Consumer Inertia and Market Power*

Alexander MacKay
Harvard University†

Marc Remer
Swarthmore College‡

January 29, 2024

Abstract

We study the pricing decisions of firms in the presence of consumer inertia. Inertia, which can arise from habit formation, brand loyalty, and switching costs, generates dynamic pricing incentives. These incentives mediate the impact of competition on market power in oligopoly settings. For example, dynamic incentives can limit the equilibrium price effects of a horizontal merger. However, the way that the merger is implemented—whether the merged firm maintains separate brands or consolidates them into a single entity—can have large effects on equilibrium prices in the presence of inertia. We develop an empirical oligopoly model to estimate consumer inertia and dynamic pricing incentives using market-level data. We apply the model to a hypothetical merger of retail gasoline companies. Our analysis shows how the static model predictions can diverge meaningfully from those obtained while 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, Robin Lee, Nathan Miller, Aviv Nevo, Ariel Pakes, 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, NYU Stern, and Harvard University for their feedback and suggestions. Researchers own analyses calculated (or derived) based in part on data from Nielsen Consumer LLC and marketing databases provided through the NielsenIQ Datasets at the Kilts Center for Marketing Data Center at The University of Chicago Booth School of Business. The conclusions drawn from the NielsenIQ data are those of the researcher(s) and do not reflect the views of Nielsen. Nielsen is not responsible for, had no role in, and was not involved in analyzing and preparing the results reported herein.

†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, or switching costs. In response to consumer inertia, profit-maximizing firms will internalize the effect of their current price on demand in future periods.

In this paper, we study the influence of consumer inertia on competition and pricing in oligopoly settings. We develop a model of inertia where a portion of consumers develop a product-specific preference, or *affiliation*, for the most recently purchased product. This formulation nests typical implementations of dynamic consumer behavior, including switching costs and brand loyalty. We then use the model to evaluate the impact of consumer inertia on equilibrium price levels and the implications for horizontal mergers.

We provide two contributions related to the analysis of mergers. First, we show that the effects of mergers on equilibrium prices depend on the presence of inertia. Using numerical simulations and an empirical model, we then show that the predictions of a static model can diverge meaningfully from the true effects, primarily because of misspecified first-order conditions for prices. The second contribution is to highlight an important feature of merger implementation. After a merger, a firm may decide to maintain separate brands or consolidate them into a single brand. As we show below, this decision, which may not matter in static models, can have substantive implications in the presence of inertia.

A primary goal of the analysis is to provide an empirical framework to identify consumer inertia and conduct counterfactuals in settings where data on prices and quantities is available, but more detailed consumer-level data is not. Thus, identification relies upon aggregate, market-level data. The demand model we introduce is a straightforward extension of the standard discrete-choice logit model with myopic consumers. In contrast to the identification problem in a typical random coefficients model, the distribution of unobserved heterogeneity in our model evolves to reflect past purchase behavior. To disentangle heterogeneity in preferences from state dependence arising from previous purchases, we impose restrictions on the demand system and rely on the panel structure of data. Intuitively, consumer inertia in our model is reflected through shocks that generate semi-persistent correlations in shares. Fully persistent features, such as brand quality, can be captured by fixed effects.

We begin by evaluating theoretical properties of the steady-state equilibrium of our model using numerical simulations. Inertia can lead to higher equilibrium prices, as consumers become locked into a particular product or brand over time. We denote the extent to which a firm sets higher prices due to the presence of state-dependent consumers as *dynamic market power*.[1] We show that dynamic market power can influence *horizontal market power*—the extent to

[1]As consumer inertia can also lead to lower equilibrium prices (Dubé et al., 2009), it is possible that dynamic market power is reduced with greater inertia.

which competitors constrain prices. Our results show that the dynamic component often has a greater impact on equilibrium prices than a reduction in competition via a merger. Therefore, to understand the competitive conditions in a market, it can be informative to assess both of these dimensions of market power.

In evaluating horizontal market power, we show that modeling the type of merger is critical. A merged firm can decide to consolidate the merged brands into a single entity or maintain them as separate brands. In practice, both types of mergers occur with some frequency. For example, in the airline industry, mergers typically result in one of the brands being eliminated.[2] Conversely, with consumer product mergers, the acquiring firm often maintains both existing and acquired brands.[3] In retail gasoline, both types of mergers have occurred, with BP eliminating the Amoco brand after its 1998 acquisition, while Exxon maintained the Mobil brand for service stations after its acquisition in 1999.[4]

Our numerical simulations illustrate how consumer dynamics can influence post-merger outcomes. In a static setting, distinguishing between types of mergers may not be relevant; we show that mergers of either type deliver the exact same outcome for symmetric firms with logit demand. However, in the presence of consumer inertia, differences in dynamic pricing incentives can lead these two merger types to deliver directionally opposite effects relative to the static model. Accounting for these differences is relevant for the decisions of firms and competition authorities, which typically employ static empirical models to estimate consumer demand and simulate counterfactual post-merger prices. Such static models could, for example, substantially over-predict the price effects of a merger. Thus, failing to account for consumer inertia may misstate the potential for horizontal market power and affect merger enforcement.

These findings highlight the need for empirical models to appropriately represent firms' dynamic pricing incentives and the structure of the post-merger firm. Real-world environments diverge from the setting of our steady-state analysis in that firms are often asymmetric and face time-series variation in marginal costs and demand. To incorporate these features, we develop an empirical model to estimate demand and evaluate dynamic pricing incentives using panel data. We then provide an approach to simulate counterfactuals with dynamic oligopoly pricing, which we use to quantify the price effects of inertia and competition (via mergers) in an application using data from retail gasoline markets. In doing so, we create a framework that accounts for both the dynamic and horizontal dimensions of market power and can address key policy questions.

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

[2] Historical examples include Delta-Northwest, United-Continental, and Southwest-AirTran mergers, leaving Delta, United, and Southwest, respectively.

[3] Coca-Cola launched Gold Peak Tea in 2006, acquired Honest Tea in 2011, and acquired Peace Tea in 2015. They maintained these three separate ready-to-drink iced tea brands through 2022.

[4] Interestingly, BP re-introduced the Amoco brand for retail stations in 2017.

Our model allows us to separately identify the choice probabilities for consumers of different unobserved types and states. Thus, we allow for endogenous unobserved heterogeneity through the presence of a serially correlated state variable for each firm. To achieve identification, we limit the degree to which static unobserved heterogeneity enters the model. Thus, we give up more flexible static substitution patterns in order to capture substitution patterns that reflect consumer dynamics, while still relying on aggregate data.

We use a series of Monte Carlo experiments to verify that our approach can correctly identify the dynamic and static components of the model with market-level data. Further, we show that our approach does not falsely attribute static consumer heterogeneity to state dependence. Specifically, we generate data using a static random coefficients demand system, and then estimate our dynamic model using the simulated data. The simulations show that, in our modeling framework, any bias from persistent unobserved heterogeneity loads onto the static parameters, and we correctly estimate zero state dependence. Thus, at a minimum, our empirical approach can be used to test for the presence of consumer inertia, even when persistent unobserved heterogeneity is more complex than we model. This test can be valuable for empirical analysis, antitrust investigations, and policy design, as our analyses show that the dynamic components of demand can be of first-order importance for equilibrium outcomes.

To conduct counterfactuals, we impose a supply-side model of price competition. In contrast to the literature, we invoke relatively weak assumptions about supply-side behavior. 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 of 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. We use forward simulations to verify that realized equilibrium payoffs are consistent with these forecasts. We then use the model to evaluate horizontal and dynamic market power in our empirical setting.

As a case study, we apply the model using data from retail gasoline markets. This setting has a relatively homogeneous product, minimal transaction costs, and, to the extent there is inertia, it may be considered relatively short-lived. Thus, to the extent that dynamic incentives matter in this market, such incentives may be important across a broad set of consumer markets. There is some existing evidence of pricing patterns in retail gasoline that are consistent with consumer affiliation, such as slow-to-adjust cost pass-through (e.g., Lewis and Noel, 2011). Furthermore, retail gasoline has a direct link to antitrust concerns. In June and December of 2017, the FTC challenged Alimentation Couche-Tard's acquisitions of Empire Petroleum Partners[5] and Holiday,[6] respectively, on the basis of overlapping retail gasoline stations in a number of states. In

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

[6]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/

2021, after 7-Eleven's acquisition of Speedway, the FTC ordered the divestiture of 192 stations to mitigate harms to competition.[7] In August 2021, FTC Chair Lina Kahn stated that the FTC would be taking more aggressive steps to deter mergers in the industry.[8]

Prior to estimating the model, we present reduced-form evidence of dynamic demand and dynamic pricing incentives in retail gasoline markets. Using consumer-level data on retail gasoline purchase histories, we find evidence of dynamics in consumer demand. For example, we find that after purchasing from the same brand at least three times in a row, a household purchases from the same brand in the next period 90 percent of the time. However, if they interrupt their spell by buying from a different brand, they return to the previous brand only 57 percent of the time. We also find that consumers are less likely to return to the previous brand if there is a longer period between purchases. These patterns are consistent with our model of state-dependent demand and are more challenging to reconcile with purely static demand.

Similarly, we find evidence that firms internalize dynamic pricing incentives. We find that new entrants initially price lower than established firms but raise prices over time. We then examine cost pass-through, and we show that firms begin raising prices in anticipation of predictable future cost increases. Finally, we note that roughly 3 percent of our data have negative price-cost margins. These findings can be rationalized by retail gasoline stations internalizing the effect of current prices on future affiliated consumers.

We estimate the demand model using a 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 64 percent of consumers have the tendency to become affiliated to the brand from which they previously purchased.[9] The remaining 36 percent are "shoppers" that are unaffected by consumer inertia. Consumers that become affiliated to a brand are not very price sensitive, with an average elasticity of -0.53. Shoppers are much more price sensitive, with an average elasticity of -6.0. Though shoppers are a minority of consumers that purchase gasoline, they are important for disciplining prices in equilibrium. Across all consumer types, firms face an average elasticity of -1.8. In addition to the contemporaneous elasticity, the dynamic incentives facing firms play an important role in disciplining equilibrium markups.

To empirically evaluate horizontal market power, 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 brand consolidation would increase prices for the merging firms by 2.4 percent post-merger. A static model of brand consolidation, on the other hand,

[7]https://www.ftc.gov/news-events/news/press-releases/2021/11/ftc-approves-final-order-requiring-divestitures-hundreds-retail-gas-diesel-fuel-stations-owned-7

[8]https://www.mitchellwilliamslaw.com/webfiles/08-25-21\%20Brian\%20Deese\%20Letter.pdf

[9]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

predicts an average price increase of 5.4 percent, which would likely result in greater antitrust scrutiny. In this case, the dynamic incentive to invest in future demand mitigates the increase in horizontal market power obtained after a merger. By comparison, the dynamic model also predicts that a merger with joint pricing for separate brands results in a price increase of over 4.9 percent for the merging firms. Thus, modeling the precise structure of the post-merger firm is important for predicting the magnitude of price effects.

To evaluate dynamic market power in the empirical model, we calculate equilibrium prices if we change the strength of inertia or the share of consumers affected by inertia. Increasing the strength of inertia by 10 percent leads to a 4.7 percent increase in equilibrium prices. Increasing the share of consumers prone to inertia by 0.10 results in prices that are 1.7 percent higher. These effects are in the same range as the price effects for the merging firms in the merger counterfactuals. We conclude that a relatively modest change in the strength or prevalence of affiliation can affect prices by as much as the elimination of a major competitor. Thus, we use the model to assess the relative magnitudes of dynamic and horizontal market power.

More generally, competition authorities often analyze mergers in markets that are likely to be characterized by consumer inertia. For example, in its lawsuit against Swedish Match-National Tobacco, the Federal Trade Commission (FTC) cited strong brand loyalty as a barrier to entry.[10] Similarly, the US Department of Justice cited customer switching as an important factor in its case against the UPM-MACtac merger.[11] 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. Consumer inertia is a central feature of many technology markets, such as those served by Amazon, Google, Apple, and Facebook, which are currently under heavy scrutiny by antitrust agencies. Yet, perhaps due to computational complexity or compressed investigative timelines, the supply-side impacts of inertia are seldom quantified in these investigations.

The paper proceeds as follows: In Section 2, we introduce the model and present results from numerical simulations. Section 3 presents the data for our empirical application and reduced-form evidence of dynamics. Section 4 provides implementation details and results for dynamic demand estimation. In Section 5, we introduce our approach to evaluate dynamic pricing incentives and construct counterfactuals. Section 6 concludes.

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 add to the literature

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

[11] Ibid.

by examining the effect of competition and changes in market structure in such settings. As we show, 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 (see, e.g., Farrell and Shapiro, 1988; Beggs and Klemperer, 1992; Bergemann and Välimäki, 2006). Such features link directly to our notion of consumer affiliation. More generally, 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). Our goal is not to distinguish among potential mechanisms that drive dynamic pricing, but to instead put forward a tractable empirical model that can be used to quantify the impacts of competition in contexts with such dynamics.

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. Bronnenberg et al. (2012) and Eizenberg and Salvo (2015) evaluate the presence of habit formation in soft drink markets. When explicitly modeled, the above papers assume that consumers are myopic when choosing among products, which we also maintain. Conceptually, our demand model shares similar features to that of Dubé et al. (2009) and Dubé et al. (2010). Relative to that model, we give up some flexibility in order to estimate demand with a panel of market-level price and quantity data, rather than requiring individual customer data. These less stringent data requirements, as well as our supply-side innovations, allow our model to be applied in a wide variety of settings where it is typically difficult to account for consumer inertia.

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, market-level data tends to be more readily available. One paper that has used aggregate data to estimate switching costs is Shcherbakov (2016), which provides an informal argument for the identification of switching costs from aggregate data in the context of television services.[12] Shcherbakov (2016) estimates a structural model with persistent observed heterogeneity and therefore richer static substitution patterns than our model. The empirical strategy of that paper is to obtain valid instruments for price and quality, and then impose that lagged values of these instruments are also orthogonal to the structural demand shock. The intuition is that these past variables influence the current market share (through state dependence) but are uncorrelated with demand shocks. By contrast, we impose directly that innovations in the demand shocks are uncorrelated,

[12] 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).

without implicitly making additional assumptions about instruments, and we demonstrate that our approach recovers the correct demand parameters with Monte Carlo exercises. Another distinction of our framework is that we allow for and estimate a proportion of consumers who are unaffected by inertia, which can vary across markets. Thus, our model is more restrictive along the static dimension but allows more flexibility in terms of consumer dynamics.

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 biased elasticities. Hendel and Nevo (2013) point out that this will matter in a merger analysis. We complement this point by showing that dynamic incentives, rather than biased elasticities, can be 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, allowing 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. Specifically, we evaluate the static demand derivatives using a standard model of demand, and we plug these estimates into the firm's first-order condition to recover the dynamic pricing incentive. This approach offers certain advantages relative to estimating the price policy function, as in Bajari et al. (2007). First, it allows us to test for the presence of dynamic pricing behavior, rather than assuming it. Second, our method of estimating firms' dynamic incentives does not require us to take a stance on the discount rate or beliefs of firms, as the role of both are captured by the dynamic pricing incentive.

In the empirical model, we estimate the derivative of the value function directly, circumventing some of the computational challenges of estimating the value function (see, e.g., Farias et al., 2012; Sweeting, 2013) and eliminating 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, or switching costs.[13] We place parametric restrictions on the form of affiliation for empirical tractability. Consumers in the model are myopic in that they maximize current period utility rather than a discounted flow of future utility. This assumption is likely a good fit for retail gasoline markets, where consumers do not choose a gas station anticipating that it will limit their future choice set. Despite this, some consumers are likely to return to the same gas station due to habit formation, brand loyalty, or inattention. Myopia is also likely a reasonable assumption for other consumer products, including many products purchased at grocery stores, where product affiliation has been documented.[14]

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.[15] We use the model to numerically and empirically examine the impact of consumer inertia on market power, in general, and in the context of horizontal mergers.

2.1 Demand

We extend the 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). Each consumer i is indexed by a discrete type, h_i, and a time-varying state, z_{it}. The first feature, h_i , captures exogenous and persistent unobserved heterogeneity.[16] The second feature, z_{it}, allows the distribution of preferences to change endogeneously over time through

[13]For certain parameter values, the model can also be interpreted as a model of search or inattention.

[14]For example, see Dubé et al. (2010).

[15]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.

[16]The discrete type assumption for the random coefficient model is made elsewhere in the literature. See, for example, Berry et al. (2006) and Berry and Jia (2010).

state-dependent utility.

Consumer i receives the following utility for choosing product j in period t:

$$u_{ijt}(z_{it}) = \delta_{jt} + \sigma_{jt}(z_{it}; h_i) + \epsilon_{ijt}. \tag{1}$$

Consumers receive an additively-separable common component δ_{jt}, a state-dependent shock that may be type-specific $\sigma_{jt}(z_{it}; h_i)$, and an idiosyncratic shock, ϵ_{ijt}. The common component will typically be a function of firm j's price, and takes the form $\delta_{jt} = \xi_{jt} + \alpha p_{jt}$ in the standard logit model (with $\alpha < 0$).

We denote the probability that a consumer of type h in state z chooses product j as $s_{jt}(z; h)$. We normalize the utility of the outside good to be zero. Given the standard assumption of a type 1 extreme value distribution on the utility shock, ϵ_{ijt}, the choice probabilities of consumers are:

$$s_{jt}(z; h) = \frac{\exp(\delta_{jt} + \sigma_{jt}(z; h))}{1 + \sum_l \exp(\delta_{lt} + \sigma_{lt}(z; h))}. \tag{2}$$

The overall share of product j, S_{jt} is given by the weighted average of choice probabilities for consumers across states and types.

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: Consumer Types Our framework allows for both endogenous and exogenous unobserved heterogeneity. We consider unobserved exogenous heterogeneity in a simple form by assuming that there are two latent consumer types, $h \in \{0, 1\}$. We assume that consumers with $h = 0$ are unaffected by state dependence, so that $\sigma_{jt}(z; 0) = \sigma_{jt}(z'; 0)\ \forall z, z'$, and we normalize $\sigma_{jt}(z; 0)$ to zero for all z. We term these consumers "shoppers."

Demand from shoppers is given by the standard logit choice probabilities, as can be seen following equation (2). It is well known that logit demand restricts consumer substitution patterns to be proportional to market share, which may be unappealing in certain settings. We impose this restriction to make progress on identifying and estimating a model with state dependence with only market-level data.[17]

Consumers with a latent type $h = 1$ can be affected by state dependence. Let the fraction of consumers of this type be given by λ.

[17] For applications, such as antitrust investigations, market-level data are common, while more detailed data about diversion are often not available (Valletti and Zenger, 2021). Moreover, the imposition that diversion is proportional to share may be palatable in more narrow product markets (Miller and Sheu, 2021), in which there are fewer dimensions of differentiation (e.g., luxury SUVs vs. automobiles). In practice, antitrust agencies have employed this assumption during investigations and in court. For example, this assumption was used by the DOJ's economic experts' in challenges to the H&R Block/TaxACT merger in 2011, the GE-Elextrolux merger in 2015, the AT&T/Time Warner merger in 2018, and by the FTC in its analysis of the Reynolds-Lorillard tabacco merger in 2015.

Assumption 3: Single-Product Affiliation We now place restrictions on the state-dependent demand shocks, $\sigma_{jt}(z;h)$, for consumers that are affected by state dependence. We assume that each consumer state corresponds to an affiliation (utility shock) to a single product. Further, we assume that there is a single state corresponding to each product. Thus, a consumer in state $z = j$ is affiliated to product j. We assume that a consumer affiliated to product j receives a perceived benefit for that product, $\overline{\sigma}_{jt}$, and the benefit is uniform relative to the other products, i.e., $\overline{\sigma}_{jt} = \sigma_{jt}(j;1) - \sigma_{jt}(z';1)\ \forall z' \neq j$.

Thus, we define *affiliation* to be a product-specific state dependence in preferences. The model can be interpreted as brand loyalty when $\overline{\sigma}_{jt}$ is a positive level shock that reflects an internal benefit for purchasing from the same brand. Alternatively, the model may be interpreted as a switching cost model when $\overline{\sigma}_{jt}$ is a level shock representing the costs (physical and psychic) of switching to another brand. These two interpretations are empirically indistinguishable because only the relative utilities affect choices in the formulation of the discrete choice model. The model can also accommodate habit formation and a special case of search.[18] Distinguishing among these different mechanisms lies outside the scope of this paper but may be important, especially when examining questions about welfare. The brand loyalty and the switching cost models can have identical outcomes but divergent welfare predictions, as $\overline{\sigma}_{jt}$ is a net benefit in the former and a net cost in the latter.

We assume that consumers become affiliated with the product they purchased in the previous period, or, if they chose no product in the previous period, they are affiliated with the outside option, $j = 0$. Hence, in our model, affiliation is only a function of a consumer's previous choice, rather than a longer purchase history.[19] Based on these assumptions, we can represent the state of each market in each period by the vector $r_t = \{r_{jt}\}$, where r_{jt} denotes the fraction of $h = 1$ consumers that are affiliated to product j in period t.

The share of state-dependent consumers affiliated to product j in period $t+1$ can be denoted as:

$$r_{jt+1} = \sum_{z \in 0,J} r_{zt} s_{jt}(z;1) \tag{3}$$

Thus, the share of consumers that are affiliated to a product depends on previous period choices (and therefore prices) and the underlying distribution of types. The evolution of states follows

[18] In the habit formation interpretation, a consumer gets either an extra benefit for repeating earlier behavior or bears a cost for adjusting behavior. 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.

In the special case where $\overline{\sigma}_{jt}$ 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 standard search models (e.g., Varian, 1980; Stahl, 1989) by allowing for a mixture of consumers that search and those that do not. The fraction that do not search is endogenous and depends on past prices.

[19] We make this simplifying assumption in order for the model to be estimated with market-level data, rather than with individual-level panel data. In Section 3 we present evidence that in retail gasoline markets, consumers may switch brand loyalty based upon their most recent purchase, and loyalty is not entirely a long-run phenomenon.

a Markov process, where the state can be expressed as a function of the joint distribution of states, types, and choices in the previous period.

Assumption 4: Demand Specification To proceed with our analysis, we make additional parametric assumptions about utility. First, in equation (1), we specify $\delta_{jt} = \xi_{jt} + \alpha p_{jt}$, where ξ_{jt} is comprised of product and time (and, later, market) fixed effects.

Next, we assume that the affiliation shock affects the utility level by a constant amount for all products, $\bar{\xi}$, while we set the affiliation shock for the outside good to 0. Thus, we chose $\overline{\sigma}_{0t} = 0$ and $\overline{\sigma}_{jt} = \bar{\xi}\ \forall j > 0$. In addition, we need take a stance on the baseline utility levels that type $h = 1$ consumers receive for products they are not affiliated to. We assume that $\sigma_{jt}(z'; 1) = 0\ \forall z' \neq j$, so that type $h = 1$ consumers receive the same utility for products they are not affiliated to as shoppers. This rules out arbitrary persistent differences in preferences between the two types of consumers and implies that a consumer of type $h = 1$ who chooses the outside option has the same choice probabilities as a shopper in the subsequent period. Because of this equivalence, we suppress types in the choice probability expression $s_{jt}(z)$ going forward, denoting the choice probabilities of shoppers as $s_{jt}(0)$. We also refer to the combined set of shoppers and consumers with $z = 0$ as *unaffiliated*.

We can thus represent the utility of a consumer i in state z for product $j > 0$ is as follow:

$$u_{ijt}(z_{it}) = \xi_{jt} + \alpha p_{jt} + \mathbb{1}[j = z_{it}] h_i \bar{\xi} + \epsilon_{ijt}. \tag{4}$$

Here, z_{it} denotes the consumer's product choice in period $t-1$. Consumers subject to affiliation always have a value of $h_i = 1$, and shoppers always have a value of $h_i = 0$. We therefore represent $\overline{\sigma}_{jt} = \mathbb{1}[j = z_{it}] h_i \bar{\xi}$. Finally, we assume the error term, ϵ_{ijt}, follows the type-1 extreme value distribution, which yields the choice probabilities specified in equations (2) and (3).

2.2 Supply

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

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

The aggregate share of product j can thus be written as a weighted sum of its share of unaffiliated consumers, $s_{jt}(0)$, and affiliated consumers, $s_{jt}(z) \forall z \neq 0$. Note that the total weight on unaffiliated consumers in any period is $(1-\lambda) + \lambda r_{0t}$, as some fraction of state-dependent consumers may have chosen the outside option in the prior period. Note also that firm j will make sales to consumers affiliated to other firms, $z \neq j$, but the probability that such consumers will choose firm j is strictly lower than the choice probability of an unaffiliated consumer when the utility shocks $\{\sigma_{jt}(z)\}$ are positive.

Assumption 5: Competition in Prices We assume that firms set prices in each period to maximize the net present value of profits from an infinite-period game. Prices are set as a best response conditional on the state and contemporaneous prices of rival products. Firms cannot commit to future prices. The state vector in each period is summarized by marginal costs, c_t, the distribution of affilation across consumers, r_t, and other variables that are captured by the vector, x_t, such as expectations about future costs. Entry is exogenous. The objective function for firm k can be summarized by the Bellman equation:

$$V_k(c_t, r_t, x_t) = \max_{p_{kt}|p_{-kt}} \big\{(p_{kt} - c_{kt})S_{kt} + \beta E(V_k(c_{t+1}, r_{t+1}, x_{t+1})|p_t, c_t, r_t, x_t)\big\}. \tag{6}$$

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

Assumption 6: Expectations 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. In contrast to the typical setup for a dynamic game, we place minimal restrictions on expectations, discount rates, and the perceived continuation value. Instead, our empirical approach is to directly estimate a reduced-form model of (the derivative of) the continuation value. We describe this approach in more detail in Section 5.

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

2.3 Theoretical and Numerical Analysis

To assess equilibrium pricing incentives in markets with consumer inertia, we first consider a deterministic setting where marginal costs are constant. We use numerical methods to analyze steady-state prices in an oligopoly game with Bertrand price-setting behavior.[20]

[20] For the monopoly case, it is possible to obtain analytical results that can highlight the role of investment and harvest incentives. We present these results in Appendix A.1.

We employ this approach to help develop an understanding about equilibrium incentives. When we turn to the empirical approach in later sections, we consider an environment with time-series variation in demand and cost shocks. These shocks generate a dynamic equilibrium with prices that are not in the long-run steady state. We therefore implement a different simulation approach to assess price effects in our empirical setting.[21]

2.3.1 Steady-State Approach

For this analysis, we specify the utility function as in (4):

$$u_{ijt}(z) = \xi_j + \alpha p_{jt} + \mathbb{1}[j = z] h_i \bar{\xi} + \epsilon_{ijt}.$$

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

$$V_k(r) = \max_{p_k | p_{-k}} \pi_k(p, r) + \beta V_k(r'). \tag{7}$$

Here, p and r are vectors of prices and affiliated customers, respectively, and r' is a vector specifying each product's affiliated customers in the next period. In accordance with the model above, an element of $r' \equiv f(p, r)$ is $r'_j = \sum_{z \in 0,J} r_{zt} s_{jt}(z) = \frac{1}{\lambda}\left(S_{jt} - (1 - \lambda) s_{jt}(0)\right)$. Static profits are $\pi_k(p, r) = \sum_{j \in J_k} (p_j - c_j) \cdot s_j(p, r)$. We drop the expectations operator, as we consider a deterministic steady state.

To find the steady-state prices and affiliated shares for each firm, we focus on Markov perfect equilibrium.[22] Firm k's profit-maximizing first-order conditions are then:

$$\frac{\partial \pi_k}{\partial p_j} + \beta \frac{dV_k(r')}{dr'} \frac{dr'}{dp_j} \ \forall j \in J_k = 0. \tag{8}$$

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 conditions, 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:

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

In this equation, $\frac{\partial \pi_k}{\partial p}$, $\frac{\partial \pi_k}{\partial r}$, $f_p(p, r)$, $f_r(p, r)$ are known, conditional on values of p and r.

[21] In the empirical application, after estimating demand, if we were to set demand and cost shocks to zero, then we could use the same approach as in this section to solve for counter-factual prices.

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

The remaining unknowns are $\frac{dp}{dr}$ and $\frac{dV_k(r)}{dr}$. To solve the model, we impose the steady-state condition governing the evolution of affiliated customers, $r' = r$. The full set of steady state conditions, provided by equation (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 the values of $\frac{dp}{dr}$ numerically using a local approximation method. First, we provide a guess for the matrix $\tilde{\frac{dp}{dr}}$ and calculate the implied steady-state values for p, r, and $\frac{dV_k(r)}{dr}$, leveraging the J first-order conditions. We then numerically differentiate p with respect to r to obtain an estimate $\hat{\frac{dp}{dr}}$. If the absolute distance between the guess $\tilde{\frac{dp}{dr}}$ and the implied estimate $\hat{\frac{dp}{dr}}$ are identical up to rounding error, a solution is found. Otherwise, we take the average of $\tilde{\frac{dp}{dr}}$ and $\hat{\frac{dp}{dr}}$ as the guess in the next iteration. For additional details, see Appendix A.2.

Consistent with Dubé et al. (2009), equilibrium prices may be increasing or decreasing for different levels of affiliation.[23] With greater affiliation, firms face less elastic demand, but the incentive to invest in future demand tends to increase. Whether or not prices increase with affiliation depends on the relative weights on these forces and underlying market parameters. Importantly for our context, these forces interact with the market structure, so that the presence of affiliation can have differential effects on prices post-merger.

Consider a merger in which firm k acquires product b and maintains it as a separate brand. In the case of static demand, the post-merger change in pricing incentives for a product $j \in J_k$ at the pre-merger equilibrium prices is given by $\sum_{l \in \{J_k, b\}} \frac{\partial \pi_l}{\partial p_j} - \sum_{l \in J_k} \frac{\partial \pi_l}{\partial p_j} = \frac{\partial \pi_b}{\partial p_j}$. In the case of substitutes, $\frac{\partial \pi_b}{\partial p_j}$ is greater than zero and the firm raises prices.

For dynamic demand, the post-merger change at the pre-merger steady-state prices is instead:

$$\textit{Change in dynamic first-order condition:} \quad \frac{\partial \pi_b}{\partial p_j} + \beta \left(\frac{d\tilde{V}_k(r')}{dr'} - \frac{dV_k(r')}{dr'} \right) \frac{dr'}{dp_j}, \tag{10}$$

where $\tilde{V}_k$ now also incorporates the discounted flow of profits from product b. The first term, $\frac{\partial \pi_b}{\partial p_j}$, is equivalent to the change in the static first-order condition, though it is evaluated at different equilibrium prices. The second term may be positive or negative, and thus consumer dynamics can either exacerbate or mitigate incentives to raise prices post merger. On the one hand, the acquiring firm will internalize the fact that an increase in price for product j will increase the number of customers affiliated to the new brand b, increasing future profits for that brand. On the other hand, such a price increase would reduce the affiliated customers of brand j. This has an indirect effect of incentivizing lower future prices for all firms in the market

[23]We provide specific examples to illustrate this in Appendix A.3.

and reducing future profits for the acquired brand.[24] The tradeoff can be conceptualized as balancing the re-capture of lost (affiliated) customers by the acquired brand versus the desire to maintain a higher stock of affiliated consumers to soften competition. The characteristics of the market and whether the acquirer maintains separate brands determines how consumer dynamics affect post-merger price increases.

2.3.2 Distinguishing Types of Mergers

In the presence of consumer inertia, it is important to define precisely the implementation of a horizontal merger. We consider two types of mergers. The first type of merger unites pricing control of two products under a single firm and the merged firm maintains these as separate products. We refer to this as a *joint pricing* merger. This is a common setting in many differentiated product mergers when the combined firms maintain separate brands/products. The second type of merger consolidates two products under a single brand and effectively offers a single product after the merger. We refer to this type of merger as *brand consolidation*. This type of merger typically occurs in settings where retail location is important and the post-merger firm rebrands all locations under one brand. This is often the case in retail gasoline mergers, but also occurs in other industries, such as wireless phone service (T-Mobile/Sprint) and airlines (American Airlines/USAir).

To implement a brand consolidation merger, we need to make an assumption about the utility that consumers of the removed product receive from buying the remaining brand. We assume that, at pre-merger prices, the consolidated brand will have the same combined share of shoppers as the separate pre-merger brands. We assume that consumers that were affiliated to the removed brand transition to state 0 following the merger. To implement this assumption, we adjust the value of ξ_j for the remaining brand of the post-merger firm.[25] We make this assumption in order to make an "apples-to-apples" comparison to joint pricing mergers. In the context of retail gasoline markets, this is akin to assuming that shoppers do not derive direct value from the brand, but from other features such as the location. One implication of our assumption is that we de-emphasize the preference for brand variety that arises in a direct interpretation of the logit model.

A second implication of this assumption is that at pre-merger prices, the share of affiliated customers increases for the remaining brand relative to the two separate brands. Given our adjustment to ξ_j, the effective choice probabilities for the affiliated consumers post-merger have an intuitive representation: they are equivalent to the sum of the choice probabilities of a pre-merger consumer that was affiliated to both of the merging brands.[26]

[24]When the elements of $\frac{d\tilde{V}_k(r')}{dr'} - \frac{dV_k(r')}{dr'}$ are positive, the tradeoff depends on the magnitudes of these elements weighed by the elements of $\frac{dr'}{dp_j}$, which are positive except for $\frac{dr'_j}{dp_j}$, which is negative.

[25]See Appendix B for technical details of the implementation.

[26]Since our model only allows for affiliation to a single brand, this is a hypothetical comparison. This can be

In practice, however, alternative assumptions on the utility provided by the post-merger brand can be incorporated into our model. For example, in a merger investigation, an antitrust authority may learn that the merging parties are planning to remove a brand from the market. Furthermore, they may also obtain an estimate of what fraction of customers are likely to transition over to the acquiring firm. In such a case, our model may be used to estimate the impact of brand consolidation. The parameters ξ_j and $\bar{\xi}$ can be calibrated so that the post-merger share of the remaining product matches the antitrust agency's estimate of the probability that a consumer switches to the new brand. This adjustment can be done more generally, even to reflect the aggregate shares of both affiliated and unaffiliated consumers.

2.3.3 Dynamic and Horizontal Market Power

We use numerical simulations to decompose the potential impacts of dynamic and horizontal market power in the presence of consumer inertia. To measure dynamic market power, we compare the three-firm oligopoly price with consumer affiliation to a baseline price where consumers have no state dependence, but the markets are otherwise identical. To measure horizontal market power, we compare the three-firm oligopoly price to the price that prevails following a merger between firms one and two.[27]

We attempt to simulate data from the support of parameters that produce reasonable outcomes for margins and shares. We employ a "shotgun" approach, generating simulations with many different parameters and selecting only the markets that meet certain criteria. We first take Halton draws of the demand parameters such that $\xi \in [0, 10]$, $\bar{\xi} \in [0, 10]$, $\alpha \in [-10, 0]$, and set each firm's marginal cost to one. For each draw of these demand parameters, we construct three-firm markets for $\lambda \in \{0.05, 0.1, 0.15, ..., 0.70\}$. We then restrict the analysis to markets where firms have shares between 0.05 and 0.30 (yielding an outside share between 0.10 and 0.85) and margins between 0.05 and 0.75.[28] Finally, to avoid composition affects, we only analyze markets with demand parameters that converged for all values of λ.[29] The data generating process yields 6,566 markets whose parameters are summarized in Appendix Table 9.

In Figure 1, we plot the effects of affiliation and reduced competition on prices. The plots employ simulation results from the 469 baseline parameter values of $\xi, \bar{\xi}$, and α that converged for all $\lambda \in [0.05, 0.70]$, or 6,566 markets. Panel (a) measures dynamic market power by plotting price effects as a function of λ. Percent changes are calculated relative to no consumer

shown by considering the following. For a merger between firms 1 and 2, we adjust the merged value of ξ_{mt} so that $\exp(\xi_{mt}) = \exp(\xi_{1t}) + \exp(\xi_{2t})$ at given price levels. It is also then the case that $\exp(\xi_{mt} + \bar{\xi}) = \exp(\xi_{1t} + \bar{\xi}) + \exp(\xi_{2t} + \bar{\xi})$. Plugging this into the standard share equations, it is straightforward to confirm the above statement.

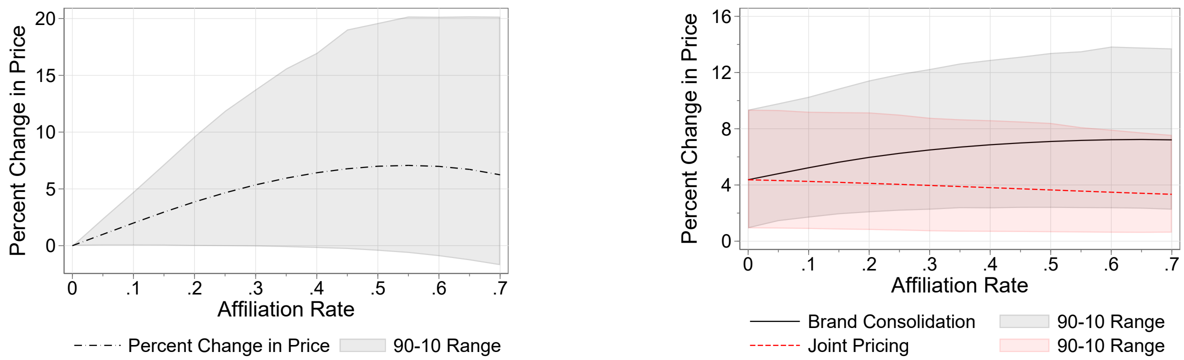
[27] Note that we use a (competitive) oligopoly price as a baseline, rather than price equal to marginal cost. Also, as marginal cost is constant across simulations, using price as a measure of market power is equivalent to the commonly used markup or margin metric.

[28] The range for each parameter is selected such that parameter values just outside the bounds of the range result in outcomes that often fall above or below our share and margin criteria.

[29] We set the upper bound of λ to be 0.70, because for higher values of λ there were often markets where the post-merger equilibrium did not converge or pre-merger margins fell above 0.90.

Figure 1: Potential Price Effects

(a) Dynamic Market Power (b) Horizontal Market Power



Notes: Panel (a) displays the mean percent price increase and for a three-firm oligopoly above the baseline model with no dynamics in consumption ($\lambda = 0$). Panel (b) displays the mean percent price increase of a merger to a duopoly for two types of mergers: joint pricing control and brand consolidation, for different values of λ. The plots reflect 469 baseline parameter values of $(\xi, \bar{\xi}, \alpha)$ that converged for all $\lambda \in \{0.05, 0.1, 0.15, ..., 0.70\}$, or 6,566 markets in total.

affiliation ($\lambda = 0$) while holding fixed the other parameters in the model. On average, prices increase with the fraction of customers prone to affiliation for values of $\lambda \leq 0.55$ and decrease thereafter. At lower values, a marginal increase in λ raises the "harvest" incentive more than the "invest" incentive, leading to higher prices. However, above some threshold, an increase in λ yields a relatively stronger incentive to invest in future demand, and prices fall. In our simulations, both market share and profits increase monotonically with λ (not shown). On average, affiliation results in moderate price increases compared to a static demand model. However, the impact of dynamic market power can be quite substantial depending on the underlying demand parameters. The 90-10 range of outcomes is plotted with the transparent area. When $\lambda > 0.4$, the 90th percentile market exhibits dynamic market power resulting in prices more than 15 percent higher than a static model.

In panel (b), we plot the price effect of horizontal market power on prices across different values of λ. We measure horizontal market power by comparing the prices in the symmetric three-firm oligopoly to the post-merger prices in both joint pricing and brand consolidation mergers. In our symmetric setting, brand consolidation mergers provide greater horizontal market power relative to joint pricing mergers across all levels of λ. This is in part due to how we specify brand consolidation mergers, as we assume the consolidated brand retains the same combined share of shoppers. Still, there is large overlap in the 90-10 percentile range across the two merger types, demonstrating that both can enable comparable levels of horizontal market power. Another interesting feature of panel (b) is that horizontal market power slightly decreases, on average, in joint pricing mergers, but it increases in brand consolidation mergers.

This highlights that the interaction between consumer inertia and a decrease in competition depends critically on the mechanism through which competition is reduced.

2.3.4 Implications for Mergers and Enforcement

The above analysis demonstrates that brand consolidation and joint pricing mergers can have diverging effects in the presence of consumer inertia. In these simulations, the average percentage merger price effect is 3.8 percent for joint pricing mergers and 6.5 percent for brand consolidation mergers.

By contrast, in a symmetric oligopoly, the joint pricing and brand consolidation mergers we consider have identical price effects when demand is characterized by the static logit model. We provide a proof in Appendix B. Thus, precisely modeling the structure of a horizontal merger may have heightened importance in markets characterized by consumer inertia, and failing to account for consumer inertia could have meaningful effects on antitrust enforcement.

To illustrate this, 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. This data is then used to recover the demand parameters of the standard logit model (without affiliation), and then the price effects of a merger are simulated. We perform this experiment for each of the numerically generated markets, and consider both joint pricing and brand consolidation mergers.

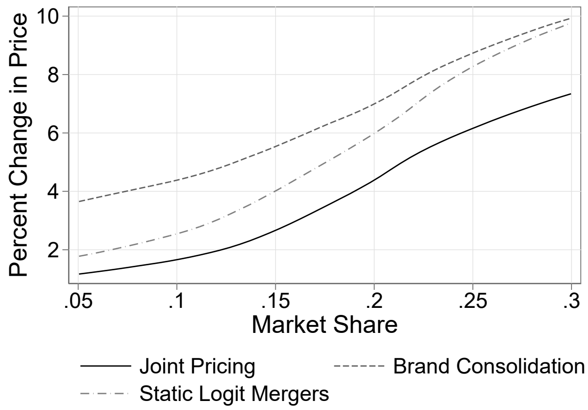
The "average" market is one that, a priori, would typically raise moderate concern from the US antitrust agencies; HHI (1067) falls in the "unconcentrated" range, but the change in HHI (712) generally warrants a thorough investigation. The average pre-merger difference between price and cost is 0.26, and the mean market share is 0.17. We use these mean values to calculate the "Upward Pricing Pressure" index, $\frac{0.17}{1-0.17} \cdot (1.26-1) = .053$, which is just over the threshold that may trigger an investigation.[30] The full range of markets spans those that would receive no scrutiny and those that almost certainly would be challenged. Thus, the simulations generate a reasonable set of markets within which to explore merger effects in the consumer affiliation model. Additional statistics are reported in Table 10 in the Appendix.[31]

Figure 2 plots the merger price effects related to the pre-merger market share of each symmetric firm. To generate the graph, we run a local polynomial regression of the merger price effect on one of the symmetric firm's pre-merger market share. We generate fitted lines for (i) the joint pricing merger effect, (ii) the brand consolidation merger price effect, and (iii) a

[30]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.

[31]Tables 11 and 12 in the Appendix explore the extent to which demand parameters can predict equilibrium prices using reduced-form regressions. On average, increasing the rate of affiliation (λ) and the strength of the affiliation ($\bar{\xi}$) tends to increase pre-merger prices, but their impact on the effect of a merger is dependent upon the type of merger. Higher values for λ and $\bar{\xi}$ decrease prices in joint pricing mergers but increase prices in brand consolidation mergers. However, these relationships do not hold in every instance.

Figure 2: 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.

misspecified static logit model (which generates the same price effect, regardless of the merger type). In line with intuition, the price effect of a merger is increasing with pre-merger market shares.

The figure indicates that, on average, the static model yields a prediction that is biased upward relative to joint pricing and biased downward relative to brand consolidation. The average prediction bias, defined as the static prediction minus the price effect, is 1.5 and -1.2 percentage points for joint pricing and brand consolidation mergers, respectively. Scaled by the magnitude of the true price effect, the mean bias is 67.3 and -19.7 percent for the two types of mergers. Thus, for both types of mergers, incorrectly assuming static demand will lead to substantially biased predictions.

These biases can primarily be explained by shifts in the "harvesting" and "investment" incentives for the merged firm and its rivals. For these simulations, we find that the magnitude of the bias increases with market share with respect to joint pricing mergers, but decreases with respect to brand consolidation mergers. This indicates that the static model fails to capture the shift in incentives toward invest for joint pricing, and this missattribution is greater at higher market shares. On the other hand, the static model fails to account for the shifts in the incentives toward harvest for the brand consolidation merger, and the bias is greater with lower market shares.

The average direction of the bias displayed in Figure 2 need not hold in every market. For brand consolidation mergers, static models tend to under-predict the true dynamic price effect, but almost 25 percent of simulations resulted in over-predictions. For joint pricing mergers, all of the included markets showed that static models resulted in over-predictions, but we were able to obtain under-predictions with other parameterizations. Appendix Figure 7 plots the distributions of bias across all 6,566 markets. Finally, we note that these results reflect a stylized case with a three-firm symmetric oligopoly. Still, these findings highlight the importance of properly accounting for dynamics when simulating the price effects of mergers.

In general, model misspecification can yield biased predictions because the underlying demand elasticities are biased. In this case, that is not the primary factor. Note that we obtain the same (misspecified) elasticity for the static model, but the direction of bias can go in opposite directions depending on the type of merger. 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.[32]

The above results suggest that affiliation has implications for counterfactual exercises, such as merger simulation. 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. These results highlight the benefit of an empirical model that can account for consumer dynamics, which we pursue in the following sections.[33]

3 Data and Reduced-Form Evidence of Dynamics

We now introduce the data used in the demand estimation and empirical application. To motivate the empirical application, we also 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.[34] 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 main empirical 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 as federal, state, and local taxes. We therefore almost perfectly observe each gas station's marginal cost changes, except for privately negotiated discounts

[32] The static model is calibrated to be more elastic, on average, than the share-weighted elasticity in the dynamic model (-4.74 vs. -3.86). More elastic demand generates smaller merger price effects in the logit model.

[33] For certain applications, an informal analysis of consumer dynamics may provide a useful indication of static model bias. Our results show how price predictions from static merger simulations could be revised downward when only pricing control is expected to change in a merger. In Appendix A.6, we provide more detail on how dynamics affect simulation bias. Therein, we again find that the relationship between the strength of dynamics and bias is a function of the type of merger under evaluation.

[34] For a review, see Eckert (2013).

per-gallon that are likely fixed over the course of a year. The data ranges from September 25th, 2013 through September 30th, 2015. The data was obtained directly from the Oil Price Information Service (OPIS), which has previously provided data for academic studies (e.g., Lewis and Noel, 2011; Chandra and Tappata, 2011; Remer, 2015).

OPIS also supplied market share data. This proprietary data is the standard used by industry participants to track local market shares.[35] The data is reported by week and county for each gasoline brand. In our analysis, we treat consumers as choosing among brands in a county.[36] Due to contractual limitations, OPIS only provided each brand's inside market shares, 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). We describe our adjustment in Section 4.3. To account for differences in demographics across markets, we merge the data with measures of income and population density from the American Community Survey.

To provide additional evidence to support the presence of consumer inertia in retail gasoline, we first document purchasing patterns using an additional NielsenIQ dataset. We employ the consumer panel data, which allows us to track individual household purchasing decisions over time. These data do not include prices or provide comprehensive market estimates, and we use this data only to document patterns of repeat purchases.

3.2 Dynamic Demand: Evidence from Consumer Data

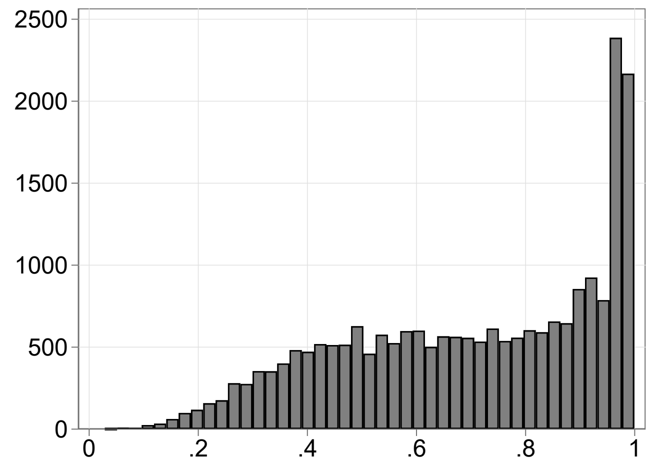
In this section, we use household-level data to provide evidence of consumer inertia in retail gasoline markets. Using the Nielsen Consumer Panel data, we analyze household purchases of gasoline from 2007 through 2018. With this data, we find patterns consistent with consumer state dependence.[37] We make the following sample restrictions. First, we exclude very small and very large purchase amounts, dropping trips where households spend less than $10 or more than $120 on gasoline. We exclude gasoline purchases from warehouse clubs and grocery stores, as gasoline is more likely to be a secondary reason for the trips in those cases. After these restrictions, we limit the analysis to households that make at least 26 purchases of gasoline in

[35] See https://www.opisnet.com/product/pricing/retail-fuel-prices/marketsharepro/. OPIS calculates these data from actual purchases that fleet drivers charge to company cards issued by Wright Express (WEX). WEX is the largest provider of fleet cards (to businesses that have fuel expenses) in the United States, with over 4 million drivers and 95 percent coverage of fuel retailers (https://www.wexinc.com/about/partner/fuel-partners/). OPIS indicates that there is a high correlation between these market share data and actual retail volumes, which they obtain directly from some retailers but cannot share due to contractual restrictions. These data may understate purchases from certain types of stations, such as supermarkets and wholesale clubs, but the fact that gasoline brands use it to track the shares of rivals makes it suitable for our purposes.

[36] 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.

[37] Our data do not allow us to distinguish among different mechanisms driving consumer inertia, such as brand loyalty or habit formation. It is also possible that preferences for a brand of gasoline may stem, in part, from gasoline being an "experience" good, especially for new brands.

Figure 3: Fraction of Repeat Purchases by Household



Notes: This figure depicts the distribution of households by fraction of repeat purchases. For each household we calculate the average fraction of gasoline purchases that are a repeat purchase, defined as returning to the same brand as the previous purchase. We restrict the sample to households with at least 26 gasoline purchases in a year.

a given year, equivalent to one trip every two weeks. 63 percent of trips are from households that meet this threshold. We obtain qualitatively similar evidence of dynamics without these sample restrictions.

Within this sample, the median time between purchases is 7 days, and the mean is 8.6 days. That is, the typical consumer makes one purchase each week. In over 90% of cases, the household only purchases gasoline during the trip. In the remaining cases, the household may additionally purchase snacks, lottery tickets, etc.

For a high-level perspective on state dependence, we first explore the propensities of households to return to the same brand. We define a repeat purchase to be when a household purchases gasoline from the same brand as they did in their previous purchase. Then, for each household, we calculate the fraction of gasoline purchases that are repeat purchases. Figure 3 plots the distribution of the fraction of repeat purchases across households. There are three characteristics of this plot that are consistent with our model of consumer inertia. First, repeat purchases are exceedingly common in retail gasoline markets. For the median household, 73 percent of their trips are repeat purchases. Second, their is a mass of households that shop almost exclusively at a single brand: the 90th percentile household has 98 percent of their gasoline purchases categorized as a repeat purchase. Finally, there is a lot of variation in the tendency to make a repeat purchase, including a large number of households where less than 50 percent of their consumption is a repeat purchase. Overall, we view this as evidence of consumer inertia, with different "types" of households; some that exhibit strong brand affiliation and others that tend to shop around.

Next, we exploit the panel data to study how inertia might affect individual households. We look at the sequence of purchase patterns, and, in particular, what happens after consumers switch retailers/brands. To frame the analysis, consider a hypothetical scenario in which a household shops from brand A two-thirds of the time and brand B one-third of the time. If there were no consumer inertia, then we would expect the probability that the household chooses brand A for a purchase would be 0.667, regardless of which brand they chose in the previous period. However, if consumer inertia is present, then we would expect the probability to depend on the identity of the previous brand. One possibility is that the household chooses A with probability 0.8 if it purchased from A previously and with probability 0.4 if it purchased from B previously; such a process yields an overall probability of choosing A about two-thirds of the time.

To show these dynamics in the data, we identify all spells where a household purchases from the same brand for at least 3 consecutive trips. This provides 156,150 distinct spells. In these spells, the probability that the next trip is a repeat purchase, excluding the first three trips, is 0.903.

We then analyze household choice probabilities after the spell has ended, and a different brand has been chosen. We define the brand chosen during spell the as brand A and the brand that was switched to after the spell as brand B. After having switched to brand B, the probability that the next purchase is from brand A is 0.565, much lower than the probability of staying with A conditional on purchasing from them in the previous period.[38] This difference is consistent with consumer inertia shifting the purchase probabilities of households and can be rationalized by a discrete choice model where the utility of a particular brand depends on the brand chosen in the previous period. The probabilities likely reflect a mix of those prone to inertia as well as those that are not, as well as other factors, such as price changes and entry.

We also find that the probability of staying with brand B, after having switched to them with the previous purchase, is 0.266, while the probability of purchasing from any another brand, C, is 0.169. The greater odds of purchasing from the new brand is also suggestive of consumer dynamics. These statistics are summarized in Table 1.

We also leverage the data on the duration between purchases to provide additional evidence on the role of consumer inertia. In our model, consumers that choose the outside option lose their affiliation to a particular brand. In the data, this would be captured by longer periods between purchases, when the consumer chooses not to purchase. We denote the number of days since the previous purchase as the "purchase interval." Longer intervals could arise due to idiosyncratic, household-specific shocks, or (e.g.,) higher prices.

Consistent with our model of consumer inertia, households are less likely to return to the same brand after a longer interval. Moreover, this effect is strongest for households that make

[38] This is also fairly close to the average share of purchases from the prior brand over the course of the year, which is 0.592.

Table 1: Evidence of Dynamic Patterns in Shopping Behavior

Variable	Value
Probability of Repeat Purchase After At Least Three Consecutive Trips	0.903
Conditional on Switch, Probability of Switching Back to Prior Brand	0.565
Conditional on Switch, Probability of Staying with New Brand	0.266
Conditional on Switch, Combined Probability of of Staying with New Brand or Switching Back	0.831

Notes: Each statistic is calculated from spells that have at least three consecutive trips to the same brand.

repeat purchases more frequently, and the effect is zero for households that are the least likely to make repeat purchases. This is consistent with our assumption that some consumers are "shoppers" and are unaffected by inertia. We provide additional details of this analysis in Appendix C.1.

These purchase patterns can be rationalized by two key features of our model. First, our model allows for latent consumer types: those that exhibit state dependence and those that do not. Second, the probability of purchasing from a brand increases if the consumer purchased from the brand in the previous period. The dynamics in behavior documented above are more challenging to rationalize with only static unobserved heterogeneity in preferences.

3.3 Dynamic Pricing

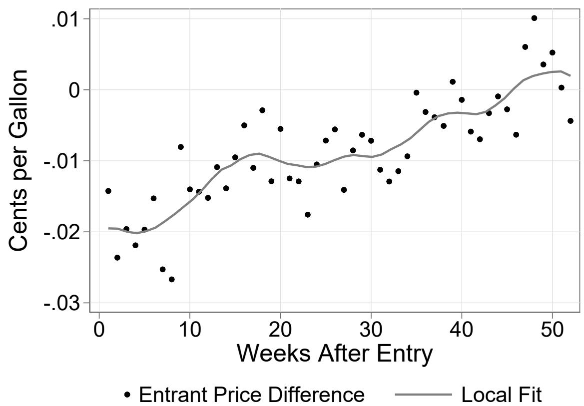
We now present reduced-form evidence of dynamic pricing. Consistent with a model where firms accumulate affiliated consumers over time, we find that new entrants price lower relative to established competitors in the same market, and that this discount dissipates over time. Second, we examine cost pass-through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

3.3.1 Dynamic Pricing of New Entrants

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

We test for and find evidence consistent with this dynamic pricing pattern in the data. To perform the analysis, we first identify a set of new entrants, defined as a gas station whose first price observation is at least six weeks after the start of the data and does not exit in the

Figure 4: 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.

remainder of the sample. To ensure there is sufficient data and to control for composition effects in the analysis, we limit the set of entrants to those with at least one year of post-entry price data. Using this filter, we identify 193 entrants. We compare prices for entrants to the 594 stations in those counties that are present for the entire sample.

Figure 4 depicts the average difference between an entrant's price and all other stations' prices 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. For the first 8 weeks following entry, new entrant prices are on average 2.1 cents per gallon lower than incumbents' (standard error: 0.24). From weeks 9 through 24, entrant prices are 1.1 cents per gallon lower (standard error: 0.17). These differences are highly significant, and, based on our empirical demand estimates, are economically meaningful for attracting unaffiliated consumers. Our model predicts this behavior, as a profit-maximizing firm would initially price lower attract shoppers and raise its price over time as consumers became affiliated and less elastic. Though this pattern could be driven by other mechanisms, it is another piece of evidence that is consistent with the model of consumer inertia.

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.[39] 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:

$$p_{nt} = \sum_{s=-50}^{50} \beta_s \hat{c}_{n(t-s)} + \sum_{s=-50}^{50} \gamma_s \tilde{c}_{n(t-s)} + \sum_{s=-50}^{50} \phi_s \tau_{n(t-s)} + \psi_n + \varepsilon_{nt}. \tag{11}$$

Here, p_{nt}, is the price observed at gas station n at time t. $\hat{c}_{n(t-s)}$ and $\tilde{c}_{n(t-s)}$ are the expected and unexpected wholesale costs observed with lag s, and $\tau_{n(t-s)}$ is the state-level sales tax.[40] 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 5 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. Estimated pass-through peaks at 0.72 after 34 days, with an average of 0.64 over days 21 through 50.

Panel (b) displays the cumulative response function for expected costs. Notably, firms begin to react to expected costs approximately 28 days in advance, with a relatively constant adjustment rate until the new long-run equilibrium passthrough is reached 21 days after the shock. The estimated pass-through rate averages 1.01 over days 21 through 50. Though the total duration of adjustment is longer compared to the unexpected cost shock, the firm incorporates the cost more quickly after it is realized. This coincides with substantial anticipation by the firm; the price already captures about 40 percent of the effect of the expected cost shock the day before it arrives.[41]

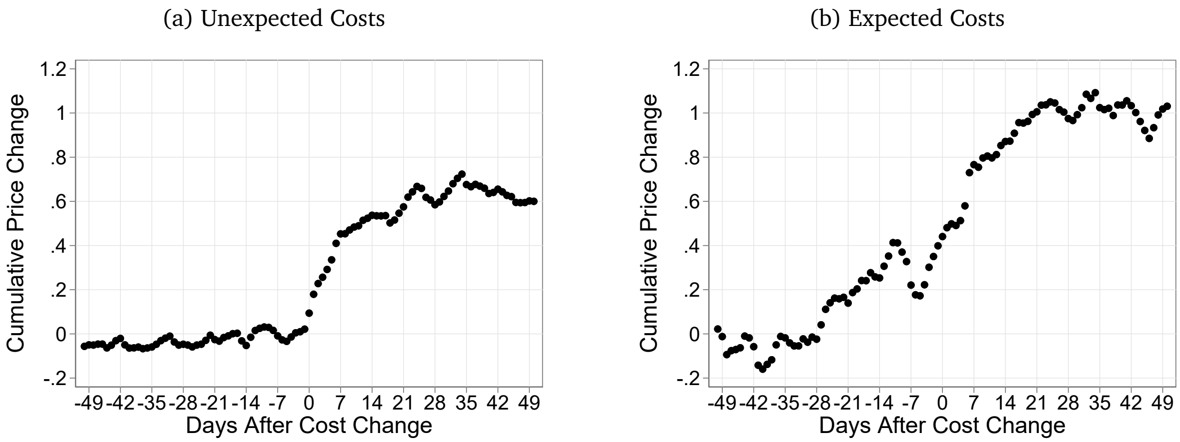
Thus, a reduced-form analysis of pricing behavior shows that retail gasoline prices adjust slowly to changes in marginal cost, and also that price changes anticipate expected changes in marginal costs. These patterns are consistent with forward-looking behavior by firms and

[39] For details, see Section C.3 in the Appendix.

[40] 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.

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

Figure 5: 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 (11).

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 $\{s_{jt}(z)\}$) conditional on observed shares S_{jt}, the share of consumers subject to state dependence λ, and the strength of the affiliation shock $\sigma_{jt}(z)$. Second, the vector $\{s_{jt}(z)\}$ allows us to recover the mean utility for unaffiliated consumers and estimate the static demand parameters. Third, we discuss the assumptions that allow us to identify the dynamic parameters.

Identification of Type-Specific Choices

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

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

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

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

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

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

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

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

To show identification, we substitute equation (13) into (14) and use the fact that $\frac{1}{s_{0t}(j)} - \frac{1}{s_{0t}(0)} = \exp(\delta_{jt})(\exp(\sigma_{jt}(j)) - 1)$ to obtain the following two relations:

$$s_{jt}(z) = s_{0t}(z)\frac{s_{jt}(0)}{s_{0t}(0)} \cdot \exp(\sigma_{jt}(z)) \tag{15}$$

$$s_{jt}(0) = \left(\frac{s_{0t}(0)}{s_{0t}(j)} - 1\right)\frac{1}{\exp(\sigma_{jt}(j)) - 1}. \tag{16}$$

Thus, we show that the $J + J^2$ unknowns $\{s_{jt}(z)\}_{|j\neq 0}$, 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 $1 - \sum_j s_{jt}(0) - s_{0t}(0) = 0$ and the observed share equations, which provide the other

J restrictions:

$$S_{jt} = (1-\lambda)\left(\frac{s_{0t}(0)}{s_{0t}(j)} - 1\right)\frac{1}{\exp\left(\sigma_{jt}(j)\right) - 1} + \lambda \sum_{z=0}^{J} r_{zt} s_{0t}(z) \frac{s_{jt}(0)}{s_{0t}(0)} \cdot \exp\left(\sigma_{jt}(z)\right) \tag{17}$$

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

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

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

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

Identification of Static Demand Parameters

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

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

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

Identification of Dynamic Demand Parameters

We have so far shown exact identification of static demand parameters conditional on the parameters governing unobserved heterogeneity. To identify λ_m and $\sigma_{jt}(j)$, we need to employ additional moments. We use fixed effects to model unobserved serial correlation in demand, and we calculate the residual demand innovations, η_{jmt}, after accounting for these fixed effects. Specifically, we include fixed effects that capture aggregate period-specific demand shocks, product-specific persistent demand, and market-specific seasonal patterns. We then assume that the residual demand innovations are uncorrelated over time. We assume that $E[\eta_{jmt} \cdot \eta_{jm(t+1)})] = 0$ holds on average within each brand (j), which provides us sufficient moments (16) to identify our four dynamic parameters.[42] The parameters are pinned down by the patterns of serial correlation in the data and the entry of brands into new markets.

In the context of our model, demand may be serially correlated due to persistent brand-market preferences, aggregate period-specific demand shocks, and market-specific seasonal patterns—all of which are captured by fixed effects. We attribute the residual serial correlation to consumer inertia, where future demand is shaped by the pricing decisions of firms. Thus, we load the systematic autocorrelation in residual demand innovations to the endogenous response of consumers, rather than treating such correlation as a feature of an exogenous stochastic process. Our assumption would be violated if brands systematically realized brand-specific transitory demand shocks within local markets. Depending on the degree to which this is the case, our results may be thought of an "upper-bound" on the impact of consumer inertia.

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

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

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

Thus, we allow the share of consumers subject to state dependence to vary with market-level measures of median household income and (log) population density. This specification allows for the possibility that consumer characteristics, as captured by income and population density, affect the prevalence of consumer inertia.[43] We assume that affiliated customers receive a constant level shock to utility $\bar{\xi}$.

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

[42] In estimation, we use the sum of the covariance and the correlation moments (squared), weighting each contribution by the number of observations in the data for that brand-market combination. We use both covariance and correlation to reduce the influence of outlier observations. 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.

[43] There are a number of studies finding that consumer behavior in retail gasoline markets are affected by income. See, for example, Nishida and Remer (2015), Levin et al. (2017), and Donna (2021). Population density can capture the differences in behavior for consumers that choose to live in more urban versus more rural areas, as well as serving as a proxy for different commuting patterns in high-traffic areas.

market. The parameters $(\theta_1, \theta_2, \theta_3)$ are identified by how these serial correlation patterns covary with demographic characteristics.

The presence of entry aids the identification of the dynamic parameters. In our model, new brands have zero affiliated consumers when they enter the market. The presence of dynamic demand allows these new entrants to initially price lower (as shown in Figure 4) and have low shares, and, over time, increase both prices and shares as they accumulate affiliated consumers. The estimated parameters reflect the magnitudes of these patterns. In the 241 markets in our data, 90 experience the entry of a new brand.

Monte Carlo Simulations

To support the identification arguments, we use Monte Carlo simulations to demonstrate that our approach indeed recovers the correct dynamic parameters when state dependence is present. We generate simulated data for 6 brands across 50 regions and 100 periods with different values for λ. The same estimation approach used in the empirical section of this paper recovers the correct values of the dynamic parameters, $\bar{\xi}$ and λ. We report the details of these exercises in Appendix D.

One potential concern is that the presence of persistent unobserved heterogeneity, which can generate autocorrelation in individual purchasing behavior, may be falsely attributed to state dependence in estimation. We use additional Monte Carlo exercises to address this concern directly. We simulate markets with random coefficients logit demand and no state dependence, and we show that our estimation routine correctly and precisely estimates zero state dependence ($\lambda = 0$). In our modeling framework, the bias from ignoring random coefficients loads onto the static parameters (price coefficient and intercept), not the dynamic parameters. Intuitively, this is because the econometric model picks up the influence of past prices on demand in future periods and can identify when this effect is absent.

Finally, we use the Monte Carlo exercises to highlight the potential impacts of model misspecification. Using the simulated data, we estimate demand elasticities using a static logit model when the underlying model has persistent unobserved heterogeneity or state dependence. As one might expect, the estimated elasticities are biased in each case. However, we find that the degree of bias is greater for the specifications with (unaccounted for) state dependence. One reason for this is that, with state dependence, the prices reflect dynamic considerations in addition to the demand elasticities.

These exercises motivate the potential importance of accounting for state dependence and the value of our modeling framework, which can correctly identify within-consumer state-dependence even when between-consumer heterogeneity is richer than our model allows.

4.2 Implementation Details

For our estimation approach, we estimate the dynamic parameters $(\theta_1, \theta_2, \theta_3, \bar{\xi})$ and static demand parameters (α, π, γ). As discussed around equation (20), the static demand parameters can be identified using an instrumental variables regression, after extracting the (latent) shares of the unaffiliated consumers. We make use of this method to efficiently estimate multilevel fixed effects, which are captured by γ.

Our estimator follows a method-of-moments approach built around the assumption that the residual shocks are uncorrelated. Specifically, our objective function is constructed from the empirical analogs of the covariance moment $E[\eta_{jmt}\eta_{jm(t+1)}] = 0$ and the correlation moment $\frac{E[\eta_{jmt}\eta_{jm(t+1)}]}{\sqrt{E[\eta_{jmt}^2]}\sqrt{E[\eta_{jm(t+1)}^2]}} = 0$. In testing, we found that using both sets of moments helped to mitigate the effect of extreme values in our empirical setting. We aggregate both of these by brand, which yields $16 \times 2 = 32$ underlying moments. We weight each moment by the number of observations when constructing the objective function.

Formally, the estimate of the dynamic parameters $\tilde{\theta} \equiv (\theta_1, \theta_2, \theta_3, \bar{\xi})$ is given by

$$\hat{\theta} = \arg\min_{\tilde{\theta}} g(\tilde{\theta})' W g(\tilde{\theta}), \qquad g(\tilde{\theta}) = \begin{bmatrix} g^{cov}(\tilde{\theta}) \\ g^{corr}(\tilde{\theta}) \end{bmatrix} \tag{23}$$

where $g^{cov}(\tilde{\theta})$ and $g^{corr}(\tilde{\theta})$ are each $J \times 1$ and are the sample analogs of $E[\eta_{jmt}\eta_{jm(t+1)}]$ and $\frac{E[\eta_{jmt}\eta_{jm(t+1)}]}{\sqrt{E[\eta_{jmt}^2]}\sqrt{E[\eta_{jm(t+1)}^2]}}$, respectively. W is a diagonal matrix where the entries are the square root of the number of observations for corresponding brand j over the square root of the total number of observations.

The static parameters are identified by a nested regression within this outer loop using standard exogeneity conditions for X_{jmt} and the exclusion and relevance conditions to instrument for p_{jmt} and $(p_{jmt} \times Income_{jmt})$. The routine proceeds in the following steps:

1. Pick a value for the parameters $(\theta_1, \theta_2, \theta_3, \bar{\xi})$, which yield corresponding values for λ_m and $\sigma_{jmt}(j)$.

2. Starting with the first period, solve for the $J + 1$ unknowns $\{s_{0t}(j)\}$ and $s_{0t}(0)$. These are obtained using the non-linear system of equations discussed in Section 4.1, conditional on the values of r_{jmt}. Use the initial value r_{jm0}^* for the first period and iterate forward using the law of motion to calculate r_{jmt} for subsequent periods.

3. Solve for $s_{jmt}(0)$ for each firm, period, and market given the $J + 1$ values above and equation (16). Calculate $\delta_{jmt} = \ln s_{jmt}(0) - \ln s_{0mt}(0)$.

4. Run the regression from equation (20) using the δ_{jmt} obtained in the previous step to solve for the linear parameters (α, π, γ). Obtain the residuals $\hat{\eta}_{jmt}$ and construct the objective function described above.

5. Repeat 1-4 to find the values for $(\theta_1, \theta_2, \theta_3, \overline{\xi})$ that minimize the (observation-weighted) sum of squared values for the empirical moments specified in equation (23).

The regression for equation (20) may involve instrumental variables and the use of panel data methods such as fixed effects. In our empirical application, we make use of both.

The estimation methodology employs two tricks to speed up the computation of the dynamic model. First, we are able to specify the $J+1$ unknowns, $\{s_{jt}(0)\}$, in each market-period in terms of only two unknowns, $s_{0t}(0)$ and $\sum_{0,z} r_{zt} s_{0t}(z)$. After obtaining these two parameters, we have an explicit formula to calculate the other $J-1$ values. We describe this result in Appendix E. This means that the non-linear solver[44] in step 2 only has to find two parameters for each market-period. Second, the linear form for the nested regression allows for a quick calculation of the inner part of the routine and allows for serial correlation in unobservables.

In models with state dependence and unobserved heterogeneity, one consideration is the initial value of the unobserved state. Because we have a long panel, with 104 separate time observations, we have a sufficient "burn-in" period where this issue does not much affect our estimates.[45] As a baseline, we set r^*_{jm0} equal to the value of the mean observed share in that market in the period prior to the first week of our sample, i.e., $E_j[S_{jm0}]$. We run robustness checks with alternative values for r^*_{jm0}, and we obtain very similar point estimates. We discuss these below in Section 4.4.

4.3 Data for Structural Model

To construct market shares that allows for the outside option ($j = 0$), we merge the market share data provided by OPIS with monthly state-level consumption provided by the EIA. We assume that the maximum observed quantity in the EIA data reflects 75 percent of the total potential market, and we scale the OPIS county-level shares according to the time variation in quantities for the corresponding state, i.e., $0.75 \times quantity_{st} / \max_t\{quantity_{st}\}$ for state s. Thus, if OPIS provides a market share for product j of 0.40, and the observed quantity in that state for that period is two-thirds of the maximum observed in the sample, we construct a market share of $0.20 = 0.40 \times 0.75 \times \frac{2}{3}$ for product j and an outside option share of 0.50.

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 two weeks. We assume that the station was not in the choice set when shares are missing or zero.[46] We drop

[44]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.

[45]We find that it takes approximately 7 weeks for different starting values to converge to same time time-series patterns for r, i.e., the choice of the initial value does not substantially affect the values of the latent variables for periods 8 through 104.

[46]After all cleaning steps, approximately 1 percent of observations have zero shares, and over 90 percent of these zero share occurrences are in spells of 6 weeks or longer.

Table 2: Summary Statistics by County

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

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

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 2 provides summary statistics of the data for the 241 counties in KY and VA. There is cross-sectional variation in wholesale prices, margins, and the number of stations in each county. To reduce the sensitivity of the analysis to brands with small shares and to make the counterfactual exercises more computationally tractable, we aggregate brands with small shares into a synthetic "fringe" brand. We designate a brand as part of the fringe if it does not appear in ten or more of the 241 markets (counties). Additionally, if a brand does not make up more than 2 percent of the average shares within a market, or 10 percent of the shares for the periods in which it is present, we also designate the brand as a fringe participant for that market. This reduces the maximum number of brands we observe in a county to 8, down from 24. Across all markets, we analyze the pricing behavior of 16 brands, including the synthetic fringe.[47]

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

We implement regression equation (20) as follows:

$$\delta_{jmt} = \alpha p_{jmt} + \pi \left(p_{jmt} \times Income_m\right) + \gamma N_{jmt} + \zeta_{jm} + \phi_t + \psi_{m,month(t)} + \eta_{jmt} \tag{24}$$

We obtain $\delta_{jmt} = \ln\left(\frac{s_{jmt}(0)}{s_{0mt}(0)}\right)$ following steps 1 through 3 from Section 4.2. We have shares and prices at the brand-county-week level. Within-county shares of unaffiliated consumers depend on prices, station amenities, and demographic characteristics of the local population. The brand-county fixed effects, ζ_{jm}, control for cross-sectional variation in the number of stations, brand amenities,and local demographic characteristics.[48] We observe station entry and exit

[47]Summary statistics by brand are presented in Table 21 in the Appendix. The fringe brand is, on average, 13 percent of the shares for the markets that it appears in. As we designate a fringe participant in nearly every market, the aggregated fringe has the highest overall share (12 percent).

[48]Station amenities include, for example, the presence of food (snack or restaurant), co-location with a super-

and include the number of stations for the brand in that market, N_{jmt}, which is identified by within-brand-county variation over time. Because we model demand and product choice at the brand-county level, the number of stations within a county can be thought of as a product feature, where more locations correlate with higher product quality.

Thus, brand-county fixed effects, which are identified by the panel, allow us to account for a first-order component of heterogeneity in preferences. Another important component of preferences in this model is price sensitivity. To account for heterogeneity in price sensitivity, we interact price with the log median household income in the county.[49]

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 $\{\psi_{m,month(t)}\}$.[50] 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 weekly US production of crude oil with the average wholesale cost for the brand in the county.[51] This gives us brand-county-specific time variation in our instrument which is (a) correlated with the wholesale cost and (b) plausibly not linked to demand. We chose this measure, rather than instrumenting directly with brand-state wholesale costs, to allow for the possibility that local variation in wholesale costs over time may reflect brand-specific demand shocks.

We interact the above instrument with $Income_m$ to create a second instrument, Z_2, to account for the endogeneity of $(p_{jmt} \times Income_m)$. Both US crude oil production and income are plausibly exogenous with respect to local, time-varying demand shocks. Figure 8 in the Ap-

market, 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. We do not directly account for commuting patterns in our analysis.

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

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

[51] 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. We obtain the quantity of crude oil produced in the U.S. each week from the EIA.

Table 3: Estimates of Static Demand Parameters

	Static Model			Dynamic Model
	(1)	(2)	(3)	(4)
Price	-0.022^{*}	-0.260^{***}	-2.315^{***}	-2.198^{***}
	(0.014)	(0.038)	(0.507)	(0.458)
Price $\times$ Income	-0.135^{***}	0.077^{***}	-0.007	0.006
	(0.021)	(0.023)	(0.032)	(0.033)
Number of Stations	0.016^{***}	0.064^{***}	0.063^{***}	0.058^{***}
	(0.004)	(0.010)	(0.010)	(0.010)
IV	No	No	Yes	Yes
Brand-County FEs		X	X	X
Week FEs		X	X	X
County-(Month of Year) FEs		X	X	X
Observations	110,844	110,844	110,844	110,844

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

pendix 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.

4.4 Results: Demand Estimation

For the empirical application, we implement the methodology described in the preceding sections. At the solution, our parameter estimates deliver an objective close to zero. The implied overall autocorrelation in shocks is -0.0007, and the overall covariance is -0.0003.

The estimates for the linear parameters are reported in Table 3. For comparison, 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. We obtain a similar price coefficient with our dynamic specification, though the economic meaning of the coefficients are different as the shares used in estimation only reflect a subset of consumers that are unaffiliated. In the static model, all consumers are assumed to be unaffiliated. We estimate essentially no relationships between income and price sensitivity for

Table 4: Estimates of Dynamic Demand Parameters

	Affiliation Rate			Strength of Affiliation
	Baseline	*Income*	*Density*	*Utility Shock*
	θ_1	θ_2	θ_3	$\bar{\xi}$
Coefficient	0.584	-0.741	0.287	5.833
95 Percent CI	[0.48, 0.75]	[-1.35, 0.27]	[0.09, 0.51]	[5.00, 6.53]

Notes: Table displays the estimated non-linear coefficients from the dynamic model. The first three parameters imply that, on average, 64 percent of consumers that purchase develop an affiliation for that brand. Brands located in areas with higher incomes and higher population densities have greater rates of affiliation, though this heterogeneity is not statistically significant. The last parameter shows the level shock for affiliated customers, which is positive, as expected. Confidence intervals are shown with brackets and are calculated via the bootstrap.

unaffiliated consumers, as the estimated coefficient of 0.006 is close to 0 with small standard errors. On the other hand, we find that an increase in the number of stations that a brand has in a market has a statistically significant positive effect on demand for unaffiliated consumers.

Table 4 reports estimates of the dynamic parameters. The parameters θ_1, θ_2, and θ_3 imply that 63.6 percent of consumers, on average, are subject to state dependence and develop an affiliation for the brand they previously purchased from (i.e., $E[\lambda_m] = 0.636$). The coefficient of -0.741 on income indicates that markets with lower-income consumers are populated by more consumers that are prone to affiliation, though this relationship is not statistically significant at the 95 percent level. On the other hand, the statistically significant coefficient of 0.287 on population density indicates there is a higher share of state-dependent consumers in more dense areas. One possible explanation for this is that consumers in urban environments may drive less frequently and thus be more prone to affiliation. Both of these demographic variables are standardized, so each coefficient corresponds to an increase of one standard deviation.

The estimated utility shock to affiliated consumers, $\bar{\xi}$, implies that the affiliated consumers are inelastic with respect to price. Across observations, the mean own-price elasticity for affiliated consumers is -0.53, and the median is -0.38. On average, an affiliated consumer re-purchases from their preferred brand 92 percent of the time. In the estimated model, affiliated consumers do not respond much to prices when the levels are low, but they do when the prices are higher. The average (absolute) weekly price change in the data is 5 cents per gallon, and the 25th and 75th percentile price changes are 1.6 cents and 7.5 cents, respectively. We therefore find that affiliated consumers do not typically switch brands within this range of price changes. The unaffiliated consumers, however, are highly elastic, with an average own-price elasticity of -5.96. This implies that for a 1 percent increase in price (roughly 3 cents), the station will lose 6 percent of its unaffiliated consumers. This high level of price sensitivity for a subset of retail gasoline consumers seems plausible, as some "shoppers" have been found to go well out of the way to save a few cents per gallon.[52]

[52]For example, the National Association of Convenience Stores found in their 2018 survey that 38 percent of

On average, roughly 76 percent of a brand's customers come from affiliated consumers in any week in equilibrium. The remaining share comes from shoppers (comprising 15 percent of purchasers for a given brand) and non-shoppers who may be unaffiilated or affiliated to a different brand. The average weighted elasticity, which weighs affiliated and unaffiliated consumers by their relative (purchasing) proportions, is -1.86. This weighted elasticity captures the effective elasticity faced by a firm and is different than the elasticity obtained when estimating a static model. A "naive" estimate using a static model would result in a value of -5.7, which implies a much greater loss in market share for a given price increase than we estimate from the dynamic model. At the market level, the parameter estimates imply an aggregate weekly elasticity of demand of -1.20. This is fairly close to the aggregate elasticity over a two-day horizon of -1.38 estimated by Levin et al. (2017). Estimates of elasticities vary in the literature due to differences in how the markets are defined. Typically, a broader scope in terms of geography and time horizon yields more inelastic demand.[53] In addition, the existing estimates do not account for consumer inertia, which, as highlighted above, can yield more inelastic demand.

We have tested the robustness of our demand estimates along a few dimensions. First, we consider the impact of the initial value of the unobserved state on our estimates. As discussed in Section 4.2, because our dataset consists of a long panel with 104 time periods, the initial choice of r^*_{jm0} may not matter much in estimation. After several periods, the system converges toward the steady-state regardless of the initial value. The estimates will be less sensitive to the initial choice if more observations reflect the steady-state values. To evaluate the potential impact, we consider two extreme choices of r^*_{jm0}. First, we assume that no affiilation exists at the start of the sample, imposing $r^*_{jm0} = 0$ for all brands and markets. At the other extreme, we scale up r^*_{jm0} from the baseline so that all non-shoppers are affiliated to a brand in the first period (i.e., we multiply the baseline values so that they sum to 1 in each market). Each extreme scenario delivers similar point estimates to those in our baseline specification. All of the dynamic parameters and reported static parameters are within the 95 percent confidence intervals of our baseline estimates. For example, we obtain estimates for α of -2.439 and -2.081, $\bar{\xi}$ of 6.01 and 5.91, and a mean value for λ_m of 0.617 and 0.637 for these two alternative specifications.

Second, we test for the uniqueness of our solution by randomly drawing initial values for the dynamic parameters in the estimation algorithm. We confirm that our estimation routine gets identical parameter estimates across random initial values (up to a negligible rounding error).

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

[53] For example, Levin et al. (2017) find an aggregate elasticity of -0.31 after 20 days, which is much more inelastic than their two-day estimate. Li et al. (2014) estimate an aggregate elasticity of -0.1. In terms of more narrow scope, the estimates of Houde (2012) of -10 to -15 reflect an elasticity at the station level, instead of the brand-county or county/aggregate level, which we would expect to result in more elastic estimates.

Table 5: Summary of Implied $\beta \frac{\partial E[V_j(\cdot)|\cdot]}{\partial p_{jt}}$

Group	Mean	Min	p25	Median	p75	Max
All	-0.122	-0.715	-0.165	-0.094	-0.051	0.040

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

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 equation (6). We assume that the price for each brand is set to maximize discounted profits at the county level. Following equation (8), the dynamic condition for optimal pricing for a single-product firm that owns brand j is:

$$\frac{\partial \pi_{jt}}{\partial p_{jt}} + \beta \frac{\partial E\left[V_j(r_{t+1}, c_{t+1}, x_{t+1})|p_t, r_t, c_t, x_t\right]}{\partial p_{jt}} = 0, \tag{25}$$

where p_{jt} is the average price for the brand in a county and $\frac{\partial \pi_{jt}}{\partial p_{jt}}$ is the derivative of the per-period profits. This derivative equals $\frac{\partial S_{jt}}{\partial p_{jt}}(p_{jt} - c_{jt}) + S_{jt}$ for single-brand firms.

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

On average, we find that $\frac{\partial \pi_{jt}}{\partial p_{jt}}$ is positive. This implies that firms are systematically pricing lower than the myopic profit-maximizing price. We interpret this as evidence of forward-looking behavior and the presence of dynamics, consistent with the reduced-form evidence of Section 3.3.2. Based on equation (25), we attribute the difference between $\sum_{l \in J_i} \frac{\partial \pi_{lt}}{\partial p_{jt}}$ and 0 to be accounted for by the derivative of the continuation value (DCV), $\beta \frac{\partial E[V_j(\cdot)|\cdot]}{\partial p_{jt}}$. That is, the dynamic incentive is the residual that rationalizes the observed pricing behavior of the firms, conditional on the demand-side assumptions, the data, and Bertrand price competition.[54] 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 5. The mean and median are negative, which implies that, typically, a reduction in price would increase the expected future return. We estimate a positive residual in only 3

[54]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.

percent of observations. The magnitudes are significant: the mean of -0.122 implies that a 1 cent increase in price would increase static profits by roughly 4 percent.[55] Intuitively, firms are lowering prices to invest in future demand. Such behavior allows firms to occasionally have negative price-cost margins, which occur in 2.7 percent of the observations in our data. 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.

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. We develop an alternative approach in which we rely on the structural demand model to calculate the static component of profits, and we use a functional approximation to capture dynamic incentives.[56]

Specifically, we approximate the dynamic component of firms' first-order conditions (the DCV) directly with a reduced-form model that is a function of state variables. Using the data and the estimated demand parameters, we obtain estimates of the DCV and project these estimates on state variables, including measures that capture expectations. We estimate the following dynamic first-order condition:

$$\frac{\partial \pi_{jt}}{\partial p_{jt}} + \Psi_j(p_t, r_t, c_t, x_t; \theta) + \zeta_{jt} = 0. \tag{26}$$

For any observed or counterfactual data point, we construct $\frac{\partial \pi_{jt}}{\partial p_{jt}}$ directly using the structural demand estimates. We use $\Psi_j(\cdot)$ to approximate $\beta \frac{\partial E[V_j(\cdot)|\cdot]}{\partial p_{jt}}$ from equation (25), 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. In general, Markovian assumptions allow for the continuation value to be expressed as a function of the state and firm actions.

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. Our approach has three key advantages. First, the static component of profits is obtained without having to make any assumptions about firm expectations and discount rates. Second, we avoid the need to make

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

[56] Note that we cannot make use of the steady-state relationship from equation (9) because time-varying demand and cost shocks imply that components like $\frac{dV_k(r)}{dr}$ may vary period-to-period and the equilibrium of the system varies accordingly.

dimension-reducing assumptions, such as constructing a limited grid for prices, that are less palatable in our setting.[57] Third, utilizing an approximation for the DCV greatly reduces the computational time needed to re-compute equilibria. For our approach to accurately represent behavior, we require 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.

This reduced-form 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. Thus, our approach can be considered as an attempt to replicate the forecasting behavior of firms that use regression analysis to predict future profits. In this case, the firms' beliefs can correspond to the econometrician's estimates. To provide a sense of how close our estimates come to rational expectations, we use forward simulations to calculate realized profits when firms set prices according to equation (26). In other words, we assume that firms' forecasts of the DCV correspond to $\Psi(\cdot)$, and we evaluate how close these forecasts are to the realized DCV when firms choose prices following these forecasts. We discuss these forward simulations below.

To estimate $\Psi_j(\cdot)$, we project the estimated DCV onto the stock of affiliated consumers ($r_{jmt} \times \lambda_m$), the derivative of own share with respect to price, marginal costs, and expectations of future costs. Our model and descriptive evidence suggests that these variable play an important role in expectations of future profits. We also include the fraction of state-dependent consumers λ_m, the number of stations, the total number of stations for all brands, and the number of brands as market-level controls. We do not include market-level or brand-level fixed effects. Instead, we use cross-market variation to quantify the relationships between the selected covariates and the DCV.

The results of estimating equation (26) are reported in Table 6. The first specification reports the coefficients from a regression of the DCV on the dependent variables. As the DCV is negative on average, a negative coefficient implies that the variable is associated with a stronger dynamic pricing incentive, or a greater deviation from the optimal static price. We flip the sign on the own-price derivative, which is also negative, to facilitate interpretation. The eight-parameter model has an R^2 of 0.97. Overall, the reduced-form approach captures the vast majority of the price variation that cannot be explained by static optimization. The high degree of explanatory power of the parsimonious model provides some confidence for reasonable interpolation and extrapolation in counterfactual analysis.

We find that a higher share of affiliated consumers, $r_{jmt} \times \lambda_m$, increases the magnitude of the DCV, corresponding to an increased investment incentive when pricing. Thus, though the

[57]Relatedly, under the policy function approach, an insufficiently flexible policy function may be incompatible with equilibrium prices.

Table 6: Dynamic Pricing Incentive: Regressions

	$\beta\frac{\partial E\left[V_j(\cdot)\mid\cdot\right]}{\partial p_{jt}}$ (1)	Sensitivity (2)
$r_{jmt} \times \lambda_m$	-0.889^{***}	6.428^{***}
	(0.001)	(0.017)
$-\frac{dS_{jmt}}{dp_{jmt}}$	-0.540^{***}	6.416^{***}
	(0.001)	(0.037)
Marginal Cost	-0.001^{***}	0.019^{***}
	(0.000)	(0.002)
Cost Change (30-Day Ahead)	0.014^{***}	-0.092^{***}
	(0.000)	(0.012)
λ_m	0.010^{***}	0.438^{***}
	(0.001)	(0.015)
Num. Stations (Brand)	0.000^{***}	0.011^{***}
	(0.000)	(0.000)
Num. Stations (Market)	-0.000^{***}	-0.001^{***}
	(0.000)	(0.000)
Num. Brands (Market)	-0.001^{***}	0.042^{***}
	(0.000)	(0.001)
Constant	Yes	Yes
Observations	110,844	110,844
R^2	0.973	0.785

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

presence of less-elastic affiliated consumers provides a direct incentive to raise prices, they also provide a dynamic incentive to keep prices high. The coefficient on the own-price derivative indicates that dynamic incentive is greater when the derivative (which is also negative) is larger in magnitude. This indicates that demand from unaffiliated consumers also play a role in dynamic incentives. One mechanism to explain this is that some unaffiliated (state-dependent) consumers become affiliated to that brand in the future, so the derivative captures the possibility of attracting more future affiliated consumers. Increases in marginal costs tend to make firms more sensitive to dynamic profit considerations, while increases in future expected costs lead firms to place relatively less weight on future profits. Finally, we find that the number of stations and brands have relatively small coefficients, after controlling for the above factors.

To help interpret how sensitive firms are to dynamic considerations, the second column of Table 6 reports a regression where we replace the value of the DCV with the logged absolute value. Thus, the coefficients reflect the semi-elasticity for the magnitude of the dynamic incentive. A positive coefficient in the second column indicates that an increase in the variable makes a firm more sensitive to dynamic considerations, whereas a negative coefficient indicates a reduced sensitivity to dynamic considerations when pricing. Typically, a negative coefficient in the first column corresponds to a positive coefficient in the second, as the average value for the DCV is negative. The results in the second column suggest that firms' dynamic considerations are most sensitive to the stock of affiliated consumers and the own-price derivative. The results show a modest marginal relationship on the overall share of state-dependent consumers in the market.

Verification with Forward Simulations

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 brand in each market and each period, we slightly perturb the price for that brand. We re-compute shares in that period and then calculate equilibrium play in future periods, while imposing that firms construct beliefs using our estimate of $\Psi(\cdot)$. We use the incremental change in profits along this simulated path to evaluate $\beta\frac{\partial E[V_j(\cdot)|\cdot]}{\partial p_{jt}}$ under rational expectations. Our simulated change in future profits are positively and significant correlated with the recovered estimates of $\Psi(\cdot)$ (the correlation coefficient is 0.549). Using an annual discount factor of 0.95, we calculate an average simulated value for the derivative of -0.064. This is roughly half of the mean estimate of -0.122 in Table 5. Thus, our approach to capturing firm expectations seems directionally well-aligned with realized equilibrium profits. If taken seriously, these results suggest that firms do not have perfect information about the effects of prices on future profits, or they do not discount according to standard models. The simulations suggest that firms may over-estimate the loss of future profits from raising prices today.[58] For additional details about the simulations, see Appendix G.

5.3 Horizontal Market Power: Merger Simulation

To evaluate the impact of dynamic pricing incentives on horizontal market power, we simulate a merger between Marathon and BP, which are the number one and number four (non-fringe) brands in terms of overall shares in our sample. Out of the 241 markets, they overlap in 75. In these 75 markets, the average (post-merger) HHI is 1511, and the mean change in HHI

[58] One of the benefits of our approach is that we do not have to take a stand on beliefs or expectations to assess dynamic behavior and conduct counterfactuals. However, this points to how our framework could be used to test different models of firm beliefs.

resulting from the merger is 383. In 8 markets, the resulting HHIs are greater than 2500, and the changes are greater than 200, meeting the typical thresholds that are presumed likely to enhance market power. The merger would change twelve markets from 3 firms to 2 firms and eighteen markets from 4 firms to 3 firms. We allow the firms to merge at the beginning of September 2014, and we calculate counterfactual prices and shares for the second half of the sample.[59]

In Section 2, we showed that the price effects of a merger can depend on the way the merger is implemented, especially in the presence of consumer inertia. To measure the potential empirical impact, we consider the two alternative merger scenarios described in that section: a joint pricing merger and a brand consolidation merger. In the joint pricing scenario, the merged firm has pricing control over both brands, which they maintain as distinct entities. In the brand consolidation scenario, the merged firm consolidates the assets under a single brand. To implement a brand consolidation merger, we need to make an assumption about how much utility consumers of the product removed from the market will receive from buying the consolidated brand. As in the theoretical analysis, we assume that, at pre-merger prices, the consolidated brand will have the same combined share of unaffiliated customers as the separate pre-merger brands. This assumption, in the context of the demand model, will also give the merged firm an advantage in retaining affiliated consumers.

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

Table 7 displays the mean effects on prices, shares, and profits of the two mergers. The first three columns report the effects from the joint pricing scenario, and the second three columns report the effects from the brand consolidation scenario. In either scenario, the overall price effects are moderate. The joint pricing scenario predicts that the merging firms will raise prices by 4.9 percent, with an 18 percent decrease in shares and a meaningful increase in profits. Rival firms in the same market see an increase in share and a small (negative) change in prices. Profits increase for merging firms and rivals, and, overall, prices in the market increase by 1.6 percent.

The brand consolidation scenario predicts a price increase for the merging firms of 2.4 per-

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

Table 7: Merger Effects

	Dynamic Model: Joint Pricing			Dynamic Model: Brand Consolidation			Static Model: Brand Consolidation		
Brand	Price	Share	Profit	Price	Share	Profit	Price	Share	Profit
Marathon-BP	4.91	-18.09	25.26	2.37	10.92	51.2	5.35	-16.74	34.44
Other	-0.33	9.44	5.85	1.53	-9.60	2.95	0.61	3.86	11.28
Overall	1.55	-2.93	15.16	2.05	-0.38	26.08	2.40	-5.40	22.39

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

cent, roughly half that of the joint pricing prediction. However, in the brand consolidation scenario, the merging firms also realize an *increase* in market share. This occurs because brand consolidation provides the merging firm with an advantage in retaining affiliated consumers and yields a greater investment incentive. Recall that our merger holds the choice probabilities for unaffiliated consumers to be identical post-merger at the pre-merger prices. In this scenario, if we hold prices fixed at the pre-merger levels, we would find that the merging firm accumulates greater shares over time due to the superior ability to retain affiliated consumers. In the brand consolidation merger, the combined share of affiliated consumers (r) for the merging firms rises from 0.39 in the observed baseline to 0.44. By contrast, in the joint pricing scenario, the average combined share falls to 0.32.[60] This highlights how the incentives to invest or harvest can vary across the two types of mergers.

Another difference between the two mergers is in the effects on rivals. In contrast to the joint pricing scenario, the brand consolidation scenario sees rivals increase prices by 1.5 percent and lose profits as consumers are attracted to the consolidated brand. Thus, a change in the relative ability to retain affiliated consumers shifts the relative incentives to invest or harvest, as shown by the differences in behavior for the merging and non-merging firms across the two merger scenarios. Empirically, this is captured by the shift in the static and dynamic components of the first-order condition. The joint pricing scenario provides an immediate incentive to raise prices to increase current-period joint profits. The brand consolidation scenario also has this static incentive, but there is also an immediate dynamic incentive to invest as the acquiring brand can now capture more state-dependent consumers that are unaffiliated, as captured by the coefficient on $-\frac{dS_{jmt}}{dp_{jmt}}$ in Table 6. The relative strength of these two forces in the longer-run equilibrium determines which of the two mergers leads to greater price changes.

For comparison, we report the results from a merger analysis using a static model in the

[60] It takes approximately 10-15 weeks to reach the new level of affiliated shares.

last three columns of Table 7. We calibrate a standard logit demand system to identical prices, margins, and shares that are used to estimate the dynamic model. We use a brand consolidation merger to illustrate the potential effects. The static model predicts price effects of 5.4 percent for the merging firms, and 2.4 for the market overall. The median price change for the merging firms is 4.6 percent in the static model. By comparison, roughly one-third of markets in the joint pricing counterfactual exceed that value, while only 9 out of 75 markets have price changes at least that large in the dynamic brand consolidation counterfactual.

Consistent with our simulations in Section 2, we find that the predictions of a dynamic model with consumer inertia can diverge from those of a static model. The dynamic incentive to invest in future demand can mitigate the short-run incentive to raise prices post-merger, dampening the exercise of horizontal market power. Further, the way that the merger is implemented can have a large impact on equilibrium prices. In the brand consolidation merger above, the merging firms have an added incentive to invest, leading to price increases that are on average less than half of that in a static model.

Despite the differences in levels, the predictions for the merging firm price changes are directionally similar. Across markets, the correlation between any two of three types of mergers ranges from 0.713 to 0.750. Across merger types, a greater combined pre-merger share correlates with a higher price increase post-merger. Thus, the standard intuition that a merger between larger firms could lead to higher price changes holds in all three models.

A key departure from the simulations in Section 2 is that the empirical setting has asymmetries across the brands within a market. Previously, we showed that, in a static setting, both types of mergers predict identical price effects in a symmetric oligopoly. With asymmetries among firms, static mergers can yield different predictions for joint pricing and brand consolidation. This is one reason why our empirical results may not track one-for-one with the examples of Section 2.

5.4 Dynamic Market Power

In the previous section, we used the empirical model to evaluate the role of horizontal market power in the presence of consumer inertia. In this section, we isolate the role of dynamic market power. To do so, we change the strength of affiliation, $\bar{\xi}$, and the share of consumers that are affected by inertia, λ.

Specifically, we simulate counterfactual scenarios where $\bar{\xi}$ increases by 10 percent and a second in which λ_m increases by 0.10 in every market. The latter corresponds to an average change in λ_m of 16 percent. Thus, we consider the impact of both the strengh of affiliation and the share of "shoppers" on equilibrium prices. We use the same 75 markets as the previous subsection to facilitate a comparison to the horizontal market power effects in Table 7.

Table 8 reports the results from the counterfactual scenarios. The first three columns report the equilibrium effects of an increase in the strength of state dependence, and the second

Table 8: Equilibrium Effects of State Dependence

	Increase $\bar{\xi}$ by 10 Percent			Increase λ by 0.10		
Brand	Price	Share	Profit	Price	Share	Profit
Marathon	4.51	-2.4	45.8	1.90	6.5	29.0
BP	4.58	-1.8	52.2	1.91	4.9	27.7
FRINGE	4.50	-0.6	63.3	1.42	6.3	29.1
Shell	4.90	-2.0	53.6	1.72	5.6	26.5
Speedway	5.04	-2.0	60.6	1.67	7.3	29.7
Exxon	5.51	-3.3	48.8	1.66	7.0	24.9
Valero	4.40	-1.9	51.5	1.77	3.0	26.7
Sheetz	5.71	0.0	78.7	1.03	10.1	26.3
Loves	4.52	-1.3	51.1	2.14	5.4	31.4
Pilot	4.42	2.2	126.1	1.04	8.9	41.5
Hucks	3.82	-2.4	67.3	1.56	5.7	36.6
Sunoco	5.11	-2.0	42.4	2.73	-3.8	19.8
Citgo	5.67	-1.0	59.3	1.36	5.9	21.5
Thorntons	4.00	-1.9	56.2	1.76	6.5	32.6
Overall	4.68	-1.8	53.6	1.73	6.1	28.1

Notes: Table displays the mean percent changes in prices, shares, and profits from counterfactual scenarios with different portions of consumers affected by state dependence. The first three columns report the equilibrium effects with a greater strength of state dependence, and the second three columns report the effects with a greater share of consumers affected by dependence.

three columns report the effects of an increase in its prevalence. The results are reported by brand, and they are sorted by the mean share in the 75 relevant markets from the previous subsection. Thus, Marathon and BP appear as the two largest brands in these markets. If the state dependent utility shock increased by 10 percent, then overall prices would increase by 4.7 percent and average shares would fall by 1.8 percent. All firms realize an increase in profits, with some smaller brands (Huck and Thorntons) realizing the lowest increases in prices. There is also variation in the change in shares, with Pilot realizing an increase in shares (and the largest gain in profits), while Sheetz having zero change in equilibrium.

Increasing λ yields smaller effects on prices (1.7 percent), though firms increase shares by 6.1 percent and profits still increase by a meaningful amount. Interestingly, the effects on prices, shares, and profits are not perfectly correlated across the two counterfactuals. For example, Sheetz has the second-largest increase in profits from a change in the strength of affiliation but the third-smallest profit change from an increase in the prevalence of state dependence. These results suggest that multiple dimensions of a firm's relative position in the market influence the impact of consumer inertia.

In terms of magnitudes, these effect on prices and profits are in the same ballpark as the effects of mergers between Marathon and BP (Table 7). In rough terms, increasing the strength of affiliation has a similar effect on prices as the joint pricing merger, though shares do not fall by as much and profits, accordingly, increase by more. By contrast, increasing the share of consumers affected by inertia by 0.10 has effects on prices, shares, and profits that are roughly three-fifths that of a brand consolidation merger. Thus, the model allows us to quantify the relative effects of dynamic incentives and horizontal competition when evaluating market power. Our findings suggest that policies that produce even modest changes in the prevalence or magnitudes of switching costs (or consumer inertia, more generally) may have a similar impact on welfare as large changes to market structure, such as those that occur through mergers.

6 Conclusion

We develop a model of consumer inertia that accounts for commonly observed dynamic pricing behavior, such as the slow adjustment of prices to changes in cost. The dynamics result from competing firms optimally setting prices to consumers that may become loyal or habituated to their current supplier. Dynamic market power reflects the incentives of firms to set different prices in response to consumer inertia, relative to static consumer demand. We show that accounting for dynamic market power is important when performing counterfactual exercises, such as merger simulation.

We present reduced-form evidence consistent with consumer inertia in retail gasoline markets. Consumers purchase histories show patterns consistent with inertia. On the supply side, new retail locations initally set prices below incumbents, then slowly raise prices over time. Prices also 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 and estimate an empirical model that can identify dynamic demand parameters using data on price, shares, and an instrument. Results suggest that 64 percent of retail gasoline consumers are prone to consumer inertia, and these consumers are relatively insensitive to price changes. Conversely, we find that unaffiliated consumers are price sensitive and play an important role in disciplining equilibrium prices. We evaluate the dynamic incentives that affect prices. We show, both theoretically and empirically, that failing to account for dynamic demand can cause significant biases when predicting post-merger price increases. Furthermore, we show the importance of distinguishing between brand consolidation and joint pricing mergers. Finally, we evaluate the relative impacts of dynamic and horizontal market power in our empirical setting. Counterfactual analyses show that a relatively small change in the number of consumers subject to inertia can impact prices as much as a change in competition arising from a merger.

References

AGUIRREGABIRIA, V. AND P. MIRA (2007): "Sequential Estimation of Dynamic Discrete Games," *Econometrica*, 75, 1–53.

ARCIDIACONO, P. AND R. A. MILLER (2011): "Conditional Choice Probability Estimation of Dynamic Discrete Choice Models With Unobserved Heterogeneity," *Econometrica*, 79, 1823–1867.

BAJARI, P., C. L. BENKARD, AND J. LEVIN (2007): "Estimating Dynamic Models of Imperfect Competition," *Econometrica*, 75, 1331–1370.

BEGGS, A. AND P. KLEMPERER (1992): "Multi-Period Competition with Switching Costs," *Econometrica*, 60, 651.

BERGEMANN, D. AND J. VALIMAKI (1996): "Learning and Strategic Pricing," *Econometrica*, 64, 1125.

BERGEMANN, D. AND J. VÄLIMÄKI (2006): "Dynamic Price Competition," *Journal of Economic Theory*, 127, 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, 1–43.

BERRY, S., J. LEVINSOHN, AND A. PAKES (1995): "Automobile Prices in Market Equilibrium," *Econometrica*, 63, 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.

BRONNENBERG, B. J., J.-P. H. DUBÉ, AND M. GENTZKOW (2012): "The Evolution of Brand Preferences: Evidence from Consumer Migration," *American Economic Review*, 102, 2472–2508.

CABRAL, L. (2011): "Dynamic Price Competition with Network Effects," *The Review of Economic Studies*, 78, 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, 681–704.

DONNA, J. D. (2021): "Measuring Long-Run Gasoline Price Elasticities In Urban Travel Demand," *RAND Journal of Economics*, 52, 945–994.

DUBÉ, J.-P., G. J. HITSCH, AND P. E. ROSSI (2009): "Do Switching Costs Make Markets Less Competitive?" *Journal of Marketing Research*, 46, 435–445.

——— (2010): "State Dependence and Alternative Explanations for Consumer Inertia," *The RAND Journal of Economics*, 41, 417–445.

ECKERT, A. (2013): "Empirical Studies of Gasoline Retailing: A Guide to the Literature," *Journal of Economic Surveys*, 27, 140–166.

EIZENBERG, A. AND A. SALVO (2015): "The Rise of Fringe Competitors in the Wake of an Emerging Middle Class: An Empirical Analysis," *American Economic Journal: Applied Economics*, 7, 85–122.

ENGLE, R. F. AND C. W. J. GRANGER (1987): "Co-Integration and Error Correction: Representation, Estimation, and Testing," *Econometrica*, 55, 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, 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*, Elsevier, vol. 3, 1967–2072.

FARRELL, J. AND C. SHAPIRO (1988): "Dynamic Competition with Switching Costs," *The RAND Journal of Economics*, 19, 123.

——— (2010): "Antitrust Evaluation of Horizontal Mergers: An Economic Alternative to Market Definition," *The BE Journal of Theoretical Economics*, 10.

GOWRISANKARAN, G. AND M. RYSMAN (2012): "Dynamics of Consumer Demand for New Durable Goods," *Journal of Political Economy*, 120, 1173–1219.

HANDEL, B. R. (2013): "Adverse Selection and Inertia in Health Insurance Markets: When Nudging Hurts," *The American Economic Review*, 103, 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, 2722–2751.

HONKA, E. (2014): "Quantifying Search and Switching Costs in the US Auto Insurance Industry," *The RAND Journal of Economics*, 45, 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, 192–226.

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

LEE, R. S. (2013): "Vertical Integration and Exclusivity in Platform and Two-Sided Markets," *American Economic Review*, 103, 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, 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, 672–682.

LI, S., J. LINN, AND E. MUEHLEGGER (2014): "Gasoline Taxes and Consumer Behavior," *American Economic Journal: Economic Policy*, 6, 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, 683–709.

MILLER, N. H., M. REMER, AND G. SHEU (2013): "Using Cost Pass-through to Calibrate Demand," *Economics Letters*, 118, 451–454.

MILLER, N. H. AND G. SHEU (2021): "Quantitative Methods for Evaluating the Unilateral Effects of Mergers," *Review of Industrial Organization*, 58, 143–177.

NISHIDA, M. AND M. REMER (2015): "The Determinants and Consequences of Search Cost Heterogeneity: Evidence from Local Gasoline Markets," *Available at SSRN 2628576*.

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, 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, 599–618.

SHCHERBAKOV, O. (2016): "Measuring Consumer Switching Costs in the Television Industry," *The RAND Journal of Economics*, 47, 366–393.

SHUM, M. (2004): "Does Advertising Overcome Brand Loyalty? Evidence from the Breakfast-Cereals Market," *Journal of Economics & Management Strategy*, 13, 241–272.

SLADE, M. E. (1998): "Optimal Pricing with Costly Adjustment: Evidence from Retail-Grocery Prices," *The Review of Economic Studies*, 65, 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, 1763–1803.

VALLETTI, T. AND H. ZENGER (2021): "Mergers with Differentiated Products: Where Do We Stand?" *Review of Industrial Organization*, 59, 179–212.

VARIAN, H. R. (1980): "A Model of Sales," *American Economic Review*, 70, 651–659.

WEINTRAUB, G. Y., C. L. BENKARD, AND B. VAN ROY (2008): "Markov Perfect Industry Dynamics With Many Firms," *Econometrica*, 76, 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.

Appendix

A Theoretical and Numerical Analysis: Additional Results

A.1 Monopoly Case

We analyze steady-state prices in a monopoly market (with an outside good) to show how habit-forming consumers affect optimal prices and markups. The monopolist sells product j and the outside good is indexed as product 0. In the steady state, $r_{jt} = r_{j(t+1)} = r_j$ and $c_{jt} = c_{j(t+1)} = c$. Given the model detailed in the body of the paper, the monopolist's steady-state number of affiliated consumers, r_j^{ss}, is:

$$\begin{aligned} r_j^{ss} &= r_0^{ss} s_j(0) + r_j^{ss} s_j(j) \\ \implies r_j^{ss} &= \frac{r_0^{ss} s_j(0)}{1 - s_j(j)}. \end{aligned}$$

In monopoly case, we also have that $r_0^{ss} = 1 - r_j^{ss}$. It follows that we can express r_j^{ss} as a function of the choice probabilities for product j.

$$r_j^{ss} = \frac{s_j(0)}{1 - s_j(j) + s_j(0)} \tag{27}$$

Using the steady-state value of affiliated consumers, the fraction of consumers subject to affiliation, λ, and the aggregate share equation, S_j we can solve for the steady-state pricing function.

The steady-state period value is:

$$\begin{aligned} V^{ss}(r^{ss}, c^{ss}) &= (p^{ss} - c^{ss})((1 - \lambda + \lambda r_0^{ss}) s_j(0) + \lambda r_j^{ss} s_j(j)) + \beta V^{ss} \\ &= (p^{ss} - c^{ss})(s_j(0) + \lambda r_j^{ss}(s_j(j) - s_j(0))) + \beta V^{ss} \\ &= \frac{p^{ss} - c^{ss}}{1 - \beta} \cdot (s_j(0) + \lambda r_j^{ss}(s_j(j) - s_j(0))). \end{aligned}$$

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 $s_j(j) = s_j(0)$), in which case profits are simply the per-unit discounted profits multiplied by the firm's market share. Because the steady-state can be expressed entirely in terms of product j choice probabilities and affiliated customers, we simplify the following notation: $s_j(j) = s^j$, $s_j(0) = s^0$, and $r_j = r$.

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

$$p^{ss} = c^{ss} + \underbrace{\frac{-s^0 - \lambda r^{ss}\left(s^j - s^0\right)}{\frac{ds^0}{dp} + \lambda \frac{dr^{ss}}{dp}\left(s^j - s^0\right) + \lambda r^{ss}\left(\frac{ds^j}{dp} - \frac{ds^0}{dp}\right)}}_{m\,=\,\text{markup of price over marginal cost}}. \tag{28}$$

The second term, m, on the right-hand side of equation (28) 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 (28) 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:

$$m_d = \frac{-s^0 - \lambda r^{ss}\left(s^j - s^0\right)}{\frac{ds^0}{dp} + \lambda \frac{dr^{ss}}{dp}\left(s^j - s^0\right) + \lambda r^{ss}\left(\frac{ds^j}{dp} - \frac{ds^0}{dp}\right)} \gtreqless -\frac{s^0}{ds^0/dp} = m_s.$$

For a given price, the terms s^0 and ds^0/dp are equivalent across the two models. First, we substitute in r^{ss} in terms of choice probabilities from equation (27) as well as it's derivative with respect to p. Then rearranging term, we obtain a simple condition relating the levels of the markup terms:

$$m_d \gtreqless m_s \iff -\frac{\partial s^0}{\partial p} \gtreqless -\frac{\partial s^j}{\partial p}. \tag{29}$$

A higher value for m_d indicates higher markups and higher prices. Thus, whether or not markups are higher in the dynamic model depends upon the price sensitivity of affiliated consumers relative to shoppers. We can further say that if the denominator of m_d is negative then $m_d > m_s \iff -\frac{\partial s^0}{\partial p} > -\frac{\partial s^j}{\partial p}$, meaning that markups are higher if affiliated customers are less price sensitive. The sign of the denominator will depend upon the specified demand system.

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 (29). 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 also 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. We solve for p and $\frac{dp}{dr}$ using steady-state conditions.

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

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

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

$$\frac{dV_k(r')}{dr'} \cdot \frac{dr'}{dp_j} = -\frac{1}{\beta} \sum_{l \in J_k} \frac{\partial \pi_l}{\partial p_j},$$

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

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

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

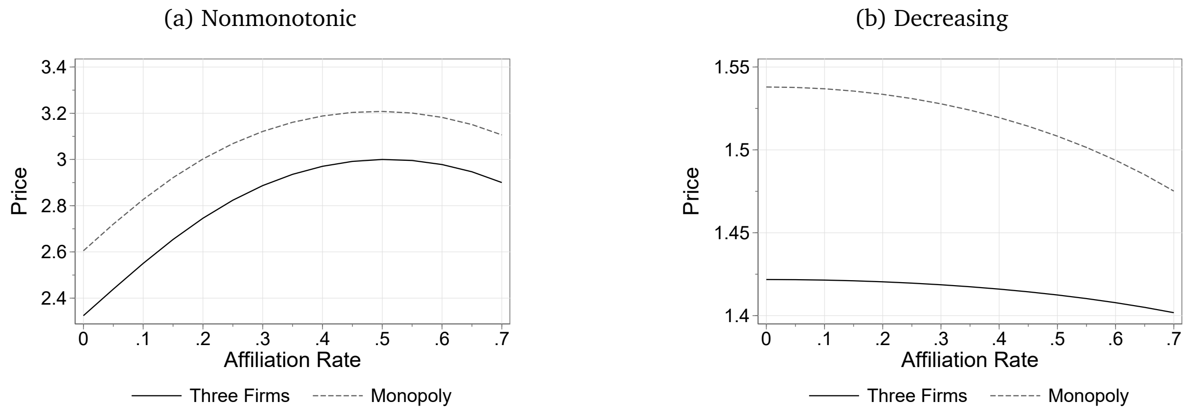
   Stack these vectors horizontally to obtain an approximation for $\frac{dp}{dr}$.

4. Calculate the absolute distance between the approximation of $\frac{dp}{dr}$ calculated in the previous step and the initial guess for $\frac{dp}{dr}$.

5. If the calculated distance falls below a critical value, then the solution is found. If not, update the guess for $\frac{dp}{dr}$ as the average between the initial guess and the approximation calculated in step 3. Repeat steps 1-4 above until a solution is found.

A.3 Simulation: Example Markets

Figure 6: Monopoly and Oligopoly Prices with Consumer Affiliation



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

To illustrate the impact of affiliation on pricing incentives, we plot the equilibrium prices for two different sets of utility parameters in Figure 6. Panel (a) plots the equilibrium prices for a monopolist and a three-firm symmetric market, for increasingly large values of λ and otherwise identical demand parameters. The equilibrium prices increase with the rate of affiliation for values of λ less than 0.5, and then decrease again. This figure highlights the potential importance of affiliation on equilibrium prices, and that the impact on price may be non-monotonic in the proportion of consumers prone to affiliation. The non-monotonicity is a result of two countervailing affects. At low levels of λ firms face increasing inelastic demand and therefore increase prices. As λ increases past 0.5, however, the incentive to invest in future demand swamps the elasticity effect and puts downward pressure on prices. Note that firms' profits continue to increase, even as prices begin to decrease.

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

A.4 Numerical Simulation Parameters

Table 9: Simulation Parameter Summary Statistics

	Mean	St. Dev.	Min	Pctl(25)	Pctl(75)	Max	N
λ	0.38	0.20	0.05	0.20	0.55	0.70	6566
α	-5.21	2.56	-9.96	-7.31	-3.12	-0.44	6566
$\underline{\xi}$	5.27	2.60	0.04	3.29	7.37	9.99	6566
$\bar{\xi}$	2.92	1.49	0.00	1.76	4.14	5.97	6566

Notes: Table displays summary statics for demand parameters for the 6,566 markets used in the numerical simulations. These markets were generated from a broader set of parameter values and selected if the resulting three-firm markets had firm shares between 0.05 and 0.30 (yielding an outside share between 0.10 and 0.85), margins between 0.05 and 0.75, and converged for all values of $\lambda \in \{0.05, 0.1, 0.15, ..., 0.70\}$. See the text for additional details.

A.5 Merger Effects and Relation to Affliation Parameters

Table 10: Simulation Summary Statistics

	Mean	St. Dev.	Min	Pctl(25)	Pctl(75)	Max	N
Pre-Merger Price	1.43	0.42	1.10	1.19	1.50	3.97	6566
Pre-Merger Margin	0.26	0.14	0.09	0.16	0.33	0.75	6566
Pre-Merger Market Share	0.17	0.07	0.05	0.11	0.24	0.30	6566
HHI: Pre-Merger	1067.35	776.19	76.28	339.49	1726.75	2695.59	6566
Δ HHI	711.57	517.46	50.85	226.33	1151.17	1797.06	6566
Joint Pricing Merger Δ	3.84	3.37	0.25	1.30	5.32	20.57	6566
Joint Pricing Non-Merging Price Δ	0.91	1.17	0.00	0.11	1.26	7.86	6566
Brand Consolidation Price Δ	6.49	4.29	0.23	3.43	8.55	25.48	6566
BC Non-Merging Price Δ	0.81	1.11	-1.14	0.07	1.17	8.64	6566
Prediction Bias: Joint Pricing	1.47	1.70	0.00	0.35	1.97	13.49	6566
Prediction Bias (pctg.): Joint Pricing	67.34	99.87	0.10	12.54	74.27	792.05	6566
Prediction Bias: Brand Consolidation	-1.17	1.46	-8.63	-1.77	-0.22	4.34	6566
Prediction Bias (pctg.): Brand Consolidation	-19.70	24.00	-76.01	-35.93	-4.16	181.04	6566
Dynamic Elasticity: Unaffiliated	-5.76	2.37	-11.96	-7.55	-3.77	-1.29	6566
Dynamic Elasticity: Affiliated	-2.13	1.90	-9.74	-3.14	-0.74	-0.08	6566
Dynamic Elasticity: Weighted	-3.86	1.98	-10.03	-5.15	-2.24	-0.59	6566
Static Elasticity	-4.74	2.08	-10.71	-6.15	-3.00	-1.34	6566

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

Table 10 provides summary statistics for the merger simulations across all 6,566 simulated markets.

Tables 11 and 12 provide results from regressions of pre-merger prices, price changes, and the absolute value of the bias in the static prediction on the demand parameters. Pre-merger prices increase with the fraction of affiliated customers (λ) and the strength of affiliation ($\bar{\xi}$). For joint pricing mergers, stronger dynamics, as represented by higher values of λ and $\bar{\xi}$, reduce merger price effects and generate a greater under-prediction that arises from a misspecified static model, which increases the size of the bias. In brand consolidation mergers, higher values of λ and $\bar{\xi}$ instead generate larger merger price changes, which also introduces absolute bias relative to the static prediction. However, these relationships do not hold in every instance, and may interact in interesting ways. As we show in Figures 6 and 1, prices are often non-monotonic in λ.

Table 11: Joint Pricing Merger Simulation: Demand Parameters

	(1) Pre-Merger Price	(2) Price Δ F1	(3) Bias F1	(4) Price Δ F3	(5) Bias F3
λ	0.128***	-1.538***	4.190***	0.156***	8.680**
	(0.017)	(0.093)	(0.060)	(0.038)	(4.287)
α	0.144***	2.553***	0.674***	0.908***	-15.754***
	(0.003)	(0.017)	(0.011)	(0.007)	(0.802)
ξ	0.026***	1.908***	0.433***	0.772***	-15.215***
	(0.003)	(0.017)	(0.011)	(0.007)	(0.792)
$\bar{\xi}$	0.040***	-0.163***	0.543***	0.186***	1.607***
	(0.002)	(0.013)	(0.008)	(0.005)	(0.588)
Constant	1.883***	8.157***	-0.451***	0.971***	29.107***
	(0.012)	(0.067)	(0.043)	(0.028)	(3.104)
N	6566	6566	6566	6566	6566

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

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

Table 12: Brand Consolidation Merger Simulation: Demand Parameters

	(1) Pre-Merger Price	(2) Price Δ F1	(3) Bias	(4) Price Δ F3	(5) Bias F3
λ	0.128***	3.609***	1.406***	1.144***	-676.899
	(0.017)	(0.115)	(0.052)	(0.039)	(973.032)
α	0.144***	2.846***	-0.282***	0.808***	-79.986
	(0.003)	(0.022)	(0.010)	(0.007)	(182.082)
ξ	0.026***	1.754***	-0.512***	0.691***	-144.450
	(0.003)	(0.021)	(0.010)	(0.007)	(179.863)
$\bar{\xi}$	0.040***	0.830***	0.410***	0.154***	9.927
	(0.002)	(0.016)	(0.007)	(0.005)	(133.556)
Constant	1.883***	8.313***	0.802***	0.505***	930.684
	(0.012)	(0.083)	(0.038)	(0.028)	(704.385)
N	6566	6566	6566	6566	6566

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

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

A.6 Predicting Merger Effects (and Bias) for Merging Firms

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

	(1) Price Δ	(2) Price Δ	(3) Price Δ	(4) Bias	(5) Bias	(6) Bias
Pre-Merger Market Share	27.834***	29.122***	10.971***	4.582***	3.264***	0.094
	(0.276)	(0.248)	(1.522)	(0.223)	(0.185)	(0.945)
Pre-Merger Margin	12.870***	13.087***	14.481***	7.184***	6.963***	12.024***
	(0.145)	(0.129)	(0.261)	(0.118)	(0.096)	(0.196)
λ		-3.612***	-2.774***		3.696***	3.225***
		(0.087)	(0.110)		(0.065)	(0.071)
α			1.013***			0.608***
			(0.100)			(0.062)
ξ			1.039***			0.763***
			(0.098)			(0.061)
$\bar{\xi}$			-0.446***			0.265***
			(0.012)			(0.008)
Price Δ						-0.252***
						(0.008)
Constant	-4.410***	-3.337***	0.250	-1.224***	-2.322***	-3.583***
	(0.058)	(0.058)	(0.276)	(0.047)	(0.043)	(0.171)
N	6566	6566	6566	6566	6566	6566

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

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

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

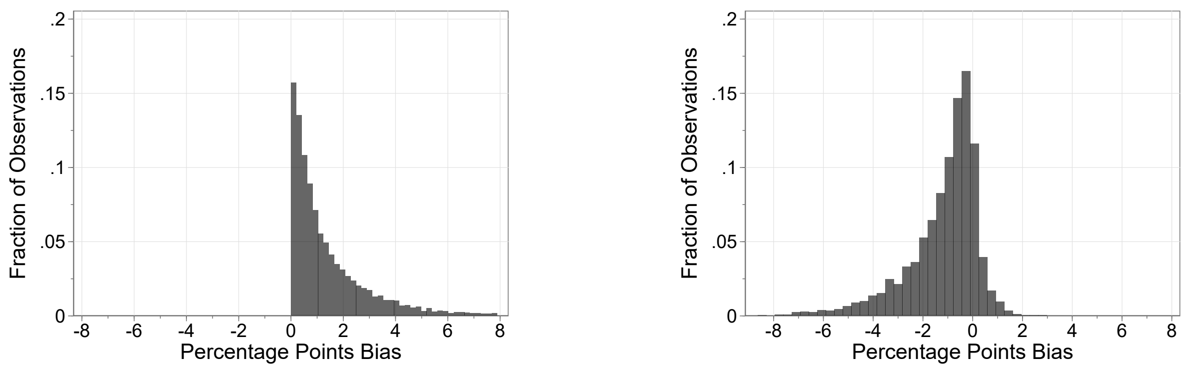
	(1) Price Δ	(2) Price Δ	(3) Price Δ	(4) Bias	(5) Bias	(6) Bias
Pre-Merger Market Share	21.874***	21.389***	28.305***	-9.580***	-10.170***	2.505***
	(0.218)	(0.215)	(1.337)	(0.174)	(0.166)	(0.930)
Pre-Merger Margin	24.937***	24.856***	25.640***	5.476***	5.377***	6.917***
	(0.115)	(0.112)	(0.229)	(0.092)	(0.087)	(0.193)
λ		1.360***	0.978***		1.656***	0.590***
		(0.076)	(0.096)		(0.059)	(0.070)
α			-0.442***			-0.244***
			(0.088)			(0.062)
ξ			-0.346***			-0.305***
			(0.086)			(0.060)
$\overline{\xi}$			0.332***			0.236***
			(0.010)			(0.008)
Price Δ						-0.235***
						(0.008)
Constant	-3.919***	-4.323***	-7.044***	1.515***	1.023***	-0.639***
	(0.046)	(0.050)	(0.243)	(0.037)	(0.039)	(0.168)
N	6566	6566	6566	6566	6566	6566

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

A.7 Distribution of Merger Prediction Bias

Figure 7: Distribution of Merger Prediction Bias

(a) Joint Pricing (b) Brand Consolidation



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

Figure 7 plots the distributions of bias across all 6,566 markets. Panel (a) depicts the distribution of bias for joint pricing mergers. In every instance the static model over predicts the true dynamic effect. In a few markets that did not meet our inclusion criteria, the static model under predicted the dynamic effect, so an upward bias will not always occur. However, our results suggest that a static logit demand model incorrectly calibrated to a demand with consumer inertia tends to exhibit upward bias. On the other hand, the bias in brand consolidation mergers may be positive or negative. While static models are more likely to under-predict the true dynamic price effect, almost 25 percent of simulations resulted in over-predictions. These findings again highlight the importance of properly accounting for dynamics when simulating the price effects of mergers. For our setting with symmetric three-firm oligopoly, we find that the sign of the bias is highly predictable for joint pricing mergers.

B Details on Brand Consolidation and Joint Pricing Mergers

B.1 Implementation of Joint Pricing Mergers

The implementation of joint pricing mergers is straightforward. We maintain the utility parameters from the demand side, and we assign the merged brands to the same firm. The merged entity sets prices for both brands, internalizing the cross-price effects on the joint profits.

B.2 Implementation of Brand Consolidation Mergers

For a brand consolidation merger, we remove one brand (the acquired brand) from the market. For this type of merger, there is a question of how the acquired assets translate into demand for the merging firm.

In this paper, we assume that the shares of unaffiliated consumers (shoppers) would remain the same if prices were maintained at the same level. Implicitly, we assume that the merged entity retains the same retail locations, and that the brand has no effect on shoppers.

To implement this, we adjust the demand shock for the remaining product of the merged firm so that the choice probabilities of unaffiliated customers are unchanged at pre-merger prices. Let a denote the acquiring brand, and b denote the acquired (and removed brand). Recall that the choice probabilities for unaffiliated consumers are given by

$$s_{jt}(0) = \frac{\exp(\delta_{jt})}{1 + \sum_k \exp(\delta_{kt})} \tag{30}$$

yielding $\delta_{jt} = \ln(s_{jt}(0)/s_{0t}(0))$. Recall that $\delta_{jt} = \xi_{jt} + \alpha p_{jt}$. For the merged brand, we adjust the utility shock fixed effect to ξ'_{at}, where

$$\xi'_{at} = \ln\left(\frac{s_{at}(0) + s_{bt}(0)}{s_{0t}(0)}\right) - \alpha \overline{p}_{ab} \tag{31}$$

Here, $s_{at}(0) + s_{bt}(0)$ is the combined pre-merger unaffiliated shares of products a and b, and $\overline{p}_{ab}$ is the share-weighted average price of products a and b. For non-merging firms, we maintain $\xi'_{jt} = \xi_{jt}$.

This adjustment ensures that, if prices are held fixed (including share-weighted prices for the merging brand), the choice probabilities of shoppers for all products, including the non-merging products, are unchanged. Then, given the adjustment, we allow the firms to price optimally.

Under these assumptions, the merging firm gets some benefit for affiliated consumers. Our adjustment implies that $\exp(\delta'_{at}) = \exp(\delta_{at}) + \exp(\delta_{bt})$, and therefore $\exp(\delta'_{at} + \overline{\xi}) = \exp(\delta_{at} + \overline{\xi}) + \exp(\delta_{bt} + \overline{\xi})$. Doing the appropriate adjustment to the choice probability equations, this implies that consumers affiliated to the merged brand choose it with the same probability as if they were affiliated to both brands a and b pre-merger. In implementation, we assume that consumers that were affiliated to brand b become unaffiliated after the merger. We found that whether or not affiliation transferred to the acquiring brand made little difference for our counterfactuals.

B.3 Equivalence of Joint Pricing and Brand Consolidation in Symmetric Logit

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

Lemma 1: Suppose the following is true,

(i) Demand is characterized by standard logit.

(ii) While holding all else equal, at a price $\bar{p}$, splitting firm m into two firms j and k yields equal shares that sum to the share of the original firm: $s_m(\bar{p}) = 2s_j(\bar{p}) = 2s_k(\bar{p})$.

It follows that $s_m(p) = 2s_j(p) \;\; \forall p$, i.e., the relation in (ii) holds for any price.

Proof: By construction:

$$s_m = \frac{e^{\xi_m + \alpha\bar{p}}}{1 + e^{\xi_m + \alpha\bar{p}} + \sum_g e^{\xi_g + \alpha p_g}} = 2\frac{e^{\xi_j + \alpha\bar{p}}}{1 + 2e^{\xi_j + \alpha\bar{p}} + \sum_g e^{\xi_g + \alpha p_g}} = 2s_j \tag{32}$$

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

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

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

$$\frac{d\Pi}{dp_1} = 1 + \alpha(p_1 - c_1)(1 - s_1) - \alpha \sum_{l=2}^{N} (p_l - c_l)s_l = 0 \tag{33}$$

Now, suppose all of the firm's products have the same marginal cost, c, and that all of it's products are symmetric, $\delta_n = \delta$ for all n. Then, equation (33) simplifies to:

$$\frac{d\Pi}{dp_n} = 1 - \alpha(p - c)(s - 1) - \alpha(N - 1)(p - c)s = 0 \tag{34}$$

Solving equation (34) for p yields the symmetric, profit-maximizing price for each product:

$$p^* = -\frac{1}{\alpha}\left[\frac{1}{1-Ns}\right] + c \tag{35}$$

Now, suppose a single product firm sets a profit-maximizing price of p^* and therefore has a market share s_l. Also, hold the number and characteristics of all other firms in the market constant. Now suppose we replace the single product firm with a two product firm with marginal cost, c and product characteristics, δ_j, such that $s_j(p^*) = \frac{s_m(p^*)}{2}$. By equation (35), the two product firm will set the same profit maximizng price, p^*. We therefore prove Lemma 2.

Lemma 1 and Lemma 2 help prove the following proposition.

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

(i) *Joint pricing*: After the merger, the firm retains both products and prices them to jointly maximize post-merger profits.

(ii) *Brand consolidation*: After the merger, the firm removes one of its brands from the market. For the consolidated product, the post-merger share would have the same market share as the sum of the two single product pre-merger firms at pre-merger prices. Post-merger the firm sets a profit-maximizing price for the one product it keeps in the market.

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

C Reduced-Form Evidence: Additional Results

C.1 Dynamic Demand: Repeat Visits and Purchase Intervals

We build on our analysis from Section 3.2 using the NielsenIQ Consumer Panel Data. Here, we leverage the data on the duration between purchases to provide additional evidence on the role of consumer inertia. In our model, consumers that choose the outside option lose their affiliation to a particular brand. In the data, choosing the outside option is captured by longer periods between purchases, when the consumer chooses not to purchase. We denote the number of days since the previous purchase as the "purchase interval." Longer intervals could arise due to idiosyncratic, household-specific shocks, or (e.g.,) higher prices.

We regress an indicator for repeat purchase on the the purchase interval, while controlling for household-year and weekly fixed effects. The estimates are reported in column (1) of Table 15. We again find evidence consistent with our model of consumer inertia. In particular, the coefficient of -0.0011 indicates that a longer purchase window is negatively related to the probability of a repeat purchase. If households delay for a week (i.e., choose the outside good) they are approximately 1% less likely to return to the same brand. This again suggests past decisions affect brand choice, and is consistent with our model where consumer lose brand loyalty when choosing the outside good.

To provide more detail on these dynamics, we subset households by the fraction of their purchases that are repeat purchases, based on the data displayed in Figure 3. Bin 1 households have less than 25% repeat purchases. Bin 2 includes those whose fraction is between 25% and 50%, Bin 3 includes those between 50% and 75%, and Bin 4 includes those between 75% and 99%. For this exercise, we exclude households with 100% repeat visits. These buckets are used to proxy for consumers that may be more or less prone to inertia.

We then repeat the regression exercise for subsets of the data using each bin. The estimates are reported in columns (2)-(5) of Table 15. The estimates indicate that the propensity of Bin 1 consumers to have a repeat purchase is not affected by the purchase interval. This is consistent with our assumption that some consumers are "shoppers" and are unaffected by inertia. The relationship between the purchase interval and the probability of a repeat purchase is stronger in columns (3), (4), and (5), which may reflect the fact that a greater proportion of households in those bins are prone to inertia/affiliation.

Finally, as a placebo test, we have also run a regression where we replace the dependent variable (the interval from the previous purchase) with the interval until the next *future* purchase. The coefficient estimate is almost exactly 0, with a t-stat of -0.01. This provides some assurance that results in column (1) are due to consumer behavior and not spurious trends in the data.

Table 15: Predicting a Repeat Visit

	(1)	(2)	(3)	(4)	(5)
Purchase Interval	-0.0011***	0.0001	-0.0006***	-0.0015***	-0.0012***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Constant	0.7329***	0.2060***	0.4054***	0.6425***	0.9285***
	(0.000)	(0.003)	(0.001)	(0.001)	(0.000)
Observations	2680548	42540	524061	766642	1347305
Household-Year FE	Yes	Yes	Yes	Yes	Yes
Week FE	Yes	Yes	Yes	Yes	Yes
Household-Year Group	All	Bin 1	Bin 2	Bin 3	Bin 4

Notes: The dependent variable is an indicator of whether a purchase is a repeat visit. Bin 1 households are those whose fraction of repeat purchases is less than 25% of their total purchases. Bin 2 includes those whose fraction is between 25% and 50%, Bin 3 includes those between 50% and 75%, and Bin 4 includes those between 75% and 99%. The sample is restricted to households with at least 26 gasoline purchases in a year and purchase intervals of less than 60 days.

C.2 Dynamic Demand: Correlation in Shares Over Time

Though ultimately the importance of demand-side dynamics in the data is estimated by the model, it is informative to examine the reduced-form relationships between key elements. The dynamic model 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 16.

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

Table 16: Regressions with Share as the Dependent Variable

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

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

that are challenging to explain with contemporaneous variables.[61]

C.3 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.[62] Using station-specific wholesale costs, we regress the 30-day lead wholesale cost on the current

[61] We have also estimated specifications that add lagged prices. Consistent with past prices and demand shocks affecting current choices, lagged prices are significant and the lagged share coefficient is similar.

[62] 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.

wholesale cost and the 30-day ahead future. In particular, we estimate the following equation.

$$c_{nt+30} = \alpha_1 c_{nt} + \alpha_2 F_t^{30} + \gamma_n + \epsilon_{nt} \tag{36}$$

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

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.

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.[63]

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

D Identification: Monte Carlo Exercise

We use Monte Carlo simulations to demonstrate that (1) our model does not incorrectly attribute persistent between-consumer heterogeneity to state dependence, and (2) the model recovers the correct parameters when state-dependent demand is present. Therefore, our method captures the presence of dynamic demand even when persistent unobserved heterogeneity is generated from more complex preferences than we model. As discussed in the paper, this is important because dynamic demand introduces bias not only through the demand elasticities but also through the firms' first-order conditions, which account for future demand. Further, the simulations help illustrate how state dependence can generate biased elasticities when not included in the model.

For the data-generating process, we assume that the indirect utility consumer i receives from purchasing product $j > 0$ in region r and period t is:

$$u_{ijrt} = (\beta + \pi_1 D_i) + (\alpha + \pi_2 D_i) p_{jrt} + \xi_j + \Delta\xi_{jrt} + \epsilon_{ijrt}, \tag{37}$$

where p_{jrt} is the retail price, ξ_j denotes product fixed effects, $\Delta\xi_{jrt}$ is a structural error term, and ϵ_{ijrt} is a consumer-specific logit error term.[64] A consumer that selects the outside good receives $u_{i0rt} = \epsilon_{i0rt}$.

Persistent consumer-specific preferences are governed by the parameters (π_1, π_2) and load onto the demographic variable D_i (e.g., "income"). This variable captures heterogeneity in consumer preferences in the propensity to buy any good (π_1) and price sensitivity (π_2).

We consider 5 different choices for the parameters (π_1, π_2), including the baseline standard logit. Specifically, for (π_1, π_2), we choose $(0, 0)$, $(4, 1)$, $(8, 0)$, $(0, 2)$, $(8, 2)$. These specifications help illustrate how different types of persistent between-consumer heterogeneity might introduce bias into our model. We assume that D_i is distributed according to the normal N(0,1) distribution. For each specification, we simulate 50 regions and 100 periods, with 6 products in each region-period. We draw 500 individuals in each region and use the standard choice probability equations to calculate shares. We choose parameter values that result in reasonable mean shares across specifications, ranging from 0.035 to 0.092. Since, for the purposes of this exercise, we are interested in estimating demand, we simulate prices by randomly drawing markups and adding these to marginal costs. The values used for the simulations are reported in Table 19.

In addition to these static specifications, we also choose five dynamic specifications. As in the paper, we assume that a fraction of consumers, λ, receive a preference shock, $\bar{\xi}$, for the product they purchased in the previous period. The remaining $1-\lambda$ consumers choose according to the standard logit. We consider $\lambda \in \{0.1, 0.3, 0.5, 0.7, 0.9\}$ and set $\bar{\xi} = 8$. Thus, aside from

[64]The product fixed effects allow for persistent (shared) preferences for specific products, but these (as might be expected) are easy to control for.

Table 17: Monte Carlo Results: Dynamic Model Estimates

	True Parameters			Elasticity	Estimate
Specification	π_1	π_2	λ	ϵ_D	$\hat{\lambda}$
1	0	0	0	-8.94	0.000
2	4	1	0	-4.29	0.000
3	8	0	0	-8.54	0.000
4	0	2	0	-1.78	0.000
5	8	2	0	-1.86	0.000
6	0	0	0.1	-5.64	0.095
7	0	0	0.3	-4.45	0.340
9	0	0	0.5	-3.85	0.496
9	0	0	0.7	-3.41	0.686
10	0	0	0.9	-3.05	0.888

the parameters π_1, π_2, and λ, all eight specifications share identical parameterizations.

Table 17 reports the 10 specifications in our simulations. To illustrate the magnitudes of the impact of consumer heterogeneity and state dependence, we report in this table the true median own-price elasticity. Because the richer preferences shift the marginal consumers, specifications 2 through 10 have less elastic demand than the baseline logit.

We estimate the model following the approach detailed in the body of the paper, assuming the structure of the dynamic demand model. We use the same estimation code for each specification, including the same (non-zero) initial parameter values, and we confirm that we reach the minimum. The last column of Table 17 reports the estimated value for $\hat{\lambda}$, the share of consumers that are affiliated. The model recovers the true share of affiliated consumers, even in the presence of random coefficients (which are not directly modeled in our empirical specification). In specifications 1 through 5, the model estimates that 0.0 percent of customers are subject to state dependence. By contrast, the estimates in specifications 6 through 10 are close to the true values, with small differences that can be explained by sampling error and our modest sample size.[65]

As an additional way to show the potential for random-coefficients demand to bias our model, we provide the estimated price coefficients and demand elastictities from a (misspecified) static logit model in Table 18. As expected, specifications 2 through 5 have substantial bias in the estimated price coefficient (the structural parameter), but the specifications have somewhat more modest bias in own-price elasticities. By contrast, the specifications with (unaccounted for) dynamic demand yield greater bias in the own-price elasticities, even for small values of λ.

We conclude two things from these simulations. First, our model is capable of correctly recovering the true degree of state dependence, even when the between-consumer heterogeneity is more complex than the standard logit model. Our dynamic model does not attribute

[65] The estimated values for $\bar{\xi}$ for specifications 6 through 10 are 8.0, 8.2, 7.9, 8.3, and 8.3, close to the true value of 8. To ensure that the state dependence shock has bite (i.e., if $\bar{\xi} = 0$, λ is irrelevant), we set the minimum value of $\bar{\xi}$ to 3 in estimation. Specifications 1-5 are above this lower bound, but it is irrelevant as λ is estimated to be zero.

Table 18: Unobserved Heterogeneity and Bias in Static Logit Estimates

	Random Coefs		Dynamics	Price Coefficient			Demand Elasticity		
Specification	π_1	π_2	λ	α	$\hat{\alpha}$	Std. Err.	ϵ_D	$\hat{\epsilon}_D$	% Bias
1	0	0	0	-3	-3.000	0.003	-8.94	-8.95	0.0
2	4	1	0	-3	-1.542	0.015	-4.29	-4.44	3.5
3	8	0	0	-3	-2.640	0.016	-8.54	-7.71	-9.7
4	0	2	0	-3	-0.679	0.014	-1.78	-1.93	8.4
5	8	2	0	-3	-0.688	0.014	-1.86	-1.94	4.7
6	0	0	0.1	-3	-2.099	0.008	-5.64	-6.21	10.0
7	0	0	0.3	-3	-1.730	0.010	-4.45	-5.02	12.7
8	0	0	0.5	-3	-1.554	0.011	-3.85	-4.43	15.3
9	0	0	0.7	-3	-1.436	0.012	-3.41	-4.04	18.4
10	0	0	0.9	-3	-1.348	0.013	-3.05	-3.73	22.2

Notes: Table 18 reports the results when the DGP has random coefficients (2-5) and has dynamic demand (6-8). The true price coefficient is α, and the estimated coefficient from the logit regression is $\hat{\alpha}$. We estimate the (misspecified) regression equation $\ln(s_{jrt}/s_{0rt}) = \beta + \alpha p_{jrt} + \xi_j + \Delta\xi_{jrt}$. We instrument for price with marginal costs plus a small error term (see Table 2 for details on z and mc). We also report, in the last three columns, the true own-price demand elasticity (ϵ_D), the estimated own-price demand elasticity ($\hat{\epsilon}_D$), and the percent bias of the estimate. The results show that both random coefficients and state dependence generate bias in static logit estimates.

persistent between-consumer heterogeneity to such dynamics in demand. Second, state dependence in demand is as capable of generating bias when omitted from the model as persistent unobserved heterogeneity (and perhaps more so, when accounting for the firms' first-order conditions). In many cases, within-consumer state dependence may be the most relevant feature of the economic environment for the question at hand. In cases where persistent between-consumer unobserved consumer heterogeneity is also important, our approach may be used as a test for the presence of dynamic demand, and the results can inform the best model to use for a more detailed analysis.

D.1 Parameters for Monte Carlo Exercises

Table 19: Data-Generating Process

Parameter/Variable	Value	Description
β	2	Demand intercept
α	-3	Mean price coefficient
ξ_j	(2, 2, 1, 1, 0, 0)	Product fixed effects
$\Delta\xi_{jrt}$	U(0,1)	Error term
D_i	N(0,1)	Demographic variable
mc_{jrt}	U(1,3)	Marginal costs
p_{jrt}	mc_{jrt} + U(0,2)	Prices
z_{jrt}	mc_{jrt} + N(0,0.01)	Instrument for price

E Estimation Routine: Reducing the Computational Burden

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.

In Section 4.1, we showed that the choice patterns can be expressed in terms of the $J+1$ parameters $\{s_{0t}(j)\}$ in each market. We now show that the system reduces to two parameters in each market, where the remaining $J-1$ parameters are solved for by a quadratic function.

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

$$\sum_z r_{zt} \cdot s_{0t}(z) \exp\left(\sigma_{jt}(z)\right) = \sum_z r_{zt} s_{0t}(z) + \left(\exp\left(\sigma_{jt}(j)\right) - 1\right) r_{jt} s_{0t}(j). \tag{38}$$

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

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

Conditional on the dynamic parameters and observables, there are only two remaining unknowns: $s_{0t}(0)$ and $\sum_{z \in 0,J} r_{zt} s_{0t}(z)$. Thus, we can solve for $\{s_{0t}(j)\}$ in each market using the quadratic formula. As $\{\delta_{jt}\}$ are identified conditional on these choice probabilities, we can obtain these mean utility parameters by solving for only two unknowns in each market, regardless of the number of products.

F Empirical Application: Additional Tables and Figures

F.1 Summary Statistics by Observation

Table 20: Retail Gasoline in Kentucky and Virginia: Oct 2013 - Sep 2015

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

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

F.2 Summary Statistics by Brand

Table 21: Summary of Brands

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

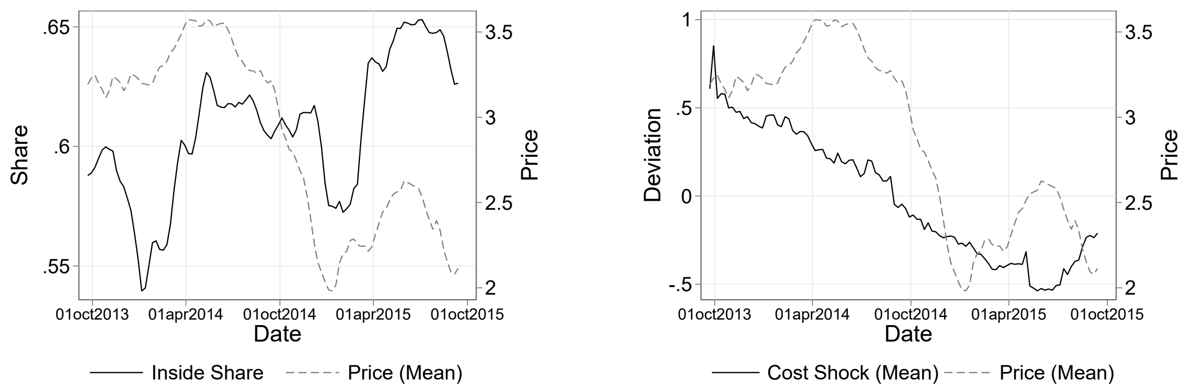
Notes: Table provides summary statistics by brand. The FRINGE brand is a synthetic brand created by aggregating brands that do not appear in 10 or more of the 241 markets in our data. Additionally, if a brand does not make up more than 2 percent of the average shares within a market, or 10 percent of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

F.3 Shares, Prices, and Instrument

Figure 8: Shares and Prices

(a) Total Market Shares and Prices

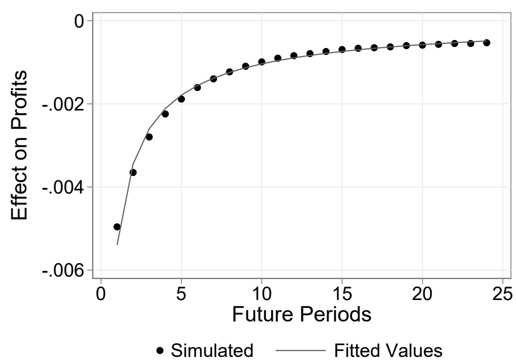
(b) Instrument and Prices



Notes: Panel (a) displays the sum of market shares for all brands, excluding only the outside option, plotted 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.

G Details on Forward Simulations

Figure 9: Simulated and Fitted Values



Notes: Figure displays the equilibrium effect of a marginal increase in price on future profits. An increase in price decreases future profits. This effect is diminishing over time. The points display simulated values from the data, and the line displays fitted values from a regression model.

To test whether firm expectations are consistent with expected profits under our supply-side approach, we simulate equilibrium prices and profits using marginal deviations for individual prices. We perturb the price of a specific brand in a specific week in a specific market by lowering that brand's price by $0.01. We re-compute shares in that period, then we compute equilibrium prices over the next 24 weeks. We repeat this experiment for all brands within the market and the period, and we iterate over markets and periods. We limit our analysis to the 75 markets and 52 weeks affected by our merger counterfactual.

We use the difference in simulated profits and realized profits to compute the marginal effect of a $0.01 change in price on profits τ periods in the future. By scaling the difference by the magnitude of the price change, we obtain an estimate of the effect on current-period profits and each period-specific component of the derivative $\frac{\partial \pi_{j(t+\tau)}}{\partial p_{jt}}$. The simulated effect on current-period profits from a price change matches the static derivative of the model. To calculate the discounted future values, we use a weekly discount rate of $\beta = 0.999$, which corresponds to an annual discount rate of 0.949. We use the discounted values of $\frac{\partial \pi_{j(t+\tau)}}{\partial p_{jt}}$ to obtain an estimate of the full value of $\beta \frac{\partial E[V_j(\cdot)|\cdot]}{\partial p_{jt}}$.

The dynamic equilibrium simulation shows that an increase in price today has a negative effect on future profits, and this effect shrinks over time. To calculate the net present value of the effect on profits, we use a fitted model to predict effects beyond 24 weeks and to mitigate the effect of simulation noise. Specifically, we fit a curve of the form $\ln\left(-\ln\left(E_t[\frac{\partial \pi_{j(t+\tau)}}{\partial p_{jt}}]\right)\right) = \omega_0 + \omega_1 \ln(t) + \epsilon_t$. We estimate $\omega_0 = 1.653$ and $\omega_1 = 0.119$ with least squares regression, using the 24 mean values for each period. Figure 9 plots the simulated effects on future profits against the fitted values. The fitted values explain 99 percent of the variation in mean values.