Consumer Inertia and Market Power

Alexander MacKay
Marc Remer

Consumer Inertia and Market Power

Alexander MacKay
Harvard Business School

Marc Remer
Swarthmore College

Working Paper 19-111

Consumer Inertia and Market Power[*]

Alexander MacKay
Harvard University[†]

Marc Remer
Swarthmore College[‡]

April 5, 2021

Abstract

We study the pricing decision of firms in the presence of consumer inertia. Inertia, which can arise from habit formation, brand loyalty, switching costs, or search, has important implications for equilibrium outcomes and affects market power. Holding market structure fixed, greater inertia can increase or decrease prices. Changes to market structure—e.g., through a merger—can have lower impacts on prices in the presence of inertia. Thus, consumer inertia plays an important role in mediating horizontal competition. We develop an empirical model to estimate consumer inertia using aggregate, market-level data. We apply the model to a hypothetical merger of two major retail gasoline companies, and we find that a static model can predict price increases greater than the price increases predicted when accounting for dynamics.

---

1 Introduction

Consumers are often more likely to buy a product if they have purchased it previously. This tendency reflects both exogenous preferences and state-dependent utility that is affected by past behavior (Heckman, 1981). Consumer state dependence, or inertia, may arise from habit formation, brand loyalty, switching costs, or search. There is a rich empirical literature that establishes the presence of such inertia in a variety of markets.[1] This behavior is typically identified by examining the choice patterns of individual consumers over time.

In response to consumer inertia, profit-maximizing firms will internalize the effect of their current price on demand in future periods. Over a long horizon, this will often lead firms to invest in future consumers by maintaining lower prices than the short-run optimum. However, these prices are typically higher than in a counterfactual world in which consumers do not exhibit state dependence in purchasing behavior. We define the extent to which prices are higher due to consumer affiliation as *dynamic market power* over consumers. Dynamic market power tends to increase with the rate consumers become affiliated. We contrast this with *horizontal market power*, which is the ability of firms to raise prices in response to a reduction in the number of independent competitors. To assess the effects of mergers or other changes to market structure, it is necessary to account for both of the dynamic and horizontal dimensions of market power.

In this paper, we study the interaction of these forces using a dynamic oligopoly model where firms compete in prices. We model consumer inertia as a product-specific preference "affiliation" for the most recently purchased product. This formulation nests typical implementations of dynamic consumer behavior, including switching costs and brand loyalty. We use the model to evaluate the impact of consumer inertia on steady-state prices while varying the degree of competition in the market. We then focus our analysis on the impact of horizontal mergers, which changes the level of competition, on equilibrium prices. We demonstrate that accounting for consumer inertia is critical to accurately predict the price effects of mergers. Given that firms and competition authorities rely on such predictions when evaluating mergers, our findings highlight the importance of accounting for dynamic pricing incentives.

Competition authorities will challenge a merger if the merging firms are expected to increase the price by a significant amount. To construct these predictions, competition authorities typically employ static empirical models to estimate consumer demand and simulate counterfactual prices. If consumer inertia is present in the market, these static models omit the first-order incentive of a firm to reduce price to increase future demand. Consequently, the static model instead attributes this effect to less differentiation among competing firms. This can cause static models to substantially over-predict the price effects of a merger. In specific settings, which we detail below, static models can also systematically under-predict the price effects of a

---

[1]See the related literature section for examples.

merger. Thus, failing to account for consumer inertia may misstate the potential for horizontal market power and affect merger enforcement.

Competition authorities often analyze mergers in markets that are likely to be characterized by consumer affiliation. For example, in its lawsuit against Swedish Match-National Tobacco, the Federal Trade Commission (FTC) cited strong brand loyalty as a barrier to entry.[2] Similarly, the US Department of Justice cited customer switching as an important factor in its case against the UPM-MACtac merger.[3] In defense of its acquisition of TaxACT, H&R Block cited the importance of dynamic incentives in exerting downward pricing pressure post-merger (Remer and Warren-Boulton, 2014). Both the FTC and DOJ routinely investigate mergers in consumer product markets, where inertia in brand choice has been documented. Yet, perhaps due to computational complexity and compressed investigative timelines, dynamics are seldom directly modeled when analyzing unilateral competitive effects. We develop a model that significantly reduces the computational and data burden of estimating the impact of these dynamics.

We begin by considering the theoretical properties of our model using numerical simulations. We demonstrate that affiliation can have a large effect on steady-state prices. In equilibrium, affiliation raises oligopoly prices and often has a greater impact on price than a reduction in competition. In evaluating horizontal market power, we show that accounting for consumer inertia and also the *type* of merger is critical. A merged firm can decide to maintain separate brands or consolidate merged brands into a single entity. In a static setting, this distinction may not be relevant; we show that mergers of either type can deliver the exact same outcome with no consumer inertia. However, these two types of mergers can deliver directionally opposite outcomes in the presence of consumer inertia. These findings highlight the need for empirical models to appropriately represent firms' dynamic pricing incentives, and for counterfactual exercises to precisely define the structure of the but-for world.

We then develop an empirical model of consumer inertia that can be estimated using aggregate, market-level data. The demand model we introduce is a straightforward extension of the standard discrete-choice logit model with myopic consumers. In contrast to the typical random-coefficients model, we allow the distribution of unobserved heterogeneity to be state-dependent, affected by past purchase behavior. We restrict the random coefficients to affect a single product for each consumer type, corresponding to our notion of product-level affiliation. To disentangle heterogeneity in preferences from state dependence arising from previous purchases, we impose a demand system and rely on the panel structure of our data. Intuitively, after adjusting for cross-sectional fixed effects, consumer inertia is captured by the residual correlation in shares over time (and their relation to prices).

An key advantage of the model is that can be estimated using aggregate market-level shares and prices, which is the typical data used in demand estimation and merger simulation. Our

---

[2]See, "Commentary on the Horizontal Merger Guidelines," 2006, https://www.justice.gov/atr/file/801216/download
[3]Ibid.

model allows us to separately identify each firm's share from unobserved consumer types and states. Thus, we allow for endogenous unobserved heterogeneity through the presence of a serially correlated state variable for each firm. This flexibility has traditionally been a challenge for the estimation of dynamic models. Importantly, the model can also be estimated independently of supply-side assumptions, and we can therefore use the estimated demand model to test for forward-looking behavior by firms.

We apply the model using data from retail gasoline markets, which have pricing patterns consistent with consumer affiliation, such as slow-to-adjust cost pass-through (e.g., Lewis and Noel, 2011). Furthermore, retail gasoline has a direct link to antitrust concerns. In June and December of 2017, the FTC challenged Alimentation Couche-Tard's acquisitions of Empire Petroleum Partners[4] and Holiday,[5] respectively, on the basis of overlapping retail gasoline stations in a number of states.

We estimate the model using a rich panel dataset of prices, shares, and costs for retail gasoline stations. In this context, the model is best interpreted as one of habit formation or consumer inattention, wherein some consumers return to the gas station from which they previously purchased without considering alternative sellers. We find evidence of strong demand dynamics. We estimate that 62 percent of consumers have the tendency to become affiliated to the brand from which they previously purchased.[6] The remaining 38 percent are "shoppers" that are unaffected by consumer inertia. Consumers that become affiliated to a brand are not very price sensitive, with an average elasticity of only -0.55. Shoppers are extremely price sensitive. Though shoppers are a minority of all consumers that purchase gasoline, they are important for disciplining prices in equilibrium. Across all consumer types, firms face an average elasticity of -3.4.

To highlight the importance of accounting for dynamics when predicting firm behavior, we impose a supply-side model of price competition. In contrast to the literature, we invoke relatively weak assumptions about supply-side behavior in order to conduct counterfactual analysis. From the estimated demand model, we obtain the derivative of static profits, which we use to infer the dynamic component of the firms' first-order conditions. We project these estimates onto state variables to construct a reduced-form approximation to the dynamic pricing incentives. This approximation is consistent with a model of Markov perfect equilibrium where firms use limited state variables to forecast their continuation value. Using this approximation, we evaluate horizontal and dynamic market power in our empirical setting.

To evaluate horizontal market power, we perform a merger analysis between two major

---

[4]https://www.ftc.gov/news-events/press-releases/2017/06/ftc-requires-retail-fuel-station-convenience-store-operator

[5]https://www.antitrustalert.com/2017/12/articles/ftc-developments/the-latest-ftc-challenges-retail-fuel-station-and-convenience-store-transaction-requires-ten-localized-divestitures-in-wisconsin-and-minnesota/

[6]Interestingly, the National Association of Convenience Stores, a retail fueling lobbying association, found in its 2018 annual survey that 57 percent of respondents have a preference for a specific brand to fill up gasoline. See, https://www.convenience.org/Topics/Fuels/Documents/How-Consumers-React-to-Gas-Prices.pdf

gasoline retailers and re-compute the price-setting equilibrium in each period. With the dynamic model, we estimate that brand consolidation would increase prices for the merging firms by 2.0 percent post-merger. A static model, on the other hand, predicts an average price increase of 5.4 percent, which would likely result in greater antitrust scrutiny. Therefore, the dynamic incentive to invest in future demand mitigates the increase in horizontal market power that arises from a merger. With the dynamic model, a merger with joint pricing for separate brands results in a price increase of 3.3 percent for the merging firms.

To evaluate dynamic market power, we calculate equilibrium prices if we increase or decrease the share of consumers affected by inertia. An increase of 7 percentage points, which shifts the population affected by inertia from 62 percent to 69 percent, would result in prices that are 2.0 percent higher.[7] The effect on prices and firm profits are similar to the merger effects with brand consolidation. Thus, a relatively small share of shoppers can reduce prices by as much as the introduction of a major competitor.

Prior to estimating the model, we present reduced-form evidence of dynamic demand and dynamic pricing in retail gasoline markets.[8] Consistent with investment in affiliated consumers, we find that new entrants initially price lower than established firms but raise prices over time. We then examine cost pass-through. Using the data to separate out expected and unexpected costs, we show that firms respond differentially to these two measures. Firms begin raising prices in anticipation of higher costs approximately 28 days prior to an expected cost shock.

Related Literature

We consider the implications of consumer state dependence on the pricing behavior of firms, building on an empirical literature that includes Dubé et al. (2009). We contribute to the literature by examining the effect of competition on price in such settings. State dependence can have a large effect on the interpretation of outcomes when studying inter-firm competition. Our analysis of mergers complements theoretical work on dynamic price competition when consumers are habit-forming or have switching costs. Such features link directly to our notion of consumer affiliation. For examples, see Farrell and Shapiro (1988), Beggs and Klemperer (1992), and Bergemann and Välimäki (2006).[9] Our empirical model can be used to assess the impacts of competition using real-world data.

We contribute to the empirical literature that estimates state dependence in consumer preferences. Meaningful switching costs, due to brand loyalty or consumer inertia, have been found in consumer packaged goods (Shum, 2004; Dubé et al., 2010), health insurance (Handel, 2013), and auto insurance (Honka, 2014). Hortaçsu et al. (2017) find that consumer inatten-

---

[7]Conversely, a decrease by 7 percentage points would result in prices that are 1.7 percent lower.

[8]By dynamic pricing we mean that there are intertemporal spillovers. This should not be confused with static pricing in response to changing market conditions, which is often colloquially referred to as "dynamic pricing."

[9]There are alternative strategic reasons for dynamic pricing, including experience goods (Bergemann and Valimaki, 1996), network effects (Cabral, 2011), learning-by-doing (Besanko et al., 2018), and search (Stahl, 1989).

tion and brand loyalty lead to substantial inertia in retail electricity markets. Conceptually, our model shares similar features to that of Dubé et al. (2009) and Dubé et al. (2010). However, the existing literature has primarily relied on consumer-specific purchase histories to document state dependence, whereas our method allows for the recovery of such state dependence using aggregate, market-level data. For an analysis of inter-firm competition, such data tends to be more readily available. One paper that has used aggregate data to estimate switching costs is Shcherbakov (2016), in the context of television services. He provides an intuitive argument for the identification of switching costs from aggregate data, placing more restrictions on the form of inertia than we allow for.[10] We provide an identification argument for a model that nests several forms of consumer inertia.

We contribute to a growing body of empirical models of dynamic demand. Existing work focuses on different contexts that drive dynamic behavior. Hendel and Nevo (2013) consider a model with storable goods and consumer stockpiling. Gowrisankaran and Rysman (2012) and Lee (2013) consider the purchase of durable goods with forward-looking behavior by consumers. In contrast to these papers, we focus on settings with positive dependence in purchasing behavior over time. The literature highlights the issue, common to our setting, that misspecified static models can produce bias elasticities. Hendel and Nevo (2013) point out that this will matter in a merger analysis. We complement this point by providing a case in which the dynamic incentives, rather than biased elasticities, are a primary concern in model misspecification.

For our empirical application, we propose a reduced-form method to approximate the dynamic incentives in supply-side pricing behavior, which allows us to side-step some of the challenges present in the estimation of dynamic games. Compared to value-function approximation methods proposed by Bajari et al. (2007) and Pakes et al. (2007), we rely more heavily on the structure of the demand model and place weaker assumptions on supply-side behavior. We circumvent some of the computational challenges with estimating the value function (see, e.g., Farias et al., 2012; Sweeting, 2013) by estimating its derivative directly, which eliminates a recursive step. We motivate our estimation of this function as either a limited-information equilibrium concept or an approximation to full-information behavior by firms, as in Weintraub et al. (2008) and subsequent work. Our focus on the pricing behavior of firms precludes the use of several developments in the conditional choice probabilities literature, which relies on discrete actions (e.g., Aguirregabiria and Mira, 2007; Arcidiacono and Miller, 2011).

---

[10]We are aware of two other papers that estimate switching costs using aggregate data: Nosal (2011) and Yeo and Miller (2018). These papers have less formal identification arguments than Shcherbakov (2016).

2 A Model of Oligopolistic Competition with Consumer Inertia

We develop a dynamic model of oligopolistic competition with product differentiation where consumers may become affiliated with the firm from which they purchased previously. Affiliation may be interpreted as habit formation, brand loyalty, switching costs, or search. Consumers in the model are myopic in that they maximize current period utility rather than a discounted flow of future utility. This assumption is likely a good fit for retail gasoline markets, where consumers do not choose a gas station anticipating that it will limit their future choice set; rather, some consumers are likely to return to the same gas station due to habit-formation, brand loyalty, or inattention.

As detailed below, we introduce consumer dynamics by allowing for endogenous unobserved heterogeneity in a differentiated product demand model. We then place the demand model into a dynamic oligopoly setting. Even though consumers are myopic, key dynamics arise when firms internalize the effect of sales today on future profits through the accumulation of affiliated consumers.[11] We use the model to numerically and empirically examine the impact of consumer inertia on market power, in general, and in the context of horizontal mergers.

2.1 Demand

We extend the standard logit discrete choice model to allow for unobserved heterogeneity that depends on past purchases. The first assumption below presents a random coefficients utility formulation with myopic choice. The second assumption restricts the random coefficients so that the type-specific utility shock affects only a single product, corresponding to our notation of consumer affiliation. The third assumption places restrictions on the evolution of consumer types over time.

Assumption 1: Myopic Discrete Choice Consumers in each market select a single product $j \in J$ that maximizes utility in the current period, or they choose the outside good (indexed by 0). When unobserved heterogeneity is exogenous, consumers may be indexed by discrete types.[12] We allow the distribution of preferences to change endogenously over time, therefore, heterogeneity in preferences is represented here by *states*. In any period, each consumer is indexed by a discrete state $z \in Z$.

A consumer n in state z receives the following utility for choosing product j:

$$u_{jt}^{(n)}(z) = \delta_{jt} + \sigma_{jt}(z) + \epsilon_{jt}^{(n)}. \tag{1}$$

---

[11]Slade (1998) estimates a model of habit-forming consumers and sticky prices. That model, however, explicitly imposes a cost of price adjustment. Our model does not rely upon a menu cost to explain dynamic price adjustments.

[12]The discrete type assumption for the random coefficient model is made elsewhere in the literature, though these types are assumed to be exogenous. See, for example, Berry et al. (2006) and Berry and Jia (2010).

Consumers receive an additively-separable common component δ_{jt}, a state-dependent shock $\sigma_{jt}(z)$, and an idiosyncratic shock, $\epsilon_{jt}^{(n)}$. The common component will typically be a function of firm j's price, and takes the form $\delta_{jt} = \xi_j + \alpha p_{jt}$ in the standard logit model (with $\alpha < 0$). The state-dependent shock, $\sigma_{jt}(z)$, may be also be a function of firm j's price, if consumers are less sensitive to the price of the product to which they are affiliated.

We denote the probability that a consumer in state z chooses product j as $s_{jt}(z)$. We normalize the utility of the outside good to be zero. Additionally, we define the state 0 consumer to be "unaffiliated" with no state-dependent preference, i.e., $\sigma_{jt}(0) = 0 \,\forall\, j$. Given the standard assumption of a type 1 extreme value distribution on the utility shock, $\epsilon_{jt}^{(n)}$, the choice probabilities of consumers are:

$$s_{jt}(0) = \frac{\exp(\delta_{jt})}{1 + \sum_k \exp(\delta_{kt})} \tag{2}$$

$$s_{jt}(z) = \frac{\exp(\delta_{jt} + \sigma_{jt}(z))}{1 + \sum_z \exp(\delta_{kt} + \sigma_{jt}(z))}. \tag{3}$$

The observed share for firm j is given by the weighted average of choice probabilities for consumers in each state: $S_{jt} = \sum_{z=0}^{Z} r_{zt} s_{jt}(z)$, where r_{zt} is the fraction of consumers in state z.

The mean utility δ_{jt} may depend on time varying-observable characteristics as well as fixed effects. In the empirical application, we makes use of this latter feature to allow for serial correlation in unobservable utility shocks over time.

Assumption 2: Single-Product Affiliation   We now place restrictions on the state-dependent demand shocks, $\sigma_{jt}(z)$. We assume that each consumer state corresponds to an affiliation (utility shock) to a single product. Further, we assume that there is a single state corresponding to each product. Thus, a consumer in state j is affiliated to product j. The state-specific demand shocks for other products are zero, i.e. $\sigma_{jt}(z) = 0 \,\forall\, z \neq j$.

Thus, we define *affiliation* to be a product-specific state dependence in preferences. We say that a consumer in state j is affiliated with product j, as this consumer has a perceived benefit of $\sigma_{jt}(j)$ relative to unaffiliated (state 0) consumers. The affiliation shock $\sigma_{jt}(j)$ has different interpretations depending on the the underlying mechanism:

- *Brand loyalty:* The model has a direct interpretation of brand loyalty when $\sigma_{jt}(j)$ is a positive level shock that reflects an internal benefit for purchasing from the same brand.

- *Switching costs:* The model may be interpreted as a switching cost model when $\sigma_{jt}(j)$ is a level shock representing the costs (physical and psychic) of switching to another brand. This interpretation is empirically indistinguishable from the brand loyalty model because only the relative utilities affect choices in the logit formulation of the discrete choice model.

- *Habit formation:* In the habit formation interpretation, a consumer gets either an extra benefit for repeating earlier behavior or bears a cost for adjusting behavior. $\sigma_{jt}(j)$ represents the net benefit. In contrast to the switching cost model, other aspects of preferences may change. For example, consumers may become less price sensitive to the affiliated product, in addition to realizing a level shock.

- *Search/Inattention:* In the special case where $\sigma_{jt}(j)$ renders affiliated consumers inelastic, the model has a search or inattention interpretation. In this case, the unaffiliated consumers are those that engage in search and realize full information about the choice set. Affiliated consumers are inattentive and simply buy the previous product. This extends a standard search model (e.g., Varian, 1980; Stahl, 1989) by (i) having non-searchers default to the previous product, rather than randomizing, and (ii) endogenizing the distribution of searchers and non-searchers.

Distinguishing among these different mechanisms lies outside of the scope of this paper but may be important, especially when examining questions about welfare. The brand loyalty model and the switching cost model can have identical outcomes but divergent welfare predictions, as $\sigma_{jt}(j)$ is a net benefit in the former and a net cost in the latter.

Assumption 3: Evolution of Consumer States   Our framework allows us to consider both endogenous and exogenous unobserved heterogeneity. We consider unobserved exogenous heterogeneity in a simple form by assuming that there are two latent consumer types. A fraction λ of consumers are affected by state dependence. Each of these consumers becomes affiliated with the product they purchased in the previous period. The remaining $1 - \lambda$ fraction of consumers are unaffected by state dependence. These represent "shoppers" that have no costs of switching among different products over time. This fraction of consumers remain in state 0 regardless of their choices.

Let $\lambda^{(n)} = 1$ if consumer n is affected by state dependence and $\lambda^{(n)} = 0$ otherwise, and let $a^{(n)}$ denote consumer n's choice in the previous period. We can write the indirect utility for consumer n of choosing product j as:

$$u_{jt}^{(n)}(z) = \delta_{jt} + \mathbb{1}[j = a^{(n)}]\lambda^{(n)}\sigma_{jt} + \epsilon_{jt}^{(n)}. \tag{4}$$

The evolution of states follows a Markov process, where the state can be expressed as a function of the joint distribution of states, types, and choices in the previous period. The coefficient $\lambda^{(n)}$ captures exogenous unobserved heterogeneity, and the coefficient $\mathbb{1}[j = a^{(n)}]$ captures unobserved heterogeneity that is endogenous and depends on past prices.

Based on these assumptions, we can represent the state *of the market* in each period by the vector $\{r_{jt}\}$, where r_{jt} denotes the share of λ-type consumers affiliated to product j in period

t. Since λ-type consumers may be unaffiliated, $\sum_{j=0}^{J} r_{jt} = 1$. λ-type consumers that choose the outside good transition into state 0 for the next period but maintain their latent type.

2.2 Supply

We assume that firms set prices to maximize the net present value of profits. We restrict attention to Markov perfect equilibria. Current-period profits are a function of shares. Firm j's aggregate share across consumer types and states is:

$$S_{jt} = (1 - \lambda)s_{jt}(0) + \lambda \sum_{z=0}^{J} r_{zt}s_{jt}(z). \tag{5}$$

Thus, a firm's total share of sales can be written as a weighted sum of its share of unaffiliated consumers, $s_{jt}(0)$, and affiliated consumers, $s_{jt}(z)$. Note that firm j will make sales to consumers affiliated to other firms $z \neq j$, but the probability that such consumers will choose firm j is strictly lower than the choice probability of an unaffiliated consumer when the utility shocks $\{\sigma_{jt}(z)\}$ are positive.

Assumption 4: Competition in Prices We assume that firms set prices in each period to maximize the net present value of profits from an infinite-period game. Prices are set as a best response conditional on the state and contemporaneous prices of rival products. Firms cannot commit to future prices. The state vector in each period is summarized by marginal costs, c_t, the distribution of affilation across consumers, r_t, and other variables that are captured by the vector, x_t, such as expectations about future costs. Entry is exogenous.[13] The firm's objective function can be summarized by the Bellman equation:

$$V_j(c_t, r_t, x_t) = \max_{p_{jt}|p_{kt}, k \neq j} \left\{ (p_{jt} - c_{jt})S_{jt} + \beta E(V_j(c_{t+1}, r_{t+1}, x_{t+1})|p_t, c_t, r_t, x_t) \right\}. \tag{6}$$

Prices in each period optimize the sum of current-period profits $(p_{jt} - c_{jt})S_{jt}$ and the continuation value. Both of these components depend upon marginal costs and the distribution of consumer states, r_t. Thus, when the perceived continuation value is non-zero, firms anticipate how price affects the future distribution of consumer states and also the impact of future changes to marginal costs. Note that the state space does not include previous period prices. We therefore exclude strategies that depend directly upon competitors' historical prices, such as many forms of collusion.

Assumption 5: Expectations Consistent with the Markov perfect framework, we make the relatively weak assumption that the continuation value function is stable conditional on the

---

[13]Entry and exit are infrequent in retail gasoline.

state and prices. In contrast to the typical setup for a dynamic game, we place minimal restrictions on expectations, discount rates, and the perceived continuation value. Instead, our empirical approach is to directly estimate a reduced-form model of (the derivative of) the continuation value. We describe this approach in more detail in Section 5.

Thus, market equilibrium is characterized by consumers making (myopic) utility-maximizing purchase decisions and firms pricing as the best response to other firm's prices, conditional on the state.

2.3 Theoretical and Numerical Analysis

To develop an understanding of pricing incentives in markets with consumer inertia, we consider a deterministic setting where marginal costs are constant. We use numerical methods to analyze steady-state prices in an oligopoly game with Bertrand price-setting behavior.[14] For the numerical exercise, we specify the utility parameters from equation (1) as $\delta_{jt} = \xi_j + \alpha p_{jt}$ and $\sigma_{jt}(j) = \bar{\xi}$, allowing for the affiliation shock to affect only the utility level. Thus, the utility of a consumer n in state z for product j is as follows:

$$u_{jt}^{(n)}(z) = \delta_{jt} + \mathbb{1}[j = z]\lambda^{(n)}\bar{\xi} + \epsilon_{jt}^{(n)}. \tag{7}$$

Here, $\mathbb{1}[j = z]$ takes a value of one if a consumer that purchases product z in period $t - 1$ purchases product $j = i$ in period t, and zero otherwise. Consumer consumers subject to affiliation always have a value of $\lambda^{(n)} = 1$, and $\lambda^{(n)}$ is zero otherwise. Assuming the error term, $\epsilon_{jt}^{(n)}$, follows the type-1 extreme value distribution, we get the market share specified in equations (2) and (3) in the previous section, which reduces to the standard logit model if $\bar{\xi} = 0$.

Each firm i sells a set of products, $j \in J_i$, and maximizes the expected discounted value of profits. Therefore, firm i's value function takes the following form:

$$V_i(r) = \max_{p_i|p_{-i}} \pi_i(p, r) + \beta V_i(r'). \tag{8}$$

Here, p and r are vectors of prices and affiliated customers, respectively, and r' is a vector specifying each firm's affiliated customers in the next period. In accordance with the model above, an element of r' is $r_i' = f(p, r) = \sum_{z \in 0, J} r_{zt} s_{jt}(z) = \frac{1}{\lambda}(S_{jt} - (1 - \lambda)s_{jt}(0))$. Static profits are $\pi_i(p, r) = \sum_{j \in J_i}(p_j - c_j) \cdot s_j(p, r)$. We drop the expectations operator, as the only source of uncertainty in the model is the realizations of marginal costs, which are fixed in the steady state. To find the steady-state prices and affiliated shares for each firm, we focus on Markov perfect equilibrium.[15] Firm i's profit-maximizing first-order condition is then:

---

[14]We provide a theoretical analysis of the monopolist's problem in the Appendix.

[15]Although we do not prove that the equilibrium is unique, the simulation results support there being a single steady-state equilibrium.

$$\sum_{l \in J_i} \frac{\partial \pi_l}{\partial p_j} + \beta \frac{dV_i(r')}{dr'} \frac{dr'}{dp_j} \; \forall j \in J_i = 0. \tag{9}$$

Next, we specify the derivatives of equation (8) with respect to r and evaluate them at the prices that solve each firm's first-order condition, which will be the prevailing prices at the steady state. These derivatives, in conjunction with the steady-state condition, $\frac{dV'}{dr'} = \frac{dV}{dr}$, yield the following system of equations:

$$\underbrace{\frac{dV_i(r)}{dr}}_{J \times 1} = \underbrace{\left[\frac{\partial \pi_i}{\partial p} \frac{dp}{dr} + \frac{\partial \pi_i}{\partial r} \right]}_{J \times 1} \underbrace{\left[I - \beta f_p(p,r) \frac{dp}{dr} - \beta f_r(p,r) \right]^{-1}}_{J \times J}. \tag{10}$$

In this equation, $\frac{\partial \pi_i}{\partial p}$, $\frac{\partial \pi_i}{\partial r}$, $f_p(p,r)$, $f_r(p,r)$ are known, conditional on values of p and r. The remaining unknowns are $\frac{dp}{dr}$ and $\frac{dV_i(r)}{dr}$. To solve the model, we impose the steady-state condition governing the evolution of affiliated customers, $r' = r$. The full set of steady state conditions, provided by equation (10) and $r' = r$, allow us to solve for steady-state prices and shares, conditional on the $J \times J$ derivative matrix, $\frac{dp}{dr}$. The values of $\frac{dp}{dr}$ are determined by the model. In our simulations, we solve for these values numerically using a local approximation method. For additional details, see the Appendix.

For the oligopoly analysis, we simulate symmetric market with three single-product firms. In this setting, symmetry is imposed by restricting the utility parameters, ξ_i, $\bar{\xi}$, and marginal cost to be the same across all firms. To illustrate the impact of affiliation on pricing incentives, we plot the equilibrium prices for two different sets of utility parameters in Figure 1.

Panel (a) of Figure 1 plots the equilibrium prices for a monopolist and a three-firm symmetric market, for increasingly large values of λ and otherwise identical demand parameters. The equilibrium prices increase with the rate of affiliation for values of λ greater than 0.5, and then decrease again. This figure highlights the potential importance of affiliation on equilibrium prices, and that the impact on price may non-monotonic in the proportion of consumers prone to affiliation. The non-monotonicity is a result of two countervailing affects. At low levels of λ firms face increasing inelastic demand and therefore increase prices. As λ increases past 0.5, however, the incentive to invest in future demand swamps the elasticity effect, and puts downward pressure on prices. Note, however, that firms' profits continue to increase, even as prices begin to decrease.

Panel (b) shows that the investment incentive can dominate and grow stronger for all values of λ. Furthermore, the investment incentive can be blunted by competition. Prices remain relatively flat for all values of λ in the firm market, whereas in the monopoly market the investment incentive is stronger and prices decrease at a faster rate. Thus, the relationship between equilibrium prices and dynamics may depend upon market structure.

Figure 1: Monopoly and Oligopoly Prices with Consumer Affiliation

(a) Prices Increasing with Affiliation

(b) Oligopoly Prices Decreasing with Affiliation
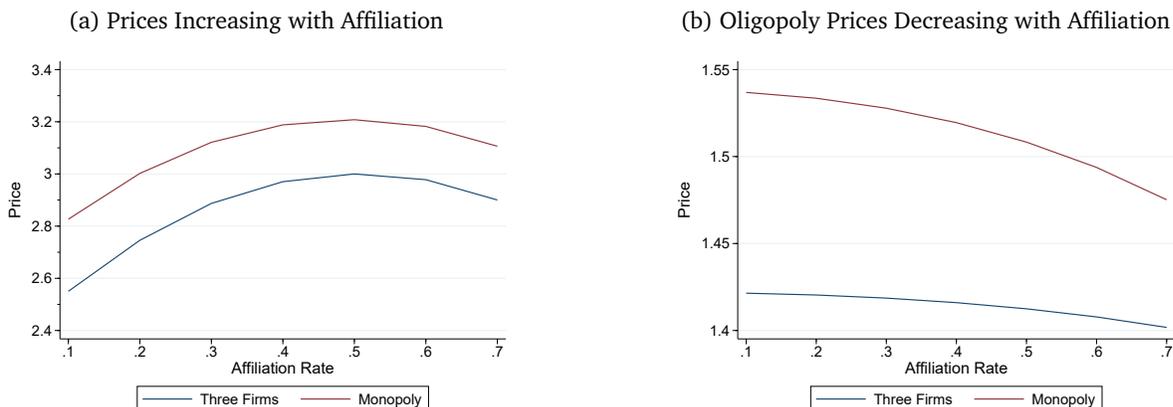
Notes: Panel (a) is generated using the following parameter values: $\xi = 0.04$, $\bar{\xi} = 4.15$, and $\alpha = -0.84$. Panel (b) is generated using $\xi = 2.33$, $\bar{\xi} = 0.04$, and $\alpha = -2.73$. Marginal cost is set to one for both figures.

2.3.1 Dynamic and Horizontal Market Power

To summarize the relationship between consumer inertia and market power in our model, we use simulations to decompose the potential impacts of dynamic and horizontal market power. To measure dynamic market power, we compare the three-firm oligopoly price with consumer affiliation to a baseline price where consumers have no state dependence, but the markets are otherwise identical. To measure horizontal market power, we simulate compare the three-firm oligopoly price to price that prevails following a merger between firms one and two.[16]

In the presence of consumer of inertia, it is important to define precisely the implementation of a horizontal merger. We consider two types of mergers. The first type of merger unites pricing control of two products under a single firm and the merged firm maintains these as separate products. We refer to this as a *joint pricing* merger. This is a common setting in many differentiated product mergers when the combined firms maintain separate brands/products. The second type of merger consolidates two products under a single brand and effectively offers a single product after the merger. We refer to this type of merger as *brand consolidation*. To implement a brand consolidation merger, we need to make an assumption about how much utility consumers of the product removed from the market will receive from buying the remaining brand. We assume that, at pre-merger prices, the consolidated brand will have the same combined share of unaffiliated customers as the separate pre-merger brands.[17] This type of merger typically occurs in settings where retail location is important and the post-merger firm

---

[16]Note that we use a (competitive) oligopoly price as a baseline, rather than price equal to marginal cost. Also, as marginal cost is constant across simulations, using price as a measure of market power is equivalent to the commonly used price-cost metric.

[17]To implement this assumption, we make an adjustment to the utility derived from consuming the post-merger brand. See the appendix for more details.

Figure 2: Potential Price Effects

(a) Dynamic Market Power

(b) Horizontal Market Power
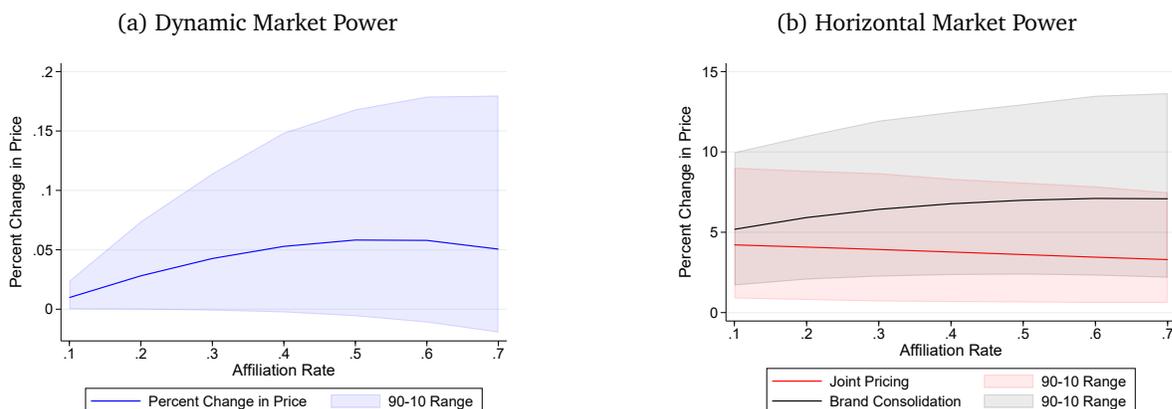
Notes: Panel (a) displays the mean percent price increase and for a three-firm oligopoly above the baseline model with no dynamics in consumption ($\lambda = 0$). Panel (b) displays the mean percent price increase of a merger to a duopoly for two types of mergers: joint pricing control and brand consolidation, for different values of λ. The plots reflect 491 baseline parameter values of $(\xi, \bar{\xi}, \alpha)$ that converged for all $\lambda \in \{0.1, 0.2, ..., 0.7\}$, or 3,437 markets in total.

rebrands all locations under one brand. This is often the case in retail gasoline mergers, but also occurs in other industries, such as wireless phone service (T-Mobile/Sprint) and airlines (American Airlines/USAir).

To provide a generalizable analysis, we attempt to simulate data from the support of parameters that produce reasonable outcomes for margins and shares. We employ a "shotgun" approach, generating simulations with many different parameters and selecting only the markets that meet certain criteria. We first take Halton draws of the demand parameters such that $\xi \in [0, 10]$, $\bar{\xi} \in [0, 10]$, $\alpha \in [-10, 0]$, and set each firm's marginal cost to one. For each draw of these demand parameters, we construct three-firm markets for $\lambda \in \{0, 0.1, 0.2, ..., 0.7\}$. We then restrict the analysis to markets where firms have shares between 0.05 and 0.30 (yielding an outside share between 0.10 and 0.85) and margins between 0.05 and 0.75.[18] Finally, to avoid composition affects, we only analyze markets with demand parameters that converged for all values of λ. The data generating process yields 3,437 markets whose parameters are summarized in Appendix Table 10.

In Figure 2, we plot the effects of affiliation and reduced competition on prices. The plots employ simulation results from the 491 baseline parameter values of $(\xi, \bar{\xi}$, and α that converged for all $\lambda \in [0, 0.7]$, or 3,437 markets. Panel (a) measures dynamic market power by plotting price effects as a function of λ. Percent changes are calculated relative to no consumer affiliation ($\lambda \approx 0$) while holding fixed the other parameters in the model. On average, prices increase with the fraction of of customers prone to affiliation for values of $\lambda \leq 0.5$ and decrease thereafter.

---

[18]The range for each parameter is selected such that parameter values toward each edge of the range result in outcomes that fall above or below our share and margin criteria.

This demonstrates that in the steady state of the model, the "harvesting" incentive tends to dominate the "investment" incentive for lower rates of affiliated customers. However, if the number of affiliated customers passes some threshold, investing in future takes precedence over harvesting. While not depicted, both market share and profits increase monotonically with λ. On average, affiliation results in moderate price increases compared to a static demand model. However, the impact of dynamic market can be quite substantial depending on the underlying demand parameters. The 90-10 range of outcomes is plotted with the transparent area. When $\lambda > 0.4$, the 90th percentile market exhibits dynamic market power resulting in prices more than 15 percent higher than a static model.

In panel (b), we plot the price effect of horizontal market power on prices across different values of λ. We measure horizontal market power by comparing the prices in the symmetric three firm oligopoly to the post-merger prices in both joint pricing and brand consolidation mergers. In our symmetric setting, brand consolidation mergers provide greater horizontal market power relative to joint pricing mergers across all levels of λ. This is in part due to how we specify brand consolidation mergers, as we assume high customer retention at consolidated brand. Still, there is large overlap in the 90-10 percentile range across the two merger types, demonstrating that both can enable comparable levels of horizontal market power. Another interesting feature of panel (b) is that horizontal market power slightly decreases, on average, in joint pricing mergers, but it increases in brand consolidation mergers. This highlights that the interaction between customer affiliation and a decrease in competition depends critically on the mechanism through which competition is reduced.

2.3.2 Consumer Inertia and Mergers

To further demonstrate the effect of consumer inertia on price competition, we examine more closely the impact on prices of horizontal mergers. We proceed by again using three symmetric, single-product firms as a baseline, and simulate both joint pricing and brand consolidation mergers. Thus, we examine how dynamic incentives affect steady-state prices and the unilateral price effects arising from mergers. We also analyze the bias that arises when a static model is calibrated to data generated by the dynamic affiliation model and used for merger simulation. We find that ignoring the presence of consumer affiliation tends to overestimate the price effects of a merger. The bias is much more pronounced, however, in joint pricing mergers.

For each parameterized market, we solve for the steady-state prices. Then, we simulate a merger between firms 1 and 2, leaving firm 3 as the non-merging firm. Table 1 summarizes the pre-merger equilibrium and the unilateral price increases of merging and non-merging firms. The "average" market is one that, a priori, would typically raise moderate concern from the US antitrust agencies; HHI (1070) falls in the "unconcentrated" range, but the change in HHI (713) generally warrants a thorough investigation. The average pre-merger difference between price and cost is 0.27, and the mean market share is 0.17. Using these mean values to calculate

Table 1: Simulation Summary Statistics

|	Mean	St. Dev.	Min	Pctl(25)	Pctl(75)	Max	N
Pre-Merger Price	1.43	0.41	1.10	1.19	1.50	3.96	3437
Pre-Merger Margin	0.27	0.14	0.09	0.16	0.34	0.75	3437
Pre-Merger Market Share	0.17	0.07	0.05	0.11	0.24	0.30	3437
HHI: Pre-Merger	1070	780	75	343	1730	2696	3437
Δ HHI	713	520	50	229	1153	1797	3437
Joint Pricing Merger Δ	3.76	3.31	0.25	1.27	5.24	20.14	3437
Joint Pricing Non-Merging Price Δ	0.89	1.16	0.00	0.11	1.23	7.85	3437
Consolidation Price Δ	6.50	4.27	0.23	3.47	8.53	25.48	3437
Consolidation Non-Merging Price Δ	0.81	1.11	-1.01	0.08	1.16	8.55	3437
Prediction Bias: Joint Pricing	1.55	1.76	0.00	0.39	2.03	13.49	3437
Prediction Bias (pctg.): Joint Pricing	73.78	110.75	0.32	13.93	79.76	890.22	3437
Prediction Bias: Consolidation	-1.19	1.52	-8.63	-1.81	-0.19	4.34	3437
Prediction Bias (pctg.): Consolidation	-19.34	25.48	-76.05	-36.39	-3.61	181.04	3437
Dynamic Elasticity: Unaffiliated	-5.76	2.36	-11.96	-7.54	-3.79	-1.29	3437
Dynamic Elasticity: Affiliated	-2.19	1.92	-9.74	-3.20	-0.76	-0.09	3437
Dynamic Elasticity: Weighted	-3.79	1.97	-10.02	-5.08	-2.17	-0.59	3437
Static Elasticity	-4.72	2.09	-10.71	-6.13	-2.98	-1.34	3437

Notes: Margin is defined as $\frac{p-c}{p}$. Δ HHI is calculated at the pre-merger shares. Merger Price Δ is the percentage price increase from the merger. Prediction Bias is the static prediction minus the dynamic prediction, in percentage points. Prediction Bias (pctg.) is the Prediction Bias divided by the dynamic Merger Price Δ. The weighted dynamic elasticity is the average of the unaffiliated and affiliated elasticities weighted by the fraction of the firm's customers of each type.

the "Upward Pricing Pressure" index, $\frac{0.17}{1-0.17} \cdot (1.27 - 1) = .055$, which is just over the threshold that may trigger an investigation.[19] The average percentage merger price effect is 3.76 percent and 6.50 percent in joint pricing and brand consolidation mergers, respectively, either of which may raise antitrust concerns. The full range of markets span those that lead to no scrutiny to those that almost certainly would be challenged. Thus, the simulations generate a reasonable set of markets within which to explore the pricing incentives of firms in the consumer affiliation model.

Tables 11 and 12 in the Appendix provide results from regressing the demand parameters on effects of the mergers. On average, increasing the rate of affiliation (λ) and the strength of the affiliation ($\bar{\xi}$) tends to increase pre-merger prices, but their impact on the effect of a merger is dependent upon the type of merger. The proportion of affiliated customers and the strength of affiliation decrease prices in joint pricing mergers, but work to increase prices in brand consolidation mergers. However, these relationships do not hold in every instance, and may interact in interesting ways. Thus, to fully understand the impact of horizontal mergers in markets with consumer inertia it is important to model the market under consideration.

[19]See, for example, Farrell and Shapiro (2010) and Miller et al. (2017). This calculation assumes that diversion is proportional to market share, which is often assumed at the early stages of an antitrust investigation.

Figure 3: Distribution of Merger Prediction Bias

(a) Joint Pricing

(b) Brand Consolidation
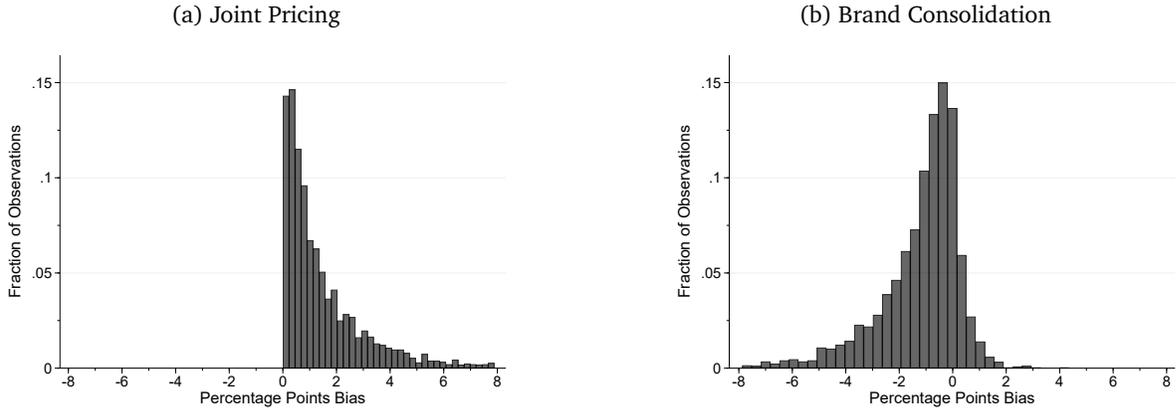
Notes: Prediction bias is defined as the the prediction of a (misspecified) static logit model minus the true price increase. Panel (a) depicts the bias when the merger consolidates pricing control of two products. Panel (b) depicts the bias when the merger consolidates two products under one brand. A brand consolidation merger is defined in the main text.

2.3.3 Model Misspecification in Merger Simulation

The above results suggest that affiliation has implications for counterfactual exercises, such as merger simulation. Failing to account for consumer inertia will result in biased elasticities from incorrect first-order conditions. Antitrust agencies often infer elasticities from markups calculated using accounting data (see Miller et al., 2013), which omit the dynamic incentive of firms. In addition to generating incorrect elasticities, failing to account for the dynamic incentives in first-order conditions can have large direct effects on post-merger predictions.

To more precisely analyze the impact of model misspecification, we measure the implications of failing to account for consumer affiliation when calibrating demand and simulating a merger. To do so, we consider the following hypothetical scenario. The true underlying model is the three-firm market with consumer state dependence due to affiliation. A practitioner observes each firms' pre-merger prices, marginal costs, and aggregate market shares (rather than separately observing its unaffiliated and affiliated shares). This data is then used to recover the demand parameters of the standard logit model, and then the price effects of a merger are simulated. We perform this experiment for each of the numerically generated markets, and consider both joint pricing and brand consolidation mergers.

In a symmetric oligopoly, the joint pricing and brand consolidation mergers we consider will have identical price effect when demand is characterized by the static logit model. We provide a proof in Appendix B. Thus, precisely modeling the structure of a horizontal model may have heightened importance in markets characterized by consumer inertia.

Table 1 summarizes the prediction bias, defined as the static prediction minus the "true" affiliation price effect. The average prediction bias is 1.55 and -1.19 percentage points for joint

Figure 4: Price Increase by Market Share
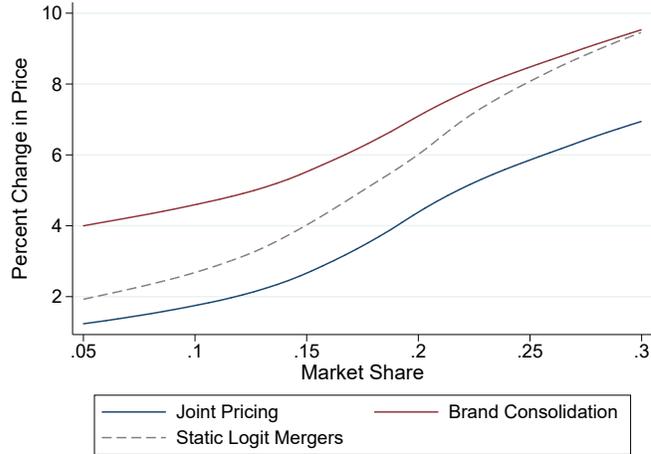
Notes: The lines depict a local polynomial regression of firm 1's merger percent price change on its pre-merger market share.

pricing and brand consolidation mergers, respectively. Dividing by the magnitude of the price change, this represents a 73.8 and -19.34 percent bias over the true price effect for the two types of mergers. Thus, for both types of mergers, incorrectly assuming static demand will lead to substantially biased predictions.

Figure 3 plots the distributions of bias across all 3,437 markets. Panel (a) depicts the distribution of bias for joint pricing mergers. In every instance the static model over predicts the true dynamic effect. In a few markets that did not meet our inclusion criteria, the static model under predicted the dynamic effect, so an upward bias will not always occur. However, our results suggest that a static logit demand model incorrectly calibrated to a demand with consumer inertia will almost always exhibit upward bias. On the other hand, the bias in brand consolidation mergers may be positive or negative. While static mergers are more likely to under-predict the true dynamic price effect, almost 25 percent of simulations resulted in over-predictions. These findings again highlight the importance of properly accounting for dynamics when simulating the price effects of mergers. Moreover, we find that the sign of the bias is only systematic in joint pricing mergers.

Figure 4 relates the simulation price effects to the pre-merger market share of each symmetric firm. To generate the graph, we run a local polynomial regression of the merger price effect on the average firm's pre-merger market share. We generate fitted lines for (i) the joint pricing merger effect, (ii) the brand consolidation merger price effect, and (iii) a misspecified static logit model (which generates the same price effect, regardless of the merger type). In line with intuition, the price effect of a merger is increasing with pre-merger market shares. Interestingly, we find that the magnitude of the bias increases with market share with respect

17

to joint pricing mergers, but decreases with respect to brand consolidation mergers. As mergers with high pre-merger market shares tend to be of greater concern to antitrust authorities, this finding highlights that the cost of misspecification may be low if the merger results in brand consolidation, but upward bias is more likely if only pricing control is expected to change.

The "harvesting" and "investment" incentive present in the dynamic model, which is absent in the static model, can explain the biases depicted in Figure 4. With respect to joint pricing mergers, the static model over-predicts the dynamic model regardless of the pre-merger market shares. Thus, as the dynamic incentive to invest in future demand increases, the upward bias of the static model becomes greater in magnitude. Interestingly, the static model is calibrated to be more elastic, on average, than the share-weighted elasticity in the dynamic model (-4.72 vs. -3.79). More elastic demand generates smaller merger price effects in the logit model. Thus, biased predictions from the static model arise primarily from the omission of dynamic incentives to invest in future demand, rather than a biased elasticity or mean utility parameters alone.

In rough terms, margins are determined in the static model by a combination of market elasticities and the degree of competition within the market. The dynamic model has a third factor disciplining margins: the dynamic incentive. When this factor is omitted in the estimation of a static model, greater weight is attributed to the degree of competition (i.e., rival products are considered to be closer substitutes). This provides some intuition as to why a static model will tend to predict greater unilateral price effects.

Yet, in brand consolidation mergers, the static model tends to under-predict the true price effect. This is possible when brand consolidation mergers present the firm with a stronger "harvesting" incentive relative to joint pricing mergers. The post-merger firm is able to transfer its loyal (i.e., inelastic) customers under to a single product. This presents the firm with ability to raise prices to a higher level then when separate products are maintained. However, in the misspecified static model, the firm is responding to the "average" elasticity of its customers, whereas in the dynamic model the firm internalizes the different types of consumers into its pricing incentive. Since affiliated customers are much less elastic, this presents a greater opportunity, relative to a static model, to increase price.

These results highlight the benefit of an empirical model that can account for consumer dynamics, which we pursue in the following section. However, for certain applications an informal analysis of consumer dynamics may provide a useful indication of static model bias. Our results show how price predictions from static merger simulations could be revised downward when only pricing control is expected to change in a merger. In Appendix A.5, we provide more detail on how dynamics affect simulation bias. Therein we again find that the relationship between the strength of dynamics and bias is a function of the type of merger under evaluation.

3 Reduced-Form Evidence of Dynamics

To motivate the empirical application, we provide evidence of dynamic demand and dynamically adjusting retail gasoline prices. A host of previous studies have found that retail gasoline prices may take multiple weeks to fully incorporate a change in marginal cost. For a review, see Eckert (2013). One innovation of our study is that we use separate measures of unexpected and expected costs to see if, consistent with forward-looking behavior, firms respond differentially to these two types of costs.

3.1 Data

The analysis relies upon daily, regular fuel retail prices for nearly every gas station in the states of Kentucky and Virginia, which totals almost six thousand stations. As a measure of marginal cost, the data include the brand-specific, daily wholesale rack price charged to each retailer as well federal, state, and local taxes. We therefore almost perfectly observe each gas station's marginal cost changes, except for privately negotiated discounts per-gallon, which are likely fixed over the course of a year. The data ranges from September 25th, 2013 through September 30th, 2015. The data was obtained directly from the Oil Price Information Service (OPIS), which has previously provided data for academic studies (e.g., Lewis and Noel, 2011; Chandra and Tappata, 2011; Remer, 2015).

OPIS also supplied the market share data, which is used by industry participants to track shares. It is calculated from "actual purchases that fleet drivers charge to their Wright Express Universal card." The data is specified at the weekly, county/gasoline-brand level.[20] Due to contractual limitations, OPIS only provided each brand's share of sales, not the actual volume. Thus, to account for temporal changes in market-level demand, we supplement the share data with monthly, state-level consumption data from the Energy Information Administration (EIA).

3.2 Dynamic Demand: Correlation in Shares Over Time

Though ultimately the importance of demand-side dynamics in the data will be estimated by the model, it is informative to examine the reduced-form relationships between key elements. The dynamic model developed in the previous section is one in which today's quantity depends on the quantity sold last period. As motivation for this model, we present the results from reduced-form regressions of shares on lagged shares in Table 2.

This exercise demonstrates that even after including rich fixed effects to capture static variation in consumer preferences, lagged shares are a significant predictor of current shares. The residual correlation in shares over time in the most detailed specification captures deviations from specific county-brand seasonal patterns. A positive correlation is consistent with state

---

[20]In some instances, the brand of gasoline may differ from the brand of the station. For example, some 7-Eleven stations in the data are identified as selling Exxon branded gasoline.

Table 2: Regressions with Share as the Dependent Variable

	(1)	(2)	(3)	(4)	(5)
Price	0.009***	0.000**	0.004	−0.005	−0.073***
	(0.001)	(0.000)	(0.002)	(0.004)	(0.017)
Lagged Share		0.973***	0.963***	0.554***	0.628***
		(0.001)	(0.001)	(0.002)	(0.003)
Price Squared			−0.000	0.001	0.010***
			(0.000)	(0.001)	(0.003)
Comp. Price (Mean)			−0.004***	−0.002	−0.108**
			(0.001)	(0.001)	(0.045)
Comp. Price (SD)			−0.000	0.003*	0.086***
			(0.001)	(0.001)	(0.024)
Comp. Stations			−0.000***	−0.000***	−0.004***
			(0.000)	(0.000)	(0.000)
Num. Stations			0.000***	0.004***	
			(0.000)	(0.000)	
Num. Brands			−0.001***	−0.003***	
			(0.000)	(0.000)	
Week FEs				X	
County-Brand FEs				X	
Brand-State-Week FEs					X
Week-County FEs					X
County-Brand-WofY FEs					X
Observations	174421	169931	169788	169770	156078
R^2	0.00	0.95	0.95	0.96	0.98

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

dependence in consumption. In specification (2), we show that lagged shares explain 95 percent of the variance in current shares, and the coefficient is close to one. In specification (3), we include measures of competition in the regressions, as well as a second-order polynomial in own price. The competition measures, which include the mean and standard deviations of competitor prices, are correlated with shares, but lagged shares still are the most important predictor of current shares. In specification (4), we include time and brand-county fixed effects. In the final specification (5), we include rich multi-level fixed effects: county-brand-(week of year), brand-state-week, and week-county. The coefficient of 0.628 on lagged shares in this specification indicates that deviations in shares are highly correlated over time, even when we condition on the most salient variables that would appear in a static analysis, adjust for brand-county specific seasonal patterns, and allow for flexible brand-state and county time trends. This finding is consistent with demand-side dynamics, as there are patterns in shares over time that are challenging to explain with contemporaneous variables.[21]

---

[21]We have also estimated specifications that add lagged prices. Though the first lag is significant, there is almost no change to the lagged share coefficient.

3.3 Dynamic Pricing

We now present reduced-form evidence of dynamic pricing. Consistent with a model where firms accumulate affiliated consumers over time, we find that new entrants price lower relative to established competitors in the same market, and that this discount dissipates over time. Second, we examine cost pass-through and show that firms are slow to adjust to marginal cost changes. Moreover, firms anticipate expected changes in future costs by raising prices in advance of the change. In the presence of consumer affiliation a firm will change it's current price in response to an expected future cost change, as it affects the current value of investing in future demand. The ability to separately estimate the response to expected and unexpected costs is a key innovation of our study.

3.3.1 Dynamic Pricing of New Entrants

When forward-looking firms price to consumers that may become affiliated, there is an incentive to initially offer prices below the static optimum. In this setting, we expect a new entrant, all else equal, to initially price below its competitors. As the new entrant builds up its share of affiliated customers, its prices will gradually converge to its competition.

We test for and find evidence consistent with this dynamic pricing pattern in the data. To perform the analysis, we first identify a set of new entrants, defined as a gas station whose first price observation is at least six weeks after the start of the data and does not exit in the remainder of the sample. To ensure there is sufficient data and to control for composition effects in the analysis, we limit the set of entrants to those with at least one year of post-entry price data. Using this filter, we identify 193 entrants. We compare prices for entrants to the 594 stations in those counties that are present for the entire sample.

Figure 5 depicts the average difference between an entrant's price and all other stations' price in the same county, sorted by the number of weeks after entry. The figure demonstrates that gas stations enter with a price that is, on average, two cents per gallon less than incumbents' prices. Entrants' prices then slowly converge over time to the market average. A series of *t*-tests confirm the statistical significance of the results. For the first 8 weeks following entry, new entrants' prices are significantly lower than the county average price at the 0.05 level. This pattern is consistent with a profit-maximizing firm building up an affiliated customer base over time, and raising its price to a gradually less elastic set of consumers.

3.3.2 Cost Pass-through

To highlight the temporal component of cost pass-though, we separately estimate how gas stations react to expected versus unexpected cost changes. Beyond motivating the structural model, these results also demonstrate the importance of capturing firms' anticipated price responses when estimating cost pass-through rates. For example, to analyze how much of a tax

Figure 5: New Entrant Prices
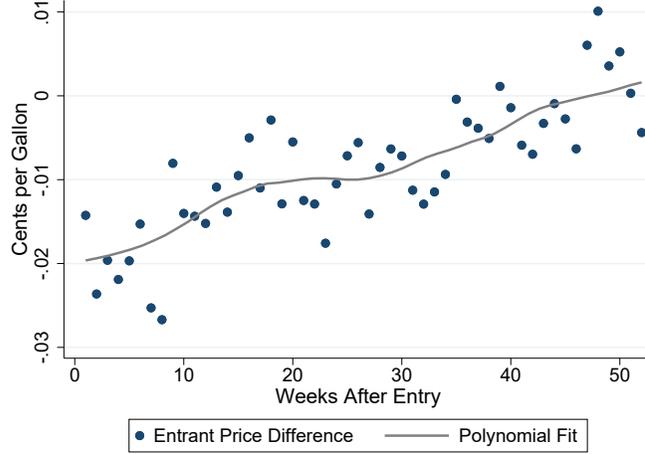
Notes: A data point measures the average difference between a new entrant's price and the county average price, for the given number of weeks after entry. The line is created using local polynomial regression.

increase firms will pass-on to consumers, it is imperative to recognize that firms may begin to adjust their prices prior to the tax increase being enacted; failure to account for this response may lead to underestimating pass-through rates.

We construct our measure of expected cost by using gasoline futures and current wholesale costs to project 30-day-ahead costs. Unexpected costs represent deviations from this projection.[22] We incorporate the main components of marginal costs for retail gasoline, which include the wholesale cost of gasoline and the per-unit sales tax. We estimate the following model:

$$p_{it} = \sum_{s=-50}^{50} \beta_s \hat{c}_{it-s} + \sum_{s=-50}^{50} \gamma_s \tilde{c}_{it-s} + \sum_{s=-50}^{50} \phi_s \tau_{it-s} + \psi_i + \varepsilon_{it}. \tag{11}$$

Here, p_{it}, is the price observed at gas station i at time t. $\hat{c}_{it-s}$ and $\tilde{c}_{it-s}$ are the expected and unexpected wholesale costs observed with lag s, and τ_{it-s} is the state-level sales tax.[23] Using the estimated coefficients on the cost measures, we construct cumulative response functions to track the path of price adjustment to a one time, one unit cost change at time $t = 0$. We incorporate 50 leads and lags to capture the full range of the dynamic response. We focus our results on unexpected and expected costs, as we do not have enough tax changes in our data to estimate a consistent pattern of response.

Figure 6 plots the cumulative response functions for unexpected and expected costs. Panel

---

[22]For details, see Section C.1 in the Appendix.

[23]To more easily incorporate future anticipated costs into the regression, we do not present an error-correction model (Engle and Granger, 1987), which is commonly used to estimate pass-through in the retail gasoline literature. As a robustness check, we estimated the price response to expected and unexpected costs using the error-correction model, and we found very similar results.

Figure 6: Cumulative Pass-through

(a) Unexpected Costs   (b) Expected Costs
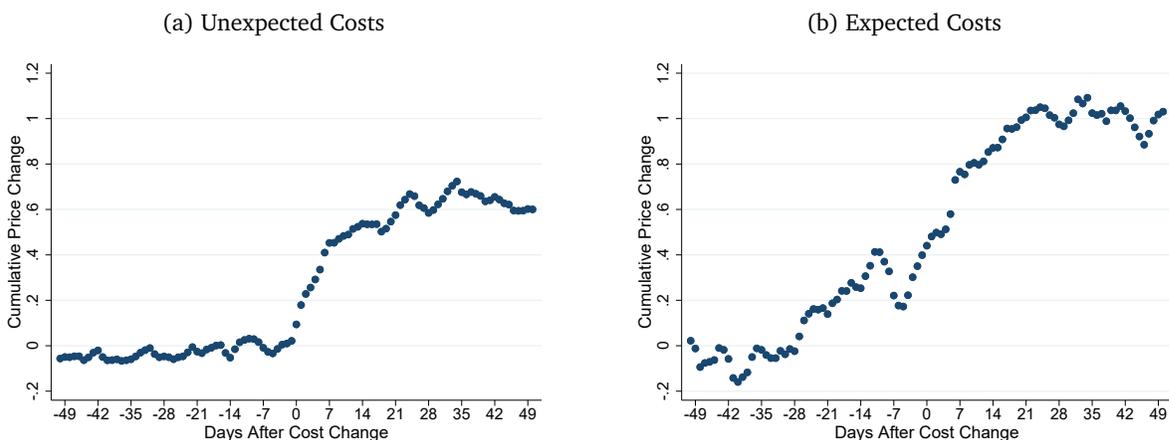
Notes: Panels (a) and (b) depict the cumulative price change in response to a one unit cost change at time = 0. Response functions are created from the estimated parameters of equation (11).

(a) displays the results for unexpected costs. Prices react suddenly and quickly at time zero, but it takes about four weeks for the prices to reach the new long-run equilibrium. Estimated pass-through peaks at 0.72 after 34 days, with an average of 0.64 over days 21 through 50.

Panel (b) displays the cumulative response function for expected costs. Notably, firms begin to react to expected costs approximately 28 days in advance, with a relatively constant adjustment rate until the new long-run equilibrium passthrough is reached 21 days after the shock. The estimated pass-through averages 1.01 over days 21 through 50. Though the total duration of adjustment is longer compared to the unexpected cost shock, the firm incorporates the cost more quickly after it is realized. This coincides with substantial anticipation by the firm; the price already captures about 40 percent of the effect of the expected cost shock the day before it arrives.[24]

Thus, a reduced-form analysis of pricing behavior shows that prices retail gasoline stations adjust slowly to changes to marginal cost and also that price changes anticipate expected changes in marginal costs. These patterns are consistent with forward-looking behavior of firms with dynamic demand arising from consumer affiliation. Readers might wonder about the relevance of asymmetric pricing, i.e., whether the price response is the same for positive and negative cost shocks. In robustness checks, we find little evidence of asymmetry. Furthermore, in our data, we do not find evidence of Edgeworth price cycles.

---

[24]A striking result from these estimates is the difference in the long-run pass-through rates. Expected costs experience approximately "full" pass-through – a cost increase leads to a corresponding price increase of equal magnitude. On the other hand, unexpected costs demonstrate incomplete pass-through, moving about only 64 cents for each dollar increase in cost.

4 Empirical Application: Demand Estimation

Given the reduced-form evidence of dynamic demand and supply behavior, we now present the empirical application of the model to the retail gasoline markets described in the previous section. First, we outline our estimation methodology. We divide it in two stages, as demand can be estimated independently of the supply-side assumptions. Our method of demand estimation relies on data that is widely used in static demand estimation: shares, prices, and an instrument. After outlining the methodology, we present results for demand estimation. In Section 5, we use the estimated demand system to analyze the dynamic incentives faced by suppliers. We use these results to consider a merger between large gasoline retailers.

4.1 Identification

We discuss the identification argument in three parts. First, we show that the structure of the model is sufficient to identify the unobserved distribution of choices (i.e., the vector $\{s_{jt}(z)\}$) conditional on observed shares S_{jt}, the share of consumers subject to state dependence λ, and the strength of the affiliation shock $\sigma_{jt}(z)$. Second, the vector $\{s_{jt}(z)\}$ allows us to recover the mean utility for unaffiliated consumers and estimate the static demand parameters. Third, we discuss the assumptions that allow us to identify the dynamic parameters.

Identification of Type-Specific Choices

A key challenge with aggregate data and unobserved heterogeneity is that we do not separately observe choice patterns by unobserved consumer type. In our context, we observe the aggregate share, S_{jt}, which is a weighted combination of the $\{s_{jt}(z)\}$ and depends on the distribution of affiliated consumers for each product $\{r_{jt}\}$. Observed shares are determined by the following:

$$S_{jt} = (1 - \lambda)s_{jt}(0) + \lambda \sum_{z=0}^{J} r_{zt}s_{jt}(z). \tag{12}$$

To separate out $\{s_{jt}(z)\}$ from S_{jt}, we leverage the structure of the model. With discrete types, we show exact identification of the choice distribution without supplemental assumptions.

Proposition 1 *With discrete types, the distribution of choice patterns is identified conditional on the distribution of types and type-specific shocks.*

Using the dynamic extension of the logit demand system detailed in section 2, we obtain the familiar expression for the log ratio of shares of unaffiliated consumers from equation (2):

$$\ln s_{jt}(0) - \ln s_{0t}(0) = \delta_{jt} \tag{13}$$

Likewise, we obtain the following relation for the shares of affiliated consumers:

$$\ln s_{jt}(z) - \ln s_{0t}(z) = \delta_{jt} + \sigma_{jt}(z). \tag{14}$$

To show identification, we combine equations (13) and (14) to obtain the following expressions:

$$s_{jt}(0) = \left(\frac{s_{0t}(0)}{s_{0t}(j)} - 1 \right) \frac{1}{\exp\left(\sigma_{jt}(j)\right) - 1} \tag{15}$$

$$s_{jt}(z) = s_{0t}(z) \frac{s_{jt}(0)}{s_{0t}(0)} \cdot \exp\left(\sigma_{jt}(z)\right). \tag{16}$$

Thus, we show that the $J + J^2$ unknowns $\{s_{jt}(z)\}_{|j\neq0}$, can be expressed in terms of the $J + 1$ unknowns $\{s_{0t}(j)\}$ and $s_{0t}(0)$. These $J + 1$ unknowns are pinned down by the adding-up condition $1 - \sum_k s_{jt}(0) - s_{0t}(0) = 0$ and the observed share equations, which provide the other J restrictions:

$$S_{jt} = (1 - \lambda) \left(\frac{s_{0t}(0)}{s_{0t}(j)} - 1 \right) \frac{1}{\exp\left(\sigma_{jt}(j)\right) - 1} \tag{17}$$

$$+ \lambda \sum_{z=0}^{J} r_{zt} s_{0t}(z) \frac{s_{jt}(0)}{s_{0t}(0)} \cdot \exp\left(\sigma_{jt}(z)\right)$$

Identification requires $\{r_{jt}\}$, which is the state describing the share of state-dependent consumers that are affiliated to each product. Given our assumptions about the evolution of demand, the value for j can be calculated from the prior period values for $S_{j(t-1)}$ and $s_{j(t-1)}(0)$:

$$r_{jt} = \sum_{z=0}^{J} r_{z(t-1)} s_{j(t-1)}(z) \tag{18}$$

$$\implies r_{jt} = \frac{1}{\lambda} \left(S_{j(t-1)} - (1 - \lambda) s_{j(t-1)}(0) \right) \tag{19}$$

Given an initial value $\{r_{j0}^*\}$, equation (19) can be used to iteratively identify the future values of the state. We discuss the choice of this initial state vector in our empirical application. Therefore, the unobserved state-dependent choice probabilities $\{s_{jt}(z)\}$ are identified conditional on λ and $\{\sigma_{jt}(z)\}$, i.e., the parameters governing unobserved heterogeneity.

Identification of Static Demand Parameters

We now allow for the observation of multiple markets, which are denoted with the subscript m. From equation (13), we obtain the utility of the unaffiliated (type 0) consumer in each market, $\{\delta_{jmt}\}$. This is analogous to recovering the mean product utility as in (Berry et al., 1995). We

make the standard assumption that the utility is linear in characteristics:

$$\delta_{jmt} = \alpha p_{jmt} + \pi \left(p_{jmt} \times Income_{jmt} \right) + X_{jmt}\gamma + \eta_{jmt}. \tag{20}$$

The utility depends on price, p, and the interaction of price with market-average income, both of which are endogenous. The exogenous covariates, X, may contain multi-level fixed effects. With valid instruments for p and $(p \times Income)$, these linear parameters are identified using standard instrumental variables arguments.

Identification of Dynamic Demand Parameters

We have so far shown exact identification of static demand parameters conditional on the parameters governing unobserved heterogeneity. To identify λ_m and $\sigma_{jt}(j)$, we need to employ additional moments. To generate these moments, we assume that the residual demand innovations, η_{jmt}, are uncorrelated over time. We construct these residuals after accounting for multi-level fixed effects, including market-specific seasonal patterns and product-specific fixed effects to account for unobserved heterogeneity in product quality. We assume that $Corr(\eta_{jmt}, \eta_{jm(t+1)}) = 0$ holds within each market, which provides us with more than sufficient moments (241) to identify our parameters.[25]

In the context of our model, serial correlation in demand arises from consumer inertia, and further, firms can shape the correlation in demand by altering prices. By imposing that the demand innovations are serially uncorrelated, we assign all systematic autocorrelation in product-specific demand to the endogenous response of consumers, rather than treating such correlation as a feature of an exogenous stochastic process. Thus, our results may be thought of an "upper-bound" on the impact of consumer inertia.

In our application, we parameterize the dynamic parameters λ_m and $\sigma_{jmt}(j)$ as follows:

$$\lambda_m = \frac{\exp(\theta_1 + \theta_2 Income_m + \theta_3 Density_m)}{1 + \exp(\theta_1 + \theta_2 Income_m + \theta_3 Density_m)} \tag{21}$$

$$\sigma_{jmt}(j) = \bar{\xi}. \tag{22}$$

Thus, we allow the share of consumers subject to state dependence to vary with market-level measures of median household income and (log) population density. The share of shoppers may be lower in areas with more affluent consumers and more congestion. This specification allows for the possibility that consumer characteristics, as captured by income and population density, affects the prevalence of consumer inertia. We assume that affiliated customers receive a constant level shock to utility $\bar{\xi}$.

---

[25]One could construct related moments by using lagged prices as instruments, under the assumption that the prices are uncorrelated with the innovation in the demand residual.

Separate identification of λ_m and $\bar{\xi}$ is made possible by the structure of the model. λ_m, the share of consumers that become affiliated, does not depend on price, whereas the impact of $\bar{\xi}$ on shares does. As can be seen by examining equations (13) and (14), a change in price affects δ_{jmt}, which shifts the relative choice patterns, holding fixed $\bar{\xi}$. Intuitively, this would be reflected in the data by how the serial correlation in shares varies price levels in the market. The parameters $(\theta_1, \theta_2, \theta_3)$ are identified by how these serial correlation patterns covary with demographic characteristics.

4.2 Implementation

Reduced Computational Complexity

Though the distribution of unobserved choices is identified, solving for the pattern of choices in estimation is another matter. The traditional approach is to "concentrate out" the distribution of unobserved heterogeneity while using a contraction mapping to solve (implicitly) for the shares of the type 0 consumers (as in Berry et al. (1995)). In our setting, the assumption of single-product affiliation allows us to reduce the computation burden, as the full distribution of choice patterns in each market can be calculated directly after solving a system of equations in two variables. Thus, we reduce the number of unknowns in each market from J to 2. This may be used to speed up estimation by implementing a non-linear equation solver or a (modified) contraction mapping.

Above, we showed that the choice patterns can be expressed in terms of the $J+1$ parameters $\{s_{0t}(j)\}$ in each market. We now show that the system reduces to two parameters in each market, where the remaining $J-1$ parameters are solved for by a quadratic function.

Under the assumption of single-product affiliation ($\sigma_{jt}(z) = 0 \,\forall\, z \neq j$), we obtain

$$\sum_z r_{zt} \cdot s_{0t}(z) \exp\left(\sigma_{jt}(z)\right) = \sum_z r_{zt} s_{0t}(z) + \left(\exp\left(\sigma_{jt}(j)\right) - 1\right) r_{jt} s_{0t}(j). \tag{23}$$

By substituting this expression into equation (17), we can obtain a quadratic equation for each of the $\{s_{0t}(j)\}$:

$$\begin{aligned}
0 = {} & \lambda \frac{1}{s_{0t}(0)} r_{jt}(\exp(\sigma_{jt}(j) - 1) \boldsymbol{s_{0t}(j)}^2 \\
& + \left([\exp(\sigma_{jt}(j) - 1] \left(S_{jt} - \lambda r_{jt}\right) + \lambda \frac{1}{s_{0t}(0)} \sum_{z \in 0, J} r_{zt} s_{0t}(z) + (1 - \lambda) \right) \boldsymbol{s_{0t}(j)} \\
& - \lambda \sum_{z \in 0, J} r_{zt} s_{0t}(z) - (1 - \lambda) s_{0t}(0)
\end{aligned}$$

Conditional on the dynamic parameters and observables, there are only two remaining unknowns: $s_{0t}(0)$ and $\sum_{z \in 0, J} r_{zt} s_{0t}(z)$. Thus, we can solve for $\{s_{0t}(j)\}$ in each market using the

quadratic formula. As $\{\delta_{jt}\}$ are identified conditional on these choice probabilities, we can obtain these mean utility parameters by solving for only two unknowns in each market, regardless of the number of products.

Estimation Routine

To implement our estimator, we use a nested regression approach with the following steps:

1. First, pick the unobserved heterogeneity parameters corresponding to λ_m and $\sigma_{jmt}(j)$.

2. Starting with the first period, solve for $s_{0mt}(0)$ and $\sum_{0,z} r_{zt} s_{0mt}(z)$ in each market using the non-linear system of equations obtained previously. Using the initial value r^*_{jm0} for the first period, then iterate forward for each period using the calculated value of r_{jmt} for subsequent periods.

3. Solve for $s_{jmt}(0)$ for each firm, period, and market.

4. Run the regression implied by equation (13) using the $\{s_{jmt}(0)\}$ obtained in the previous step to solve for the linear parameters (α, π, γ). Calculate the correlation of the residuals $Corr(\hat{\eta}_{jmt}, \hat{\eta}_{jm(t+1)})$ within each market.

5. Repeat 1-4 to find the unobserved heterogeneity parameters that minimize the sum of squared correlations.

The regression for equation (13) may involve instrumental variables and the use of panel data methods such as fixed effects. In our empirical application, we make use of both.

The estimation methodology employs two tricks to speed up the computation of the dynamic model. First, the explicit formula for $\{s_{jt}(0)\}$ means that the non-linear solver[26] only has to find two parameters, $s_{0t}(0)$ and $\sum_{0,z} r_{zt} s_{0t}(z)$, for each market-period. The quadratic form for the remaining unknowns results in fast calculation. Second, the linear form for the nested regression allows for a quick calculation of the inner part of the routine and allows for serial correlation in unobservables.

In models with state dependence and unobserved heterogeneity, one consideration is the initial value of the unobserved state. Because we have very long panel, with 104 separate time observations, we have a sufficient "burn-in" period where this issue does not materially affect our estimates.[27] As a baseline, we set r^*_{jm0} equal to the value of the aggregate share S_{jm0} in the period prior to the first week of our sample. This value would be consistent with the steady state if the affiliation shock $\bar{\xi} = 0$. Because we find $\bar{\xi} > 0$, we typically find that $r_{jmt} > S_{jm(t-1)}$ in our estimated model.

---

[26]To solve for these unknowns, we use a modified contraction mapping that uses the average of the previous guess and the implied solution for the two parameters in each market. This modification improves stability.

[27]We find that it takes approximately 6 weeks for these values to converge to the steady state.

Table 3: Summary Statistics by County

Statistic	Mean	St. Dev.	Min	Pctl(25)	Pctl(75)	Max	N
Num. Brands	4.52	1.46	1	3	6	8	241
Price	2.87	0.11	2.53	2.77	2.96	3.14	241
Wholesale Price	2.25	0.06	2.02	2.21	2.28	2.44	241
Margin	0.21	0.06	0.06	0.17	0.24	0.44	241
Num. Stations	22.33	27.88	1.88	7.69	25.59	239.13	241

Notes: Table displays summary statistics averaged across each of the 241 markets in the sample.

4.3 Data for Structural Model

We supplement the EIA-adjusted weekly brand-county share measures with the average prices for the brand in a week-county. To reduce the occurrence of zero shares, which do not arise in the logit model, we use a simple linear interpolation for gaps up to four weeks. For any gap greater than four weeks, we assume the station was not in the choice set for that gap. We drop any observations that have missing prices, missing shares, or missing shares in the previous week. This includes dropping the first week of data, for which we do not have previous shares.

Table 3 provides summary statistics of the data for the 241 counties in KY and VA. There is cross-sectional variation in wholesale prices, margins, and the number of stations in each county. To reduce the sensitivity of the analysis to brands with small shares and to make the counterfactual exercises more computationally tractable, we aggregate brands with small shares into a synthetic "fringe" brand. We designate a brand as part of the fringe if it does not appear in ten or more of the 241 markets (counties). Additionally, if a brand does not make up more than 2 percent of the average shares within a market, or 10 percent of the shares for the periods in which it is present, we also designate the brand as a fringe participant for that market. This reduces the maximum number of brands we observe in a county to 8, down from 24. Across all markets, we analyze the pricing behavior of 16 brands, including the synthetic fringe.[28]

We also take steps to reduce measurement error in the number of stations in our data. We assume that stations exist for any gaps in our station-specific data lasting less than 12 weeks. Likewise, we trim for entry and exit by looking for 8 consecutive weeks (or more) of no data at the beginning or end of our sample. After cleaning, we retain 110,844 observations in our sample.

---

[28]Summary statistics by brand are presented in Table 16 in the Appendix. The fringe brand is, on average, 13 percent of the shares for the markets that it appears in. As we designate a fringe participant in nearly every market, the aggregated fringe has the highest overall share (12 percent).

4.4 Results: Demand Estimation

For the empirical application, we implement the methodology described in Section (4.1). Conditional on dynamic parameters, we extract the unobserved shares for all unaffiliated type consumers, obtaining the baseline utility δ_{jmt}. We then estimate static demand parameters using instrument variables regression following equation (20). Our chosen dynamic parameters minimize the average correlation in brand-market shocks over time (between contemporaneous and a single-period lag), where the correlation is calculated within each market.[29]

Our regression equation takes the following form:

$$\ln\left(\frac{s_{jmt}(0)}{s_{0mt}(0)}\right) = \alpha p_{jmt} + \pi\left(p_{jmt} \times Income_m\right) + \gamma N_{jmt} + \zeta_{jm} + \phi_t + \psi_{m,month(t)} + \eta_{jmt}$$

Here, the subscript m denotes the market (county). We have shares and prices at the brand-county-week level. Within-county shares of unaffiliated consumers depend on prices, station amenities,[30] and demographic characteristics of the local population. The brand-county fixed effects, ζ_{jm}, control for variation in the number of stations, brand amenities, and local demographic characteristics. Because we observe station entry and exit, we also include the number of stations for the brand in that market, N_{jmt}, to capture within-brand-county variation over time in this variable.

Thus, brand-county fixed effects, which are identified by the panel, allow us to account for a first-order component of heterogeneity in preferences. Another important component of preferences in this model is price sensitivity. To account for heterogeneity in price sensitivity, we interact price with the log median household income in the county.[31]

In addition to the brand-county fixed effects, we employ panel data methods to address other unobservables. We allow for the fact that δ_{jmt} may be correlated over time in ways not dependent on (p, N, ζ). We let the time-varying unobserved components of demand be specified as $\phi_t + \psi_{m,month(t)} + \eta_{jmt}$. That is, we estimate period (weekly) fixed effects $\{\phi_t\}$ and county-specific (monthly) seasonal demand shocks $\{\psi_{m,month(t)}\}$.[32] Once we incorporate these fixed effects, the identifying restriction for the dynamic parameters is that the brand-market-period specific shock η_{jmt} is uncorrelated across periods, after accounting for aggregate period-specific shocks, county-level seasonal patterns, and brand-county level differences. Thus, our model attribute the residual brand-specific correlation in demand over time within a market to

---

[29]In the estimated model, the mean market-level correlation in shocks is -0.01.

[30]Station amenities include, for example, the presence of food (snack or restaurant), co-location with a supermarket, car services, and proximity to an interstate. Demographic characteristics might include median household income, population, population density, and commute percent. These do not vary much over time in our sample; when a new station enters or exits, the averages within a brand do change slightly.

[31]We do not control for unobserved heterogeneity in price sensitivity, which would add a significant computational burden. Despite a wide range in market-level median income in our data (from \$20,000 to \$124,000, or a spread of 1.8 log points) we find modest effects of income on price sensitivity.

[32]We benefit from the size of our dataset. 95 percent of county-months have at least 18 observations, and 99 percent of county-brands have at least 40 observations.

unobservable consumer types arising from affiliation.

We allow for endogeneity in pricing behavior by instrumenting for p_{jmt} with deviations in wholesale costs arising from crude oil production in the US. The instrument (z_1) is constructed from a regression of deviations of wholesale costs (from the brand-county average) on the interaction of US production of crude oil with the average wholesale cost for the brand in the county.[33] This gives us brand-county-specific time variation in our instrument which is (a) correlated with the wholesale cost and (b) plausibly not linked to demand. We chose this measure, rather than instrumenting directly with brand-state wholesale costs, to allow for the possibility that local variation in wholesale costs over time may reflect brand-specific demand shocks.

We interact the above instrument with $Income_m$ to create a second instrument, z_2, to account for the endogeneity of ($p_{jmt} \times Income_m$). Both US crude oil production and income are plausibly exogenous with respect to local, time-varying demand shocks. Figure 7 in the Appendix summarizes the time-series variation by plotting mean total market shares and mean prices during our sample in panel (a). In panel (b), we plot the mean instrument z_1 against the mean price. As the figure shows, there is a strong correlation with the instrument, constructed from US production of crude oil, and prices. Prices display seasonal patterns, reflecting demand, while our instrument does not.

The estimates for the linear parameters are reported in Table 4. The first three columns report coefficient estimates from a logit demand regression using observed shares. The fourth column reports the results for unaffiliated customers from our dynamic model. In the static model, all consumers are assumed to be unaffiliated. We obtain a larger (in magnitude) price coefficient with our dynamic specification, as the model has isolated unaffiliated consumers (including shoppers) from affiliated consumers with lower price sensitivities. As expected, higher income corresponds to a lower price sensitivity, though this effect is small. A 1 log point (or 271 percent) increase in income changes the price coefficient from -7.615 to -7.285. We find that an increase in the number of stations that a brand has in a market has a statistically significant positive effect on demand for unaffiliated consumers.

Table 5 reports estimates of the dynamic parameters. The parameters θ_1, θ_2, and θ_3 imply that 62 percent of consumers, on average, are subject to state dependence and develop an affiliation for the brand they previously purchased from. The coefficient of 0.049 on income indicates that higher-income consumers are more likely to develop an affiliation, consistent with a habit-formation model where switching costs are increasing in wages. Likewise, the coefficient of 0.019 on population density indicates that consumers in more dense areas are more likely to become affiliated. We interpret this to reflect that urban environments have higher driving costs, which also increase switching costs. Both of these demographic variables

---

[33]Our measure of the average brand-county wholesale cost is the fixed effect obtained by a regression of wholesale costs on brand-county and weekly fixed effects, thereby accounting for compositional differences across time.

Table 4: Demand Regressions: Unaffilated Customers

| | Static Model | | | Dynamic Model |
	(1)	(2)	(3)	(4)
Price	−0.022*	−0.260***	−2.315***	−7.615***
	(0.014)	(0.038)	(0.507)	(1.862)
Price × Income	−0.135***	0.077***	−0.007	0.329***
	(0.021)	(0.023)	(0.032)	(0.129)
Number of Stations	0.016***	0.064***	0.063***	0.241***
	(0.004)	(0.010)	(0.010)	(0.040)
IV	No	No	Yes	Yes
Brand-County FEs		X	X	X
Week FEs		X	X	X
County-(Month of Year) FEs		X	X	X
Observations	110,844	110,844	110,844	110,844

Notes: Significance levels: * 10 percent, ** 5 percent, *** 1 percent. Table displays the estimated coefficients for a logit demand system, where the dependent variable is the log ratio of the share of the brand to the share of the outside good. For the first three models, the dependent variable uses observed, aggregate shares. For the fourth model, the dependent variable uses the shares of free agent customers, which are calculated based on the estimated dynamic parameters. Standard errors are clustered at the county level. For the dynamic model, standard errors are cacluated via the bootstrap.

are standardized, so each coefficient corresponds to an increase of one standard deviation. Neither is statistically significant at the 95 percent level.

The estimated utility shock to affiliated consumers, $\bar{\xi}$, implies that the affiliated consumers are inelastic with respect to price. Across observations, the mean own-price elasticity for affiliated consumers is $−0.55$, and the median is zero. The utility shock generates behavior so that affiliated consumers do not respond to prices when the levels are low, but they do when the prices are high enough. The average (absolute) weekly price change in the data is 5 cents per gallon, and the 25th and 75th percentile price changes are 1.6 cents and 7.5 cents, respectively. We therefore find that affiliated consumers do not typically switch brands within this range of price changes.[34] The unaffiliated consumers, however, are highly elastic, with an average own-price elasticity of $−21.1$. This is large in magnitude, and it implies that for a 1 percent increase in price (roughly 3 cents), the station will lose 21.1 percent of the unaffiliated consumers. This high level of price sensitivity for a subset of retail gasoline consumers seems plausible, as some "shoppers" have been found to go well out of the way to save a few cents per gallon.[35]

---

[34]A common limitation of empirical models is that it is only possible to capture local variation in the data. We imagine that these consumers would be more elastic if subject to price changes of a much greater magnitude.

[35]For example, the National Association of Convenience Stores found in their 2018 survey that 38 percent of people would drive 10 minutes out of their way to save 5 cents per gallon. See, `https://www.convenience.org/Topics/Fuels/Documents/How-Consumers-React-to-Gas-Prices.pdf`

Table 5: Estimated Dynamic Parameters

	Baseline θ_1	Affiliation Rate Income θ_2	Density θ_3	Strength of Affiliation Utility Shock $\bar{\xi}$
Coefficient	0.479	0.049	0.019	18.682
95 Percent CI	[0.45, 0.50]	[-0.05, 0.08]	[-0.01, 0.04]	[18.68, 19.06]

Notes: Table displays the estimated non-linear coefficients from the dynamic model. The first three parameters imply that, on average, 62 percent of consumers that purchase develop an affiliation for that brand. Brands located in areas with higher incomes and higher population densities have greater rates of affiliation, though this heterogeneity is not statistically significant. The last parameter shows the level shock for affiliated customers, which is positive, as expected. Confidence intervals are shown with brackets and are calculated via the bootstrap.

On average, roughly 26 percent of a brand's customers come from unaffiliated consumers in any week in equilibrium. The average weighted elasticity, which weighs affiliated and unaffiliated consumers by their relative (purchasing) proportions, is -3.4. This weighted elasticity captures the effective elasticity faced by a firm and is different than the elasticity obtained when estimating a static model. A "naive" estimate using a static model would result in a value of -5.7, which implies a much greater loss in market share for a given price increase than we estimate from the dynamic model. Our brand-county elasticity estimate of -3.4 is more inelastic than some other estimates in the literature, which may be due to the fact that the existing estimates do not account for consumer inertia.[36] At the market level, the parameter estimates imply an aggregate weekly elasticity of demand of -2.5.[37]

5 Empirical Application: Supply-Side Analysis

5.1 Dynamic Pricing Behavior

Given the demand estimates, we construct the components in each firm's Bellman equation from (6). The dynamic condition for optimal pricing for brand j is:

$$\frac{\partial \pi_{jt}}{\partial p_{jt}} + \beta \frac{\partial E\left[V_j(r_{t+1}, c_{t+1}, x_{t+1}) | p_t, r_t, c_t, x_t\right]}{\partial p_{jt}} = 0, \tag{24}$$

---

[36]Additionally, the estimates of Houde (2012) of -10 to -15 reflect an elasticity at the station level, rather than at the brand-county level, which should result in more elastic estimates.

[37]The literature has typically estimated the aggregate market elasticity for gasoline to be highly inelastic. For example, Levin et al. (2017) find an aggregate elasticity of -0.30 and Li et al. (2014) estimate it to be -0.1. An important distinction between our estimate and previous aggregate elasticities is that ours is obtained using more narrowly defined markets and over a shorter time horizon (one week). Both of these features should produce more elastic estimates. Levin et al. (2017) do estimate a two-day (aggregate) elasticity of -1.38, which is closer to our estimate.

Table 6: Summary of Implied $\beta \frac{\partial E[V_j(\cdot)|\cdot]}{\partial p_{jt}}$

Group	Mean	Min	p25	Median	p75	Max
All	-0.109	-0.761	-0.152	-0.083	-0.042	0.628

Notes: Table displays the estimated derivative of continuation value. A finding of zero would indicate the absence of forward-looking behavior by firms. Negative values indicate that firms are pricing lower in that period than the optimal myopic price.

where $\frac{\partial \pi_{jt}}{\partial p_{jt}}$ is the derivative of the per-period profits. This derivative equals $\frac{\partial S_{jt}}{\partial p_{jt}}(p_{jt} - c_{jt}) + S_{jt}$ for single-product firms.

The estimation of dynamic parameters, along with our measures of marginal costs, allow for a direct estimate of the derivative of the static profit with respect to price: $\frac{\partial \pi_{jt}}{\partial p_{jt}}$. If this were zero, it would imply that firms are pricing myopically in the context of the model, as they are simply maximizing the current-period profits. When it is non-zero, it implies that dynamic considerations are affecting a firm's pricing decision.

On average, we find that $\frac{\partial \pi_{jt}}{\partial p_{jt}}$ is positive. This implies that firms are systematically pricing lower than the myopic profit-maximizing price. We interpret this as evidence of forward-looking behavior and the presence of dynamics, consistent with the reduced-form evidence of Section 3.3.2. Based on equation (24), we attribute the difference between $\sum_{l \in J_i} \frac{\partial \pi_{lt}}{\partial p_{jt}}$ and 0 to be accounted for by the derivative of the continuation value (DCV), $\beta \frac{\partial E[V_j(\cdot)|\cdot]}{\partial p_{jt}}$. That is, the dynamic incentive is the residual that rationalizes the observed pricing behavior of the firms, conditional on the demand-side assumptions, the data, and Bertrand price competition.[38] After estimating demand in an independent step, we are able to recover these residuals directly.

Summary statistics for the value of the derivative of the continuation value (DCV) are presented in Table 6. The mean and median are negative, which implies that, typically, a reduction in price would increase the expected future return. We estimate a positive residual in only 3 percent of observations. The magnitudes are significant: the mean of -0.109 implies that a 1 cent increase in price would increase static profits by roughly 4 percent.[39] Intuitively, firms are lowering prices to invest in future demand. Such behavior allows firms to occasionally have negative price-cost margins, which occur in 2.7 percent of the observations in our data. These result, combined with our reduced-form findings of anticipatory pricing for expected costs, provides consistent evidence of forward-looking pricing behavior in retail gasoline.

---

[38]Other explanations may be plausible. For example, a component of this residual may be profits obtained by complementary products, such as food sold at retail gasoline stations.

[39]The average (scaled) profit in our data is 0.029. Price-cost margins are approximately 21 cents per gallon.

5.2 Supply-Side Estimation

To estimate counterfactual pricing behavior by firms, it is necessary to estimate how dynamic incentives vary with state variables and firm actions. Two approaches are possible. The first is to take a stance on the beliefs of firms and, via forward simulation, solve for the equilibrium continuation value function. Alternatively, one can approximate the DCV with a reduced-form model that is a function of state variables. We pursue the second approach. This greatly reduces computational time to re-compute the price equilibria and avoids the need to make dimension-reducing assumptions (such as constructing a limited grid for prices) that are less palatable in our setting. To accurately represent behavior, this approach requires that the state variables included in the reduced-form approximation capture the payoff-relevant states (including market structure) and also that the counterfactual states can be reasonably interpolated from the data.

Using the data and the estimated demand parameters, we obtain estimates of the DCV and project these estimates on prices and state variables, including measures that capture expectations. In general, Markovian assumptions allow for the continuation value to be expressed as a function of the state variables and actions of the firms. We estimate the following dynamic first-order condition:

$$\frac{\partial \pi_{jt}}{\partial p_{jt}} + \Psi_j(p_t, r_t, c_t, x_t; \theta) + \zeta_{jt} = 0 \tag{25}$$

Thus, we use $\Psi_j(\cdot)$ to approximate $\beta \frac{\partial E[V_j(\cdot)|\cdot]}{\partial p_{jt}}$, and ζ_{jt} is the unobserved error. We can use this function to approximate how the dynamic incentives change with the state and the endogenous pricing decisions by firms, allowing for counterfactual analysis. This approach is an alternative to that of Bajari et al. (2007), who use an approximation to the policy function, and, based on this, leverage model structure to estimate the dynamic incentives and static parameters. Conversely, we use structural modeling to obtain static parameters and calculate a reduced-form approximation to the dynamic incentives. One advantage of our approach is that it is not necessary to take a stance on the discount rate or the beliefs of firms; both of these are absorbed into the reduced-form model.

This approach is consistent with a structural model (and solving for the equilibrium DCV) under the assumption that (i) the information set of firms matches the information set of the econometrician and (ii) firms perform limited forecasts of the evolution of the future profits, consistent with the approximation used in estimation. In equilibrium, if firms use a limited set of state variables and a simplified functional form to estimate the dynamic incentives, then the econometrician may be able to replicate the regression (or machine-learning procedure) implemented by firms. In this case, the firms' beliefs can correspond to the econometrician's estimates.[40]

---

[40]To get a sense of how close our estimates come to rational expectations, we use forward simulations to check if our estimate of $\Psi(\cdot)$ is consistent with the actual DCV, conditional on firms' choosing price according to $\Psi(\cdot)$. We discuss in greater detail in the following pages.

	$\beta \frac{\partial E[V_j(\cdot)|\cdot]}{\partial p_{jt}}$	Sensitivity
	(1)	(2)
r_{jmt}	-0.625^{***}	5.864^{***}
	(0.001)	(0.017)
$r_{jmt} \times \frac{\partial S_{jmt}}{\partial p_{jmt}}$	-0.147^{***}	2.329^{***}
	(0.002)	(0.044)
Marginal Cost	-0.003^{***}	0.042^{***}
	(0.000)	(0.004)
Cost Change (30-Day Ahead)	0.019^{***}	-0.203^{***}
	(0.001)	(0.019)
λ_m	-0.181^{***}	3.070^{***}
	(0.012)	(0.268)
Num. Stations (Brand)	-0.000^{***}	0.010^{***}
	(0.000)	(0.000)
Num. Stations (Market)	-0.000^{***}	0.000^{**}
	(0.000)	(0.000)
Num. Brands (Market)	-0.002^{***}	0.051^{***}
	(0.000)	(0.002)
Constant	Yes	Yes
Observations	110,844	110,844
R^2	0.912	0.625

Notes: Significance levels: * 10 percent, ** 5 percent, *** 1 percent. Table displays the estimated coefficients from a regression of the dynamic pricing incentive on state variables. The second column reports the regression with a measure of sensitivity, which is the log absolute value of the dynamic pricing incentive. In general, a negative coefficient in the first column implies a greater sensitivity to dynamics when pricing, generating a positive coefficient in the second column.

To estimate $\Psi_j(\cdot)$, we project the residual DCV onto the share of affiliated consumers r_{jmt}, marginal costs, and expectations of future costs. Our model and descriptive evidence suggests that these variable play an important role in expectations of future profits. We also include the interaction of r_{jmt} with the derivative of shares with respect to price, motivated by the fact that future considerations may be directly be influenced by the derivative of current period shares. We also include the fraction of state-dependent consumers λ_m, the number of stations, the total number of stations for all brands, and the number of brands as market-level controls.

The results of estimating equation (25) are reported in Table 7. The first specification reports the coefficients from a regression of the DCV onto the dependent variables. As the

DCV is negative on average, a negative coefficient implies that the variable is associated with a stronger dynamic pricing incentive, or a greater deviation from the optimal static price.

To show more directly how sensitive firms are to dynamic considerations, the second column reports a regression where we replace the value of the DCV with the logged absolute value. Thus, the coefficients reflect the semi-elasticity for the magnitude of the dynamic incentive. A positive coefficient in the second column indicates that an increase in the variable makes a firm more sensitive to dynamic considerations, whereas a negative coefficient indicates a reduced sensitivity to dynamic considerations when pricing. Typically, a negative coefficient in the first column corresponds to a positive coefficient in the second, as the average value for the DCV is negative.

We find that a higher share of affiliated consumers at the firm level r_{jmt} and the market level λ_m increases the magnitude of the DCV, corresponding to an increased investment incentive when pricing. These variables have the largest impact on the DCV. The effect of affiliated consumers on pricing behavior is exacerbated when the consumers are more responsive to price in the current period, as indicated by the coefficient of -0.147 on $r_{jmt} \times \frac{\partial S_{jmt}}{\partial p_{jmt}}$. Consistent with forward-looking behavior, increases in marginal costs today, which tend to reduce current profits, make firms more sensitive to dynamic profit considerations. Conversely, expected increases in future costs lead firms to place relatively more weight on current profits. Relative to the impact of affiliated consumers, we find that other market-level controls have modest effects. Overall, our nine-parameter model captures over 91 percent of the variation of the residual that rationalizes observed prices.

5.3 Horizontal Market Power: Merger Simulation

To evaluate the impact of dynamic pricing incentives on horizontal market power, we simulate a merger between Marathon and BP, which are the number one and number four (non-fringe) brands in terms of overall shares in our sample. Out of the 241 markets, they overlap in 75. In these 75 markets, the average (post-merger) HHI is 1511, and the mean change in HHI resulting from the merger is 383. In 8 markets, the resulting HHIs are greater than 2500, and the changes are greater than 200, meeting the typical thresholds that are presumed likely to enhance market power. The merger would change twelve markets from 3 firms to 2 firms and eighteen markets from 4 firms to 3 firms. We allow the firms to merge at the beginning of September 2014, and we calculate counterfactual prices and shares for the second half of the sample.[41]

In Section 2, we showed that the price effects of a merger can depend on the way the merger is implemented, especially in the presence of consumer inertia. To measure the potential

---

[41]Because our inelastic affiliated customers might technically purchase at very high prices, we impose a choke price of $5 in demand and impose a penalty for prices that exceed this value. The baseline functional form of demand may not be reasonable for extreme out-of-sample values. Across all merger counterfactuals, only 8 observations approach the choke price.

Table 8: Merger Effects

Brand	Dynamic Model: Joint Pricing			Dynamic Model: Brand Consolidation			Static Model		
	Price	Share	Profit	Price	Share	Profit	Price	Share	Profit
Marathon-BP	3.14	-11.07	20.12	1.97	0.11	30.43	5.35	-16.74	34.44
Other	0.05	4.64	5.29	1.82	-3.59	17.21	0.61	3.86	11.28
Overall	1.26	-2.42	12.4	1.90	-1.93	23.55	2.40	-5.40	22.39

Notes: Table displays the mean effects from counterfactual mergers between two brands in our data. The first six columns provide estimates from dynamic models that account for consumer inertia. The first three columns report a counterfactual joint pricing merger and the second three columns report a counterfactual brand consolidation merger. The last three columns provide the estimates for a brand consolidation merger from a static model that is calibrated to match prices, margins, and shares from the same data. Price effects are weighted by share.

empirical impact, we consider the two alternative merger scenarios described in that section: a joint pricing merger and a brand consolidation merger. In the joint pricing scenario, the merged firm has pricing control over both brands, which they maintain as distinct entities. In the brand consolidation scenario, the merged firm conslidates the assets under a single brand. To implement a brand consolidation merger, we need to make an assumption about how much utility consumers of the product removed from the market will receive from buying the consolidated brand. As in the theoretical analysis, we assume that, at pre-merger prices, the consolidated brand will have the same combined share of unaffiliated customers as the separate pre-merger brands.

For the joint pricing scenario, we also need to make an assumption about the cross-price effects on the continuation value, i.e., $\beta \partial E [V_k] / \partial p_{jt}$ when k and j are owned by the merged firm. For our baseline results, we assume that the effects are proportional to the diversion ratios D_{kj}, so that $\beta \partial E [V_k] / \partial p_{jt} = D_{jk} \partial E [V_k] / \partial p_{kt}$. The diversion ratios capture the relative effects on shares, which should be correlated with the effects on profits. Our counterfactual results are qualitatively similar with moderate changes in this scaling factor, such as assuming the cross-price effects are zero. As expected, if we reduce the impact of the cross-price effects on the continuation value, we get a lower impact on post-merger prices.

Table 8 displays the mean effects on prices, shares, and profits of the two mergers. The first three columns report the effects from the joint pricing scenario, and the second three columns report the effects from the brand consolidation scenario. In either scenario, the overall price effects are moderate. The joint pricing scenario predicts the merging firms will raise prices by 3.1 percent, with an 11 percent decrease in shares and a meaningful increase in profits. Rival firms in the same market see an increase in share and a small effect on prices, which results in higher profits. Overall, prices in the market increase by 1.3 percent.

The brand consolidation scenario predicts a lower price increase by the merging firms, about

2.0 percent. However, in the brand consolidation scenario, the merging firms realize a slight *increase* in market share. This occurs because of asymmetries among firms. When merging firms have different brand values, the affiliated consumers of the weaker brand become more attached to the consolidated brand post-merger. In other words, affiliated consumers benefit from becoming locked in to a better brand. Recall that our merger holds the choice probabilities for unaffiliated consumers to be identical post-merger at the pre-merger prices. Thus, if we hold prices fixed at the pre-merger levels, we observe the merging firm to accumulate greater shares over time due to the superior strength of affiliation.

Another difference between the two mergers is in the effects on rivals. In contrast to the joint pricing scenario, the brand consolidation scenario sees rivals increase prices by 1.8 percent, which is almost as much as the merging firms. This leads to a greater overall price increase of 1.9 percent across the affected markets and higher firm profits.

Overall, these effects are modest relative to what might be expected from a merger with the corresponding market shares and HHIs. For comparison, we report the results from a merger analysis using a static model in the last three columns of Table 8. We calibrate a standard logit demand system to identical prices, margins, and shares that are used to estimate the dynamic model. We use a brand consolidation merger to illustrate the potential effects.[42] The static model predicts price effects of over 5 percent for the merging firms, which is meaningfully higher than the predictions from the dynamic models.

Consistent with our simulations in Section 2, we find that the dynamic model with consumer inertia can predict smaller price increases than a static model. The dynamic incentive to invest in future demand can mitigate the short-run incentive to raise prices post-merger, dampening the exercise of horizontal market power.

5.4 Dynamic Market Power

In the previous section, we used the empirical model to evaluate the role of horizontal market power in the presence of consumer inertia. In this section, we isolate the role of dynamic market power. To do so, we change the share of consumers that are affected by inertia, λ.

We simulate counterfactual scenarios where the market-specific λ increases or decreases by 0.07. These changes correspond approximately to an 11 percent change in λ, as the average value across markets is 0.62. Thus, we consider the impact of the share of "shoppers" on equilibrium prices. Though we do not consider a merger in this section, we report results for Marathon and BP separately from other brands and only for the markets where both firms are present. This facilitates a comparison to the horizontal market power effects in Table 8.

Table 9 reports the results from the counterfactual scenarios. The first three columns report the equilibrium effects of an increase in the prevalence of state dependence, and the second

---

[42]Though we showed in Section 2 that both types of mergers predict identical price effects in a symmetric oligopoly, they can have different predictions with asymmetries.

Table 9: Equilibrium Effects of State Dependence

	Increase λ by 0.07			Decrease λ by 0.07		
Brand	Price	Share	Profit	Price	Share	Profit
Marathon-BP	2.03	6.66	33.81	-1.73	-7.82	-28.5
Other	1.40	7.99	26.44	-1.64	-8.24	-26.27
Overall	1.68	7.39	29.97	-1.68	-8.05	-27.34

Notes: Table displays the mean effects from a counterfactual scenarios with different portions of consumers affected by state dependence. The first three columns report the equilibrium effects with a greater prevalence of state dependence, and the second three columns report the effects with less state dependence.

three columns report the effects of a decrease. If an additional 7 percent of consumers in the market were affected by state dependence, then overall prices would increase by 1.7 percent. Marathon and BP would benefit slightly more than other firms, increasing prices by 2 percent and realizing a 34 percent increase in profits. The effects of reducing λ by the same amount are quite similar, but in the opposite direction.

Overall, these effect on prices and profits are similar to the effects of a brand consolidation merger between Marathon and BP (Table 8). In rough terms, reducing the share of consumers that are shoppers by 7 percentage points has an equilibrium price effects on par with the effects of a merger. Thus, the model allows us to quantify the relative effects of dynamic incentives and horizontal competition when evaluating market power. Our findings suggest that policies that produces even modest changes in switching costs (or consumer inertia, more generally) may have a similar impact on welfare than large changes to market structure, such as those that occur through mergers.

6 Conclusion

We develop a model of consumer inertia that accounts for commonly observed dynamic pricing behavior, such as the slow adjustment of prices to changes in cost. The dynamics result from competing firms optimally setting prices to consumers that may become loyal or habituated to their current supplier. Dynamic market power reflects the ability of firms to raise price in response to consumer inertia relative to static consumer demand, though consumer inertia can lead to lower prices in equilibrium, depending upon whether the incentive to harvest affiliated consumers dominates the incentive to invest in future demand. High levels of dynamic market power may correspond with lower levels of horizontal market power, i.e., the impact of competition from rival products on prices.

Using data from retail gasoline markets, we first present reduced-form evidence consistent with consumer inertia. Firm-level shares are highly correlated over time, prices slowly adjust to

changes in marginal costs, and prices anticipate future changes in expected costs. The evidence suggests that, even in a relatively competitive market with a homogeneous product, accounting for dynamics in consumer behavior and firm behavior may be important.

We develop an empirical model that can identify dynamic demand parameters using data on price, shares, and an instrument. Results suggest that 62 percent of retail gasoline consumers become affiliated to the firm from which they currently purchase on a week-to-week basis, and that these consumers are price insensitive. Conversely, we find that unaffiliated consumers are quite price sensitive and play an important role in disciplining equilibrium prices. We evaluate the dynamic incentives that affect prices. We show, both theoretically and empirically, that failing to account for dynamic demand can cause significant biases when predicting post-merger price increases.

References

Aguirregabiria, V. and P. Mira (2007). Sequential estimation of dynamic discrete games. *Econometrica 75*(1), 1–53.

Arcidiacono, P. and R. A. Miller (2011). Conditional Choice Probability Estimation of Dynamic Discrete Choice Models With Unobserved Heterogeneity. *Econometrica 79*(6), 1823–1867.

Bajari, P., C. L. Benkard, and J. Levin (2007). Estimating dynamic models of imperfect competition. *Econometrica 75*(5), 1331–1370.

Beggs, A. and P. Klemperer (1992, May). Multi-Period Competition with Switching Costs. *Econometrica 60*(3), 651.

Bergemann, D. and J. Valimaki (1996, September). Learning and Strategic Pricing. *Econometrica 64*(5), 1125.

Bergemann, D. and J. Välimäki (2006, March). Dynamic price competition. *Journal of Economic Theory 127*(1), 232–263.

Berry, S., M. Carnall, and P. Spiller (2006). Airline Hubs: Costs, Markups and the Implications of Customer Heterogeneity. *Advances in Airline Economics 1*.

Berry, S. and P. Jia (2010, August). Tracing the Woes: An Empirical Analysis of the Airline Industry. *American Economic Journal: Microeconomics 2*(3), 1–43.

Berry, S., J. Levinsohn, and A. Pakes (1995, July). Automobile Prices in Market Equilibrium. *Econometrica 63*(4), 841.

Besanko, D., U. Doraszelski, and Y. Kryukov (2018). How Efficient is Dynamic Competition? The Case of Price as Investment. Working Paper w23829, National Bureau of Economic Research, Cambridge, MA.

Cabral, L. (2011, January). Dynamic Price Competition with Network Effects. *The Review of Economic Studies 78*(1), 83–111.

Chandra, A. and M. Tappata (2011). Consumer search and dynamic price dispersion: An application to gasoline markets. *The RAND Journal of Economics 42*(4), 681–704.

Dubé, J.-P., G. J. Hitsch, and P. E. Rossi (2009). Do switching costs make markets less competitive? *Journal of Marketing Research 46*(4), 435–445.

Dubé, J.-P., G. J. Hitsch, and P. E. Rossi (2010). State dependence and alternative explanations for consumer inertia. *The RAND Journal of Economics 41*(3), 417–445.

Eckert, A. (2013). Empirical studies of gasoline retailing: A guide to the literature. *Journal of Economic Surveys 27*(1), 140–166.

Engle, R. F. and C. W. J. Granger (1987, March). Co-Integration and Error Correction: Representation, Estimation, and Testing. *Econometrica 55*(2), 251.

Farias, V., D. Saure, and G. Y. Weintraub (2012, June). An approximate dynamic programming approach to solving dynamic oligopoly models. *The RAND Journal of Economics 43*(2), 253–282.

Farrell, J. and P. Klemperer (2007). Chapter 31 Coordination and Lock-In: Competition with Switching Costs and Network Effects. In *Handbook of Industrial Organization*, Volume 3, pp. 1967–2072. Elsevier.

Farrell, J. and C. Shapiro (1988). Dynamic Competition with Switching Costs. *The RAND Journal of Economics 19*(1), 123.

Farrell, J. and C. Shapiro (2010). Antitrust evaluation of horizontal mergers: An economic alternative to market definition. *The BE Journal of Theoretical Economics 10*(1).

Gowrisankaran, G. and M. Rysman (2012, December). Dynamics of Consumer Demand for New Durable Goods. *Journal of Political Economy 120*(6), 1173–1219.

Handel, B. R. (2013). Adverse selection and inertia in health insurance markets: When nudging hurts. *The American Economic Review 103*(7), 2643–2682.

Heckman, J. J. (1981). Heterogeneity and State Dependence. *Studies in Labor Markets*, 91–140.

Hendel, I. and A. Nevo (2013, December). Intertemporal Price Discrimination in Storable Goods Markets. *American Economic Review 103*(7), 2722–2751.

Honka, E. (2014). Quantifying search and switching costs in the US auto insurance industry. *The RAND Journal of Economics 45*(4), 847–884.

Hortaçsu, A., S. A. Madanizadeh, and S. L. Puller (2017, November). Power to Choose? An Analysis of Consumer Inertia in the Residential Electricity Market. *American Economic Journal: Economic Policy 9*(4), 192–226.

Houde, J.-F. (2012, August). Spatial Differentiation and Vertical Mergers in Retail Markets for Gasoline. *American Economic Review 102*(5), 2147–2182.

Lee, R. S. (2013, December). Vertical Integration and Exclusivity in Platform and Two-Sided Markets. *American Economic Review 103*(7), 2960–3000.

Levin, L., M. S. Lewis, and F. A. Wolak (2017). High Frequency Evidence on the Demand for Gasoline. *American Economic Journal: Economic Policy 9*(3), 314–347.

Lewis, M. and M. Noel (2011). The speed of gasoline price response in markets with and without Edgeworth cycles. *Review of Economics and Statistics 93*(2), 672–682.

Li, S., J. Linn, and E. Muehlegger (2014). Gasoline Taxes and Consumer Behavior. *American Economic Journal: Economic Policy 6*(4), 302–342.

Miller, N., M. Remer, C. Ryan, and G. Sheu (2017). Upward Pricing Pressure as a Predictor of Merger Price Effects. *International Journal of Industrial Organization 52*, 216–247.

Miller, N. H., M. Remer, C. Ryan, and G. Sheu (2016, December). Pass-Through and the Prediction of Merger Price Effects: Pass-Through in Merger Analysis. *The Journal of Industrial Economics 64*(4), 683–709.

Miller, N. H., M. Remer, and G. Sheu (2013, March). Using cost pass-through to calibrate demand. *Economics Letters 118*(3), 451–454.

Nosal, K. (2011). Estimating Switching Costs for Medicare Advantage Plans. Working paper.

Pakes, A., M. Ostrovsky, and S. Berry (2007). Simple estimators for the parameters of discrete dynamic games (with entry/exit examples). *The RAND Journal of Economics 38*(2), 373–399.

Remer, M. (2015, September). An empirical investigation of the determinants of asymmetric pricing. *International Journal of Industrial Organization 42*, 46–56.

Remer, M. and F. R. Warren-Boulton (2014). United States v. H&R Block: Market definition in court since the 2010 Merger Guidelines. *The Antitrust Bulletin 59*(2), 599–618.

Shcherbakov, O. (2016). Measuring consumer switching costs in the television industry. *The RAND Journal of Economics 47*(2), 366–393.

Shum, M. (2004). Does advertising overcome brand loyalty? Evidence from the breakfast-cereals market. *Journal of Economics & Management Strategy 13*(2), 241–272.

Slade, M. E. (1998). Optimal pricing with costly adjustment: Evidence from retail-grocery prices. *The Review of Economic Studies 65*(1), 87–107.

Stahl, D. O. (1989). Oligopolistic pricing with sequential consumer search. *The American Economic Review*, 700–712.

Sweeting, A. (2013). Dynamic Product Positioning in Differentiated Product Markets: The Effect of Fees for Musical Performance Rights on the Commercial Radio Industry. *Econometrica 81*(5), 1763–1803.

Varian, H. R. (1980). A Model of Sales. *American Economic Review 70*(4), 651–659.

Weintraub, G. Y., C. L. Benkard, and B. Van Roy (2008). Markov Perfect Industry Dynamics With Many Firms. *Econometrica 76*(6), 1375–1411.

Yeo, J. and D. P. Miller (2018, November). Estimating switching costs with market share data: An application to Medicare Part D. *International Journal of Industrial Organization 61*, 459–501.

A More Details on Theoretical and Numerical Steady-State Analysis

A.1 Monopoly

We analyze steady-state prices in a monopoly market (with an outside good) to show how habit-forming consumers affect optimal prices and markups.

To simplify notation in the monopoly case, let the monopolist's share of affiliated and unaffiliated consumers be s_j and s_0, respectively, and its number of affiliated consumers be r. Consumers become affiliated at rate λ. We assume positive dependence in purchase behavior, so that $s_j > s_0$. In the steady state, $r_{jt} = r_{j(t+1)} = r_j$ and $c_{jt} = c_{j(t+1)} = c_j$. The steady-state number of affiliated consumers, r^{ss}, is:

$$r^{ss} = \lambda\big((1 - r^{ss})s_0 + r^{ss} \cdot s_j\big)$$

$$\implies r^{ss} = \frac{\lambda s_0}{1 - \lambda(s_j - s_0)}.$$

The steady-state number of affiliated consumers is increasing in the probability of becoming affiliated, λ, and the difference between the choice probabilities of affiliated consumers and unaffiliated consumers, $s_j - s_0$. Using the steady-state value of affiliated consumers, we can solve for the steady-state pricing function.

The steady-state period value is:

$$V^{ss}(r^{ss}, c^{ss}) = (p^{ss} - c^{ss})((1 - r^{ss})s_0 + r^{ss}s_j) + \beta V^{ss}$$

$$= \frac{p^{ss} - c^{ss}}{1 - \beta} \cdot \frac{s_0}{1 - \lambda(s_j - s_0)}.$$

This equation represents the monopolists discounted profits, conditional on costs remaining at its current level. Thus, profits are increasing in both λ and the difference in choice probabilities of affiliated and unaffiliated consumers. These results are straightforward: affiliated consumers are profitable. Also, note that a model with no affiliation is embedded in this formulation ($\lambda = 0$ and $s_j = s_0$), in which case profits are simply the per-unit discounted profits multiplied by the firm's market share.

Maximizing the steady-state value with respect to p^{ss} yields the firm's optimal pricing function:

$$p^{ss} = c^{ss} + \underbrace{\frac{-s_0\left(1 - \lambda s_j + \lambda s_0\right)}{\frac{ds_0}{dp}\left(1 - \lambda s_j\right) + \frac{ds_j}{dp}\lambda s_0}}_{m \,=\, \text{markup of price over marginal cost}}. \tag{26}$$

The second term, m, on the right-hand side of equation (26) captures the extent to which the firm prices above marginal cost (in equilibrium). As this markup term depends upon choice probabilities, it is implicitly a function of price. Thus, as in the standard logit model, we cannot derive an analytical solution for the steady-state price. Nonetheless, we derive a condition below to see how markups are impacted by consumer affiliation. In the usual case, m will be declining in p, ensuring a unique equilibrium in prices.

Are markups higher or lower in the presence of affiliation? When affiliation is absent, $\lambda = 0$ and $s_j = s_0$, equation (26) reduces to the first-order condition of the static model, $p^{ss} = c^{ss} - \frac{s_0}{ds_0/dp}$. Denoting the markup term with affiliation as m_d and the markup term from the static model as m_s, we compare these two terms at the solution to the static model:

$$m_d = -\frac{s_0 \left(1 - \lambda s_j + \lambda s_0\right)}{\frac{ds_0}{dp} \left(1 - \lambda s_j\right) + \frac{ds_j}{dp} \lambda s_0} \gtreqless -\frac{s_0}{ds_0/dp} = m_s.$$

For a given price, the terms s_0 and ds_0/dp are equivalent across the two models. Rearranging terms, we obtain a simple condition relating the levels of the markup terms:

$$m_d > m_s \iff -\frac{\partial s_0}{\partial p} > -\frac{\partial s_j}{\partial p}. \tag{27}$$

A higher value for m_d indicates higher markups and higher prices. Thus, if affiliated consumer quantities are relatively less sensitive to changes in price, then markups are higher.

This is an intuitive result. However, there is a nuanced point to this analysis, stemming from the fact that there is not a direct mapping between our assumption of positive dependence and the condition in (27). Given our extension of the logit formulation, $\frac{\partial s_0}{\partial p} = \frac{\partial \delta}{\partial p} s_0 (1 - s_0)$ and $\frac{\partial s_j}{\partial p} = (\frac{\partial \delta}{\partial p} + \frac{\partial \sigma}{\partial p}) s_j (1 - s_j)$. Thus, whether or not markups are higher depends on the derivative of the type-specific shock with respect to price and the relative distance of s_0 and s_j from 0.5 (at which point $s(1-s)$ is maximized). Therefore, steady-state markups may be higher or lower with the presence of consumer affiliation. If we make the additional assumption that affiliated consumer utility is less sensitive to price, i.e. $-\frac{\partial \delta}{\partial p} > -(\frac{\partial \delta}{\partial p} + \frac{\partial \sigma}{\partial p})$, we might expect that markups are higher in the presence of consumer affiliation. However, the results show that it is still ambiguous whether markups are higher in the steady state, as s_j may be close enough to 0.5 relative to s_0 to flip the inequality.

Thus, the presence of positively affiliated consumers may, counter-intuitively, lower the steady-state price, relative to the static model. The intuition for this result is akin to those summarized in Farrell and Klemperer (2007); with dynamic demand and affiliation, firms face a trade-off between pricing aggressively today and "harvesting" affiliated consumers in future periods. In the steady state, our model shows that either effect may dominate.

A.2 Simulation Methodology

The number of unknowns in the system is $J + J + J \times J$, for p, r, and $\frac{dp}{dr}$. The law of motion in the steady state gives us J restrictions ($r = f(p, r)$). This allows us to solve for r given p. p and $\frac{dp}{dr}$ need to be solved for.

We implement the following procedure to solve numerically for the steady state:

1. Provide an initial guess for the matrix $\frac{dp}{dr}$.

2. Solve for steady-state values of p, r, and $\frac{dV_i(r)}{dr}$ given the guess for $\frac{dp}{dr}$. Use the J restrictions implied by the first-order conditions (one for each product j)

$$\frac{dV_i(r')}{dr'} \cdot \frac{dr'}{dp_j} = -\frac{1}{\beta} \sum_{k \in J_i} \frac{\partial \pi_k}{\partial p_j},$$

to solve for $\frac{dV_i(r)}{dr}$, where $\frac{dV_i(r)}{dr} = \frac{dV_i(r')}{dr'}$ in the steady state. Note that π_k, in this notation, is equal to the sum of profits from all products by a firm.

3. Take the numerical derivative of p with respect to r. Approximate the numerical derivative by slightly perturbing r: $\tilde{r}_j = r + \epsilon_j$, where j indicates a perturbation in the jth element. Re-solve for p using the first order condition. Calculate

$$\frac{dp}{dr_j} \approx \frac{p^*(r + \epsilon_j) - p^*(r - \epsilon_j)}{2 |\epsilon_j|}$$

Stack these vectors horizontally to obtain an approximation for $\frac{dp}{dr}$.

4. Calculate the absolute distance between the approximation of $\frac{dp}{dr}$ calculated in the previous step and the initial guess for $\frac{dp}{dr}$. If this distance falls below a critical value, then the solution is found. If not, update the guess for $\frac{dp}{dr}$ and repeat steps 1-4 above until a solution is found.

A.3 Numerical Simulation Parameters

Table 10: Simulation Parameter Summary Statistics

	Mean	St. Dev.	Min	Pctl(25)	Pctl(75)	Max	N
λ	0.40	0.20	0.10	0.20	0.60	0.70	3437
α	-5.20	2.56	-9.96	-7.31	-3.05	-0.44	3437
ξ	5.24	2.62	0.04	3.21	7.39	9.99	3437
$\underline{\xi}$	2.88	1.54	0.00	1.67	4.14	6.18	3437

Notes: Table displays summary statics for demand parameters for the 3,437 markets used in the numerical simulations. These markets were generated from a broader set of parameter values and selected if the resulting three-firm markets had firm shares between 0.05 and 0.30 (yielding an outside share between 0.10 and 0.85), margins between 0.05 and 0.75, and converged for all values of $\lambda \in \{0.1, 0.2, ..., 0.7\}$. See the text for additional details.

A.4 Outcomes and Demand Parameters

Table 11: Joint Pricing Merger Simulation: Demand Parameters

	(1) Pre-Merger Price	(2) Price Δ F1	(3) Bias F1	(4) Price Δ F3	(5) Bias F3
λ	0.108***	-1.542***	4.161***	0.018	28.299***
	(0.023)	(0.128)	(0.086)	(0.053)	(7.942)
α	0.146***	2.519***	0.727***	0.907***	-18.767***
	(0.004)	(0.024)	(0.016)	(0.010)	(1.484)
ξ	0.029***	1.889***	0.475***	0.774***	-18.294***
	(0.004)	(0.023)	(0.016)	(0.010)	(1.453)
$\bar{\xi}$	0.043***	-0.158***	0.573***	0.192***	-1.557
	(0.003)	(0.017)	(0.011)	(0.007)	(1.052)
Constant	1.875***	8.038***	-0.475***	0.988***	33.754***
	(0.017)	(0.093)	(0.062)	(0.038)	(5.742)
N	3437	3437	3437	3437	3437

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: The dependent variables are outcomes from a merger simulation that consolidates pricing control of products 1 and 2. F1 and F3 refer to firms 1 and 3, respectively. Pre-Merger price is for firm 1. Price Δ and Bias are the merger price change and the absolute value of the simulation bias, respectively. Parameters correspond to the dynamic demand model detailed in section 2. Standard errors in parentheses. * p < 0.10, ** p < 0.05, *** p < 0.01

Tables 11 and 12 provide results from regressions of pre-merger prices, price changes, and the absolute value of the bias in the static prediction on the demand parameters. Pre-merger prices increase with the fraction of affiliated customers (λ) and the strength of affiliation ($\bar{\xi}$). For joint pricing mergers, stronger dynamics reduce merger price effects and generate a greater under-prediction that arises from a misspecified static model, which increases the size of the bias. In brand consolidation mergers, stronger dynamics also generate larger merger price changes and absolute bias. However, these relationships do not hold in every instance, and may interact in interesting ways. As we show in Figures 1 and 2, prices are often non-monotonic in λ.

Table 12: Brand Consolidation Merger Simulation: Demand Parameters

	(1) Pre-Merger Price	(2) Price Δ F1	(3) Bias	(4) Price Δ F3	(5) Bias F3
λ	0.108***	3.089***	1.036***	1.070***	265.694*
	(0.023)	(0.160)	(0.073)	(0.054)	(147.404)
α	0.146***	2.865***	-0.278***	0.817***	-129.089***
	(0.004)	(0.030)	(0.014)	(0.010)	(27.539)
ξ	0.029***	1.783***	-0.510***	0.699***	-134.551***
	(0.004)	(0.029)	(0.013)	(0.010)	(26.969)
$\bar{\xi}$	0.043***	0.904***	0.431***	0.172***	-57.268***
	(0.003)	(0.021)	(0.010)	(0.007)	(19.522)
Constant	1.875***	8.220***	0.909***	0.467***	255.319**
	(0.017)	(0.116)	(0.053)	(0.039)	(106.571)
N	3437	3437	3437	3437	3437

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: The dependent variables are outcomes from a merger simulation that consolidates products one and two under one brand. F1 and F3 refer to firms 1 and 3, respectively. Pre-Merger price is for firm 1. Price Δ and Bias are the merger price change and simulation bias, respectively. Parameters correspond to the dynamic demand model detailed in section 2. Standard errors in parentheses. * p < 0.10, ** p < 0.05, *** p < 0.01

A.5 Numerical Results for Merging Firms

Table 13: Joint Pricing Simulation: Merger Price Change and Bias

	(1) Price Δ	(2) Price Δ	(3) Price Δ	(4) Bias	(5) Bias	(6) Bias
Pre-Merger Market Share	27.661***	29.041***	13.757***	5.029***	3.609***	2.654**
	(0.378)	(0.342)	(1.973)	(0.314)	(0.264)	(1.276)
Pre-Merger Margin	12.497***	12.642***	14.102***	7.552***	7.402***	12.320***
	(0.200)	(0.179)	(0.362)	(0.166)	(0.138)	(0.279)
λ		-3.578***	-2.819***		3.684***	3.160***
		(0.122)	(0.152)		(0.094)	(0.102)
α			0.820***			0.462***
			(0.130)			(0.084)
ξ			0.841***			0.622***
			(0.127)			(0.082)
$\bar{\xi}$			-0.451***			0.278***
			(0.016)			(0.011)
Price Δ						-0.250***
						(0.011)
Constant	-4.371***	-3.219***	-0.094	-1.335***	-2.521***	-4.168***
	(0.081)	(0.082)	(0.359)	(0.067)	(0.063)	(0.231)
N	3437	3437	3437	3437	3437	3437

Notes: The dependent variables are outcomes from a merger simulation that consolidates pricing control of products 1 and 2. Observations are for firm 1. Price Δ and Bias are the merger price change and absolute value of the static prediction bias, respectively. Market share is the aggregate market share. Standard errors in parentheses. * p < 0.10, ** p < 0.05, *** p < 0.01

To provide directional guidance on static model bias in mergers, we look at price changes as a function of pre-merger margins and shares, which are often used to simulate or approximate unilateral merger price increases (Miller et al., 2016). Columns (1)-(3) of Tables 13 and 14 explore how the percentage price change from a merger relates to pre-merger margins and market shares, which are often directly observed, as well as to primitives of the demand model. As is typically the case in static models, both pre-merger shares and margins are positively related to the size of the price change. Conditional on these observables, however, the dynamic parameters dampen the effect of a joint pricing merger but increase the effect of a brand consolidation merger. Correspondingly, the absolute value of the static model bias, which is the dependent variable in columns (4)-(6) of Tables 13 and 14, increases in joint pricing mergers but decreases in brand consolidation mergers as the pre-merger market share increases. Therefore, even if affiliation cannot be directly estimated, price change estimates should be revised accordingly if affiliation is expected to play an important role.

Table 14: Brand Consolidation Simulation: Merger Price Change and Bias

	(1) Price Δ	(2) Price Δ	(3) Price Δ	(4) Bias	(5) Bias	(6) Bias
Pre-Merger Market Share	21.813*** (0.309)	21.433*** (0.308)	32.328*** (1.770)	-9.905*** (0.241)	-10.437*** (0.234)	1.365 (1.240)
Pre-Merger Margin	24.952*** (0.163)	24.912*** (0.161)	25.472*** (0.325)	5.601*** (0.127)	5.545*** (0.123)	6.627*** (0.271)
λ		0.986*** (0.110)	0.383*** (0.136)		1.380*** (0.084)	0.362*** (0.099)
α			-0.676*** (0.117)			-0.189** (0.082)
ξ			-0.582*** (0.114)			-0.256*** (0.080)
$\bar{\xi}$			0.375*** (0.014)			0.258*** (0.011)
Price Δ						-0.225*** (0.011)
Constant	-3.927*** (0.066)	-4.245*** (0.074)	-7.590*** (0.322)	1.574*** (0.051)	1.130*** (0.056)	-0.343 (0.224)
N	3437	3437	3437	3437	3437	3437

Notes: The dependent variables are outcomes from a merger simulation that consolidates products one and two under one brand. Observations are for firm 1. Price Δ and Bias are the merger price change and the absolute value of the static prediction bias, respectively. Market share is the aggregate market share. Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

B   Additional Details on Merger Analysis

B.1   Implementation of Brand Consolidation Mergers

To implement the brand consolidation merger we adjust the ξ of the remaining product of the merging firm, such that it will retain the same share of unaffiliated customers if it were to keep the same pre-merger price. Then, given the adjustment, ξ_a, we allow the new single brand firm to price optimally. We therefore make the following adjustment:

$$\xi_a = ln(\frac{s_1 + s_2}{s_0}) - \alpha p \tag{28}$$

Here, $s_1 + s_2$ is the combined pre-merger shares of products 1 and 2, s_0 is the share of the outside good, and p is the symmetric price of products 1 and 2.

B.2   Joint Pricing and Brand Consolidation Mergers in a Symmetric Static Logit Model

We now prove that in a static logit model, a joint pricing and brand consolidation merger (as defined in the previous subsection) will produce the same price effect if the merging firms are symmetric.

Lemma 1: Suppose the following is true,

  i. Demand is characterized by standard logit

  ii. While holding all else equal, at a price $\bar{p}$, splitting firm l into two firms yields equal shares for those two firms: $s_l(\bar{p}) = 2s_j(\bar{p})$

It follows that $s_l(p) = 2s_j(p) \;\; \forall p$.

Proof: By construction:

$$s_l = \frac{e^{\delta_l - \alpha\bar{p}}}{1 + e^{\delta_l - \alpha\bar{p}} + \sum_g e^{\delta_g - \alpha p_g}} = 2\frac{e^{\delta_j - \alpha\bar{p}}}{1 + 2e^{\delta_j - \alpha\bar{p}} + \sum_g e^{\delta_g - \alpha p_g}} = 2s_j \tag{29}$$

Equation (1) is true if and only if $e^{\delta_l - \alpha\bar{p}} = 2e^{\delta_j - \alpha\bar{p}}$. Dividing both sides by $e^{\delta_j - \alpha\bar{p}}$ and then taking logs, we have $\delta_l - \alpha\bar{p} - \delta_j + \alpha\bar{p} = \log(2)$. Therefore, $\delta_l - \delta_j = \log(2)$. It follows that if (i) and (ii) are true then $s_l(p) = 2s_j(p) \;\; \forall p$.

Lemma 2: Suppose demand is characterized by logit. Suppose a single product firm has a marginal cost, c, a market share, s_l, and a profit maximizing price p^*. Then a two product firm with marginal cost, c, a market share of $s_j(p) = \frac{s_l(p)}{2}$, will set the same profit-maximizing price, p^*.

Proof: With logit demand, the first-order condition of one product for an N product firm is:

$$\frac{d\Pi}{dp_1} = 1 + \alpha(p_1 - c_1)(s_1 - 1) + \alpha \sum_{k=2}^{N}(p_k - c_k)s_k = 0 \tag{30}$$

Now, suppose all of the firm's products have the same marginal cost, c, and that all of it's products are symmetric, $\delta_n = \delta$ for all n. Then, equation (2) simplifies to:

$$\frac{d\Pi}{dp_n} = 1 + \alpha(p - c)(s - 1) + \alpha(N - 1)(p - c)s = 0 \tag{31}$$

Solving equation (3) for p yields the symmetric, profit-maximizing price for each product:

$$p^* = \frac{1}{\alpha}\left[\frac{1}{1 - Ns}\right] + c \tag{32}$$

Now, suppose a single product firm sets a profit-maximizing price of p^* and therefore has a market share s_l. Also, hold the number and characteristics of all other firms in the market constant. Now suppose we replace the single product firm with a two product firm with marginal cost, c and product characteristics, δ_j, such that $s_j(p^*) = \frac{s_l(p^*)}{2}$. By equation (4), the two product firm will set the same profit maximizng price, p^*. We therefore prove Lemma 2.

Lemma 1 and Lemma 2 prove the following proposition.

Proposition: Let demand be characterized by logit. Consider the following two mergers of symmetric single product firms with marginal cost, c:

i Post-merger the firm retains both products and prices them to jointly maximize post-merger profits.

ii Post-merger the firm removes one product from the market, say product 2. At the pre-merger price, the post-merger share of the first product would have the same market share as the sum of the two single product pre-merger firms. Post-merger the firm sets a profit-maximizing price for the one product it keeps in the market

The post-merger price for the two products in Merger (i) is the same as the price for the one remaining product in Merger (ii).

C   Reduced-Form Evidence: Supplemental Results

C.1   Cost Pass-through: Identifying Expected and Unexpected Costs

We now analyze gas stations' dynamic reactions to expected and unexpected costs. To disentangle the reaction to anticipated and unanticipated cost changes, we leverage data on wholesale gasoline futures traded on the New York Mercantile Stock Exchange (NYMEX). The presence of a futures market allows us to project expectations of future wholesale costs for the firms in our market.

To make these projections, we assume that firms are engaging in regression-like predictions of future wholesale costs, and we choose the 30-day ahead cost as our benchmark.[43] Using station-specific wholesale costs, we regress the 30-day lead wholesale cost on the current wholesale cost and the 30-day ahead future. In particular, we estimate the following equation.

$$c_{it+30} = \alpha_1 c_{it} + \alpha_2 F_t^{30} + \gamma_i + \epsilon_{it} \tag{33}$$

Here, c_{it+30} is the 30-day-ahead wholesale cost for firm i, F_t^{30} is the 30-day ahead forward contract price at date t, and γ_i is a station fixed effect. We use the estimated parameters to construct expected 30-day ahead costs for all firms: $\hat{c}_{it+30} = \hat{\alpha}_1 c_{it} + \hat{\alpha}_2 F_t^{30} + \hat{\gamma}_i$. The unexpected cost, or cost shock, is the residual: $\tilde{c}_{it+30} = c_{it+30} - \hat{c}_{it+30}$.

For robustness, we construct a number of alternative estimates of expected costs, including a specification that makes use of all four available futures. However, we found that these alternative specifications were subject to overfit; the estimates performed substantially worse out-of-sample when we ran the regression on a subset of the data. Our chosen specification is remarkably stable, with a mean absolute difference of one percent when we use only the first half of the panel to estimate the model. Expected costs constitute 74.6 percent of the variation in costs (R^2) in our two-year sample, which includes a large decline in wholesale costs due to several supply shocks in 2014.

In subsection 3.3.2, we consider only the simple cut between unexpected and expected costs to focus attention on this previously unexplored dimension of pass-through. In retail gasoline markets, costs are highly correlated, with common costs tending to dominate idiosyncratic costs at moderate frequencies. For robustness, we have estimated the cost pass-through (i) using only common costs and (ii) controlling for the mean cost of rival brands (in the same county). In either scenario, we find estimates that are very similar.

---

[43]Futures are specified in terms of first-of-the-month delivery dates. To convert these to 30-day ahead prices, we use the average between the two futures, weighted by the relative number of days to the delivery date.

A Note on 30-Day Ahead Expectations

One of the challenges in discussing expectations is that they change each day with new information. News about a cost shock 30 days from now may arrive anytime within the next 30 days, if it has not arrived already. Therefore, any discussion of an "unexpected" cost shock must always be qualified with an "as of when." Given previous findings in the gasoline literature indicating that prices take approximately four weeks to adjust, a 30-day ahead window seems an appropriate one to capture most of any anticipatory pricing behavior. Additionally, our findings support this window as being reasonable in this context. We see no relationship between unexpected costs or expected costs and the price 30 days prior.[44]

---

[44]We interpret slight deviations from a zero as arising from an underlying correlation in unobserved cost shocks.

D Empirical Application: Supplemental Tables and Figures

D.1 Summary Statistics by Observation

Table 15: Retail Gasoline in Kentucky and Virginia: Oct 2013 - Sep 2015

Statistic	Mean	St. Dev.	Min	Pctl(25)	Pctl(75)	Max	N
Share	0.137	0.103	0.0003	0.060	0.183	0.688	110,844
Price	2.871	0.529	1.715	2.384	3.311	4.085	110,844
Wholesale Price	2.257	0.527	1.245	1.754	2.673	3.545	110,844
Wholesale FE	2.261	0.031	2.206	2.230	2.293	2.366	110,844
Margin	0.206	0.114	−0.440	0.132	0.272	1.048	110,844
Num. Stations	5.050	6.780	1	2	6	79	110,844

Notes: Table provides summary statistics for the observation-level data in the analysis. The greatest number of stations a brand has in a single county in our data is 79. The 25th percentile is 2, and we have several observations of a brand with only a single station in a market. The variable Wholesale FE is the average wholesale price for a brand within a county. We interact this variable with the US oil production data to generate an instrument for price in the demand estimation. Negative price-cost margins occur in 2.7 percent of observations.

D.2 Summary Statistics by Brand

Table 16: Summary of Brands

	Brand	Cond. Share	Share	Num. Markets	Num. Stations	Margins
1	Marathon	0.18	0.10	134	5.30	0.21
2	Sheetz	0.18	0.03	37	1.70	0.17
3	Speedway	0.17	0.03	39	3.70	0.18
4	Wawa	0.16	0.01	22	3.20	0.12
5	Exxon	0.15	0.07	116	4.60	0.25
6	7-Eleven	0.13	0.02	41	6.80	0.18
7	Shell	0.13	0.09	163	4.20	0.22
8	FRINGE	0.13	0.12	233	8.70	0.19
9	Pilot	0.12	0.01	21	1.40	0.13
10	BP	0.12	0.06	124	3.30	0.21
11	Loves	0.11	0.01	15	1.00	0.20
12	Valero	0.11	0.03	59	3.50	0.21
13	Thorntons	0.11	0.00	9	5.90	0.14
14	Hucks	0.11	0.00	10	1.90	0.15
15	Sunoco	0.09	0.01	32	4.70	0.28
16	Citgo	0.07	0.01	34	4.00	0.25

Notes: Table provides summary statistics by brand. The FRINGE brand is a synthetic brand created by aggregating brands that do not appear in 10 or more of the 241 markets in our data. Additionally, if a brand does not make up more than 2 percent of the average shares within a market, or 10 percent of the shares for the periods in which it is present, we also designate the brand as a fringe participant for that market.

D.3 Shares, Prices, and Instrument

Figure 7: Shares and Prices

(a) Total Market Shares and Prices

(b) Instrument and Prices
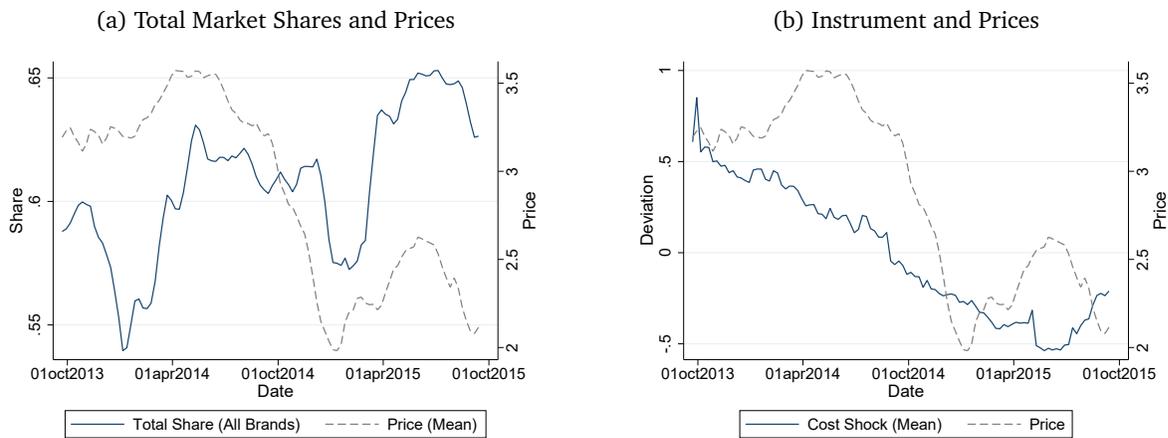
Notes: Panel (a) displays the average firm share plotted along with the average price over the period in our sample. Both lines indicate seasonality, with peaks occurring during the summer. Panel (b) plots the constructed instrument against the average price in our sample. Overall, there is a strong negative correlation between the instrument and average prices.