Consumer Inertia and Market Power

Alexander MacKay
Marc Remer

Working Paper 19-111

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 29, 2019

Abstract

We study the pricing decision of firms in the presence of consumer inertia. Inertia can arise from habit formation, brand loyalty, switching costs, or search, and it has important implications for the interpretation of equilibrium outcomes and counterfactual analysis. In particular, consumer inertia affects the scope of market power. We show that the effects of competition on prices and profits are non-monotonic in the degree of inertia. Further, a model that omits consumer inertia tends to overstate the marginal effect of competition on price, relative to a benchmark that accounts for consumer dynamics. 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 predicts price increases greater than the price increases predicted when accounting for dynamics.

JEL Codes: D12, D43, L13, L41, L81

Keywords: Consumer Inertia, Market Power, Dynamic Competition, Demand Estimation

*An earlier version of this paper was entitled "An Empirical Model of Consumer Affiliation and Dynamic Price Competition." We are grateful to Ali Hortaçsu, Nathan Miller, Karl Schurter, and Nathan Wilson for their comments. We thank participants at the IIOC conference and the EEA-ESEM conference, and also seminar participants at the University of Toronto, Haverford College, Swarthmore College, Harvard Business School, the University of Florida, Boston College, Bowdoin College, US Department of Justice, and NYU Stern for their feedback and suggestions.

[†]Harvard University, Harvard Business School. Email: amackay@hbs.edu.

[‡]Swarthmore College, Department of Economics. Email: mremer1@swarthmore.edu.

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 this paper, we develop an empirical model of consumer inertia that can be estimated using aggregate, market-level data. Our specific model captures a product-specific "affiliation" for the most recently purchased product and nests typical implementations of dynamic consumer behavior. 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). We consider the implications of this model on the pricing behavior of firms and, in particular, on the potential for market power.

In response to consumer affiliation, profit-maximizing firms will internalize the effect of their current price on the distribution of affiliated consumers 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. Our analysis allows us to measure these two forces using aggregate data and study the interaction between the two. Our simulations show that horizontal market power increases at low rates of inertia and decreases at high rates of inertia, generating a non-monotonic relationship between dynamic and horizontal market power.

Given these effects, it is natural to wonder whether (static) demand models that are typically employed in empirical work provide accurate counterfactual approximations to a world with consumer state dependence. We explore this question in the context of horizontal mergers, which allows us to examine the effect of competition on prices. Antitrust authorities will challenge a merger if the merging firms are expected to increase the price by a significant amount. The ability to accurately predict the price effect of a merger hinges on an appropriate representation of the firms' pricing incentives. In the presence of consumer affiliation, a first-order

[1] See the related literature section for examples.

determinant of price is its effect on future demand. A static model, which omits this incentive, will incorrectly attribute this effect to consumer elasticities and the degree of competition. As we demonstrate, static models consistently over-predict the price effects of a merger, compared to an analysis that accounts for the dynamic incentive to invest in future demand. Thus, failing to account for consumer inertia may overstate the potential for horizontal market power and affect merger enforcement.

Antitrust 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 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 current 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. In our empirical application, we find that a static demand model generates larger counterfactual price increases from a horizontal merger than does a model that incorporates consumer inertia.

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

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

[3] Ibid.

[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/

To provide intuition about the model, we consider price competition in an oligopoly setting using numerical simulations. We demonstrate that affiliation can have a large effect on steady-state prices. In equilibrium, affiliation can raise oligopoly prices higher than the monopoly price with no affiliation. Second, we show that the percentage price effect of a merger may decrease with the rate of affiliation, implying that the usual intuition about effects of a merger may be misleading. If consumers are sufficiently price insensitive, then affiliation allows firms to capture a large portion of the monopoly rents pre-merger, resulting in a smaller increase in post-merger prices.

Next, we consider the empirical implications of failing to account for dynamic demand in a merger analysis. In this exercise, we calibrate a static demand model to simulated data from the dynamic model and perform a merger simulation. Compared to the true impact of the merger, the (misspecified) static model systematically over-predicts merger price effects. In the dynamic model, the incentive to invest in future demand pushes prices down, and this effect remains after the merger. Due to the suppressed margins pre-merger, the static model falsely considers the products less differentiated; when competition is reduced post-merger, the static model then generates higher prices.

The data needed to estimate the empirical model are commonly used in the demand estimation literature, namely: prices, shares, and an instrument for prices. Even with only firm-level shares (aggregated across all consumer types), our model allows us to separately identify each firm's share from affiliated and unaffiliated consumers, as well as the probability that a consumer becomes affiliated after purchasing. 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.

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 60 percent of consumers become affiliated to the brand from which they previously purchased on a week-to-week basis, and therefore effectively do not consider competitors.[6] Affiliated consumers display a much lower price sensitivity. Unaffiliated consumers have an average elasticity of -7.7, whereas affiliated consumers have a near-zero average elasticity. Though unaffiliated consumers 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 -3.1.

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 relati-

[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

vely 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 perform a merger analysis between two major gasoline retailers and re-compute the price-setting equilibrium in each period. With the dynamic model, we estimate that prices would increase by 3.3 percent post-merger. A static model, on the other hand, predicts an average price increase of 5.9 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.

Prior to estimating the model, we present reduced-form evidence of dynamic demand and dynamic pricing in retail gasoline markets.[7] 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).[8] 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 inattention 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,

[7] 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."

[8] 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).

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.[9] 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).

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

[9]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).

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.[10] We use the model to numerically 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). For the empirical model, we make the additional assumption that utility follows the random-coefficient logit setup. Consumers are indexed by discrete types $i \in I$.[11] We allow the distribution of types to change endogenously over time.

A consumer n of type i receives the following utility for choosing product j:

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

Consumers receive an additively-separable common component δ_{jt}, a type-specific shock $\sigma_{jt}(i)$, and an idiosyncratic shock, $\epsilon_{ijt}^{(n)}$. 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 type-specific shock, $\sigma_{jt}(i)$, 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, as is the case in our empirical application.

[10]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.

[11]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).

We denote the probability that a consumer of type i chooses product j as $s_{jt}(i)$. We normalize the utility of the outside good to be zero. Additionally, we define the type 0 consumer to be "unaffiliated," and we assume that this type has 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 conditional on type are:

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

$$s_{jt}(i) = \frac{\exp(\delta_{jt} + \sigma_{jt}(i))}{1 + \sum_{k/i} \exp(\delta_{kt} + \sigma_{jt}(i))}. \tag{3}$$

The observed share for firm j is given by the weighted average of choice probabilities by types: $S_{jt} = \sum_{i=0}^{I} r_{it} s_{jt}(i)$, where r_{it} is the fraction of consumers of type i.

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 type-specific demand shocks, $\sigma_{jt}(i)$. We assume that each consumer type has an affiliation (utility shock) to a single product. Further, we assume that there is a single type corresponding to each product. Thus, a consumer of type j is affiliated to product j. The type-specific demand shocks for other products are zero, i.e. $\sigma_{jt}(i) = 0 \,\forall\, i \neq j$.

Thus, we define *affiliation* to be a product-specific state dependence in preferences. We say that a consumer of type j is affiliated with product j, as this consumer has a perceived benefit of $\sigma_{jt}(j)$ relative to unaffiliated (type 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, $i \neq j$. 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 Types Types are not fixed for a consumer, but they may depend dynamically on previous behavior. We assume that the evolution of types follows a Markov process, where the state can be expressed as a function of the joint distribution of types and choices in the previous period.

Given the previous assumption that there is a one-to-one mapping between types and an affiliation for each product, we implicitly assume that consumers are symmetric within a market. Thus, we can express the distribution of consumer types in any period as a function of the distribution of choices in the previous period. For example, the share of consumer type j, r_{jt}, might be expressed as $r_{jt} = f(S_{j(t-1)})$, where $S_{j(t-1)}$ is the aggregate share of consumers (across all types) that chose product j in the previous period.

We make the assumption that consumers that purchase product j become affiliated with product j in the next period at a given rate: $r_{jt} = \lambda_{jt} S_{j(t-1)}$. Under the habit-formation interpretation, a (random) fraction λ_{jt} of consumers that purchased j in the previous period are habituated to product j. We allow this fraction to vary with observable characteristics of the firm and market. The remaining $(1 - \lambda_{jt})S_{j(t-1)}$ consumers transition to type 0, along with any consumers that chose the outside good.

2.2 Supply

We assume that firms set prices to maximize the net present value of profits. In contrast to the typical setup for a dynamic game, we make relatively weak assumptions about the perceived continuation value, which depends on expectations and discount rates. We restrict attention to Markov perfect equilibria.

First, we define firm j's aggregate share as:

$$S_{jt} = (1 - \sum_{i \neq 0} r_{it}) s_{jt}(0) + \sum_{i \neq 0} r_{it} s_{jt}(i). \quad (4)$$

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, $s_{jt}(i)$. Note that firm j will make sales to consumers affiliated to other firms $j \neq i$, 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 $\{\sigma_{jt}(i)\}$ 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 consumer types, r_t, and other variables that are captured by the vector, x_t, such as expectations about future costs. Entry is exogenous.[12] The firm's objective function can be summarized by the Bellman equation:

$$V_j(c_t, r_t, x_t) = \max_{p_{jt}|p_{kt,k\neq j}} \big\{ (p_{jt} - c_{jt}) S_{jt} + \beta E(V_j(c_{t+1}, r_{t+1}, x_{t+1}) | p_t, c_t, r_t, x_t) \big\}. \quad (5)$$

Prices in each period optimize the sum of the continuation value and current-period profits, $(p_{jt} - c_{jt}) S_{jt}$. Both of these components depend upon marginal costs and the distribution of consumer types, r_t. Thus, firms anticipate both future marginal costs and how price affects the future distribution of consumer types. 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 We place minimal restrictions on the expectations and discount factors for each firm. Consistent with the Markov perfect framework, we make the relatively weak assumption that the continuation value function is stable conditional on the state and prices.

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

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

analyze steady-state prices in an oligopoly game[13] with Bertrand price-setting behavior. For the numerical exercise, we specify the utility parameters from equation (1) as $\delta_{jt} = \xi_j + \alpha p_{jt}$ and $\sigma_{jt}(j) = \bar{\xi} + \overline{\alpha} p_{jt}$, allowing for the affiliation shock to affect the consumer's price sensitivity in addition to the utility level. Thus, the utility of a consumer n of type i for product j is as follows:

$$u_{jt}^{(n)}(i) = \xi_j + \alpha p_{jt} + \mathbb{1}[i = j](\bar{\xi} + \overline{\alpha} p_{jt}) + \epsilon_{jt}^{(n)}. \tag{6}$$

Here, $\mathbb{1}[i = j]$ takes a value of one if a consumer of type i purchases from firm $j = i$, and zero otherwise. For unaffiliated consumers, the indicator function equals zero for all j. Assuming the error term, $\epsilon_{jt}^{(n)}$, follows the type-1 extreme value distribution, we get the market share specified in equation (2) in the previous section, which reduces to the standard logit model if $\bar{\xi} = \overline{\alpha} = 0$.

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:

$$V_i(r) = \max_{p_i | p_{-i}} \pi_i(p, r) + \beta V_i(r'). \tag{7}$$

Here, p and r are vectors of prices and affiliated customers, respectively, and $r' = f(p, r) = \lambda \cdot s(p, r)$ is a vector specifying each firm's affiliated customers in the next period. 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.[14] Firm i's profit-maximizing first-order condition is then:

$$\sum_{l \in J_i} \frac{\partial \pi_l}{\partial p_j} + \beta \frac{dV_i(r')}{dr'} \frac{dr'}{dp_j} \; \forall j \in J_i = 0. \tag{8}$$

Next, we specify the derivatives of equation (7) 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:

$$\underbrace{\frac{dV_i(r)}{dr}}_{J \times 1} = \underbrace{\left[\frac{\partial \pi_i}{\partial p} \frac{dp}{dr} + \frac{\partial \pi_i}{\partial r} \right]}_{J \times 1} \underbrace{\left[I - \beta f_p(p, r) \frac{dp}{dr} - \beta f_r(p, r) \right]^{-1}}_{J \times J}. \tag{9}$$

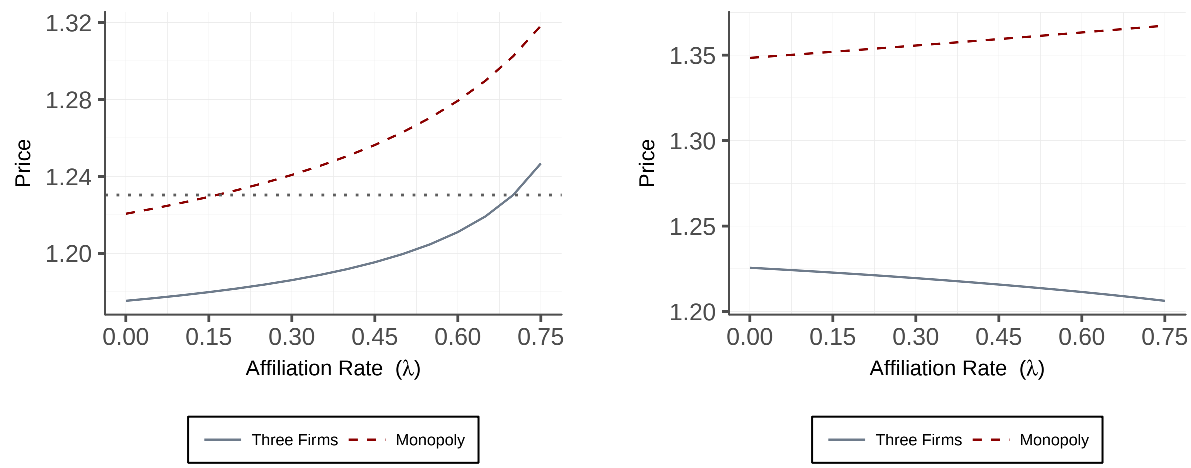
In this equation, $\frac{\partial \pi_i}{\partial p}$, $\frac{\partial \pi_i}{\partial r}$, $f_p(p, r)$, $f_r(p, r)$ are known, conditional on values of p and r. The remaining unknowns are $\frac{dp}{dr}$ and $\frac{dV_i(r)}{dr}$. To solve the model, we impose the steady-state

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

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

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: $\xi = 6$, $\bar{\xi} = 5$, $\alpha = -6.5$, and $\overline{\alpha} = 5.85$. Panel (b) is generated using $\xi = 6.6$, $\bar{\xi} = 0.9$, $\alpha = -5.7$, and $\overline{\alpha} = 0.6$. Marginal cost is set to one for both figures.

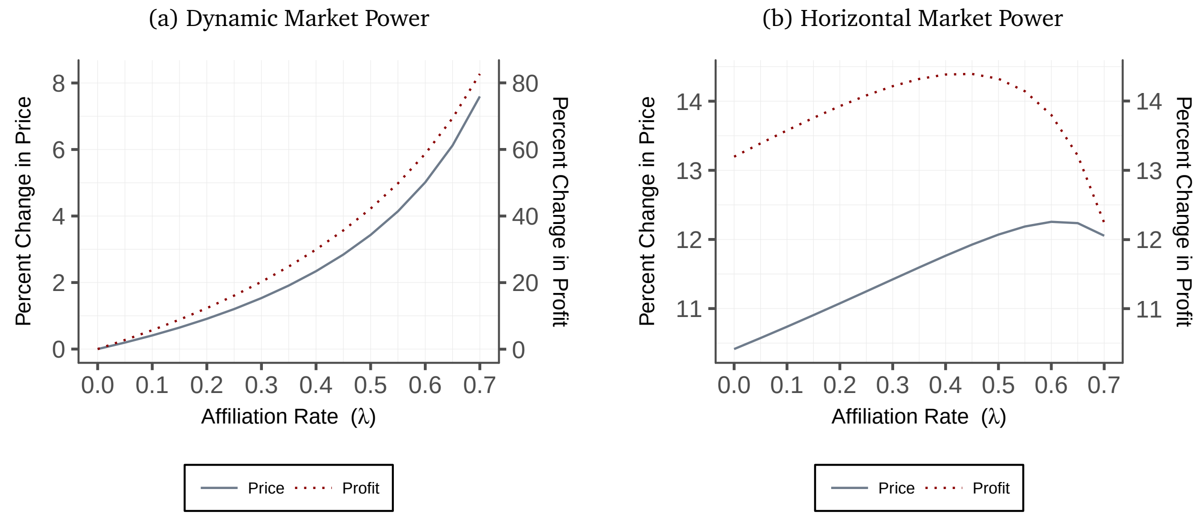
condition governing the evolution of affiliated customers, $r' = r$. The full set of steady state conditions, provided by equation (9) and $r' = r$, allow us to solve for steady-state prices and shares, conditional on the $J \times J$ derivative matrix, $\frac{dp}{dr}$. The values of $\frac{dp}{dr}$ 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 three-firm markets. In this setting, symmetry is imposed by restricting the utility parameters, ξ_i, $\bar{\xi}$, and $\overline{\alpha}$, 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. This figure highlights the potential importance of affiliation. In the three-firm market, when consumers have high rates of affiliation ($\lambda \geq 0.70$), the equilibrium prices are higher than the monopoly price for static demand $(\lambda = 0)$. Thus, intuition about the ability of competition to discipline prices should be paired with an understanding of consumer dynamics, as affiliation can offset or even overwhelm the horizontal effects of competition.

Panel (b) shows that the interaction of competition and affiliation can have surprising qualitative results. For a different baseline set of demand parameters from those of panel (a), the oligopoly price is decreasing in the rate of affiliation, even though the monopoly price is increasing. Dynamic market power, defined by the price increase relative to the price with no inertia,

Figure 2: The Scope of Market Power



Notes: Panel (a) displays the mean percent price increase and mean percent profit increase for a three-firm oligopoly above the baseline model with no dynamics in consumption ($\lambda = 0$). Note that the profits are scaled by a factor of 10, which is labeled on the right axis. Panel (b) displays the marginal effects, in terms of mean percent price increase and mean percent profit increase, of a merger to monopoly for different values of λ. The plots reflect 275 baseline parameter values of $(\xi, \bar{\xi}, \alpha, \overline{\alpha})$ that converged for all $\lambda \in [0, 0.7]$, or 4,125 markets in total.

may be negative. Thus, the relationship between equilibrium prices and dynamics may depend upon market structure.

2.3.1 The Scope of Market Power

To summarize 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 oligopoly price with consumer affiliation to a baseline price where consumers have no state dependence. To measure horizontal market power, we compare the monopoly price to the oligopoly price, conditional on the demand parameters.[15]

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 $\xi \in [-3, 17]$, $\bar{\xi} \in [0, 20]$, $\alpha \in [-15, 0]$, $\overline{\alpha} \in [0, 15]$, $\alpha + \overline{\alpha} < 0$, and set each firm's marginal cost to one. For each draw of these demand parameters, we construct three-firm markets for

[15]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.

$\lambda \in \{0, 0.05, 0.10, ..., 0.95\}$. Finally, we 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 0.10 and 0.75. The data generating process yields 8,025 markets whose parameters are summarized in Appendix Table 10. 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.

In Figure 2, we plot the effects of affiliation and reduced competition on prices and profits. The plots employ simulation results from the 275 baseline parameter values of $(\xi, \bar{\xi}, \alpha, \overline{\alpha})$ that converged for all $\lambda \in [0, 0.7]$, or 4,125 markets out of the 8,025 that meet the above qualifications. Panel (a) shows that, on average, prices and profits increase with the rate at which consumers become affiliated. This demonstrates that in the steady state of the model, the "harvesting" incentive tends to dominate the "investment" incentive. Perhaps surprisingly, the effect on profits is roughly ten times the effect on prices (note the right axis). This highlights the importance of the investment incentive that firms face in this setting. As the affiliation rate increases, there is a modest increase in price but a large increase in the share of affiliated customers in the following period, generating correspondingly large increases in profits. In total, dynamic market power tends to increase with consumer inertia, but it is dampened by the incentive to invest in future demand.

In panel (b), we plot the effect on prices and profits that a monopolist realizes, relative to the "competitive" three-firm oligopoly, as the rate of consumer affiliation increases. These effects represent the scope for horizontal market power, either explicitly via merger or potentially through collusion. The gain in profits from moving to the monopoly outcome is non-monotonic on average and decreases with λ for $\lambda > 0.45$. Correspondingly, we also find that the price increases are non-monotonic. The non-monotonicity is due to dynamic market power decreasing the potential for horizontal market power as consumer inertia becomes more prevalent. As the rate of affiliation increases, consumers are less likely to view oligopoly firms as substitutes. If the strength of affiliation is sufficiently strong then firms can extract the majority of the consumer surplus from their own customers, which limits the scope of horizontal market power.

2.3.2 Consumer Inertia and Mergers

To further demonstrate the effect of consumer inertia on price competition, we examine the impact on prices of a horizontal merger. We proceed by simulating symmetric three-firm markets as a baseline and allowing two of the firms to merge. The merged firm maintains both pre-existing firms as separate entities but maximizes their joint profits. Thus, we examine how dynamic incentives affect steady-state prices and the unilateral price effects arising from a merger. 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

Table 1: Simulation Summary Statistics

	Mean	St. Dev.	Min	Pctl(25)	Pctl(75)	Max	N
Pre-Merger Price	1.36	0.39	1.11	1.15	1.39	3.99	8025
Pre-Merger Margin	0.23	0.14	0.10	0.13	0.28	0.75	8025
Pre-Merger Market Share	0.19	0.08	0.05	0.12	0.26	0.30	8025
HHI: Pre-Merger	1287.33	848.69	75.15	460.21	2078.63	2699.35	8025
Δ HHI	858.22	565.79	50.10	306.80	1385.75	1799.57	8025
Merger Price Δ	4.54	3.73	0.34	2.11	5.64	25.15	8025
Non-Merging Price Δ	0.52	0.78	-0.19	0.03	0.73	6.55	8025
Prediction Bias	0.39	0.68	-2.47	0.04	0.48	7.73	8025
Prediction Bias (pctg.)	9.47	15.99	-18.52	1.38	12.33	330.73	8025
Dynamic Elasticity: Unaffiliated	-7.02	3.28	-17.26	-9.45	-4.32	-1.37	8025
Dynamic Elasticty: Affiliated	-3.60	2.78	-13.45	-5.27	-1.34	-0.00	8025
Dynamic Elasticty: Weighted	-5.37	2.62	-13.65	-7.31	-3.20	-0.77	8025
Static Elasticity	-5.67	2.42	-10.00	-7.72	-3.60	-1.33	8025

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 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.

presence of consumer affiliation tends to overestimate the price effects of a merger.

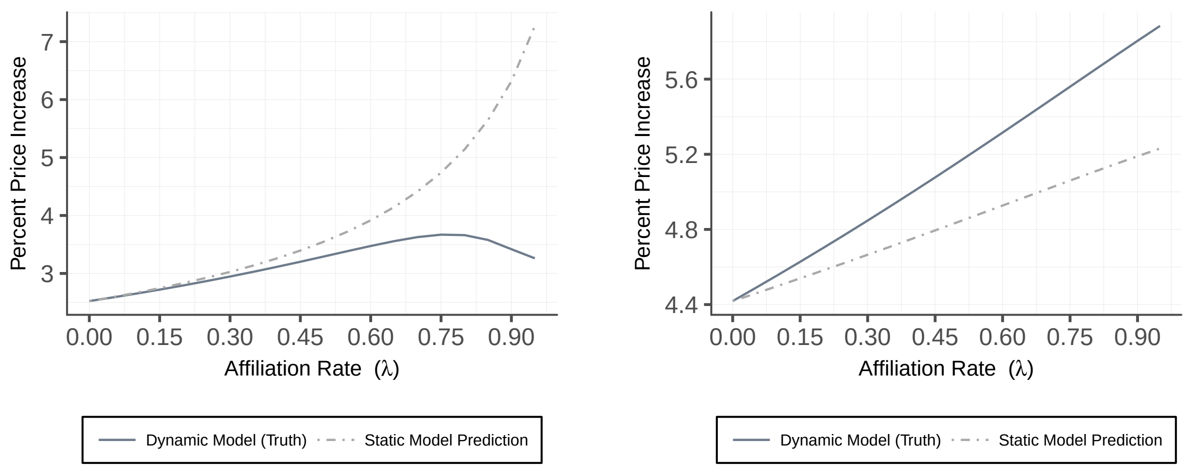
For each parameterized market, we solve for the steady-state prices. Then, we simulate a merger between firms 1 and 2, leaving firm 3 as the non-merging firm. Table 1 summarizes the pre-merger equilibrium and the unilateral price increases of merging and non-merging firms. The "average" market is one that, a priori, would typically raise moderate concern from the US antitrust agencies; HHI (1287) falls in the "unconcentrated" range, but the change in HHI (858) generally warrants a thorough investigation. The average pre-merger difference between price and cost is 0.36, and the mean market share is 0.19. Using these mean values to calculate the "Upward Pricing Pressure" index, $\frac{0.19}{1-0.19} \cdot (1.36 - 1) = .084$, yields a value in the midpoint of 0.05-0.10, which are heuristic thresholds sometimes used to trigger an investigation.[16] The average percentage merger price effect, 4.5 percent, would typically 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.

On average, increasing the rate of affiliation (λ) and the strength of the affiliation ($\overline{\alpha}$ and $\bar{\xi}$) tend to increase both pre-merger prices and the price effects arising from a merger. Table 11 in the Appendix provides results from regressions of these outcomes on the demand parameters. However, these relationships do not hold in every instance, and may interact in interesting ways. As shown in Figure 1, prices may increase or decrease in the rate of affiliation. In a small number of markets, we find that the relationship between λ and prices is non-monotonic.

[16] 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: Simulated Merger Price Increases

(a) Static Model Biased Upward (b) Static Model Biased Downward



Notes: Panel (a) is generated using the following parameter values: $\xi = 9.1$, $\bar{\xi} = 4.8$, $\alpha = -8.6$, and $\overline{\alpha} = 0.9$. Panel (b) is generated using $\xi = 6.6$, $\bar{\xi} = 0.9$, $\alpha = -5.7$, and $\overline{\alpha} = 0.6$. Marginal cost is set to one for both figures.

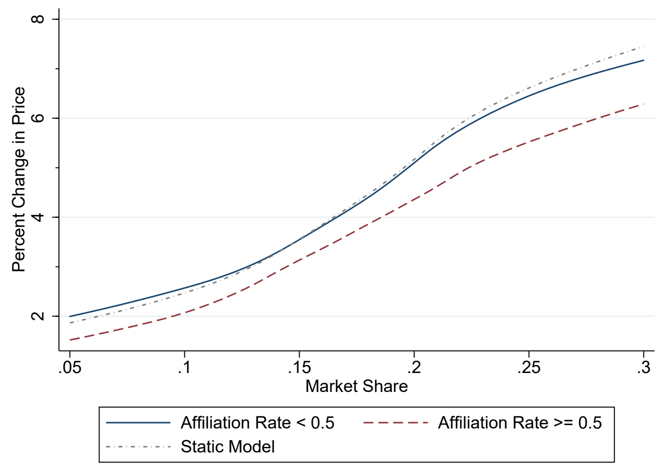
2.3.3 Model Misspecification in Merger Simulation

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.

Table 1 summarizes the prediction bias, defined as the static prediction minus the "true" affiliation price effect. The average prediction bias is 0.39 percentage points; dividing by the magnitude of the price change, this represents a 9.47 percent increase over the true price effect. While the average prediction bias is moderate, there is a fair amount of dispersion in the size of the bias; the 25th and 75th percentile of the bias, as a percentage of the price increase, is 1.38 percent and 12.3 percent, respectively. Usually, the static model over-predicts the true price

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. Markets where $\lambda < 0.5$ and $\lambda >= 0.5$ are separately estimated. All markets are used for the static model.

change, but in 5.48 percent of simulations it predicts a smaller effect.

Figure 3 presents the results of the experiment for two individual markets. In panel (a), the "true" dynamic model and the misspecified static model generate similar price effects for affiliation rates (λ) less than 0.5. As λ increases above 0.5, however, the dynamic model generates lower price effects, as the post-merger investment effect increases relative to the pre-merger market structure. The static model, however, cannot account for this incentive, resulting in increasingly biased price predictions. By contrast, panel (b) demonstrates a parameterization where the static model consistently under-predicts the dynamic model. Moreover, for values of λ between 0.4 and 0.65, the true model yields price effects greater than 5 percent, while the biased prediction from the static model is below that threshold. Taking 5 percent as a hypothetical threshold for enforcement by antitrust agencies, panels (a) and (b) demonstrate that not accounting for affiliation can lead to errors in enforcement action in either direction.

On average, the static model tends to overstate the unilateral price effects of a merger. This is highlighted in Figure 4, which relates the merger price effect to the pre-merger market share of each symmetric firm. In line with intuition, the price effect of a merger is increasing with pre-merger market shares. To generate the graph, we run a local polynomial regression of the merger price effect on the firm's market share; we generate regression lines for (i) the affiliation model with $\lambda < 0.5$, (ii) the affiliation model with $\lambda >= 0.5$, and (iii) the static model. The dynamic model and the static model yield similar predictions for low values of the affiliation rate. For higher affiliation rates, however, 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. The figure also highlights that for the dynamic model, conditional on observed market share, a greater affiliation rate suppresses merger price effects.

Interestingly, the static model is calibrated to be more elastic, on average, than the share-weighted elasticity in the dynamic model (-5.67 vs. -5.37). In the 60 percent of markets where the static model is calibrated to be more elastic, it still over-predicts the dynamic price effect. 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.

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 that the price predictions from static merger simulations should be revised downward for greater degrees of consumer affiliation. We provide more detailed results in support of this conclusion in Appendix A.5.

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),

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

Notes: Regression estimates indicate the correlation structure of shares over time.
Standard errors in parentheses.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

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.[17] 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.

[17] 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.

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 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.[18]

3.3 Dynamic Pricing

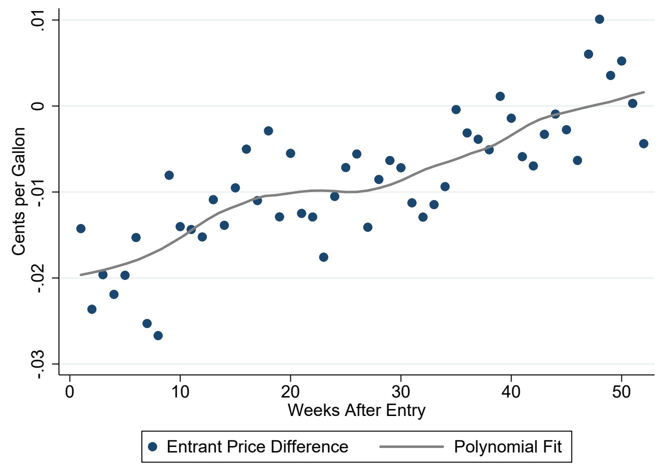
We now present reduced-form evidence of dynamic pricing. Consistent with a model where firms accumulate affiliated consumers over time, we find that new entrants price lower relative to established competitors in the same market, and that this discount dissipates over time. Second, we examine cost pass-through 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

[18]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.

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.

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.

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 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.[19] 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:

$$p_{it} = \sum_{s=-50}^{50} \beta_s \hat{c}_{it-s} + \sum_{s=-50}^{50} \gamma_s \tilde{c}_{it-s} + \sum_{s=-50}^{50} \phi_s \tau_{it-s} + \psi_i + \varepsilon_{it}. \tag{10}$$

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.[20] 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 (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, reaching a peak of 0.71 after 34 days.

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 is reached 21 days after the shock. 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 a third of the effect of the expected cost shock the day before it arrives.[21]

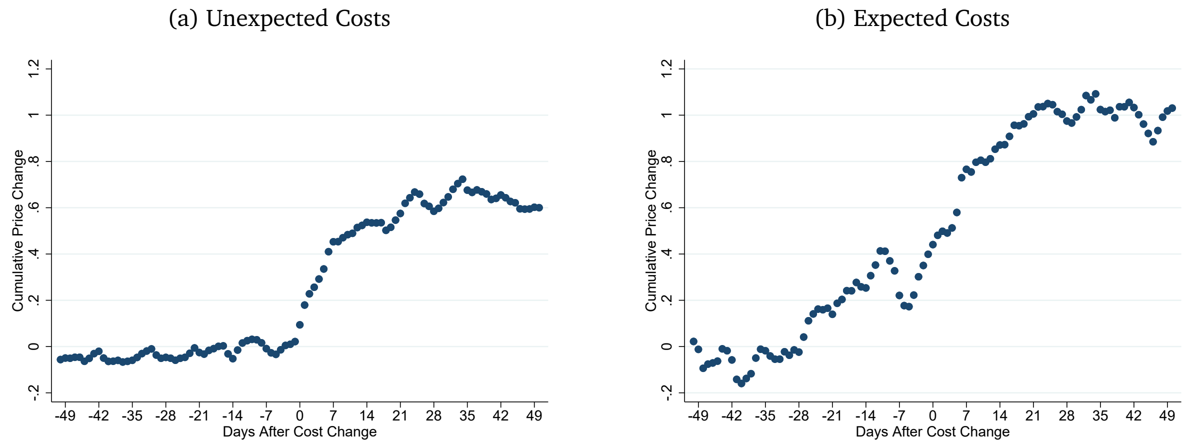
Thus, a reduced-form analysis of pricing behavior shows that prices retail gasoline stations

[19] For details, see Section B.1 in the Appendix.

[20] 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.

[21] 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 66 cents for each dollar increase in cost.

Figure 6: Cumulative Pass-through



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 (10).

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.

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 $\{s_{jt}(i)\}$) conditional on observed shares S_{jt} and the dynamic parameters governing λ_{jt} and $\sigma_{jt}(i)$. Second, the vector $\{s_{jt}(i)\}$ 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 $\{s_{jt}(i)\}$ and depends on the distribution of affiliated consumers for each product $\{r_{jt}\}$. Observed shares are determined by the following:

$$S_{jt} = (1 - \sum_k r_{kt}) \cdot s_{jt}(0) + \sum_i r_{it} \cdot s_{jt}(i). \tag{11}$$

To separate out $s_{jt}(i)$ 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):

$$\ln s_{jt}(0) - \ln s_{0t}(0) = \delta_{jt} \tag{12}$$

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

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

To show identification, we use equations (12) and (13) to obtain the following expressions:

$$s_{jt}(0) = \left(\frac{s_{0t}(0)}{s_{0t}(j)} - 1\right) \frac{1}{\exp\left(\sigma_{jt}(j)\right) - 1} \tag{14}$$

$$s_{jt}(i) = s_{0t}(i) \frac{s_{jt}(0)}{s_{0t}(0)} \cdot \exp\left(\sigma_{jt}(i)\right). \tag{15}$$

That is, the $J + J^2$ unknowns $\{s_{jt}(i)\}$, 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 provides the other J restrictions:

$$S_{jt} = (1 - \sum_k r_{kt}) \cdot \left(\frac{s_{0t}(0)}{s_{0t}(j)} - 1 \right) \frac{1}{\exp\left(\sigma_{jt}(j)\right) - 1} + \frac{s_{jt}(0)}{s_{0t}(0)} \cdot \sum_i r_{it} \cdot s_{0t}(i) \exp\left(\sigma_{jt}(i)\right) \tag{16}$$

The observed share equation requires $\{r_{jt}\}$, which is calculated by $r_{jt} = \lambda_{jt} S_{j(t-1)}$. Therefore, the unobserved type-specific shares $\{s_{jt}(i)\}$ are identified conditional on λ_{jt} and $\{\sigma_{jt}(i)\}$, i.e., the unobserved heterogeneity parameters.

Identification of Static Demand Parameters

We now allow for the observation of multiple markets, which are denoted with the subscript m. From equation (12), 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:

$$\delta_{jmt} = \alpha p_{jmt} + \pi \left(p_{jmt} \times Income_{jmt} \right) + X_{jmt} \gamma + \eta_{jmt}. \tag{17}$$

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 dynamic (non-linear) parameters. To identify the parameters that govern λ_{jt} 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. We assume that $Corr(\eta_{jmt}, \eta_{jm(t+1)}) = 0$ holds within each market, which provides us with more than sufficient moments (252) to identify our parameters.[22]

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

[22] 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. Using the first three lags of prices produces similar empirical results.

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_{jmt} and $\sigma_{jmt}(j)$ as follows:

$$\lambda_{jmt} = \frac{\exp(\theta_1 + \theta_2 Income_{jmt} + \theta_3 Density_{jmt})}{1 + \exp(\theta_1 + \theta_2 Income_{jmt} + \theta_3 Density_{jmt})} \tag{18}$$

$$\sigma_{jmt}(j) = \bar{\xi} + \overline{\alpha} p_{jmt}. \tag{19}$$

Thus, we allow the rate of affiliation to depend on median household income and (log) population density, which we take a proxy for lifestyle.[23] Affiliated customers may receive a level shock to utility $\bar{\xi}$, and they may become less (or more) price sensitive for the affiliated brand. This change in price sensitivity is captured by $\overline{\alpha}$. This specification allows for the possibility that consumer characteristics, as captured by income and population density, affects the propensity for consumer inertia.

Separate identification of λ_{jmt} and $\sigma_{jmt}(j)$ is made possible by the structure of the model. λ_{jmt}, the rate at which consumers become affiliated, does not depend on price, whereas the impact of $\sigma_{jmt}(j)$ on shares does. As can be by examining equations (12) and (13), a change in price affects δ_{jmt}, which shifts the relative choice patterns, holding fixed $\sigma_{jmt}(j)$. Intuitively, this would be reflected in the data by how the serial correlation in shares varies with average prices in the market. The parameter $\overline{\alpha}$ allows the model to calibrate how responsive affiliated consumers are to price levels. 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.

[23] We construct these variables by taking the population weighted average across stations by using the local measures in the Census tract.

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 ($\sigma_{jt}(i) = 0 \, \forall \, i \neq j$), we obtain

$$\sum_i r_{it} \cdot s_{0t}(i) \exp\left(\sigma_{jt}(i)\right) = \sum_i r_{it} s_{0t}(i) + \left(\exp\left(\sigma_{jt}(j)\right) - 1\right) r_{jt} s_{0t}(j). \tag{20}$$

By substituting this expression into equation (16), we can obtain a quadratic equation for each of the $\{s_{0t}(j)\}$:

$$\begin{aligned} 0 =& s_{0t}(j)^2 \left[\exp\left(\sigma_{jt}(j)\right) - 1\right] \frac{r_{jt}}{s_{0t}(0)} \\ &+ s_{0t}(j) \left(\left[\exp\left(\sigma_{jt}(j)\right) - 1\right] \left(S_{jt} - r_{jt}\right) + \frac{1}{s_{0t}(0)} \sum_{0,i} r_{it} s_{0t}(i) \right) \\ &- \sum_{0,i} r_{it} s_{0t}(i). \end{aligned}$$

These allow us to solve for $\{s_{0t}(j)\}$ in each market as a quadratic function of dynamic parameters, observables, and the two market-level parameters: $s_{0t}(0)$ and $\sum_{0,i} r_{it} s_{0t}(i)$. 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 values for the non-linear parameters that govern λ_{jt} and $\sigma_{jt}(j)$.
2. Calculate $r_{jt} = \lambda_{jt} S_{jt-1}$ for all periods except the first.
3. In each market, solve for $s_{0t}(0)$ and $\sum_{0,i} r_{it} s_{0t}(i)$ using the non-linear system of equations obtained previously. Find $s_{jt}(0)$ for each firm.
4. Run the regression implied by equation (12) using the $\{s_{jt}(0)\}$ obtained in the previous step to solve for the linear parameters (α, π, γ). Calculate the correlation of the residuals $Corr(\hat{\eta}_{jt}, \hat{\eta}_{j(t+1)})$ within each market.
5. Repeat 1-4 to find the non-linear parameters that minimizes the sum of squared correlations.

The regression for equation (12) may involve instrumental variables and the use of panel data methods such as fixed effects. In our empirical application, we make use of both.

Table 3: Summary Statistics by County

Statistic	Mean	St. Dev.	Min	Pctl(25)	Pctl(75)	Max	N
Num. Brands	4.42	1.51	1	3	5	8	252
Price	2.87	0.11	2.69	2.77	2.95	3.14	252
Wholesale Price	2.25	0.05	2.12	2.21	2.28	2.44	252
Margin	0.21	0.06	0.06	0.17	0.24	0.44	252
Num. Stations	22.56	28.33	1.12	7.00	25.48	243.47	252

Notes: Summary statistics calculated at the market level.

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[24] only has to find two parameters, $s_{0t}(0)$ and $\sum_{0,i} r_{it} s_{0t}(i)$, 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 practice, λ_{jt} may not be parameterized in a way that is sufficiently flexible to match the data. For example, if affiliated consumers are highly inelastic, then we should expect $S_{jt} \geq r_{jt}$, as a firm can expect to (at least) retain all of its affiliated customers. If the data have a few instances where shares for a certain product fall precipitously from one week to the next ($S_{jt} \ll S_{j(t-1)}$), then the implied value of λ_{jt} is low and a parsimonious representation of λ_{jt} may not be able to capture these dynamics. One could consider adding a stochastic term to the affiliation rate or adding a measurement error component to shares to account for this variability. In our application, we impose an ad hoc assumption, adjusting r_{jt} in such instances to the implied value based on the realized share and the average ratio of r_{jt} to S_{jt} across other observations. We add a penalty to the objective function for observations that need this adjustment.

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 252 counties in KY and VA. There is cross-sectional variation in wholesale prices, margins, and the number of stations in each

[24] 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.

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 252 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. These steps reduce the number of observations from 194,275 down to 112,929. Additionally, 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.[25]

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.

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 demand using the typical logit regression. 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.[26]

Our regression equation takes the following form:

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

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,[27] 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.

[25] Summary statistics by brand are presented in Table 15 in the Appendix. The fringe brand is, on average, 14 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).

[26] In the estimated model, the implied correlation in shocks is -0.01.

[27] 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.

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 median household income of consumers near a brand's stations.[28] We do not control for unobserved heterogeneity in price sensitivity, which would add a significant computational burden. Our results suggest that accounting for such heterogeneity, in addition to what we capture, may not be important.

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 $\{\psi_{m,month(t)}\}$.[29] 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 attributes the residual brand-specific correlation in demand over time within a market to 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.[30] 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 $Income_{jmt}$ to create a second instrument, z_2, to account for the endogeneity of $(p_{jmt} \times Income_{jmt})$. Both US crude oil production and income are plausibly exogenous with respect to local, time-varying demand shocks. Figure 9 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.

[28] More explicitly, we take the median household income for the Census tract of each station, and we average these values across stations within a brand. The average is weighted by the population in each Census tract.

[29] We benefit from the size of our dataset. 95 percent of county-months have at least 16 observations, and 98.5 percent of county-brands have at least 40 observations.

[30] 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	-0.021^{***}	-0.205^{***}	-2.755^{***}	-2.952^{***}
	(0.006)	(0.009)	(0.284)	(0.516)
Price × Income		0.056^{***}	-0.006	-0.009
		(0.009)	(0.006)	(0.012)
Number of Stations	0.017^{***}	0.017^{***}	0.068^{***}	0.066^{***}
	(0.001)	(0.001)	(0.006)	(0.014)
IV	No	Yes	Yes	Yes
Week FEs			X	X
County-(Month of Year) FEs			X	X
Brand-County FEs			X	X
Observations	112,929	112,929	112,929	112,929
R^2	0.044	0.030	0.835	0.648

Notes: Significance levels: * 10 percent, ** 5 percent, *** 1 percent. The 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.

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 we separate out affiliated consumers with lower price sensitivities. The number of stations appears to matter about the same for attracting unaffiliated consumers as it does in the static model. After instrumenting and including rich fixed effects, we find that the coefficient on the interaction of price and income is small and not significant. We attribute this finding to the fact the brand-county fixed effects capture first-order variation in consumer heterogeneity.

We note that the R-squared for the regression with unaffiliated consumers falls in the dynamic model. This is to be expected, as the transformation of the log ratio of observed shares to unaffiliated shares is a non-affine transformation and will not preserve the measure of R-squared. Specifically, the shares for unaffiliated consumers fall closer to zero, from 14 percent (observed) to 9 percent in the dynamic model, and the majority of the variance in aggregate

Table 5: Estimated Dynamic Parameters

	Affiliation Rate			Strength of Affiliation	
	Baseline	*Income*	*Density*	*Level*	*Price*
	θ_1	θ_2	θ_3	$\bar{\xi}$	$\overline{\alpha}$
Coefficient	0.400	0.135	0.078	15.707	2.533
95 Percent CI	[0.31, 0.61]	[-0.05, 0.34]	[-0.01, 0.28]	[15.39, 19.93]	[1.02, 2.56]

Notes: The table displays the estimated non-linear coefficients from the dynamic model. The first three parameters imply that, on average, 60 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. The next three parameters show that the level shock to utility for affiliated customers is positive, as expected. The last parameter shows that affiliated customers are less price sensitive. Standard errors are in parentheses and are calculated via the bootstrap.

shares is attributed to the variance in unaffiliated shares. This serves to reduce the R-squared under the log-ratio transformation.

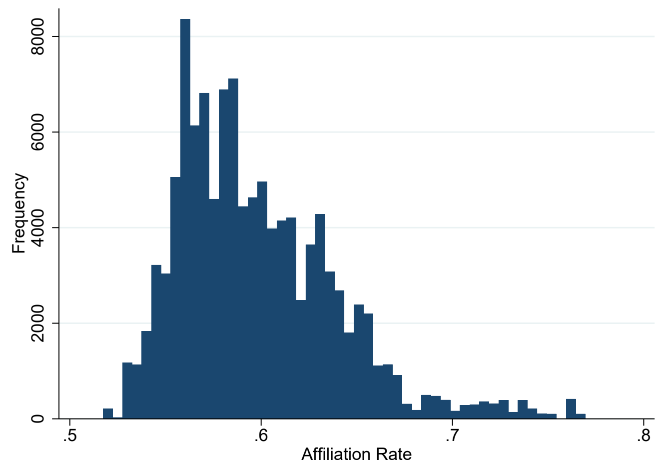
Table 5 reports estimates of the dynamic parameters. The parameters θ_1, θ_2, and θ_3 imply that 60 percent of consumers, on average, develop an affiliation for the brand they previously purchased from. The coefficient of 0.135 on income indicates that higher-income consumers are more likely to develop an affiliation, consistent with a habit-formation model where switching costs are increasing in wages. Likewise, the coefficient of 0.078 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 are standardized, so each coefficient corresponds to an increase of one standard deviation.

The estimated utility shocks to affiliated consumers, $\bar{\xi}$ and $\overline{\alpha}$, imply that the affiliated consumers are almost completely inelastic with respect to price. To interpret these coefficients, we summarize the implied elasticities in Table 6. The affiliated consumers of our model are very inelastic, with a near-zero response to price effects. 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 a subset of consumers almost never switch brands within this range of price changes.[31] The unaffiliated consumers, however, are highly elastic, with an average own-price elasticity of -7.7. This is large in magnitude, and it implies that for a 1 percent increase in price (roughly 3 cents), the station will lose 7.7 percent of the unaffiliated consumers. This high level of price sensitivity for a subset of retail gasoline consumers seems plausible, some "shoppers" have been found to go well out of the way to save a few cents per gallon.[32]

[31]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.

[32]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`

Figure 7: Estimated Heterogeneity in Affiliation Rates



Notes: Histogram represents the empirical distribution of the estimated λ_{jt} parameters.

Table 6: Implied Elasticities from Dynamic Model

Group	Mean	Pctl(25)	Median	Pctl(75)
Unaffiliated	−7.704	−9.040	−7.779	−6.358
Affiliated	−0.000	−0.000	−0.000	−0.000
Weighted	−3.090	−3.697	−2.969	−2.345
Naive (Static)	−6.792	−8.089	−6.780	−5.602

Notes: Table displays the estimated elasticities for affiliated and unaffiliated consumers. The weighted average elasticity is −3.09, which is less elastic than the estimated elasticity of −6.79 obtained by a static model.

Thus, our model captures consumer inertia in the gasoline data by distinguishing a subset of inelastic consumers from other price-sensitive consumers. The inelastic consumers are affiliated to their previously purchased brand. Our model captures heterogeneity in the rate at which consumers become affiliated. Figure 7 provides a histogram of the affiliation rates across the observations in our sample. There is heterogeneity in our estimates, with high-income, high-density areas realizing affiliation rates as high as 0.75, and low-income, low-density areas having affiliation rates closer to 0.5.

On average, roughly 40 percent of a brand's customers come from the unaffiliated pool of consumers in any week in equilibrium. The average weighted elasticity, which weighs affiliated and unaffiliated consumers by their relative (purchasing) proportions, is -3.09. This weighted elasticity is starkly different than the elasticity obtained by estimating the static model. A "naive" estimate using a static model (and supposing the true model were dynamic) would result in a value of -6.79, which implies a much greater loss in market share for a given price increase than we estimate from the dynamic model. Indeed, the static model obtains an elasticity closer to the unaffiliated elasticity than the overall elasticity from the dynamic model. Our brand-county elasticity estimate of -3.09 is more inelastic than some other estimates in the literature. In part, this difference may be due to the fact that the existing estimates do not account for consumer inertia. When we estimate a model with no inertia, we estimate much more elastic demand.[33] At the market level, the parameter estimates imply an aggregate elasticity of -2.14.[34]

We have tested the robustness of our demand estimates on two important dimensions. First, we examine the effect of the penalty to account for week-to-week measurement error in S_{jt}, which is discussed in Section 4.2. Because our estimates of $\overline{\xi}$ and $\overline{\alpha}$ imply very inelastic affiliated consumers, a single observation with $S_{jmt} \ll S_{jm(t-1)}$ would place a low upper bound on λ_{jmt}, resulting in a corner solution. Our solution is to adjust r_{jt} in such instances to the implied value based on the observed share, S_{jmt}, and the average value of $\frac{r_{jmt}}{S_{jmt}}$ across other observations. For n_a observations that need this adjustment, we add a penalty of $\left(\frac{n_a}{\mu}\right)^2$ to the objective function. In estimation, we use $\mu = 500$, but we obtain identical results if we use any value of $\mu \in [400, 1000]$, giving us confidence that we are picking up idiosyncratic errors. At the solution, the adjustment is applied to 2,705 observations, which is roughly 2 percent of the sample.

Second, we test for the uniqueness of our solution by randomly drawing 500 initial values for the estimation algorithm. We confirm that our solution minimizes the objective function, and our estimation routine gets nearly-identical results from most starting values.[35] Though we match the θ parameters, the data are only able to strongly identify a combination of $\overline{\xi}$ and $\overline{\alpha}$. Since we find that affiliated consumers are not sensitive to price, different combinations of $\overline{\xi}$ and $\overline{\alpha}$ can generate inelastic demand for those consumers, resulting in the same dynamics and coming very close to the solution that minimizes the objective function.

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 (5). Using the estimated demand parameters, we are able to recover the (derivative of)

[33]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.

[34]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 much closer to our estimate.

[35]The standard deviation of the mean λ is 0.009 and the standard deviation for the affiliated elasticity is 0.00006.

Table 7: Summary of Implied $\beta \frac{\partial E[V_j(\cdot)|\cdot]}{\partial p_{jt}}$

Group	Mean	Min	p25	Median	p75	Max
All	-0.111	-0.760	-0.148	-0.084	-0.046	0.096

Notes: The 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.

the continuation value. The dynamic condition for optimal pricing is:

$$\frac{\partial \pi_{jt}}{\partial p_{jt}} + \beta \frac{\partial E\left[V_j(r_{t+1}, c_{t+1}, x_{t+1})|p_t, r_t, c_t, x_t\right]}{\partial p_{jt}} = 0 \tag{21}$$

Where $\frac{\partial \pi_{jt}}{\partial p_{jt}}$ is the derivative of the per-period profits, which 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.

On average, we find that $\frac{\partial \pi_{jt}}{\partial p_{jt}}$ is positive, meaning 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, which is consistent with the reduced-form evidence of Section 3.3.2. Based on equation (21), we attribute the difference between $\frac{\partial \pi_{jt}}{\partial p_{jt}}$ 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.[36] 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 7. The mean and median are negative, which implies that, typically, a reduction in price would increase the expected future return. The magnitudes are significant: the mean of -0.111 implies that a 1 cent increase in price would increase static profits by roughly 4 percent.[37] Intuitively, firms are lowering prices to invest in future demand. This result, combined with our reduced-form findings of anticipatory pricing for expected costs, provides consistent evidence of forward-looking pricing behavior in retail gasoline.

[36] 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.

[37] The average (scaled) profit in our data is 0.029. Recall that 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:

$$\frac{\partial \pi_{jt}}{\partial p_{jt}} + \Psi_j(p_t, r_t, c_t, x_t; \theta) + \zeta_{jt} = 0 \tag{22}$$

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 soaked up by 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.[38]

[38]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 actual DCV, conditional on firms' choosing price according to $\Psi(\cdot)$. We discuss in greater detail in the following pages.

To estimate $\Psi_j(\cdot)$, we project the residual DCV onto the derivative of shares with respect to prices, the affiliation rate parameter λ_{jt}, marginal costs, expectations of future costs, competition and market structure variables, and the stock of current affiliated consumers. To control for competitive factors, we include the mean stock of affiliated consumers of rivals and the rivals' mean price; both of these are constructed relative to respective variables of the firm. We also include the number of stations, the total number of stations for all brands, and the number of brands as market-level controls. Finally, we interact $\frac{\partial S_{jt}}{\partial p_{jt}}$ with all of the other variables.

Constructing $\Psi_j(\cdot)$ as a function of $\frac{\partial S_{jt}}{\partial p_{jt}}$ and λ_{jt} is motivated by the fact that the marginal effect of price on the stock of affiliated customers in the next period is determined by the interaction of these variables. Thus, to a first-order approximation, we should expect that the derivative of the continuation value moves with the derivative of current period shares. We use λ_{jt} to control for the future value of a marginal consumer today.

The results of estimating equation (22) are reported in Table 8. The first specification reports the coefficients from a regression of the DCV onto the listed variables, additional market controls, and the first-order interactions with $\frac{\partial S_{jt}}{\partial p_{jt}}$. The variables are de-meaned so that the interpretation of the reported coefficient is the marginal effect at the mean of the other 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. 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 the derivative of the continuation value with respect to price is positively correlated with the derivative of shares with respect to prices (coefficient of 0.250), which is consistent with our model of consumer affiliation. We also estimate that a higher value of λ is associated with a continuation value that is less sensitive to price. Firms with a higher stock of affiliated customers are more sensitive to dynamic incentives, as indicated by the coefficient of -0.958. Overall, we find our model, with 20 parameters including a constant, captures over 95 percent of the variation of the residual that rationalizes observed prices.

In general, dynamic models can be sensitive to the choice of functional form and the variables employed in estimation. To verify that our estimate of $\Psi(\cdot)$ is consistent with realized profits, we use forward simulations to calculate the discounted present value of per-period profits when a firm unilaterally deviates its price. For each market and each period, we slightly perturb the price for a single brand. We re-compute shares in that period and then, using our estimate of $\Psi(\cdot)$, calculate equilibrium play in future periods. We use the estimated change in profits to calculate $\beta\frac{\partial E[V_j(\cdot)|\cdot]}{\partial p_{jt}}$ under rational expectations. Our estimates are positively correla-

Table 8: Dynamic Pricing Incentive: Regressions

	$\beta\frac{\partial E[V_j(\cdot)\|\cdot]}{\partial p_{jt}}$	Sensitivity
	(1)	(2)
$\frac{\partial S_{jt}}{\partial p_{jt}}$	0.250^{***}	-2.474^{***}
	(0.001)	(0.019)
λ_{jt}	0.021^{***}	-0.601^{***}
	(0.002)	(0.030)
Marginal Cost	-0.004^{***}	0.046^{***}
	(0.000)	(0.002)
Cost Change (30-Day Ahead)	0.018^{***}	-0.108^{***}
	(0.001)	(0.010)
Mean Price (Rivals)	-0.113^{***}	1.418^{***}
	(0.001)	(0.019)
Mean Affliated (Rivals)	-0.008^{***}	-0.084^{*}
	(0.003)	(0.043)
Affliated Customers	-0.958^{***}	12.117^{***}
	(0.004)	(0.064)
Market Controls	Yes	Yes
$\frac{\partial S_{jt}}{\partial p_{jt}}$ Interactions	Yes	Yes
Observations	112,929	112,929
R^2	0.952	0.869

Notes: Significance levels: * 10 percent, ** 5 percent, *** 1 percent. The table displays the estimated coefficients from a regression of the dynamic pricing incentive on state variables and the firm's price. The regression includes first-order interactions of all of the displayed variables. The variables are de-meaned, so the coefficient is interpreted as the marginal association at the mean of the other 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.

ted with the implied continuation value under rational expectations.[39]

[39]We simulate forward for 12 periods. We find that the effect of a marginal change in price on future profits dissipates after approximately 9 periods. We use a weekly discount rate of 0.999, which corresponds to an annual discount rate of 0.949. The mean DCV under rational expectations is 28 percent of the magnitude of the implied residual. It is plausible that profits from complementary goods, which we cannot measure directly and might include sales of items such as food, make up the difference. Our reduced-form estimate will capture the profit incentives from complementary goods, which could be considered a feature for the purposes of counterfactual simulation.

Table 9: Merger Effects

(a) Dynamic Model

Brands	Price	Share	Profit
Marathon-BP	3.33	-13.75	7.33
Other	-0.32	10.37	-6.78
Overall	1.34	-0.47	0.01

(b) Static Model

Brands	Price	Share	Profit
Marathon-BP	5.87	-21.29	29.77
Other	0.53	4.79	11.75
Overall	2.96	-6.93	20.43

Notes: Table displays the mean effects from a counterfactual merger between two brands in our data. Panel (a) provides the estimates from a dynamic model that accounts for consumer inertia. Panel (b) provides the estimates from a static model that is calibrated to match prices, margins, and shares from the same data. Prices are weighted by the shares observed in the data, i.e., the "no merger" shares.

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 252 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[40] at the beginning of September 2014, and we calculate counterfactual prices and shares for the second half of our sample.[41]

Our setting, with dynamic brand-specific effects, puts an emphasis on modeling specifically what a merger will be in practice. This can be a nuanced question. Do the merging firms retain separate brands, or do they convert all stations to a single brand? Do affiliated consumers retain their affiliation to the merged company? In our merger counterfactual, we assume that

[40]We combine the merging firms into a single entity and adjust the utility shocks so that the market shares would remain constant at the no merger (share-weighted) prices.

[41]Because our inelastic affiliated customers would 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. In our merger counterfactuals, none of the prices approach the choke price.

consumer habits are tied to gas station locations and are unaffected by the brand name, and therefore the merging firms retain affiliated consumers from both Marathon and BP.[42]

Panel (a) in Table 9 displays the mean effects of the merger. The effects are modest, with an average price increase for the merging firms of 3.3 percent. Shares for these firms fall by 14 percent, and profits increase. Prices for competitors decline slightly, which is not a typical outcome in static price setting models, as prices are usually strategic complements. Prices are strategic substitutes due to affiliation dynamics. When the merging firms raise price, the pool of unaffiliated consumers increases. This increases the average consumer elasticity faced by the non-merging firms, which tends to dampen prices. This effect can outweigh the usual incentive to raise price (arising from the higher price of the merging firms), resulting in lower prices. In aggregate, prices rise by 1.3 percent, and profits barely increase, as most of the extra surplus captured by the merging firms comes at the expense of the rivals.

Overall, these effects are modest relative to what one might expect from a merger with the corresponding share structure (and HHI values). For comparison, we report the results from a merger analysis using a static model in panel (b) of Table 9. We calibrate a standard logit demand system to identical prices, margins, and shares that are used to estimate the dynamic model. The static model predicts price effects of over 5 percent for the merging firms. Consistent with the usual merger logic, prices in the static model are strategic complements, as the prices for rivals rise. Overall prices increase by 3 percent, and industry profits increase by 20 percent, which is in stark contrast to the predictions from the dynamic model.

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

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.

[42]We consider an alternative scenario where the acquiring firm loses the affiliated consumers from the acquired brand: all of the BP consumers become unaffiliated. Since the one-year post-merger window is sufficient to reach the new long-run equilibrium, the results look similar.

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 60 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 affecting prices, and we show, both theoretically and empirically, that failing to account for dynamic demand will typically cause merger simulations to over-predict 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). Multi-Period Competition with Switching Costs. *Econometrica 60*(3), 651.

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

Bergemann, D. and J. Välimäki (2006). 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). 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). 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). 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). Co-Integration and Error Correction: Representation, Estimation, and Testing. *Econometrica 55*(2), 251.

Farias, V., D. Saure, and G. Y. Weintraub (2012). 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). 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). 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). 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). Spatial Differentiation and Vertical Mergers in Retail Markets for Gasoline. *American Economic Review 102*(5), 2147–2182.

Lee, R. S. (2013). 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). 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). 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). 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). 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:

$$
\begin{aligned}
r^{ss} &= \lambda\big((1 - r^{ss})s_0 + r^{ss} \cdot s_j\big) \\
\implies r^{ss} &= \frac{\lambda s_0}{1 - \lambda(s_j - s_0)}.
\end{aligned}
$$

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:

$$
\begin{aligned}
V^{ss}(r^{ss}, c^{ss}) &= (p^{ss} - c^{ss})((1 - r^{ss})s_0 + r^{ss}s_j) + \beta V^{ss} \\
&= \frac{p^{ss} - c^{ss}}{1 - \beta} \cdot \frac{s_0}{1 - \lambda(s_j - s_0)}.
\end{aligned}
$$

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:

$$
p^{ss} = c^{ss} + \underbrace{\frac{-s_0\,(1 - \lambda s_j + \lambda s_0)}{\frac{ds_0}{dp}\,(1 - \lambda s_j) + \frac{ds_j}{dp}\lambda s_0}}_{m\,=\,\text{markup of price over marginal cost}}. \tag{23}
$$

The second term, m, on the right-hand side of equation (23) 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 (23) 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:

$$m_d > m_s \iff -\frac{\partial s_0}{\partial p} > -\frac{\partial s_j}{\partial p}. \tag{24}$$

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 (24). 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. Using the J restrictions implied by the first-order conditions (one for each product j)

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

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

3. Plug the value of $\frac{dV_i(r)}{dr}$ into the first-order conditions and solve for the steady-state value of p.

4. 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 j^{th} 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 $\frac{dp}{dr}$.

5. 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	0.43	0.23	0.05	0.25	0.60	0.95	8025
$\lvert\alpha\rvert$	6.50	3.56	0.38	3.48	9.26	14.97	8025
$\overline{\alpha}$	3.24	2.79	0.01	0.99	4.90	14.08	8025
ξ	6.54	4.00	-1.96	3.37	9.35	16.59	8025
$\bar{\xi}$	9.86	5.85	0.01	4.69	14.98	19.97	8025

Notes: Table displays summary statics for demand parameters for the 8,025 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) and margins between 0.10 and 0.75. See the text for additional details.

A.4 Outcomes and Demand Parameters

Table 11: Simulation: Demand Parameters

	(1) Pre-Merger Price	(2) Price Δ F1	(3) Bias F1	(4) Price Δ F3	(5) Bias F3
λ	0.172***	2.840***	1.839***	0.232***	0.865***
	(0.014)	(0.091)	(0.026)	(0.025)	(0.013)
$\lvert\alpha\rvert$	-0.135***	-2.730***	-0.285***	-0.526***	-0.200***
	(0.003)	(0.018)	(0.005)	(0.005)	(0.003)
$\overline{\alpha}$	0.001	0.016*	0.009***	-0.001	0.006***
	(0.001)	(0.009)	(0.003)	(0.003)	(0.001)
ξ	0.055***	1.992***	0.241***	0.432***	0.167***
	(0.002)	(0.016)	(0.004)	(0.004)	(0.002)
$\overline{\xi}$	0.001**	0.004	0.013***	-0.005***	0.007***
	(0.001)	(0.003)	(0.001)	(0.001)	(0.000)
Constant	1.359***	4.544***	0.392***	0.525***	0.238***
	(0.003)	(0.020)	(0.006)	(0.005)	(0.003)
N	8025	8025	8025	8025	8025

Notes: 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. All dependent variables are demeaned. Parameters correspond to the dynamic demand model detailed in section 2. Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 11 in provides results from regressions of pre-merger prices, price changes, and the bias in the static prediction on the demand parameters. On average, these outcomes are increasing in the rate of affiliation and strength of affiliation, as captured by the dynamic paramters. However, these relationships do not hold in every instance, and may interact in interesting ways. As show in Figure 1, prices may be increasing or decreasing in the rate of affiliation. In a small number of markets, we find that the relationship between λ and prices is non-monotonic.

A.5 Numerical Results for Merging Firms

Table 12: 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	25.526^{***} (0.198)	25.487^{***} (0.194)	25.481^{***} (0.194)	3.392^{***} (0.087)	3.442^{***} (0.073)	3.450^{***} (0.072)
Pre-Merger Margin	20.596^{***} (0.113)	20.523^{***} (0.111)	20.512^{***} (0.111)	1.188^{***} (0.049)	1.279^{***} (0.041)	1.297^{***} (0.041)
λ		-1.126^{***} (0.065)	-1.148^{***} (0.065)		1.413^{***} (0.024)	1.446^{***} (0.024)
$\lvert\frac{\overline{\alpha}}{\alpha}\rvert$			-0.087^{*} (0.052)			0.118^{***} (0.019)
$\bar{\xi}$			-0.008^{***} (0.003)			0.013^{***} (0.001)
Constant	4.544^{***} (0.015)	4.544^{***} (0.015)	4.544^{***} (0.015)	0.392^{***} (0.007)	0.392^{***} (0.006)	0.392^{***} (0.006)
N	8025	8025	8025	8025	8025	8025

Notes: Observations are for firm 1. Price Δ and Bias are the merger price change and static prediction bias, respectively. All dependent variables are demeaned. 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 Table 12 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 the merger. The affiliation rate (λ), the relative affiliation price sensitivity effect, ($\lvert\frac{\overline{\alpha}}{\alpha}\rvert$), and the utility boost from affiliation ($\bar{\xi}$), all decrease the unilateral price impact for the merging firm.[43] Correspondingly, the static model bias, which is the dependent variable in columns (4)-(6) of Table 12, increases with the dynamic parameters, conditional on the pre-merger shares and margins. Therefore, even if affiliation cannot be directly estimated, price change estimates should be revised downward if affiliation is expected to play an important role.

[43]Results are directionally the same, but smaller in magnitude, for the non-merging firm. Results are presented in Appendix A.6. We also find that in 3.8 percent of simulations, the non-merging firm lowers its price in response to the merger. This is consistent with our empirical application, where we find that the non-merging firm tends lower its price in the counterfactual simulations.

A.6 Numerical Results for Non-Merging Firm

Table 13: Simulation: Merger Price Change and Bias for Non-Merging Firm

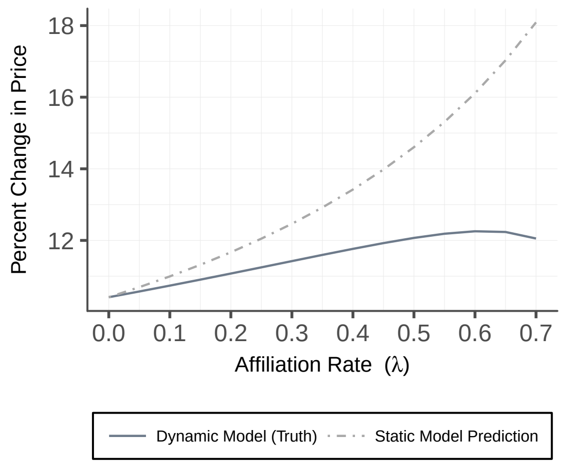
	(1) Price Δ	(2) Price Δ	(3) Price Δ	(4) Bias	(5) Bias	(6) Bias
Pre-Merger Market Share	6.159^{***}	6.141^{***}	6.136^{***}	2.357^{***}	2.377^{***}	2.382^{***}
	(0.072)	(0.070)	(0.069)	(0.041)	(0.036)	(0.035)
Pre-Merger Margin	2.565^{***}	2.532^{***}	2.523^{***}	0.852^{***}	0.889^{***}	0.899^{***}
	(0.041)	(0.040)	(0.040)	(0.023)	(0.021)	(0.020)
λ		-0.514^{***}	-0.531^{***}		0.568^{***}	0.587^{***}
		(0.023)	(0.023)		(0.012)	(0.012)
$\lvert\frac{\overline{\alpha}}{\alpha}\rvert$			-0.064^{***}			0.073^{***}
			(0.018)			(0.009)
$\overline{\xi}$			-0.006^{***}			0.007^{***}
			(0.001)			(0.000)
Constant	0.525^{***}	0.525^{***}	0.525^{***}	0.238^{***}	0.238^{***}	0.238^{***}
	(0.006)	(0.005)	(0.005)	(0.003)	(0.003)	(0.003)
N	8025	8025	8025	8025	8025	8025

Notes: Observations are for firm 3. Price Δ and Bias are the merger price change and simulation bias, respectively. All dependent variables are demeaned. Market share is the aggregate market share. Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 13 presents results for the non-merging firm in the numerical simulations. Results are qualitatively the same as the merging firm, but the estimates are smaller in magnitude. This reflects that the non-merging firm reacts strategically to the merger, but the size of its reaction is less pronounced as it is responding only to the residual demand shock from the merged firm's price change. In the numerical simulations, we find that in 3.8 percent of observations, the non-merging firm lowers its price in response to the merger. Thus, as in the empirical counterfactual, dynamic affiliation may cause prices to be strategic substitutes.

A.7 Monopoly Prices from Calibrated Static and Dynamic Models

Figure 8: Monopoly Market Power in Static and Dynamic Models



Notes: The plot displays the marginal effects, in terms of mean percent price increase, of a merger to monopoly for different values of λ. The solid line represents the true price increase arising in the dynamic model. The dashed line represents the price increase arising in a static model that is calibrated to the same data. The plot reflects 275 baseline parameter values of $(\xi, \bar{\xi}, \alpha, \overline{\alpha})$ that converged for $0 \leq \lambda \leq 0.7$, or 4,125 markets in total.

B Reduced-Form Evidence: Supplemental Results

B.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.[44] 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.

$$c_{it+30} = \alpha_1 c_{it} + \alpha_2 F_t^{30} + \gamma_i + \epsilon_{it} \tag{25}$$

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.

[44] 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.[45]

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

C Empirical Application: Supplemental Tables and Figures

C.1 Summary Statistics by Observation

Table 14: Retail Gasoline in Kentucky and Virginia: Oct 2013 - Sep 2015

Statistic	Mean	St. Dev.	Min	Pctl(25)	Pctl(75)	Max	N
Share	0.141	0.110	0.0003	0.061	0.187	0.688	112,929
Price	2.870	0.529	1.715	2.383	3.310	4.085	112,929
Wholesale Price	2.256	0.527	1.245	1.754	2.673	3.545	112,929
Wholesale FE	2.261	0.031	2.207	2.231	2.293	2.366	112,929
Margin	0.206	0.115	−0.440	0.132	0.273	1.048	112,929
Num. Stations	5.235	7.093	1	2	6	83	112,929
Food	0.859	0.259	0.000	0.803	1.000	1.000	112,431
Supermarket	0.019	0.093	0.000	0.000	0.000	1.000	112,431
Car Service	0.146	0.261	0.000	0.000	0.200	1.000	112,431
Interstate	0.004	0.049	0.000	0.000	0.000	1.000	112,431

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 83. 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. For the 297 observations that are missing station-specific amenities, we impute the values as the market-period mean.

C.2 Summary Statistics by Brand

Table 15: Summary of Brands

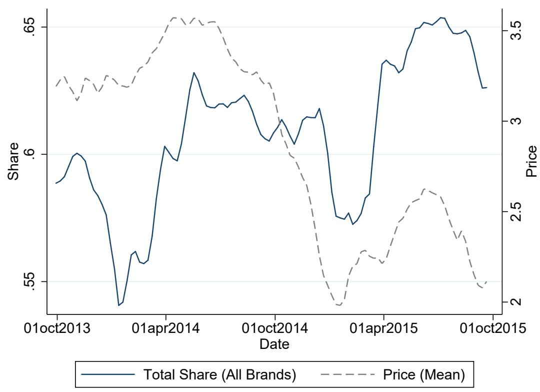
	Brand	Cond. Share	Share	Num. Markets	Num. Stations	Margins
1	Marathon	0.18	0.10	135	5.20	0.21
2	Sheetz	0.18	0.03	37	1.70	0.17
3	Speedway	0.17	0.03	39	3.70	0.18
4	Wawa	0.16	0.01	22	3.20	0.12
5	Exxon	0.16	0.07	119	4.60	0.25
6	Hucks	0.15	0.01	11	1.80	0.15
7	7-Eleven	0.15	0.02	42	6.70	0.18
8	FRINGE	0.14	0.12	244	9.80	0.19
9	Shell	0.13	0.09	165	4.20	0.22
10	Pilot	0.12	0.01	21	1.40	0.13
11	BP	0.12	0.06	127	3.30	0.21
12	Loves	0.11	0.01	15	1.00	0.20
13	Valero	0.11	0.02	59	3.50	0.21
14	Thorntons	0.11	0.00	9	5.90	0.14
15	Sunoco	0.10	0.01	35	4.30	0.29
16	Citgo	0.08	0.01	35	3.90	0.24

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 252 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.

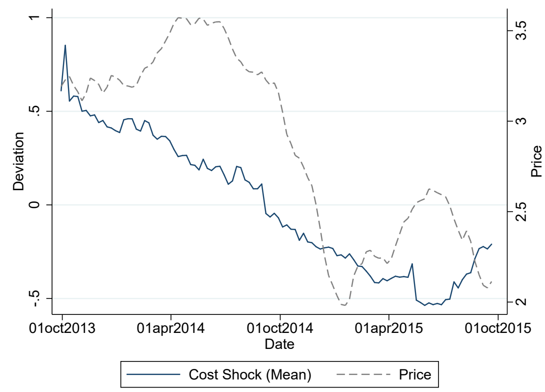
C.3 Shares, Prices, and Instrument

Figure 9: 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.