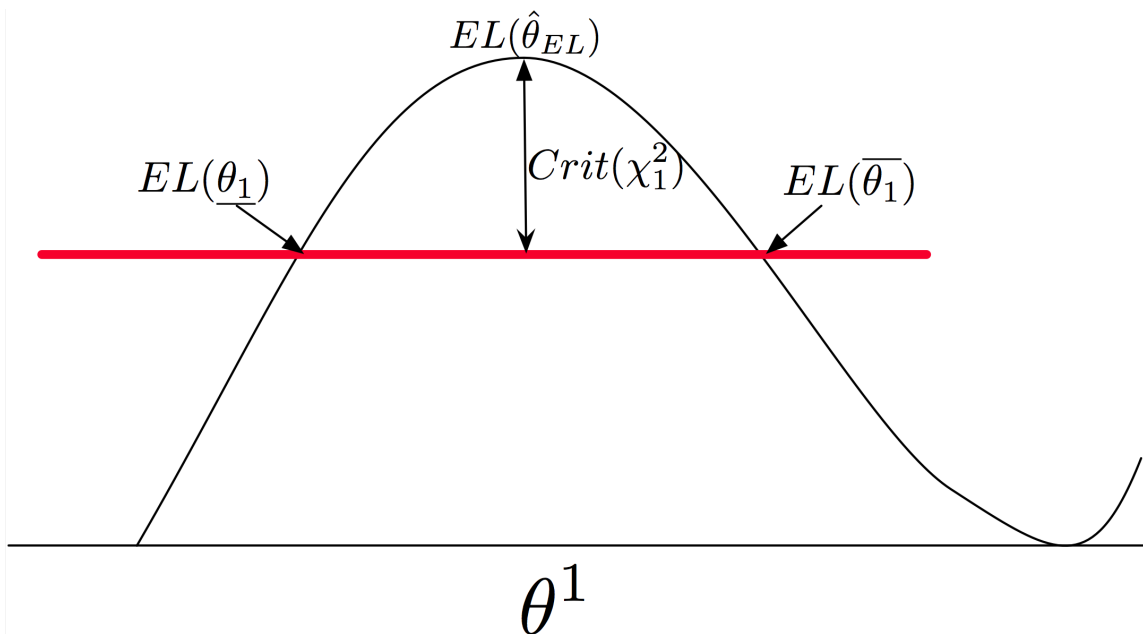


Figure 5: Empirical Likelihood Confidence Interval Construction

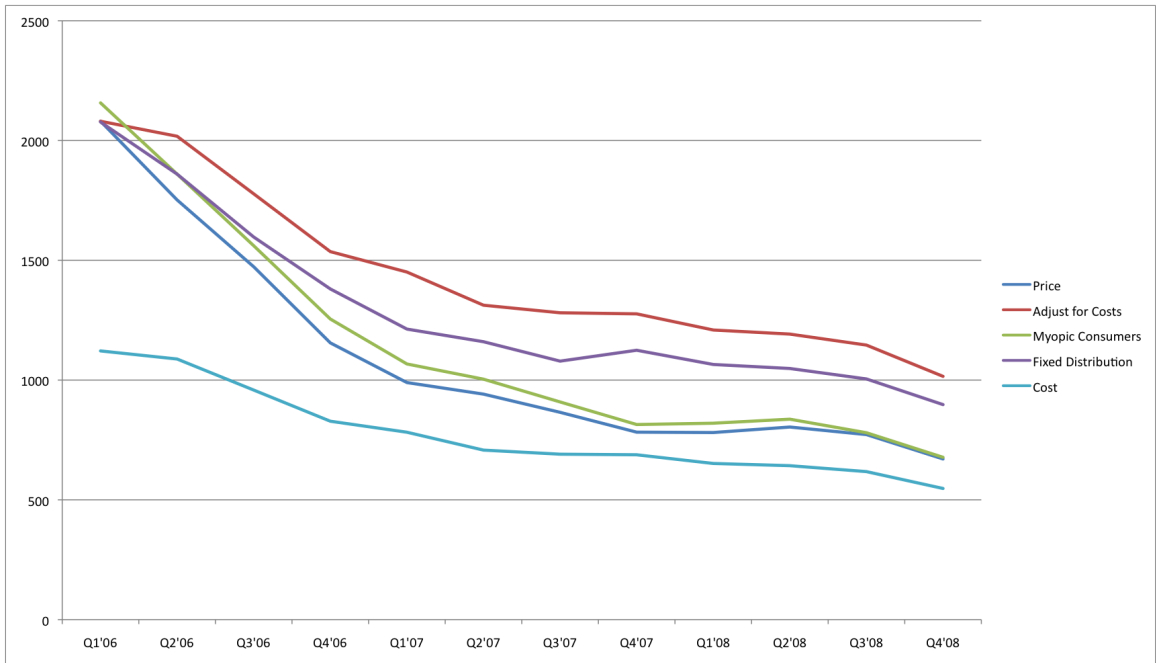


Figure 6: Counterfactual Pricing Under Different Scenarios

Appendix

Static Model

We could imagine a world where in every period consumers choose to purchase the product that gives them the highest utility, where d_{it} denotes the decision of consumer i in period t and u_{ijt} denotes the utility consumer i receives from choosing product j in period t . Because utility is ordinal, not cardinal, it is standard to normalize the mean utility of the outside good to be zero.

$$\begin{aligned}d_{it} &= \arg \max_{j \in \{0, J_t\}} u_{ijt} \\u_{ijt} &= \alpha_i^x x_{jt} - \alpha_i^p p_{jt} + \xi_{jt} + \varepsilon_{ijt} \\u_{i0t} &= \varepsilon_{i0t}\end{aligned}$$

The key addition is the random ε_{ijt} term, this represents an idiosyncratic shock to the consumer's utility. There are two ways to understand this, it might incorporate actual shocks to an individual's utility (it might fit perfectly in some space, the one the consumer is standing in front of when they make a decision, etc.), or it may be some purely statistical error that provides smoothness. When the random term is Type I Extreme Value and IID over all (i, j, t) then consumer choice probabilities take on the well known logit form:

$$s_{ijt} = \frac{\exp[\alpha_i^x x_{jt} - \alpha_i^p p_{jt} + \xi_{jt}]}{1 + \sum_{k \in J_t} \exp[\alpha_i^x x_{kt} - \alpha_i^p p_{kt} + \xi_{kt}]}$$

The goal of estimation is typically to identify the distribution of the tastes α_i by matching observed shares to those predicted by the model. We're typically interested in some parametric distribution $f(\alpha_i|\theta)$ subject to the nuisance parameters ξ

$$s_{jt} = \int s_{ijt}(\alpha_i, \xi) f(\alpha_i|\theta)$$

There are several problems in using a traditional static differentiated products model to study the purchases of TV's. The core problem is that the static model fails to account for the fact that there are tradeoffs between purchasing a product today, and waiting until later.

One aspect is that the utility, u_{ijt} of buying a product may continue for several periods after the purchase is made, rather than all enjoyed immediately. We might consider u_{ijt} to be the present discount value of the flow of future utility from owning the product. But simple adjustments like this are not sufficient. When we consider the outside option 0, as representing "waiting until later", then it should be clear that the assumption $E[u_{i0t}] = 0$ is incorrect, especially if we know tomorrow's products will be better and less expensive than today's products, or if a consumer has already purchased a product in the previous period.¹²

¹²One approach in the literature for handling this problem is adding a time trend to the consumer's utility function. An example of this is (CITE). This is ad-hoc at best unless we really believe that consumers are

The other problem is that in order to identify $f(\alpha_i|\theta)$ we typically make the assumption that the distribution is constant over time, and use repeated observations of the cross section for identification. This might be appropriate for studying durable consumer goods in markets with stable characteristics and prices such as washing machines or refrigerators where replacement is driven on a fixed interval or by random failures. However, in markets where prices and product characteristics are rapidly evolving, we should worry that the individuals who purchase the product in early periods when prices are high look markedly different from the individuals who purchase the product in later periods when prices are low.

Empirical Likelihood MPEC

The typical exposition follows Kitamura (2006), which provides an extensive and accessible survey of the EL literature. Assume data $\{z_i\}_{i=1}^n$ are distributed IID according to unknown measure μ . The econometric model gives us some moment conditions. The moment conditions, $g(z, \theta) \in \mathbb{R}^q$ may be scalar valued ($q = 1$) or vector valued ($q > 1$).

$$E[g(z_i, \theta)] = \int g(z, \theta) d\mu = 0, \quad \theta \in \Theta \in \mathbb{R}^k$$

We can assume (w.l.o.g.) the existence of a nonparametric (multinomial) (log) likelihood function. Then, we search for a measure (set of probability weights) $(p_1, \dots, p_n) \in \Delta$ within the unit simplex, where each p_i corresponds to weight of each observed z_i in the data. These weights are then manipulated so that the moment conditions hold exactly. The empirical likelihood estimator is defined as the solution to the following optimization problem:

$$\hat{\theta}_{EL} = \arg \max_{\theta, p_1, \dots, p_n} l_{NP}(\theta) = \sum_{i=1}^n \log p_i \text{ s.t. } \sum_{i=1}^n p_i g(z_i, \theta) = 0, \text{ and } (p_1, \dots, p_n) \in \Delta$$

EL offers a different interpretation of moment condition models. While GMM asks, “How close (in a space defined by some weight matrix) can I get the moments of my model to match the moments of my data?”; EL asks, “How unusual a draw from the multinomial distribution of data would my data represent if my moment conditions held exactly?”. EL and GMM produce the same asymptotic distribution. As a limiting case, in a model without over-identifying restrictions (where the maximized GMM objective is zero), the optimal weights are the empirical weights $\frac{1}{n}$ and the EL objective function would be $-n \log n$.

EL has a number of advantages that make it attractive to applied researchers. Most of these advantages are due to the fact that EL avoids estimating the GMM weighting matrix. There is a large literature on why estimating the weight matrix can be difficult, and when it can lead to problems. An well-known example is Altonji and Segal (1996) who show that the optimal weighting matrix used in Abowd and Card (1989) exhibits bias as the number of moments increases and as moments become more correlated. Newey and Smith (2004) provide formal results to show that the source of the bias comes from the correlation

becoming unhappier over time.

of moments with their Jacobian (especially in the case of endogeneity), and the weighting matrix. They also show that generalized empirical likelihood estimators do not exhibit this bias and are higher-order efficient. Conlon (2012) shows that in Monte-Carlo experiments for the static demand problem of Berry, Levinsohn, and Pakes (1995) EL exhibits less bias and tighter confidence intervals than GMM.

The principal reason that empirical likelihood methods are not more popular, is that they are believed to be computationally difficult relative to GMM. Previous approaches generally focus on the unconstrained dual problem, which is appropriate for solving linear models, but becomes challenging when solving nonlinear models:

$$\hat{\theta}_{EL} = \arg \max_{\theta \in \Theta} l_{NP}(\theta) = \arg \max_{\theta \in \Theta} \min_{\lambda \in \mathbb{R}^q} - \sum_{i=1}^n \log(1 + \lambda' g(z_i, \theta))$$

An additional advantage of the MPEC method is that the constrained version of EL is no more difficult to estimate than GMM, and in many cases easier. Conlon (2012) also demonstrates how the MPEC approach of Su and Judd (2008) and Dube, Fox, and Su (2009) can be combined with the EL estimator and applied to the demand estimation problem of Berry, Levinsohn, and Pakes (1995). The resulting constrained optimization problem can be written as:

$$\arg \max_{p, \xi, \theta, w, R} \sum_{j,t} \log p_{jt} \text{ s.t. } \sum_{j,t} p_{jt} \xi_{jt} z_{jt} = 0 \text{ and } S_{jt} = s_{jt}(\xi, \theta) \text{ and } C(R, w, \theta, \xi) = 0 \quad (21)$$

By combining this objective and constraints with the constraints from the dynamic durable goods problem $C(R, w, \theta, \xi) = 0$ in (8)-(10), we can construct a one-step EL estimator.

Empirical Likelihood/MPEC Inference

Since the MPEC method represents a different computational algorithm, rather than a different statistical estimator, inference can proceed exactly as it does in the fixed point case. In the example of the static demand estimation problem, it is well known that the asymptotic standard errors depend on $D = E \frac{\partial}{\partial \theta} g(z_{jt}, \theta)$ and $S = E[g(z_{jt}, \theta)g(z_{jt}, \theta)']^{-1}$

$$\begin{aligned} \sqrt{n}(\hat{\theta} - \theta) &\rightarrow^d N(0, D'S^{-1}D) \\ g(z_{jt}, \xi, \theta) &= \xi_{jt}(\theta) \times z_{jt} \\ \frac{dg}{d\theta} &= \frac{\partial g}{\partial \theta} + \frac{\partial g}{\partial \xi} \cdot \frac{\partial \xi(\theta)}{\partial \theta} = \frac{\partial \xi(\theta)}{\partial \theta} \times z_{jt} \end{aligned} \quad (22)$$

A small challenge is that $\frac{\partial \xi(\theta)}{\partial \theta}$ is not directly defined in the constrained problem, because ξ does not directly depend on θ . In the traditional fixed-point/contraction mapping approach this is the gradient of the contraction mapping, but must generally be computed numerically. The fixed point approach provides the derivative of ξ along the manifold of the share constraint. To perform inference in the MPEC framework, there are two options. One is

to use the fixed-point approach to compute standard errors after estimation is done this is what Dube, Fox, and Su (2009) implement in their MPEC-GMM routine, the other is to use the results of MPEC estimation and the implicit function theorem to compute the appropriate objects for computing standard errors as suggested by Su and Judd (2008). As an alternative, Su and Judd (2008) recommend a bootstrap procedure for maximum likelihood problems.

As the model becomes more complicated and the parameter space expands, implementing the fixed point approach for a dynamic model such as Gowrisankaran and Rysman (2009) can become more complicated. One way to see this is that it requires computing $\nabla_{\theta} g = \frac{\partial g(z_{jt}, \theta)}{\partial \theta}$ along the manifold defined by a number of additional constraints. Likewise, it might require several iterative applications of the implicit function theorem to define the standard errors.

Empirical Likelihood admits the same form of asymptotic distribution as GMM, where $V = D'S^{-1}D$, except that D and S are computed under the empirical likelihood weights p_{jt} rather than the empirical weights $\frac{1}{n}$, in order to obtain the higher order efficiency property. It is possible to construct the standard errors using the same methodology described above for the GMM case. However, empirical likelihood presents an alternate solution to this problem, whereby we can use the empirical likelihood ratio test statistic. The ELR test statistic is computed as follows. Estimate an unrestricted model and compute the corresponding empirical likelihood $l_{EL}(\hat{\theta})$. Then estimate a model with a linear restriction on the parameter $R(\theta) = 0$ and call this $l(\hat{\theta}_0)$. Then we can construct the ELR test statistic, where r is the dimension of R .

$$-2(l(\hat{\theta}_0) - l(\hat{\theta}_{EL})) \sim \chi_r^2$$

We can directly test for significance using this test statistic. Since the test statistic is just a χ^2 with r degrees of freedom we can easily compute the critical value c_{α} for some significance level, and construct a confidence interval for θ^k an element of θ by inverting the test:

$$[\underline{\theta}^k, \bar{\theta}^k] = \min \theta^k / \max \hat{\theta}^k \in N(\hat{\theta}_{EL}) \text{ s.t. } l(\theta) \geq l(\hat{\theta}_{EL}) - 2c_{\alpha}$$

We search for the largest and smallest value of the element of interest θ^k that produces an EL objective value close enough to the unrestricted value. This is demonstrated graphically in Figure 6. This procedure must be repeated for each parameter for which we want to construct a confidence interval. An advantage is that constructing the confidence interval by inverting the test often has better statistical properties than confidence regions constructed from the asymptotic distribution (Kitamura 2006)¹³.

¹³I need to track down the correct citation here!