NBER WORKING PAPER SERIES

PHARMACEUTICAL ADVERTISING IN DYNAMIC EQUILIBRIUM

Pierre Dubois
Ariel Pakes

Working Paper 35025
http://www.nber.org/papers/w35025

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
April 2026, Revised June 2026

We thank Steven Berry, Myrto Kaloupstidi, Aureo de Paula, Johannes Horner, John Rust, Bernard Salanié, Philipp Schmidt-Dengler and participants in many seminars and conferences. We gratefully acknowledge the financial support of the ANR under grants no. ANR17-EURE-0010 (Investissements d'Avenir program) as well as the European Research Council (ERC) and the European Union (grant agreement No 101141890-EHCI). This research was partly carried out at TSE's Data & Computing Center, whose help we gratefully acknowledge. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

At least one co-author has disclosed additional relationships of potential relevance for this research. Further information is available online at http://www.nber.org/papers/w35025

NBER working papers are circulated for discussion and comment purposes. They have not been peer-reviewed or been subject to the review by the NBER Board of Directors that accompanies official NBER publications.

© 2026 by Pierre Dubois and Ariel Pakes. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

Pharmaceutical Advertising in Dynamic Equilibrium
Pierre Dubois and Ariel Pakes
NBER Working Paper No. 35025
April 2026, Revised June 2026
JEL No. I0, I1, L2, M37

ABSTRACT

Direct-to-consumer advertising (DTCA) of prescription drugs may expand treatment access but also risks promoting overuse and business stealing without generating welfare gains. Among developed nations, only the United States and New Zealand permit DTCA, whereas detailing - promotion aimed at prescribers - is widely practiced. This paper analyzes the impact of DTCA on profits by modeling a counterfactual environment in which DTCA is banned. This is implemented through a dynamic equilibrium framework that adapts the Experience-Based Equilibrium (Fershtman and Pakes, 2012) for empirical analysis. EBE incorporates constraints on the cognitive abilities of decision-makers and mitigates researchers' computational concerns. Using data from four therapeutic markets, we first validate the EBE's ability to replicate observed advertising patterns, then simulate counterfactual DTCA bans. Both the data and our empirical work indicate that DTCA and detailing are strong complements in generating current profits, and our results illuminate the need to account for this when evaluating the ban. The ban leads firms to reduce detailing and has a negative effect on profits in all markets, but the magnitude of the effect varies from under 5% in the market for Ulcer to 27.5% for Asthma medications.

Pierre Dubois
Toulouse School of Economics
Pierre.Dubois@tse-fr.eu

Ariel Pakes
Harvard University
Department of Economics
and NBER
apakes@fas.harvard.edu

1 Introduction

The United States represents the world's largest pharmaceutical market, with annual revenues approaching $600 billion and promotional expenditures exceeding $20 billion. Pharmaceutical promotion in the U.S. primarily takes two forms: detailing, which targets prescribers through visits and free samples, and direct-to-consumer advertising (DTCA), approximately three-quarters of which is allocated to television. Among developed nations, only the United States and New Zealand permit DTCA for prescription drugs. In the U.S., DTCA was first authorized in print media in 1985 and subsequently expanded to include television in 1997, with expenditures increasing substantially since then (Rosenthal et al., 2002; Donohue et al., 2007).

This paper examines the impact of pharmaceutical advertising on firm profitability. Profits serve as the primary incentive for privately funded pharmaceutical R&D, which drives innovation (Acemoglu and Linn, 2004; Blume-Kohout and Sood, 2013; Dubois et al., 2015). Approximately 80% of pharmaceutical R&D is privately funded (Ho and Pakes, 2025), and the social benefits of pharmaceutical innovation appear to far outweigh their costs[1]. However the effect of DTCA on profits is not clear. While DTCA may generate market expansion effects that enhance profitability[2], substitution effects between branded and generic drugs - or between competing brands - are likely to produce more ambiguous outcomes, particularly given the high costs of marketing.

We empirically investigate four major therapeutic markets - hypercholesterolemia, depression, asthma, and ulcers - which collectively account for approximately 30% of DTCA expenditures and 23% of detailing expenditures (averaging $11.8 billion annually during our 2002–2014 sample period). Using IMS (IQVIA) data, we estimate separate models for each market focusing on the relationship between detailing, DTCA, and profits.

Our analysis begins with the estimation of demand models following Berry et al. (1995) for each therapeutic market. The demand system generates period-specific *quality* variables, which we extract and model

[1] Buxbaum et al. (2020) report that between 1990 and 2015, life expectancy increased by 1.32 years per decade, with pharmaceuticals accounting for approximately 35% of this increase (0.46 years). During the 2005-2015 period, there were 48.9 million live births in the U.S. Valuing each life-year at $100,000 generates approximately $2.25 trillion in economic value. This calculation excludes the impact of pharmaceutical research on morbidity (see, e.g., Chernew et al. (2016)), quality of life improvements, benefits to immigrant populations, and spillovers to other countries.

[2] DTCA could also potentially improve consumer welfare. The welfare impact depends on whether the expansion reaches consumers for whom the drug would be effective, as DTCA may also encourage excessive drug utilization.

as a controlled Markov process. The increments of this process depend on both DTCA and detailing expenditures. The estimates from the demand system and this Markov process serve as two of the primitives used to compute profits. The third is marginal cost which we come back to below.

The framework for analyzing advertising investments is novel and comprises two complementary models. The first is an empirical model that examines advertising decisions as a function of observable variables in our data and two serially correlated unobservables - one for DTCA and one for detailing. These unobservables capture information available to firms but not to the econometrician. The empirical model assumes that investment decisions are made to maximize the discounted value of future net cash flows, conditional on firms' perceptions of likely market evolution. Importantly, this model *does not* assume that these perceptions are either correct or consistent with equilibrium conditions.

The first-order conditions from the empirical model serve three purposes: (1) estimating marginal costs, (2) identifying which observable variables drive our two investment variables, and (3) characterizing the properties of the serially correlated unobservable determinants of investments. The determinants of investments are subsequently used as state variables in the equilibrium model employed to evaluate our counterfactual scenario - a ban on DTCA.

Our equilibrium analysis employs the Experience-Based Equilibrium (EBE) framework developed in Fershtman and Pakes (2012). EBE accommodates asymmetric information between firms in a dynamic environment and imposes consistency between firms' perceptions of market evolution and observed outcomes at frequently visited states, while mitigating the demanding cognitive requirements of a full Markov Perfect (or Bayesian Perfect) Equilibrium. EBE conditions on the state variables that firms condition on when making their investment decisions, but is silent on what those state variables are. This is the source of the complementarity between the two models; we use the empirical model to identify the relevant state variables without needing to assume equilibrium perceptions. A novel algorithm enables efficient computation of equilibrium policies conditional on those state variables. We compare the in-sample fit of the equilibrium policies to both the empirical model and actual data before moving to our counterfactual.

Both our data and empirical results indicate that a realistic assessment of a DTCA ban requires modeling how detailing expenditures would respond to such a policy. An advantage of using an equilibrium model for

this evaluation is that it allows us to analyze detailing's response without requiring data from an environment without DTCA, enabling the policy to be evaluated before it is implemented. Our counterfactual analyzes what profits would have been during our sample period in the absence of DTCA. Our findings reveal that profits would decline in all four markets, though the magnitude of this decline varies across markets and depends critically on differences in detailing responses. Moreover the profit reduction would not affect all products within a market equally, and these intra-market differences would influence both R&D incentives and the evolution of market structure.

Relationship to theory and applied work on market dynamics. Much of the applied industrial organization literature that computes or empirically analyzes counterfactual market dynamics relies on the concept of Markov Perfect Equilibrium (MPE). That framework was introduced into Industrial Organization by Maskin and Tirole (1988a,b) and was operationalized for applied analysis by Pakes and McGuire (1994) and Ericson and Pakes (1995). However both theoretical and empirical research have raised concerns about the cognitive demands it places on decision-makers and the computational demands it places on researchers. Incorporating asymmetric information within a Bayesian Perfect framework further exacerbates these challenges.

In response, subsequent theoretical work has developed alternative weaker equilibrium concepts. These include: Self-Confirming Equilibrium, which relaxes assumptions on off-path beliefs (Fudenberg and Levine, 1993); models of procedurally rational players, where agents observe only their own actions but infer consequences consistent with aggregate behavior (Osborne and Rubinstein, 1998); analogy based equilibria by Jehiel (2005) and Jehiel and Samet (2007) which coarsens the state space; and Berk-Nash Equilibrium, in which agents' prior beliefs need not assign positive probability to the true model, but learning processes guide them toward close approximations (Esponda and Pouzo, 2016).

An Experience-Based Equilibrium is related in different ways to each of these papers. The unifying difference is that the theory literature is primarily concerned with learning about competitors actions in an environment where the payoff-relevant state variables are constant over time. EBE focuses on learning in a dynamic environment where payoff-relevant states evolve over time as a consequence of firms' actions. In

this context EBE coarsens perceptions, and only requires them to be consistent with outcomes at states that are visited repeatedly.

Two key aspects of the applied literature underpin our approach: the estimation of investment policy functions and the computation of dynamic market equilibria. The closest paper to ours in estimation of investment policies is Bajari et al. (2007). We differ from Bajari et al. (2007) in two ways: we do not impose consistency between the perceptions that led to the observed investment decisions and the outcomes we see in the data, and we allow for serially correlated state variables (for a recent review of estimation in dynamic games, see Aguirregabiria et al. (2021)).

A growing literature addresses the computation of equilibria in complex dynamic games using advanced computational techniques and approximation methods. In economics this was preceded by work on computational tools for single agent dynamic models (Rust, 1996), and methods to approximate dynamic programming using randomization (Rust, 1997). Subsequent work dealt explicitly with dynamic games (Pakes and McGuire (2001); Benkard et al. (2006); Doraszelski and Judd (2012); Aguirregabiria and Nevo (2013); Benkard et al. (2015); Aguirregabiria et al. (2021)). The primary focus has been on computing Markov (or Bayesian) Perfect Equilibria[3]. However, few studies have tackled the dynamic modeling of complex sectors like pharmaceuticals, with notable exceptions such as the calibrated simulation in Filson (2012). Our equilibrium calculations depart from this literature by not attempting to generate policies for the entire state space. Instead, we focus on states that are visited repeatedly (or the "recurrent class" of the underlying Markov process). While this approach has limitations, it significantly simplifies computation and can be rationalized through a learning process.

Relationship to the literature on pharmaceutical advertising. A large literature examines the demand and health consequences of pharmaceutical advertising (Arcidiacono et al., 2013; Shapiro, 2018; Sinkinson and Starc, 2019). Arcidiacono et al. (2013) assesses welfare effects of generic entry in the antiulcer market, Grennan et al. (2024) show promotional payments can improve prescribing outcomes, and prior work highlights the challenges of modeling advertising decisions (Villas-Boas, 1993; Dubé et al., 2005).

[3]See, for example, (Ryan, 2012; Sweeting, 2013; Dubé et al., 2005; Doraszelski and Markovich, 2007). In some cases, such as Abi-Rafeh et al. (2026), industry-specific institutional structures help reduce the dimensionality of strategic interactions.

Ling et al. (2002) analyze how DTCA and detailing jointly shape prescription and OTC drug sales. Bhattacharya and Vogt (2003) predict declining advertising over a drug's lifecycle, Lakdawalla et al. (2013) find advertising intensity increases with market size after Medicare Part D, de Frutos et al. (2013) show high-quality drugs are more heavily advertised, and Ridley (2015) shows firms offset copayment sensitivity through physician-targeted advertising.

Dave (2013) concludes DTCA is informative and market-expanding while physician-targeted advertising is more persuasive. Anderson et al. (2013) find advertising is more informative for high-quality, small-share brands. Iizuka and Jin (2005, 2007), Wosinska (2005), and Iizuka (2004) document effects on doctor visits, adherence, and drug quality, while Ching and Ishihara (2010), Ching et al. (2016), and Liu et al. (2017) show advertising complements scientific information and generates cross-product spillovers.

Lakdawalla and Philipson (2011) find advertising declines post-expiry reducing consumer welfare, Lakdawalla (2018) argues it complements innovation, Anderson et al. (2016) find comparative advertising harms targeted competitors, and David et al. (2010) find detailing may worsen patient-drug matching. Sinkinson and Starc (2019), Shapiro (2018), and Shapiro (2022) document own-brand and spillover demand effects, with the latter estimating meaningful productivity gains from increased prescriptions.

Our focus and modeling framework differ from much of the prior literature. First we do not examine the health consequences of pharmaceutical advertising; our data is not rich in the details needed for that analysis[4]. Instead we focus on the extent to which policies targeting pharmaceutical advertising affect pharmaceutical profits—a consideration of interest due to its implications for R&D incentives and the health benefits of the drugs that R&D produces (see the literature cited in Ho and Pakes (2025) or Dubois (2025)). Second both our data and empirical results demonstrate that a product's detailing expenditures are closely tied to its direct-to-consumer advertising (DTCA) expenditures. These two forms of advertising are jointly determined, so any policy affecting one type is likely to influence the other. Also current advertising shapes future demand, and these effects depend on competitor advertising. These are all features of the environment

[4]In particular we do not see the implications of drug use, and hence cannot distinguish overuse from justifiable use. Also, as we note below, the data we have for calculating the fractions of patients who use an outside alternative rather than a drug to treat our diseases is based on relatively small samples.

that we incorporate in our analytic framework.

The remainder of the paper proceeds as follows. Section 2 introduces the data and defines the market structure. Section 3 presents the demand model and estimates the effects of advertising on perceived product quality. Section 4 outlines the empirical model of advertising behavior and the estimation of marginal costs. Section 5 describes the Experience-Based Equilibrium (EBE) framework and the computation of equilibrium policies. The readers not interested in the details of the estimation algorithm should be able to omit section 4.2, and those not interested in the details of how to compute equilibrium should be able to omit section 5.1, and find the rest of the paper understandable. Section 6 presents the results of our counterfactual policy analysis. Section 7 concludes.

2 Market and Descriptive Statistics

2.1 Data and Market Definitions

We use data from IMS Health (IQVIA) MIDAS on revenues and quantities by drug, country and quarter for the period from 2002 to 2014. These data provide wholesale transactions values and quantities disaggregated at form and strength level of each drug, together with information on the molecule, the patent expiration date and whether it's a generic version of the initial branded drug. We use it for aggregate sales values in US dollars and volumes in standard units[5].

The IMS Health (IQVIA) Global Promotional Track Advertising data provides monthly advertising expenditures for each drug by media (Detailing, Direct To Consumer Advertising, others) for the US but it is only available from 2005 to 2014. Table 10 in appendix A.2 shows that the total DTCA in our markets represents between 24 and 37% of total DTCA and 20 and 25% of total detailing across the years studied.

Finally, we use the Medical Expenditures Panel Survey (MEPS) to obtain both the size of the U.S. market and the fraction being treated with medication for each indication studied. Appendix A.3 describes

[5]Separate data is provided for: drugstores, mail service, food stores, clinics, federal hospitals, non-federal hospitals, long term care houses, HMO, and home health care segments. We analyze four markets which are the most advertising intensive during the time period studied (2005-2014): Ulcer, Hypercholesterolemia, Depression and Asthma. Appendix A.1 provides a brief description of each of these diseases and the drugs that treat and mitigate their effects. For these indications, some of the health care segments are less relevant and not affected by promotional strategies. We thus do not account for sales in federal hospitals, health maintenance organizations (HMOs) and home health care for our four markets.

this data in more detail and provides estimates of both the population of patients with each of the four illnesses and the fraction being treated. The population in need of treatment in anticholesterol increases from about 60 million in 2007 to 70 million in our period, in anti-depression it grows from 30 to more than 40 million, in antiasthma it increases from 23 millions to 25 millions people, but in antiulcer it decreases rather dramatically from a bit more than one million to half a million.

2.2 Descriptive Statistics

All four of our markets are highly concentrated with four firm sales concentration ratios ranging from .61 (in antiasthma) to .77 (in antiulcer). Also the large sales firms are typically also the firms that do the most advertising; in all markets at least three of the top four sales firms are also the top performers in DTCA or detailing, and nine of the top four firms in sales in the four markets are in the top four firms in both detailing and DTCA.

Table 1: Detailing and DTC Advertising per product ($US1000/year)

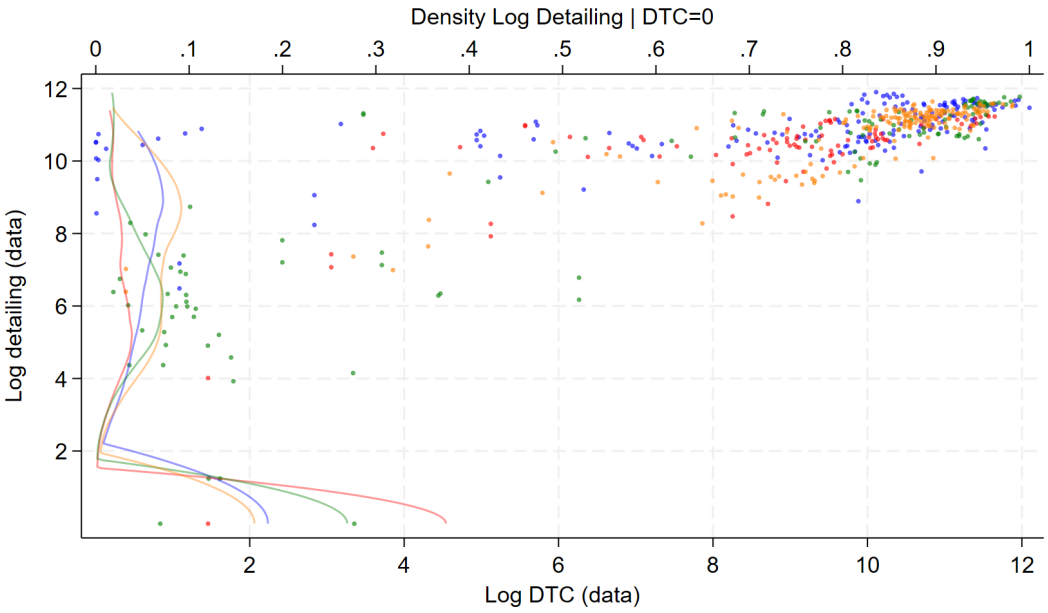
	Detailing			DTC Advertising		
($US1000/year/product)	All	Branded	Top 4	All	Branded	Top 4
Anticholesterol	29,437	40,884	113,713	12,454	17,391	26,278
Antiasthma	18,468	24,309	110,186	7,509	10,599	79,331
Antidepressants	13,844	22,464	136,679	4,808	8,251	20,642
Antiulcer	9,974	14,954	138,048	2,561	3,995	58,139

Note: Mean per product for the 2005-2014 period. Multiple generics of a molecule are aggregated and counted only once. They have zero advertising. Remark that the top four advertisers for detailing and DTCA are the same for Antiasthma while one is different for Antiulcer and Anticholesterol and two are different for Antidepressants. In terms of drug sales, the top 4 sellers in each market represent between 60 and 77% of total sales in the market.

Table 1 provides the average advertising per product per quarter for all products, for branded products only, and for the mean of the four largest advertised products. The overall averages include generic products who very rarely advertise, so we ignore generic advertising below[6]. The rank correlation across markets between average detailing and DTCA per branded product is perfect, with anticholesterol drugs doing the most advertising and antiulcer drugs the least.

[6]There is no generic DTCA in three of the four markets and it is 2.7% of total DTCA in the fourth. Generic detailing is between 2.6 and 6.6% of total detailing.

Figure 1: Detailing versus DTCA distribution in the data



Note: Blue: Anticholesterol, Red: Antiulcer, Green: Antidepressants, Orange: Antiasthma. The density distribution for detailing when DTCA is zero is plotted using the left vertical axis for log detailing and the top horizontal axis for the density values.

Figure 1 shows that the within-market correlation between detailing and DTCA is also very high in all four markets. Each point is a data point corresponding to the amount[7] of detailing on the vertical axis and the amount of DTCA on the horizontal axis *when both are positive*. The colors differentiate the four markets. DTCA is never positive when detailing is zero. However sometimes detailing is positive when DTCA is zero. So to provide a complete picture of the relationship between the two types of advertising the figure provides the density distribution of detailing when DTCA is zero (the curves on the left axis). The detailing expenditures of drugs that do not do DTCA is typically lower than those that do.

Table 2 provides details on the variation in advertising within and across firms. Throughout the paper, d index detailing and D indexes DTCA. Some products never advertise, so columns 2 and 3 provides the share of products that do some form of advertising in a given quarter. Overall 29-63% of observations do detailing but only 7-18% do DTCA. Columns 4 and 5 provide the share of drugs doing advertising in two adjacent periods, while columns 6 and 8 provide the share of drugs that have zero advertising between two

[7]Log values are log of (1+value).

periods where they do advertise. It is clear that firms doing either form of advertising periodically stop and then restart their advertising expenditures. Moreover if they stop and then restart, there is an extended period in the interim (particularly noticeable for DTCA). This underlies the need for a selection correction in our empirical work.

Table 2: Advertising changes over time

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
						$t \in \{t_d^0, t_d^1\}$		$t \in \{t_D^0, t_D^1\}$	
Market	N	$a_{d,t} > 0$	$a_{D,t} > 0$	$a_{Dt} > 0$ $a_{Dt-1} > 0$	$a_{dt} > 0$ $a_{dt-1} > 0$	$a_{dt} = 0$	$a_{dt} = 0$, $a_{dt-1} > 0$	$a_{Dt} = 0$	$a_{Dt} = 0$, $a_{Dt-1} > 0$
Antiasthma	1409	0.627	0.140	0.118	0.601	0.051	0.026	0.218	0.023
Antiulcer	1314	0.292	0.078	0.062	0.271	0.182	0.032	0.187	0.028
Anticholesterol	902	0.560	0.186	0.161	0.539	0.076	0.017	0.136	0.045
Antidepressants	1325	0.432	0.106	0.087	0.413	0.054	0.022	0.185	0.040

Note: t_h^0 and t_h^1 denote the first and last period in the data where we observe positive advertising of type $h \in \{d, D\}$.

We now turn to the description of demand and how advertising impacts the demand model.

3 Demand and the Impact of Advertising

3.1 Aggregate Demand for Prescription Drugs

For each market we use a different random coefficient logit à la Berry et al. (1995) demand function that allows for unobserved product-by-quarter specific unobservables (that we refer to as perceived quality). If we index individuals by i, time by t, and drugs by $j = \{0, 1, \ldots J\}$, where $j = 0$ is the outside good (the person is not medicated), the underlying individual utility function for those with the given conditions is [8]

$$U_{ijt} = \beta_p^i(X_j)p_{jt} + \beta_{m(j)} + \beta_x X_{jt} + \xi_{jt} + \varepsilon_{ijt}, \tag{1}$$

where p_{jt} is the price of drug j at time t, β_m is a molecule fixed effect, X_{jt} are controls that vary by market and include dummy variables for; whether the drug is generic, whether it can be sold over the counter (OTC) without prescription (which happens only for the Antiulcer market), and time interacted with dummies for

[8] To define our products, we aggregate across strengths, formats and forms of administration. Thus our products are differentiated by molecule and brand name. When there is a molecule that has gone off patent there can be several products with the same molecule due to the introduction of generics.

sub classes of drugs within a market. $\beta_p^i(X_j)$ is a random coefficient assumed to be normally distributed $(\beta_p^i \sim N(\beta_p(X_j), \sigma_p^2(X_j)))$, and ε_{ijt} is an i.i.d. extreme value error.

The ξ_{jt} terms are our product-by-time specific effects that affect demand but our model cannot directly control for and insure that the market shares predicted by our model exactly equal those in our data. Among the factors we cannot control for is the current impact of current and past advertising investments. As in Berry et al. (1995) we allow the ξ_{jt} to be correlated with price, and model their evolution as a controlled Markov process whose increments are affected by advertising.

Given the i.i.d. extreme value distribution of ε_{ijt}, market shares are obtained by integration as

$$s_{jt}\left(p_t, \xi_t\right) = \int \frac{\exp(\beta_p^i(X_j)p_{jt} + \beta_{m(j)} + \beta_x X_{jt} + \xi_{jt})}{1 + \sum_k \exp(\beta_p^i(X_k)p_{kt} + \beta_{m(k)} + \beta_x X_{kt} + \xi_{kt})} dF(\beta_p^i(X)), \tag{2}$$

and demand for j is obtained as

$$D_{jt}\left(p_t, \xi_t\right) = M_t s_{jt}\left(p_t, \xi_t\right),$$

where M_t is market size (the number of people with the disease) at time t and is obtained directly using the MEPS data estimates of the outside good market share or s_{0t} (see Appendix A.3). The MEPS data show that the outside good market share for ulcers is declining over time and is the smallest of our four conditions, going from 30% in 2002 to around 20% in 2015. That share is highest for asthma (more than 60%), while the outside share averages about 40% for cholesterol and depression after 2010 and is slightly higher before.

3.2 The Impact of Advertising on "Quality" (i.e. the ξ process)

The $\{\xi_{jt}\}_{j,t}$ insure that the demand model fits exactly for each product in each period. They account for perceptions of drug quality that can change over time due to promotional activity or other factors, so we assume it evolves as a Markov process whose increments are impacted by advertising, or

$$\xi_{jt} = \rho_\xi \xi_{jt-1} + f(a_{D,jt}, a_{d,jt}) + z_{jt}\beta_z + \mu_{jt}, \tag{3}$$

where $a_{D,jt}$ is DTCA for product j in the period prior to t and $a_{d,jt}$ is the analogous detailing variable. The z_{jt} allow the evolution of quality to differ for generic or over the counter drugs, and included time trends. The $\{\mu_{jt}\}$ are innovations in the quality measure that are assumed mean independent of variables known at period $t-1$. So we are allowing advertising to affect the evolution of demand in a way to be determined by the data.

Note that equation (3) allows the impact of current advertising on future ξ to decay at the rate ρ_ξ, which is a parameter we estimate. Since the role of advertising is central to our analysis we focus on the functional form for $f(a_{D,jt}, a_{d,jt})$. Our final estimates use the specification

$$f(a_{d,jt}, a_{D,jt}) \equiv \beta_{a_d} \log(1 + a_{d,jt}) + \beta_{a_D} \log(1 + a_{D,jt}) + \beta_{a_{d,D}} \log(1 + a_{d,jt}) \log(1 + a_{D,jt}), \tag{4}$$

but we also tested for: (i) separate serial correlation coefficients (i.e. ρ_ξ) for generics and patented alternatives (they were not needed), (ii) whether z_{jt} include time effects (not significantly different from zero, though we should keep in mind that there are time effects in the demand functions), dummies for generics and over the counter (OTC) drugs, (iii) whether μ_{jt} is allowed to be correlated with $a_{D,jt}$ or $a_{d,jt}$. The sign and magnitude of the coefficient $\beta_{a_{d,D}}$ will determine the extent of complementarity between current detailing and DTCA in the production of $\xi_{j,t}$, and through the impact of the $\xi_{j,t}$ on demand, their complementarity in the production of profits.

3.3 Estimation Method

As we have a longer time series on demand than on advertising, we first estimate the demand function (equation 2) and then use the estimates of $\{\xi\}_{j,t}$ it generates to estimate the advertising equation (3) for the periods when advertising data is available[9]. Both estimations rely on instrumental variables Z_{jt} and $\tilde{Z}_{jt}$ satisfying the following moment conditions (where $(\boldsymbol{\beta})$ denotes the vector of demand parameters)

$$E(\xi_{jt}(\boldsymbol{\beta}) Z_{jt}) = 0 \text{ and } E(\mu_{jt}(\rho_\xi, \beta_a, \beta_z) \tilde{Z}_{jt}) = 0.$$

[9]The joint estimation of demand and ξ equation is also possible and leads to overall similar parameters. We use the sequential method as it uses all of the demand data in a natural two step way.

Estimation of demand follows the two-step estimation procedure introduced in Berry et al. (1995), though we have to be careful about the choice of instruments. Below we allow a firm's advertising to be a function of the advertising and other characteristics of competing products, and since advertising impacts ξ we cannot use competing product characteristics as instruments. Accordingly we use functions of the current changes in prices in countries that do not advertise as instruments, as they should partially reflect changes in marginal costs of producing the drug[10]. Stacking those instruments with z_{jt} (which are assumed exogenous) into the vector Z_{jt}, the demand estimation uses a method of moments estimation procedure that minimizes a weighted quadratic norm of $\sum_{j,t} \xi_{jt}(\boldsymbol{\beta}) \otimes Z_{jt}$.

These estimates of $\{\xi_{jt}\}_{j,t}$ are then used to estimate the parameters of equation (3), the equation that determines the impact of the advertising variables on demand. That equation is estimated with two stage least squares using lagged advertising variables as instruments (as $\tilde{Z}_{jt}$).

Timing of advertising dynamics on perceived quality ξ. Our demand, and hence $\xi_{j,t}$ estimates, are from quarterly data (i.e. t indexes quarters). We investigated three alternative specifications for the distribution of the timing of the impact of advertising on perceived quality in equation (3): quarterly, semi-annually, and annually. We use the semi-annual specification as it had a better fit. However since this complicates the exposition of the derivations below, we relegate the development of the model that embodies this difference to the Appendix A.9, and proceed in the text as if the impact of advertising on the innovation in ξ is only from the last quarter's advertising expenditures.

3.4 Estimation Results for the Demand and ξ Equations

We have estimated several models for each market with varying specifications for the interactions between price and observed and unobserved product characteristics, and with different subsets of dummy variables. Table 3 shows the preferred models for each market[11].

[10] More precisely, we use functions of the change in price (of $p_{jct} - p_{jct-1}$) of the drug as well as of drugs in the same therapeutic classes in Canada, UK, France, Germany, Australia, and Mexico as instruments. The functions of other drugs are the means and sum within the ATC4 or ATC3 classes, where ATC refers to Anatomical Therapeutic Chemical classes.

[11] The estimates use the market sizes we calibrated from the MEPS data and led to outside good market shares shown in Figure 10 in Appendix A.3.

Table 3: BLP demand models

	Anticholesterol	Antiasthma	Antidepressants	Antiulcer
Price	-2.012***	-1.634***	-0.554***	-2.426***
	(0.230)	(0.331)	(0.144)	(0.441)
Price × Generic	-1.077**		-0.732***	
	(0.371)		(0.190)	
Price × OTC				-4.617***
				(1.007)
Generic			2.133***	1.893***
			(0.365)	(0.416)
OTC				-1.682***
				(0.347)
Standard Deviation Price Coefficient				
Price	0.823***	0.764***	0.230***	0.968***
	(0.134)	(0.142)	(0.061)	(0.220)
Controls				
Drug fixed effects	Molecule	Molecule ×Patent status	Molecule	Product
Year dummies			✓	
ATC4×Year dummies		✓		✓
Generic dummy			✓	✓
OTC dummy				✓
N	1049	1684	1592	1555
Parameters	20	46	30	34
Nb products	32	58	46	45
ATC(4) classes	4	14	3	5

Note: A molecule corresponds to the ATC5 level. A product fixed effect corresponds to the brand-molecule combination, a molecule and patent status fixed effect corresponds to the combination of the molecule fixed effect with a dummy for patent expiration. * for $p < .05$, ** for $p < .01$, and *** for $p < .001$.

The number of drugs marketed varies by quarter but totaled between 32 and 58 over the sample period for our markets. Within each market drugs are classified into ATC4 therapeutic groups by their active ingredients[12], and the different markets have between 3 and 14 of these classes within them.

There are between 20 and 46 parameters estimated for the different markets. All markets except Anticholesterol required year fixed effects, sometimes interacted with ATC4 dummies. Fixed effects were also added for either; products, molecules, or molecules×patent status (depending on where the different effects

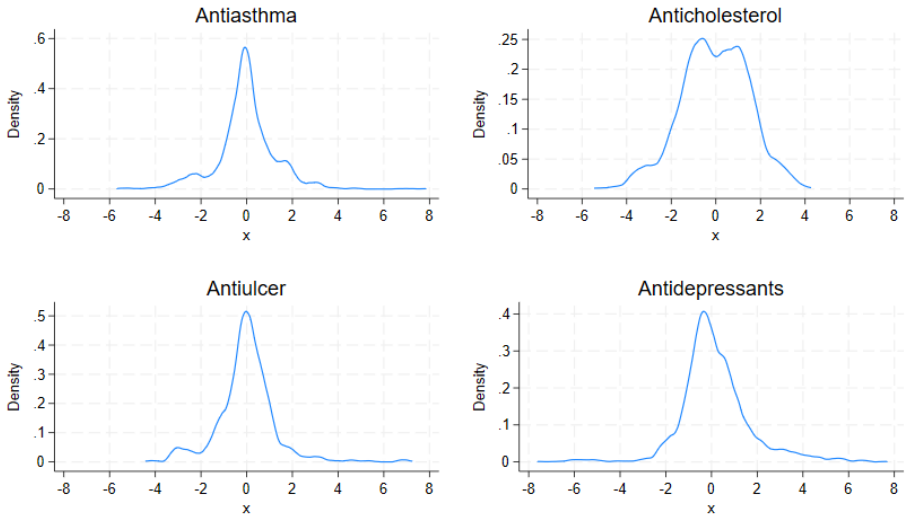
[12]In the Anatomical Therapeutic Chemical (ATC) classification system, the active substances are divided into different groups according to the organ or system on which they act and their therapeutic, pharmacological and chemical properties; see https://www.who.int/tools/atc-ddd-toolkit/atc-classification.

were statistically significant). The generic dummy interaction with price was significant for Antidepressants and Antiulcer, and the OTC dummy was significant in the Antiulcer market; they all indicated that generics and OTC drugs are more price elastic.

The other price coefficients are also significant at .1% level in all markets, as are the standard deviations of the random coefficients interacted with price. The magnitude of the price coefficients do vary across markets, but so does the range of observed prices, the severity of illnesses (primarily between but also within markets), and other characteristics of the market discussed below.

Figure 2 presents kernel density estimates of the distribution of ξ. Though different for the different markets, they all seem to be approximately normal and we will assume conditional normality of μ_{jt} (conditional on the information available at the beginning of the period) in our counterfactual calculations.

Figure 2: Density distribution of ξ_{jt}



Note: Kernel Density distributions.

The Markov process for ξ. Appendix A.5 presents a set of estimates of equation (3) for each market. There you will see that the time dummies were not needed (recall that they are in the demand system), the generic and OTC dummies were significant in some equations, and the constraint $\beta_{a_d} = \beta_{a_D} \equiv \beta_a$ was clearly accepted in all markets. So Table 4 presents results that accord with these findings. The serial correlation

coefficients (ρ) vary between .929 and .963. They are for quarterly data and imply annual coefficients between .74 and .86.

Table 4 presents two sets of results for each market. Both impose $\beta_{a_d} = \beta_{a_D} \equiv \beta_a$, but one imposes $\beta_{a_{d,D}} = 0$ and one does not. The cross-product term appears insignificant in all markets but Antiasthma and its value in that market is less then a tenth of the value of the linear effect. As a result we constrain $\beta_{a_{d,D}} = 0$ throughout the analysis that follows. This implies that detailing and DTCA are neither complements nor substitutes in the production of ξ. However since ξ enters the demand function as an exponent, current detailing and DTCA are complements in the production of profits. Note that the advertising coefficients we use, that is β_a, range between .006 and .014 and are all significant.

Table 4: Regression of ξ for all markets

	Anticholesterol		Antiasthma		Antidepressants		Antiulcer	
$\xi_{j,t-1}$	0.929***	0.917***	0.963***	0.951***	0.929***	0.922***	0.934***	0.935***
	(0.014)	(0.021)	(0.023)	(0.024)	(0.015)	(0.016)	(0.019)	(0.019)
$\log((1+a_{dj,t-1})(1+a_{Dj,t-1}))$	0.007**	0.006	0.013***	0.034***	0.014***	0.020**	0.006*	0.014
	(0.003)	(0.006)	(0.002)	(0.010)	(0.002)	(0.006)	(0.003)	(0.013)
$\log(1+a_{dj,t-1})\log(1+a_{Dj,t-1})$		0.001		-0.003**		-0.001		-0.002
		(0.001)		(0.001)		(0.001)		(0.003)
Generic		0.120*	0.153**	0.262**	0.177***	0.207***	0.073*	0.087*
		(0.055)	(0.055)	(0.083)	(0.025)	(0.040)	(0.034)	(0.044)
Constant	0.101***	0.150***	0.141***	0.251***	0.155***	0.185***	0.084**	0.098*
	(0.027)	(0.044)	(0.035)	(0.067)	(0.023)	(0.040)	(0.029)	(0.040)
N	645	645	1005	1005	948	911	930	930

Note: 2SLS instrumenting by two quarters lags the advertising variables. $a_{d,jt-1}$ is total detailing advertising and $a_{D,jt-1}$ is total DTC advertising. * for $p < .05$, ** for $p < .01$, and *** for $p < .001$.

4 Advertising Choices

We require a dynamic model for the choice of the advertising variables. When applied to markets as complex as those we study, the cognitive assumptions that underlie the "perfectness" assumption in Markov Perfect or Bayesian Perfect equilibria make the usefulness of those models questionable. So we adapt recent theoretical developments that use weaker assumptions. The theoretical models condition behavior on the information the firms have at their disposal when their decisions are made, but are silent of what that

information set is. The empirical model we now describe is designed to uncover these information sets.

The empirical model assumes agents choose policies to maximize their perceptions of the expected discounted values that emanate from their actions, but does not require those perceptions to be equilibrium perceptions, or accurate in any other sense. The goal of the empirical model is to determine the variables in our data that advertising responds to and the properties of the disturbances in the advertising equations. The disturbances represent the variables that the firm conditions its advertising decisions on and that are not in our data. The empirical model differs in two ways from the empirical models used previously to analyze dynamic market behavior that are central to how we proceed[13]; i) we do not assume that agents' perceptions of the outcomes from their actions are consistent with the outcomes in the data, and ii) we allow the disturbances in the advertising equation to be serially correlated.

The empirical model cannot be used to analyze counterfactuals as it approximates decisions that are not generated by the counterfactual environment. To analyze counterfactuals we require theory and use the Experience Based Equilibrium concept (henceforth EBE) introduced in Fershtman and Pakes (2012). This adapts prior theoretical work that weakens the perfectness assumption in environments in which the payoff-relevant random variables do not change over time to an environment where agents' actions impact their future demand and/or cost primitives. As in the empirical model, it assumes that agents make choices to maximize their perceptions of expected discounted values. It differs from the empirical model in assuming that those actions are at least consistent with what the agent observes. At states that are visited repeatedly the firm can learn the distribution of outcomes from actual competitor play, and this enables the firm to make choices that imply that their perceptions of the outcomes from their actions are accurate.

After estimating the empirical model we introduce an algorithm which computes an EBE. We first apply that algorithm to compute equilibrium policies for the in-sample environment. This allows us to compare its predictions to both the data and to the predictions of the empirical model. We then introduce our counterfactual environment, compute an EBE for it, and consider what would have happened if the counterfactual had been in place in the in-sample period.

[13]See Bajari et al. (2007); Ackerberg et al. (2007) and the review of the literature that followed in Aguirregabiria et al. (2021)

4.1 The Empirical Model

We assume advertising is determined at the beginning of the period, so the firm conditions its decision on variables that were known at the end of the last period, say J_{jt}. The ξ_{jt} obtained from the demand system is in J_{jt}, but it will be convenient to keep separate notation for it. The firm's perception of the expected discounted value of future net profit conditional on a choice of $a_{jt} = (a_{D,jt}, a_{d,jt})$ will be denoted by $W(a|\xi_{jt}, J_{jt})$, so

$$W(a_{jt}|\xi_{jt}, J_{jt}) \equiv \mathcal{E}\left[\sum_{\tau=0}^{\infty} \beta^{\tau}(\pi_j(\cdot)_{t+\tau} - a_{d,jt+\tau} - a_{D,jt+\tau})|\xi_{jt}, J_{jt}, a_{jt}\right] \tag{5}$$

where $\mathcal{E}(\cdot|\cdot)$ is the agent's expectations operator, $\pi_j(\cdot)_{t+\tau}$ is the period $t+\tau$ profit of j, and (J_{jt}, ξ_{jt}) are the set of variables the firm conditions on when making its advertising decisions.

Since the firm chooses optimal advertising given its perceptions, and Table 2 indicates that it sometimes chooses not to advertise, we assume the choice of advertising satisfies the Kuhn-Tucker first order conditions at the optimal $\{a_{h,jt}\}_{h\in\{d,D\}}$. That is either $a_{h,jt} = 0$ or it sets advertising to satisfy the first order condition derived from maximizing (5), formally:

$$a_{h,jt}\frac{\partial W(a_{h,jt}|\xi_{jt}, J)_{jt}}{\partial a_{h,jt}} = a_{h,jt}\mathcal{E}\Big[\sum_{\tau=0}^{\infty}\beta^{\tau}\frac{\partial \pi_j(\cdot)_{t+\tau}}{\partial a_{h,jt}} - 1|J_{jt}, \xi_{jt}\Big] = 0. \tag{6}$$

The firm has an expectation of the derivative of the increment in current profit with respect to $a_{h,jt}$ which does not depend on the current actions of its competitors (it is a simultaneous move game), and equals

$$\mathcal{E}\left[\frac{\partial \pi_{jt}}{\partial \xi_{jt}} \times \frac{\partial \xi_{jt}}{\partial a_{h,jt}}\Big| J_{jt}, \xi_{jt}\right] = \mathcal{E}\left[\frac{\partial \pi_{jt}}{\partial \xi_{jt}}\Big| J_{jt}, \xi_{jt}\right] \times \frac{\beta_a}{1 + a_{h,jt}}.$$

We assume that the firm knows how advertising affects demand, that is $\partial \xi_{jt}/\partial a_{h,jt} = \beta_a/(1 + a_{h,jt})$, and our empirical investigation shows that $\partial \pi_{jt}/\partial \xi_{jt}$ can be predicted extremely accurately by the analogous derivative in period $t-1$, so we use the $t-1$ derivative directly into this equation.

Using

$$\frac{\partial \pi_j(\cdot)_{t+\tau}}{\partial a_{h,jt}} = \frac{\partial \pi_{jt+\tau}}{\partial \xi_{jt+\tau}}\frac{\partial \xi_{jt+\tau}}{\partial \xi_{jt}}\frac{\partial \xi_{jt}}{\partial a_{h,jt}} = \frac{\partial \pi_{jt+\tau}}{\partial \xi_{jt+\tau}}\rho_{\xi}^{\tau}\frac{\beta_a}{1 + a_{h,jt}}$$

the Kuhn-Tucker condition (6) becomes

$$a_{h,jt}\mathcal{E}\left[\sum_{\tau=0}^{\infty}(\beta\rho_{\xi})^{\tau}\frac{\partial\pi_{jt+\tau}}{\partial\xi_{jt+\tau}}\frac{\beta_a}{1+a_{h,jt}}-1\Big|J_{jt},\xi_{jt}\right]=0. \tag{7}$$

All components of the sum in equation (6) but the first depend on the firm's perceptions of their competitors' future choices and how those choices respond to the firm's own advertising. The goal of the empirical model is to determine the variables that the firm conditions when it forms the expectation of the impact of the choices of its competitors on the profitability of its own advertising investments. Some of these variables will be in our data. They will be labeled $\{w_{h,jt}\}_{h\in d,D}$ and be determined empirically. The perceived impact of the variables that the firm conditions on but we do not observe are disturbances $\{\omega_{h,jt}\}_{h\in\{d,D\}}$ allowed to be serially correlated. With these understandings we assume that advertising depends on $\dfrac{\partial\pi_{jt}}{\partial\xi_{jt}}$, $\{w_{h,jt}\}_{h\in d,D}$, and $\{\omega_{h,jt}\}_{h\in\{d,D\}}$, as

$$\mathcal{E}\Big[\sum_{\tau=0}^{\infty}(\beta\rho_{\xi})^{\tau}\frac{\partial\pi_j(\cdot)_{t+\tau}}{\partial a_{h,jt}}\Big|J_{jt},\xi_{jt}\Big]=\tilde{\theta}_{0,h}\left(\mathcal{E}\left[\frac{\partial\pi_j(\cdot)_t}{\partial\xi_{jt}}\Big|J_{jt}\right]\right)^{\theta_{1,h}}\frac{\beta_a}{1+a_{h,jt}}\exp[w_{h,jt-1}\theta_{w,h}+\omega_{h,jt}], \tag{8}$$

for $h\in\{d,D\}$, where $(\omega_{d,t},\omega_{D,t})$ is modeled as the Markov process

$$\omega_{h,jt}\equiv\rho_h\omega_{h,jt-1}+\nu_{h,jt},\ \text{ with }\ \nu_{h,jt}|_{J_t}\sim\mathcal{N}(0,\sigma_h^2). \tag{9}$$

Notice that this implies that the impact of current DTCA and detailing on future values of ξ decays at the rates ρ_D and ρ_d, respectively.

Substituting equation (8) into (6) and solving for optimal advertising, we find that if $a_{h,t}>0$ then

$$\log(1+a_{h,jt})=\theta_{0,h}+\theta_{1,h}\log\left(\mathcal{E}\left[\frac{\partial\pi_j(\cdot)_t}{\partial\xi_t}|J_{jt}\right]\right)+w_{h,jt-1}\theta_{w,h}+\omega_{h,jt}, \tag{10}$$

where $\theta_{0,h}\equiv\log(\beta_a\tilde{\theta}_0)$, while if $a_{h,jt}=0$

$$\theta_{0,h}+\theta_{1,h}\log\left(\mathcal{E}\left[\frac{\partial\pi_j(\cdot)_t}{\partial\xi_t}|J_{jt}\right]\right)+w_{h,jt-1}\theta_{w,h}+\omega_{h,jt}\leq 0. \tag{11}$$

We let $\theta_h\equiv\{\theta_{0,h},\theta_{1,h},\theta_{w,h},\rho_h,\sigma_h\}$ denote the parameters to be estimated in these equations for $h\in\{d,D\}$.

There is a question of whether, when the firm chooses advertising, it takes account of the impact of advertising on price through advertising's impact on ξ, and ξ's impact on price. If it does we should compute the impact of ξ on profits as

$$\frac{\partial \pi(\cdot)_{jt-1}}{\partial \xi_{jt-1}} = \frac{\partial D(\cdot)_{jt-1}}{\partial \xi_{jt-1}}(p_{jt-1} - c_j) + \frac{\partial p_{jt-1}}{\partial \xi_{jt-1}} D(\cdot)_{jt-1}, \tag{12}$$

while if it does not we should set the last term to zero. Our reduced form analysis of the advertising equation discussed next indicated that the impact of ξ on price was at least marginally significant. Hence, we chose the specification in (12). Finally notice that equation (12) depends on the marginal cost of each product, our $\{c_j\}_j$, which we will have to estimate along with θ_h.

The observables that impact advertising. Before going to the estimation of the system in equations (10) and (11) we did exploratory reduced form analysis of the observable determinants of advertising[14]. This was done separately both for each market, and for detailing and DTCA within the markets. The results were consistent with both the theory and prior empirical work.

The largest and most significant coefficient in all eight cases (two types of advertising in four markets) was the estimate of the derivative of log profits with respect to advertising calculated as in equation (12). To understand what underlies this recall that if the value function were known the optimal advertising would be determined by its derivative with respect to advertising. The value function, though too difficult to compute, is the discounted sum of iterates of the profit function. So using the derivative of the profit function to approximate the derivative of the value function seems eminently sensible.

Time to loss of patent exclusivity was positive and significant in all equations; advertising expenditures tend to decline at the end of the patent life. While advertising of all competitors enters the advertising equation through its impact on the demand function, we found that advertising of competitors in the same therapeutic class had an additional independent effect in many of the advertising equations (therapeutic

[14]More precisely we did one set of linear OLS regressions with the log of advertising as the dependent variable and no serial correlation correction, and one which allowed for a first order Markov process for the disturbance and used a quasi-first differenced estimator. The regressions included our estimates of: the log of the derivative of revenue w.r.t. ξ, the derivative of price with respect to ξ times demand, the time to loss of exclusivity for patented drugs, and year dummies. They also included separate sums of competitors in the same ATC4 class values of; ξ, DTCA, and detailing expenditures. Once we obtained estimates of marginal costs, we redid the regression using log of the derivative of profit w.r.t. ξ instead of revenue.

class is defined as the ATC4 class)[15].

All equations indicated that the estimated disturbances were highly serially correlated. So the variables that the firm conditions its advertising on and we do not observe exhibit significant persistence over time.

4.2 Estimation Method of the Empirical Model

To estimate the model using equations (10) and (11), we need to account for the fact that ω is serially correlated and past ω is a determinant of current expected profits. This implies that the disturbance term in both equations is correlated with the regressors. To deal with this issue we use the Markov assumption in equation (9), and consider the implications of the model for the "quasi-first difference"

$$\log(1+a_{h,jt}) - \rho_h \log(1+a_{h,jt-1}) = \\ \theta_{0,h}(1-\rho_h) + \theta_{1,h}\left[\log(\frac{\partial\pi(\cdot,c_j)_{jt-1}}{\partial\xi_{jt-1}}) - \rho_h \log(\frac{\partial\pi(\cdot,c_j)_{jt-2}}{\partial\xi_{jt-2}})\right] + \theta_{w,h}[w_{hjt} - \rho_h w_{hjt-1}] + \nu_{h,jt}$$

where $\nu_{h,t}|_{J_t} \sim \mathcal{N}(0,\sigma_h^2)$, and we have explicitly indexed the profit function by marginal cost.

Defining

$$\underline{v}_{h,jt}(\theta,c_j) \equiv \theta_{0,h}(1-\rho_h)+\theta_{1,h}\left[\log(\frac{\partial\pi(\cdot,c)_{jt-1}}{\partial\xi_{jt-1}}) - \rho_h \log(\frac{\partial\pi(\cdot,c)_{jt-2}}{\partial\xi_{jt-2}})\right]+\beta_{w,h}[w_{jt}-\rho_h w_{jt-1}]+\rho_h \log(1+a_{h,jt-1}), \tag{13}$$

we obtain[16]

$$P(a_{h,jt}=0|a_{h,jt-1}>0) = Pr\left\{\nu_{h,jt} \leq -\underline{v}_{h,jt}(\theta,c_j)\right\}, \quad \text{and} \tag{14}$$

$$E(\log(1+a_{h,jt})|a_{h,jt}>0, a_{h,jt-1}>0) = \underline{v}_{h,jt}(\theta,c_j) + E\left[\nu_{h,jt}|\nu_{h,jt} > -\underline{v}_{h,jt}(\theta,c_j), J_{jt}\right]. \tag{15}$$

[15] We also tested whether an entry dummy matters in these advertising equations. We found close to zero and insignificant effects of these dummies, so the effect of time to loss of exclusivity seems to capture the reasons for time varying advertising expenditures.

[16] Notice that these equations condition on $a_{h,jt-1}>0$. In some periods $a_{h,jt-1}$ is not available. This occurs when a firm stops advertising and does not restart immediately and when the firm stops but does restart. For simplicity we drop all observations when $a_{h,jt-1}=0$. This does not generate a selection problem as we are conditioning on all relevant variables in estimating these equations but $\nu_{h,jt}$, and $\nu_{h,jt}$ is independent of all variables in the firm's information set when it is making its advertising decisions. We could have kept those observations and condition only on the last positive advertising but this complicates estimation unnecessarily.

where we use the independence and normality assumption of $\nu_{h,jt}$ for $E\left[\nu_{h,jt}|\nu_{h,jt} > -\underline{v}_{h,jt}(\theta, c_j), J_{jt}\right]$[17].

Marginal costs. Marginal costs are not directly observed and need to be estimated. Since the derivative of profits with respect to advertising depends on the markup the firm earns from an extra unit of demand, marginal cost will impact optimal advertising. This enables us to estimate marginal costs for each product from the dynamic first order condition for advertising, and avoids the need to specify a pricing equation. Since our data aggregates across buyers and transaction prices may vary across buyers, we would need additional assumptions to back out costs from a pricing equation derived from the revenue and quantity data we have. For comparison we also compare our cost estimates to those obtained from assuming a static Nash-Bertrand pricing game with one price for each market.

Estimation algorithm. The parameters to estimate are $\theta \equiv \{\theta_{0,h}, \theta_{1,h}, \theta_{w,h}, \rho_h, \sigma_h\}_{h=d,D}$ and marginal costs (assumed molecule-specific and fixed over time). The estimation algorithm has an inner and an outer loop:

Inner loop: At $\theta = \theta_0$, the true value of θ, the average of the first order condition for the periods with positive advertising should equal zero. So the inner loop calculates, for each trial value of θ, the values of $c_j(\theta_d)$ that solve[18]

$$\sum_t 1_{\{a_{d,jt}>0\}} \Big(\log(1 + a_{d,jt}) - \underline{v}_{d,jt}(\theta, c_j) - E[\nu_{d,jt}|a_{d,jt} > 0, J_t, \theta]\Big) = 0. \tag{16}$$

Outer Loop: This searches over θ values. We use moments using the probability that $a_{h,j,t} = 0$ given by equation (14) and the positive values of $a_{h,j,t}$ given by (15). Details are provided in Appendix A.6.

4.3 Empirical Estimates

Table 5 provides our estimates of the average marginal costs and markups of branded drugs for each of our markets. The assumption underlying these estimates are those of the empirical model. For comparison,

[17] With $\Phi(\cdot)$ the standard normal c.d.f. and $\phi(\cdot)$ its density, $E\left[\nu_{h,jt}|\nu_{h,jt} > -\underline{v}_{h,jt}(\theta, c_j), J_{jt}\right] = \sigma_h \dfrac{\phi(-\underline{v}_{h,t}(\theta, c_j)/\sigma_h)}{1 - \Phi(-\underline{v}_{h,t}(\theta, c_j)/\sigma_h)}$.

[18] Table 2 indicates that there are many more detailing than DTCA observations, so we only assume the average of the detailing observations satisfy this equation, and only use detailing data in the inner loop.

the same table provides the estimates obtained from a static Nash-Bertrand competition in prices. Generics do very little advertising so we cannot estimate generic costs and markups from the advertising equation but, since we do have generic prices, we can estimate them from a Nash pricing assumptions. The results for generics are not shown in the table but are summarized below.

The advertising equation generates markups which average 76% to 92% depending on the market. As noted our price variable is an average across different markets. Not surprisingly then the static Nash-Bertrand pricing assumption generates average markups of between 27% and 49% for the patented drugs, and between 46% to 72% for generics; results that do not accord with industry lore. Accordingly we use the cost estimates generated by the advertising equations.

Table 5: Marginal Costs Estimates from Detailing First Order Condition

		Adverting FOC		Price FOC	
	Price	Cost	Margin	Cost	Margin
Anticholesterol	2.95	0.44	0.79	1.60	0.44
Antiasthma	1.85	0.31	0.76	-0.53	0.79
Antidepressants	5.49	0.69	0.92	0.87	0.75
Antiulcer	3.01	0.28	0.81	1.59	0.38

Note: Mean prices, marginal costs and margins relative to price using cost estimates from advertising equation and those from the static Nash-Bertrand pricing conditions FOC. Margins are calculated as $(p - c)/p$. In this table we remove observations with negative costs or margins in the case of the price FOC. Cost estimates on advertising equation are based on 16 molecule level parameters for Anticholesterol, 32 for Antiasthma, 25 for Antidepressants and 14 for Antiulcer. We used a grid search for the dynamic equation between zero cost and cost equal to price.

Estimates of θ. The top and bottom panels of Table 6 present the parameter estimates for the detailing equation and the DTCA equation respectively, in each of the four markets.

We consider the detailing equation first. Both the serial correlation coefficient of the disturbance process (ρ_d), and the standard error of the innovation in that process (σ_d), are highly significant and have similar values in different markets. The quarterly serial correlation values ranged from .891 to .978, implying annual correlation rates of .63 to .91. So the variables agents use to determine their adverting expenditures that are not in our data are serially correlated. The coefficient of the derivative of log profits ($\theta_{1,d}$) is also positive and highly significant in all markets. The time to loss of exclusivity coefficient ($\beta_{2,d}$) is always positive and

is significant in three of the four markets. The coefficient of detailing of competitors in the same ATC4 group (β_d) is negative in three markets, significant in two of those, and near zero in the Antiasthma market.

Perhaps because there are far less observations on DTCA than detailing, the DTCA coefficients differ much more across markets. Both the serial correlation (ρ_D) and the standard error of the innovation in the disturbance process (σ_D) are positive and significant in all markets. The serial correlation varied a lot implying annual values from .96 in the Antidepressant to .59 in the Antiulcer market, while the standard error ranged from .12 to .52. The coefficient of the derivative of the log of profit ($\theta_{1,D}$) is always positive. It is highly significant in two markets, large but estimated imprecisely in Antidepressants market, and smaller in the Antiulcer markets. The time to loss of exclusivity coefficient ($\beta_{2,D}$) is positive in three markets, significantly so in two of them, and zero in the Antiasthma market. The coefficient of the DTCA of competitors in the same ATC4 group (β_D) is either near zero or negative, but only significantly so for Anticholesterol.

The fact that the coefficient of time to loss of exclusivity is positive attests to the importance of patent rights, and is reflective of the fact that advertising decisions are made with the future in mind. The largely negative effects of the advertising of competitors in the same ATC4 class compounds the negative effects of competitor's advertising that is built into the demand functions.

From this table only $\{\rho_h, \sigma_h\}_{h\in\{d,D\}}$ are used in the EBE calculations, and only $\{\rho_d, \sigma_d\}$ are used in the counterfactual calculations, so it is noteworthy that these parameters are estimated quite precisely.

Fit of the Empirical Model. Figures 3 and 4 illustrate the fit of both the detailing and the DTCA equations for the periods when advertising was positive. Figure 3 uses the $t-1$ information from the data to predict period t while for figure 4, we condition on the first quarter of advertising of each product and use the empirical model estimates of the advertising equation to predict the second quarter, the prediction for the second quarter to predict the third quarter and so on until either the product stops advertising or the sample ends, giving us a maximum prediction length of thirty eight quarters. Given the high serial correlation coefficients we expect Figure 3 to fit well, but there is no guarantee that it will have prediction errors which are symmetrically distributed about the 45 degree line, and it does. There is no symmetry

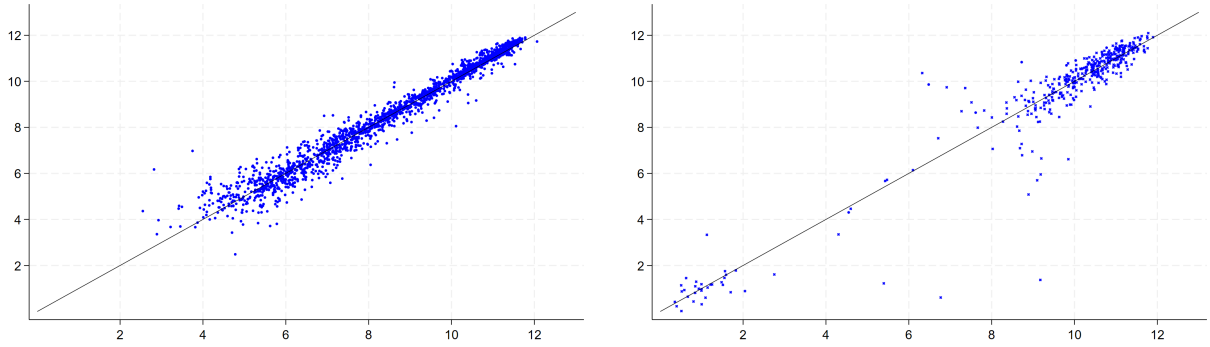
Table 6: Parameters Estimated from the Detailing and DTCA Equations

	Anticholesterol	Antiasthma	Antidepressants	Antiulcer
ρ_d	0.962**	0.978**	0.891**	0.946**
	(0.028)	(0.006)	(0.013)	(0.011)
$\theta_{1,d}$	0.419**	0.172**	0.808**	0.248**
	(0.140)	(0.035)	(0.095)	(0.058)
β_d (Other detail)	-0.079**	0.007	-0.094**	-0.357**
	(0.025)	(0.018)	(0.026)	(0.126)
$\beta_{2,d}$ (time to loe)	-0.019	0.045**	-0.023*	0.074**
	(0.085)	(0.013)	(0.012)	(0.021)
$\theta_{0,d}$	2.777	5.233**	-0.665	7.118**
	(1.881)	(0.729)	(0.904)	(1.045)
ρ_D	0.927**	0.789**	0.968**	0.592**
	(0.032)	(0.059)	(0.015)	(0.053)
$\theta_{1,D}$	1.131**	0.887**	0.790	0.241
	(0.564)	(0.276)	(0.602)	(0.267)
β_D (Other DTCA)	-0.061**	0.016	0.002	-0.055
	(0.026)	(0.033)	(0.110)	(0.050)
$\beta_{2,D}$ (time to loe)	0.245*	0.001	0.182	0.036**
	(0.129)	(0.019)	(0.338)	(0.018)
$\theta_{0,D}$	-9.366	-1.379	-7.965	6.201*
	(7.752)	(3.827)	(9.682)	(3.494)
σ_d	0.299**	0.208**	0.255**	0.343**
	(0.028)	(0.028)	(0.028)	(0.042)
σ_D	0.528**	0.146**	0.260**	0.124
	(0.108)	(0.066)	(0.065)	(0.078)

Note: Estimates where $\frac{\partial p}{\partial \xi}$ is approximated empirically in $\frac{\partial \Pi}{\partial \xi}$. Standard errors in parenthesis. * for $p < .05$, ** for $p < .01$, and *** for $p < .001$.

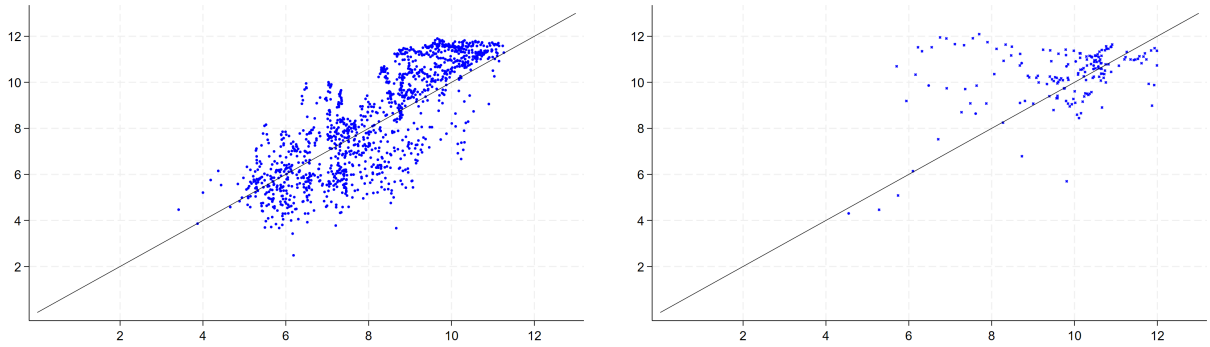
requirement for Figure 4.

Figure 3: Fit of the Empirical Model using expectation equation (15) from $t-1$ to predict t



Note: Scatter plot of observed log advertising on vertical axis against predicted log advertising on horizontal axis. Detailing on the left, DTCA on the right. All four markets together.

Figure 4: Fit of the Empirical Model using expectation equation (15) from $t=0$ to predict t



Note: Scatter plot of observed log advertising on vertical axis against predicted log advertising on horizontal axis. Detailing on the left, DTCA on the right. All four markets together.

As noted we used the determinants of advertising policies in the empirical model as state variables in the equilibrium calculations used to determine the likely implications of banning DTCA. We now turn to those equilibrium calculations.

5 Experience Based Equilibrium

The empirical model captures firms' behavior in the sample period, a period in which DTCA was allowed. We want to perform a counterfactual simulation of the equilibrium that would prevail had DTCA not been allowed. Both the raw data and our analysis of demand indicate that abandoning DTCA is likely to induce

changes in detailing. So to approximate the counterfactual behavior we will need a model for what detailing is likely to be in the absence of DTCA. To do this we turn to the Experience Based Equilibrium (EBE) framework introduced in Fershtman and Pakes (2012) and use the information sets uncovered in the empirical model as the state variables the agents condition on.

We evaluate the ban in two steps. First we compare the data to an EBE which allows for DTCA, and then compare both to an EBE which does not allow for DTCA[19].

Multiplicity. EBE accommodates asymmetric information as agents only condition their actions on variables in their own information set. The individual policies it produces generate a Markov process for the full vector of all agents' state variables, which we refer to as the market's state. Since the Markov process for the market state is finite dimensional, that process will eventually enter a recurrent subset of the possible states, and then stay within that class forever. States in the recurrent class are visited repeatedly, enabling the firm to learn the outcomes from its actions in those states. In an EBE each firm's actions are optimal conditional on the agent's information set for all states in the recurrent class.

Full information and Bayesian versions of Markov Perfect equilibrium are special cases of EBE. Except in unusual circumstances both of the earlier Markov models exhibit multiple equilibria, and the weaker requirements of an EBE will typically increase the number of possible equilibria. This because some of the points in the recurrent class can transit to points outside the recurrent class if feasible, but non optimal, policies are followed. The points outside the recurrent class need not be evaluated correctly, and different perceptions of the possible returns from those points can lead to different recurrent classes.

The algorithm we use to compute and simulate the policies of an EBE applies two restrictions designed to limit the possible set of equilibria. First, we fix the initial state of the system at the empirical data's initial state. This anchors the equilibrium to observed history and rules out equilibria that are not consistent with where the process starts. It also enables direct comparisons to both the data and the empirical model; that is our counterfactuals ask what would have happened during the in-sample period were DTCA not allowed, and then compares it to what did happen.

[19] An alternative would be to embed the state variables uncovered in the empirical model in a learning model, as is done in Fershtman and Pakes (2012). If this converged it would necessarily have to converge to an EBE equilibrium.

Second we use a refinement proposed in Asker et al. (2020). They divide the states that are in the recurrent class into two subsets. Optimal policies insure that all states in the recurrent class can only communicate with other states in the recurrent class. However some of those states could communicate to a state outside the recurrent class if a feasible, though non-optimal, policy was chosen. They call states where this can happen boundary states, and refer to states where this can not happen as interior states. That paper introduces a refinement, labeled "boundary consistency" which the firms could implement by experimentation. This insures that at boundary points the value that would be obtained by any feasible action and the perception of the value of the sequence of actions that would follow is less than the perception of the value from the optimal action[20].

Estimates used in the equilibrium calculations. We use the estimates of the demand system and of the model for the evolution of ξ obtained from estimating equations (2) and (3). We use the marginal cost estimates and the estimates of the process generating the unobservable states that impact advertising policies obtained from estimating equations (14) and (15). That is we are assuming that none of these parameters change in the counterfactual world. We initially hold prices fixed but allow them to vary in our robustness analysis.

5.1 An Algorithm to Compute an EBE

Recall that the estimates of the policy functions from the empirical model are not necessarily equilibrium policy functions. The role of the computational algorithm is to find the equilibrium policy functions. The algorithm is iterative, but differs from prior iterative algorithms used in the literature on industry dynamics in two ways designed to limit the two aspects of computing equilibrium to dynamic games that can generate a "curse of dimensionality". First, the algorithm does not compute policies at each feasible state, rather it computes policies for each product in each time period. This enables it to use a continuous state space and limits the number of policies computed to at most the number of products times the number of time periods. Second, instead of computing continuation values directly at each iteration, we simulate sample

[20]The Asker et al. (2020) paper deals with a finite set of feasible policies. When advertising expenditures are positive our set of feasible policies are continuous. As a result formally we are only boundary consistent with respect to local (in contrast to global) perturbations.

paths and then average the simulated paths to approximate continuation values. These two modifications make it computationally feasible to analyze markets as large as those we are dealing with. The major cost of proceeding in this way is that we only obtain optimal policies at the states that the model predicts would be observed in an EBE equilibrium.

Details. The iterations of the algorithm are indexed by l and condition on the policies generated by that iteration's parameter estimates. Each iteration has two distinct steps. The first simulates many sample paths of profits given the iteration-specific policy functions. It then uses the simulated continuation values to compute the derivatives of the average of the discounted simulated future returns with respect to advertising. These are substituted into the Kuhn-Tucker conditions in equation (6) which is used to estimate the advertising equations implied by the simulated sample paths. This second step produces the $(l+1)^{th}$ iteration policy functions. This second step produces the $(l+1)^{th}$ iteration policy functions. We iterate on this process until the difference in parameters from adjacent iterations satisfies our convergence criteria.

We begin by computing all random draws needed for the calculations. These are held constant over iterations, otherwise the algorithm will not converge for a sharp tolerance. For each $k = 1, .., K$ and $h \in \{d, D\}$:

- Draw[21] μ_{jt}^k from a $\mathcal{N}(0, \sigma_\mu^2)$ and $\nu_{h,jt}^k$ from a $\mathcal{N}(0, \sigma_h^2)$, the innovations in the ξ and ω_h processes, for all products j and $t = 1, .., T$.
- Compute all $\omega_{h,jt}^k$ for all j, t using: $\omega_{h,jt}^k = \rho_h \omega_{h,jt-1}^k + \nu_{h,jt}^k$.

We now describe the two steps of the algorithm, where the first step consists of simulating K sample paths and computing their implications for advertising choices and the second step consists of updating the policy parameters.

Simulating K sample paths from iteration l's policy functions.

[21] For in-sample comparisons we let μ_{jt}^k equal the estimates of μ_{jt}.

- Iterate from $t = 0$ to T using the initial conditions in the data, to compute

$$\xi_{jt}^{kl} = \rho_\xi \xi_{jt-1}^{kl} + \beta_a \log((1 + a_{D,jt}^{kl})(1 + a_{d,jt}^{kl})) + z_{jt}\beta_z + \mu_{jt}^{k} \tag{17}$$

  where the advertising variables are determined according to the iteration's policy function and the realizations of the determinants of advertising from the prior period. Then compute profit derivatives $\frac{\partial \pi_{jt}^{kl}(p_t, \xi_{jt}^{kl})}{\partial \xi_{jt}^{kl}}$ using the aggregate demand equation

$$s_{jt}^{kl}(p_t, \xi_t^{kl}) = \int \frac{\exp(\beta_p^i(X_j)p_{jt} + \beta_{m(j)} + \beta_x X_{jt} + \xi_{jt}^{kl})}{1 + \sum_{j'} \exp(\beta_p^i(X_j')p_{j't} + \beta_{m(j')} + \beta_x X_{j't} + \xi_{j't}^{kl})} dF(\beta_p^i(X))$$

  and the implied policies[22] $a_{h,jt}^{kl}$, using

$$\log(1 + a_{h,jt}^{kl}) = \max\left\{0, \theta_{0,h}^{l} + \theta_{1,h}^{l} \log \frac{\partial \pi_{jt-1}^{kl}}{\partial \xi_{jt-1}} + w_{jt}^{kl}\beta_{w,h}^{l} + \omega_{h,jt}^{k}\right\}. \tag{18}$$

Advertising choices implied by the l^{th} iteration's simulations.

- Using the simulated profits, calculate for each j, t, the solution $a_{h,jt}^{*l}$ of the Kuhn-Tucker condition $a_{h,jt}^{*l} \frac{\partial W_{jt}^{l}}{\partial a_{h,jt}^{*l}}(a_{h,jt}^{*l}) \approx 0$. That is

$$W_{jt}^{l} \equiv \pi_{jt}^{l} - a_{d,jt}^{*l} - a_{D,jt}^{*l} + \beta \mathcal{E}_k \left[\sum_{\tau=1}^{\infty} \beta^\tau \left(\pi_{jt+\tau}^{kl} - a_{d,jt+\tau}^{kl} - a_{D,jt+\tau}^{kl}\right)\right]$$

  and thus

$$\frac{\partial W_{jt}^{l}}{\partial a_{h,jt}^{*l}}(a_{h,jt}^{*l}) = -1 + \frac{\beta_a}{1 + a_{h,jt}^{*l}} \left(\frac{\partial \pi_{jt}^{l}}{\partial \xi_{jt}^{l}} + \mathcal{E}_k \left[\sum_{\tau=1}^{\infty} (\beta\rho_\xi)^\tau \frac{\partial \pi_{jt+\tau}^{kl}}{\partial \xi_{jt+\tau}^{kl}}\right]\right).$$

- The advertising expenses $a_{h,jt}^{*l}$ are zero if the expected discounted sum of marginal return $\frac{\partial W_{jt}^{l}}{\partial a_{h,jt}^{*l}}(a_{h,jt}^{*l})$

[22] The variables included in the vector w_{jt}^{kl} depend on simulation draw k because they include the lag mean advertising of other companies which depends on the simulation draw.

is less than one and otherwise equal the marginal return, that is, if positive, $a^{*l}_{h,jt}$ is calculated as

$$\begin{aligned}\log(1+a^{*l}_{h,jt}) &\equiv \log\left(\beta_a \frac{\partial \pi^l_{jt}}{\partial \xi^l_{jt}} + \beta_a \mathcal{E}_k\left[\sum_{\tau=1}^{\infty}(\beta\rho_\xi)^\tau \frac{\partial \pi^{kl}_{jt+\tau}}{\partial \xi^{kl}_{jt+\tau}}\right]\right) \\ &\approx \log\left(\beta_a \frac{\partial \pi^l_{jt}}{\partial \xi^l_{jt}} + \beta_a \frac{1}{K}\sum\nolimits_k \sum_{\tau=1}^{\infty}(\beta\rho_\xi)^\tau \frac{\partial \pi^{kl}_{jt+\tau}}{\partial \xi^{kl}_{jt+\tau}}\right).\end{aligned} \tag{19}$$

Updating the policy function.

- Substitute the estimates of a^{*l} for the actual advertising in the estimation algorithm of the empirical section, condition on the parameters taken from the empirical model, and update the parameters for a new iteration.
- This updates the θ parameters to $(\theta^{l+1}_{0,h}, \theta^{l+1}_{1,h}, \beta^{l+1}_{w,h})$ with the maximum likelihood estimates of the model

$$\log(1+a^{*l}_{h,jt}) = \max\left\{0, \theta^{l+1}_{0,h} + \theta^{l+1}_{1,h}\left(\frac{1}{K}\sum_k \log\frac{\partial \pi^{kl}_{h,jt-1}}{\partial \xi_{jt-1}}\right) + \beta^{l+1}_{w,h}\left(\frac{1}{K}\sum_k w^{kl}_{j,ht-1}\right) + \rho_h \frac{1}{K}\sum_k \omega^{kl}_{h,jt-1} + \nu_{h,jt}\right\} \tag{20}$$

  where $\nu_{h,jt} \sim \mathcal{N}(0, \sigma^2_h)$.
- If θ^l are the estimated parameters we compute $\Upsilon^l \equiv \left|\left|\frac{\theta^l - \theta^{l-1}}{\theta^{l-1}}\right|\right|$ where $\|\cdot\|$ represents the Euclidean norm, and keep iterating until $\Upsilon^l < 10^{-5}$.

Initial Conditions, discount factor β, and Perceptions of Post Sample Profits. Entry periods for each product and for the two types of advertising, as well as initial period of advertising and initial advertising, are all taken directly from the data. This enables a straightforward comparison of the EBE profits and policies to the data. Though the empirical model does not depend on the discount rate, the EBE does. Varying values of β between .9 and .98, we selected the value that makes detailing policies closest to the data[23].

We treated post sample profits in two different ways but found that they made little difference. The first simply simulated forward 300 periods. This assumes that firm's perception of the post sample profits

[23]The data being quarterly, this generates annual discount rates between between .7 and .9, in line with Gormsen and Huber (2025) that shows that such discount rates are prevalent in industry, especially for firms with market power.

follow the same process as in sample profits. In order to test the robustness of results with respect to this assumption, we also did calculations using the last observed period of advertising to proxy the firm's expectation of post-sample returns.

More precisely, if $T_{h,j}$ denotes the last date with $a_{h,T_{h,j}} > 0$, the expectation in that period is

$$1 + a_{h,jT_{h,j}} = \beta_a \mathcal{E}_k \left[\sum\nolimits_{\tau=0}^{\infty} (\beta\rho_\xi)^\tau \frac{\partial \pi^{kl}_{j,T_{h,j}+\tau}}{\partial \xi_{jt+\tau}} \right].$$

We can then replace the post $T_{h,j}$ terms in (19) using $1 + a_{h,jT_{h,j}}$ to obtain

$$\mathcal{E}_k \left[\sum_{\tau=0}^{T_{h,j}-t-1} (\beta\rho_\xi)^\tau \left[\frac{\partial \pi^{kl}_{j,t+\tau}}{\partial \xi_{jt+\tau}} \right] + (\beta\rho_\xi)^{T_{h,j}-t} (1 + a_{h,jT_{h,j}}) \right].$$

Doing so, we pick up a mean zero error which is the difference between expectations of post $T_{h,j}$ profits in period $T_{h,j}$ from that expectation in earlier years. That error does impact the advertising choices in (19) in a non linear way, so will have some impact on the choice of advertising in earlier years, but we assume this is small enough to be ignored.

5.2 Experience Based Equilibrium with DTCA

Table 7 presents the EBE parameter estimates, including a market specific estimate of the discount rate[24]. The coefficients ($\theta_{1,d}$) of the derivative of log profits in the detailing policy function are more similar across markets than those estimated in the empirical model. There are also differences in the coefficients ($\theta_{1,D}$) of the derivative of log profits in the DTCA policy equation, especially in the two markets where the estimates from the empirical model were imprecise (Antidepressants and Antiulcer).[25] However the largest differences between the parameters estimated in the EBE and those estimated in the empirical model are in the constant terms ($\theta_{0,d}, \theta_{0,D}$). As a result, we computed an equilibrium which held the constant terms equal to those from the empirical model and re-estimated the rest of the parameters. The results are summarized in Table 15 in appendix A.4, and do not differ substantively from those we focus on.

[24] Table 13 in Appendix A.4 shows the EBE parameters estimates when price is allowed to change in the EBE.

[25] Figure 11 and Figure 12 in Appendix A.4 show the scatter plots of the EBE policies against the empirical model's policies. The scatter plots of the EBE policies against the data are similar to the scatter plots of the empirical model against the data of Figures 3 and 4.

Table 7: Experience Based Equilibrium parameters

	Anticholesterol	Antiasthma	Antidepressants	Antiulcer
Quarterly discount rate (β)	.97	.914	.922	.97
$\theta_{1,d}$	.9267	.9304	1.016	.7892
$\beta_{w,d}$ (Other detail)	-.102	.1036	-.031	-.080
$\beta_{w,d2}$ (time to loe)	-.0096	.00615	-.0048	.00509
$\theta_{0,d}$	-.530	-1.80	-2.01	.7269
$\theta_{1,D}$	.8213	.5896	.0628	1.1032
$\beta_{w,D}$ (Other DTCA)	.1097	.0159	.1465	-.0217
$\beta_{w,D2}$ (time to loe)	.0665	-.0104	-.0104	-.0002
$\theta_{0,D}$	-1.7141	3.4758	8.6121	-3.4934

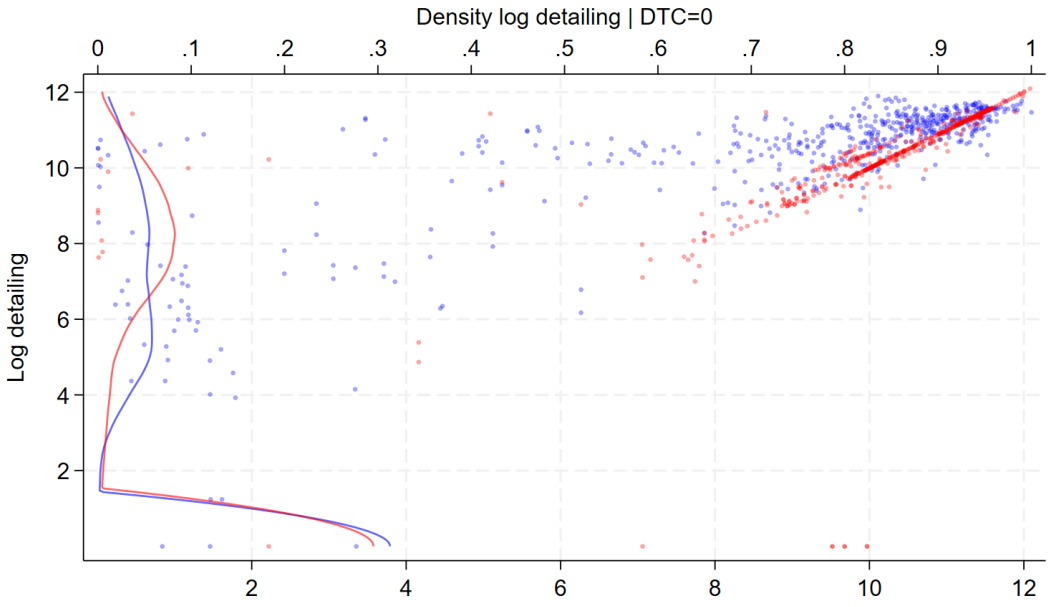
Figure 1 shows that the data indicated a strong interdependence between detailing and DTCA expenditures within each market. Figure 5 provides the analogous relationships for the predictions generated by the EBE (in red) to those in the data (in blue)[26]. The figure also shows the kernel density estimates of the EBE's prediction of log detailing when the predicted DTCA is zero[27].

As one might expect the EBE policies show a tighter relationship between detailing and DTCA than does the data, but the graphs suggest that the firms were not that far off in any of the markets. We return to a numeric comparison of the EBE profits and detailing to the data on these variables in the next section.

[26]Figure 13 in Appendix A.4 shows the EBE plot of detailing against DTCA for each of the four industries separately.

[27]Table 16 provides a detailed comparison of the probability of zero detailing between the model and the data. It shows that the EBE approximates the data quite well.

Figure 5: Detailing versus DTCA in data (blue) vs predicted by the EBE (in red)



Note: Detailing on the vertical axis versus DTCA on horizontal. The density distribution of Detailing when DTCA is zero is plotted horizontally on the left vertical axis and the associated probabilities can be read off the top of the table.

6 Counterfactual Profits and Detailing with and without DTCA

This section asks what would have happened during the data period were firms not allowed to engage in DTCA. As noted, DTCA of prescription drugs is not allowed in all developed economies but the U.S. and New Zealand, and has only been allowed on TV in the U.S. since the late 1990's. Moreover the current U.S. administration is attempting to regulate DTCA more aggressively[28].

We focus on the effects of DTCA on profits. Both the data and our empirical results suggest that banning DTCA would likely change detailing policies, and an advantage of evaluating a ban through the lens of our dynamic equilibrium model is that it allows detailing to respond to the ban. Since the effects of the ban may differ across markets, we present results for each market separately. The literature provides suggestive evidence on what the sources of these differences might be, and we will note some of them below, but we do

[28]On September 9, a presidential memorandum directed the FDA to strengthen enforcement against misleading prescription drug advertising (https://www.whitehouse.gov/presidential-actions/2025/09/memorandum-for-the-secretary-of-health-and-human-services-the-commissioner-of-food-and-drugs/). In response, the FDA and HHS announced a major enforcement campaign: thousands of warning letters and roughly 100 cease-and-desist letters targeting deceptive or unbalanced ads (https://www.hhs.gov/press-room/hhs-fda-drug-ad-transparency.html).

not have the data required to investigate their impacts in the detail they deserve.

Table 8 uses our estimates of the demand and cost systems to compare the implications of the advertising in the data to; (i) the advertising policies generated by an EBE that allows for DTCA, and to (ii) the policies from an EBE that bans DTCA. For the calculation of the EBE counterfactual without DTCA, we assume the change was instituted in the second quarter of the data, and initiate the counterfactual process at the values of the state variables in the first period[29].

Panel A of Table 8 provides the net profits, i.e. profits minus advertising expenditures, generated by the different policies. Its first three rows list: the profits generated by the policies in the data, the profits generated by an EBE that allows for DTCA and detailing and keeps prices fixed at the prices in the data, and the profits from the EBE that allows for both types of advertising but allows prices to vary with those policies. As noted the nature of our data made it difficult to compute a reliable pricing model, so we use a reduced form model for prices to examine the robustness of our results to endogenous prices[30].

The industry profits generated by the EBE policies are within 1% of the profits generated from the policies in the data for two industries, and differ by 3.1 and 5.5% in the other two. We get very similar results when we allow prices to vary in the equilibrium calculation. So one takeaway from Table 8 is that the equilibrium profits emanating from the EBE are quite close to those computed from the policies in the data for all four markets.

On the other hand the fit of the detailing equation varies more across markets. Panel B of Table 8 lists the detailing values; in the data, from an EBE that allows for DTCA, and an EBE that does not allow for DTCA. The difference between the detailing predicted by the equilibrium model that allows for DTCA and that in the data is only 1.5% for Antidepression medications and -2.5% for Antiulcer. However the EBE policies for Anticholesterol drugs under-predict the detailing data by a more noticeable 9.2%, and in Antiasthma drugs they underpredict by a full 24.2%. Possible reasons for the differences include error in

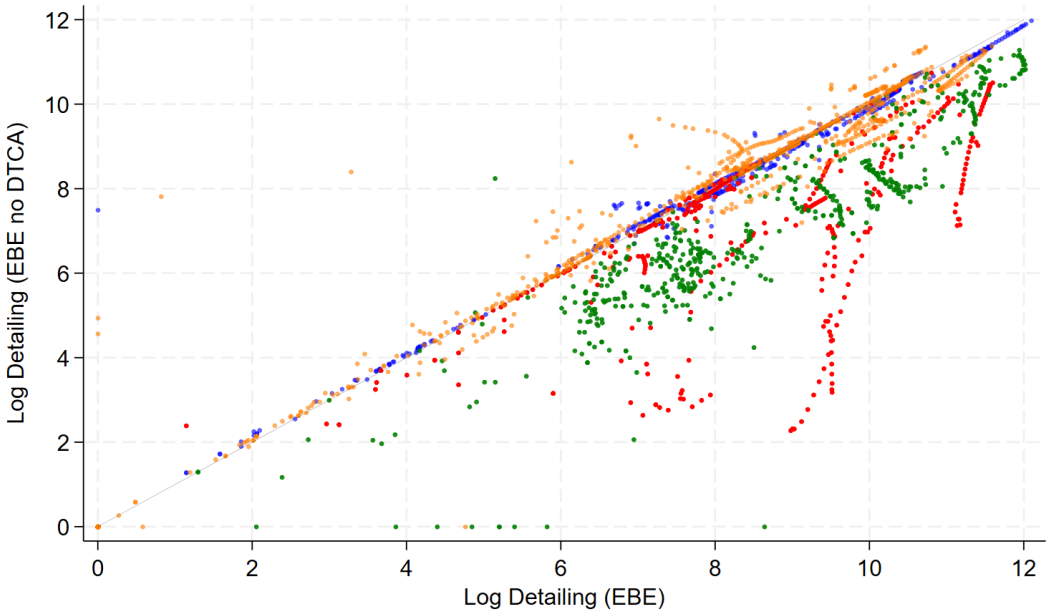
[29]The EBE calculation when DTCA is banned uses the algorithm described above imposing that DTCA is zero. The parameters this generates are similar to the analogous parameters in Table 7 and so are relegated in Table 14 in Appendix A.4.

[30]Recall that the prices in the data are averages over submarkets. We estimated a partially linear model which focused on the effect of changes in quality (or ξ) on prices controlling for the other state variables (which is similar to what is done in Berry et al. (2004)). The R^2 from that model when applied to the data was between .91 and .93 in three of the markets and .84 in the Antiasthma market.

the specification of the EBE and/or that the participants are playing policies that are inconsistent with an EBE equilibrium. The Antiasthma market had the most patented drug entry during our sample period, so one reason for non EBE policies might be that it takes time to adjust to changing market conditions[31].

We now turn to the analysis of the likely impacts of banning DTCA. Our interest in the impact of a ban on DTCA on R&D incentives leads us to our focus on a comparison of equilibrium profits with and without DTCA. Figure 6 indicates why getting a realistic picture of the ban on profits requires us to consider the implication of the ban on detailing. The figure plots equilibrium detailing after the ban on DTCA against equilibrium detailing before the ban. As Sinkinson and Starc (2019) stress, one role of DTCA is to "ask your doctor" for the advertised pharmaceutical product. A role of detailing is to insure that, when asked, the doctor is aware of that pharmaceutical and appreciates its properties. Accordingly when we ban DTCA, detailing falls in all markets. However, the extent of the decline differs across markets.

Figure 6: Equilibrium Detailing if DTCA banned versus Equilibrium Detailing when DTCA allowed



Note: Detailing if no DTCA on the vertical axis versus Detailing if DTCA on horizontal. Colors for each market (Blue for Anticholesterol, Red for Antiulcer, Green for Antidepressants, Orange for Antiasthma).

[31]The Antiasthma market saw the introduction of Leukotrienne Receptor Antagonists (LTRA's) in the late 1990's (Hoxha et al., 2017) which attained wide spread use in the early 2000's, and the first biologic for severe Asthma was approved in 2003 (Omalizumab, or Xolair, which foreshadowed the biologic era). Relatedly it is the only market where our EBE policies indicate that the detailing of different firms are strategic complements, so a decrease in one firm would cause decreases in others. This also points to another possible reason for the discrepancy; multiple equilibria.

The fall in detailing is dramatic in the Antidepressant (66.2%) and the Antiulcer (72.1%) markets. The usefulness of DTCA is likely to be higher in markets where the need for, and/or the perceptions of the usefulness of, a drug is least self evident. Both are considered a problem in patients who may be in need of Antidepressants. So it is perhaps not surprising that the ban on DTCA has a large impact on detailing in that market. A combination of medical developments and pharmaceutical products have made ulcers a disappearing disease[32], and both detailing and DTCA expenditures in Antiulcer drugs were falling in the data in the period we study. Our results indicate that banning DTCA would have accentuated the fall.

Banning DTCA causes EBE policies for detailing among Anticholesterol drugs to fall by 9.2%, and by only 7% in Antiasthma drugs. There are reasons why the ban might have less of an impact on detailing policies in these two markets. A patient typically will not know about a cholesterol problem until after a lab test suggested by a doctor, and the doctor might suggest the medication directly. Asthma is a disease whose symptoms lead to self-perception that the problem exists, and this might make DTCA less central to inducing doctor visits. Also as noted Antiasthma was the market with the most entry of patented drugs, and this might generate a need for detailing regardless of DTCA incentives.

We evaluate the impact of the ban on DTCA by comparing the EBE profit predictions when DTCA is allowed to those predictions when it is banned (in percentages this is equivalent to subtracting the percentage difference in the second row in panel A in table 8 from the fourth row in that panel). If DTCA were banned all markets would have saved what they spent on DTCA as well as a fraction of their detailing that varied across markets. Despite these savings, all markets would have lost profits. However, in the Antiulcer market the fall in DTCA and detailing induced by the ban almost compensated for the loss in sales, and the market's profits only fall by 5%. Though the Antidepressants market had a similar fall in detailing, the ban on DTCA caused a larger fall in profits (13%), consistent with the importance of DTCA in that market. The percentage profit loss in Antidepressants was similar to that in the Anticholestrol market (12.4%) where the fall in detailing expenditures was much smaller.

Antiasthma, the market with the smallest percentage fall in detailing, has the largest percentage fall in

[32] Age adjusted mortality decreased by about 40% between 2005 and 2019, see (ncbi.nlm.nih.gov) and hospitalizations fell accordingly.

profits when DTCA was banned; a fall of 30.6%. Antiasthma was the only market where the detailing of the different products were strategic complements in the EBE equilibrium (see Table 7). This will generate equilibrium detailing responses that impact the inside shares of the products (see below), and accentuate the decrease in profits.

Next we consider changes in profitability among products in the same market. The last four rows of panel A of Table 8 illustrate that we are obtaining large differences in the impact of the ban between patented and generic drugs. We estimate that the ban would lead to an increase in the profits of the generic products in all markets as consumers substitute away from patented brands that appear in ads to generics that are not advertised. So the percentage loss in profits of the patented brands is higher than the percentage loss in the market as a whole. Most of the R&D in the drug industry is done by the firms who develop patented products, so the shift in profits to generics is likely to hurt R&D incentives. Relatedly, if we order products by their profits, the lower 50% of the distribution increase their profits as a result of the ban in all markets.

Table 8: Total quarterly market level detailing, DTCA and net profit with and without DTCA

	Anticholesterol		Antiasthma		Antidepressants		Antiulcer	
All	Total	Change	Total	Change	Total	Change	Total	Change
Quarterly discount rate (β)	.97		.914		.922		.97	
A. Net Profit								
Data	3,405,022		3,387,941		1,786,078		2,685,483	
EBE	3,441,257	1%	3,495,210	3.1%	1,884,881	5.5%	2,692,134	.2%
EBE with price change	3,406,281	0%	3,471,521	2.4%	1,885,224	5.5%	2,730,894	1.6%
EBE no DTCA	3,015,383	-11.3%	2,426,027	-27.5%	1,639,701	-7.7%	2,557,469	-4.7%
EBE branded drugs	3,149,684	1.2%	3,223,753	3.5%	1,436,299	7.5%	2,473,538	.3%
EBE no DTCA branded drugs	2,699,270	-14.3%	2,138,788	-33.6%	1,174,861	-18.2%	2,327,706	-5.8%
EBE generic drugs	291,573	-.5%	271,457	-1.4%	448,582	-.4%	218,596	-.9%
EBE no DTCA generic drugs	316,113	8.4%	287,239	5.8%	464,840	3.6%	229,763	5.1%
B. Detailing								
Detailing data	179,036		217,271		160,495		86,708	
Detailing EBE	149,441	-16.5%	164,494	-24.2%	163,048	1.5%	84,489	-2.5%
Detailing EBE no DTCA	135,615	-9.2%	152,861	-7%	55,021	-66.2%	23,545	-72.1%

Note: Totals in 1,000 US$ per quarter. All % changes of EBE allowing DTCA are compared to the data values. All % changes of EBE without DTCA are compared to the EBE allowing DTCA values.

Table 9 provides the average, over the sample period, of EBE predictions of the market share of branded

and generics before and after the ban, and implicitly the market share of the outside alternative. Patented drugs lose between 2.3% and 3.2% of market share over the period, but generics gain market share, so on net the fall in the inside share is smaller, ranging from 0.4 percentage points in Antiulcer medications to 2.7 percentage points in Antiasthma medications. We note that we cannot tell whether those who stop taking the drug are in need of the drug, and that these numbers are impacted by the MEPS data uncertainty estimates as samples for particular diseases are not always large (see Appendix A.3).

Table 9: EBE Market Shares of Branded and Generic Drugs with and without DTCA

	Anticholesterol	Antiasthma	Antidepressants	Antiulcer
Share generics	.231	.122	.343	.338
Share generics (no DTCA)	.240	.127	.354	.345
Share branded	.195	.189	.105	.270
Share branded (no DTCA)	.172	.157	.076	.259

Note: The sum of market share of generics and branded drugs equal one minus the market share of the outside good. The Table shows the mean over our sample data period of these shares.

7 Conclusion

Our empirical analysis focuses on the implications of banning direct-to-consumer advertising (DTCA) on pharmaceutical profits, and hence on their incentives to engage in R&D. This, as well as the impact of regulating DTCA on health outcomes, should inform policies designed to regulate DTCA.

The data and our demand results indicate that any realistic assessment of a DTCA ban must account for its impact on detailing expenditures. This requirement underlies our use of a dynamic equilibrium model of advertising investments. The cognitive demands of Markov (or Bayesian Markov) perfect equilibrium models call into question their realism for modeling markets as complex as those we study. So we adopted an Experience-Based Equilibrium (EBE) framework. While EBE is designed to mitigate the cognitive and computational demands of Markov perfection, it had not previously been adapted for empirical applications. Our methodological contribution lies in developing a framework to operationalize EBE for empirical analysis.

The methodology consists of two components. First, we estimate an empirical model of advertising investments using our data. This model assumes that firms' actions were taken to maximize their perceived

expected discounted value of future net cash flows, but does not require that these perceptions be either correct or consistent with equilibrium conditions. The empirical model serves to identify: (1) the observable variables that influence advertising expenditures, and (2) the properties of the serially correlated disturbances that capture unobserved determinants of advertising. The equilibrium model then uses these outputs to define the state variables that condition advertising decisions. It assumes that firms maximize the expected discounted value of future net cash flows conditional on these state variables, and that expectations are correct at frequently observed states. We develop a novel algorithm that computes equilibrium policies for these states relatively easily.

Our empirical results first compare the advertising policies generated by an EBE that incorporates both detailing and DTCA to those observed in the data. The EBE successfully replicates the major features of observed policies across all four markets, though it predicts a tighter relationship between detailing and DTCA than what appears in the data. When examining market averages, we find that the difference between average quarterly profits predicted by the EBE and those observed in the data is minimal across all four markets. However, the EBE's predictions for average detailing expenditures show noticeable discrepancies from the data in two of the four markets.

When we ban DTCA both detailing expenditures and profits decline in all markets. The magnitude of the declines in both variables varies significantly across markets. The most dramatic detailing reductions occur in the Antiulcer and Antidepressant medication markets (approximately 70%). These are the two markets where EBE predictions for detailing expenditures when DTCA is permitted closely match the observed data on detailing. The Antiulcer market presents a unique case during our sample period, as peptic ulcer disease was becoming less prevalent and this market exhibited the lowest levels of both detailing and DTCA expenditures (both per product and in aggregate). As a result in Antiulcers the cost savings from reduced detailing following a DTCA ban nearly offset the loss in sales, resulting in a profit decline of less than 5%.

While the antidepressant market experiences a similar reduction in detailing, the profit impact differed. Depression represents a condition where low self-awareness is considered problematic, and we estimate that the cost savings from reduced advertising in the Antidepression market fails to compensate for the decline in markups. Consequently profits fall by 13%, about the same as the fall in the Anticholesterol market which

had a much smaller percentage fall in detailing (9.2%).

The largest profit decline occurs in the Antiasthma market (30.6%), the market which experienced the smallest reduction in detailing (7%). This result could be attributable to several factors: the introduction of new drugs during our sample period, the finding that this is the only market where detailing expenditures of different products are estimated to be strategic complements, and the observation that actual detailing expenditures do not seem to have converged to our EBE equilibrium predictions.

We conclude that banning DTCA would cause a fall in profits in all markets, but the extent of the fall differs in a non-trivial way across markets. This does complicate policy design, as would consideration of the health effects of DTCA. The results also highlight that an understanding of the magnitude of the impact of regulating DTCA on profitability, and likely also on health, requires an understanding of how the regulation would affect detailing.

Bibliography

Abi-Rafeh, R., P. Dubois, G. Rachel, and M. O'Connell (2026). The effects of sin taxes and advertising restrictions in a dynamic equilibrium. *American Economic Journal: Microeconomics forthcoming*.

Acemoglu, D. and J. Linn (2004). Market Size in Innovation: Theory and Evidence from the Pharmaceutical Industry. *Quarterly Journal of Economics 119*, 1049–1090.

Ackerberg, D., L. Benkard, S. Berry, and A. Pakes (2007). *Econometric Tools for Analyzing Market Outcomes*, Volume 6A, pp. 4171–4276. Amsterdam: North-Holland.

Aguirregabiria, V., A. Collard-Wexler, and S. P. Ryan (2021). Chapter 4 - dynamic games in empirical industrial organization. In K. Ho, A. Hortacsu, and A. Lizzeri (Eds.), *Handbook of Industrial Organization, Volume 4*, Volume 4 of *Handbook of Industrial Organization*, pp. 225–343. Elsevier.

Aguirregabiria, V. and A. Nevo (2013). *Recent Developments in Empirical IO: Dynamic Demand and Dynamic Games*, pp. 53–122. Econometric Society Monographs. Cambridge University Press.

Anderson, S. P., F. Ciliberto, and J. Liaukonyte (2013). Information content of advertising: Empirical evidence from the OTC analgesic industry. *International Journal of Industrial Organization 31*(5), 355–367.

Anderson, S. P., F. Ciliberto, and J. Liaukonyte (2016). Push-me pull-you : comparative advertising in the OTC analgesics industry. *The RAND Journal of Economics 47*(4), 1029–1056.

Arcidiacono, P., P. B. Ellickson, P. Landry, and D. B. Ridley (2013). Pharmaceutical followers. *International Journal of Industrial Organization 31*(5), 538–553.

Asker, J., C. Fershtman, J. Jeon, and A. Pakes (2020). A computational framework for analyzing dynamic auctions: The market impact of information sharing. *The RAND Journal of Economics 51*(3), 805–839.

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

Benkard, C. L., P. Jeziorski, and G. Y. Weintraub (2015). Oblivious equilibrium for concentrated industries. *The RAND Journal of Economics 46*(4), 671–708.

Benkard, C. L., G. Y. Weintraub, and B. Van Roy (2006). Oblivious equilibrium: A mean field approximation for large scale dynamic games. In *Advances in Neural Information Processing Systems*, Volume 18. MIT Press.

Berry, S., J. Levinsohn, and A. Pakes (2004). Differentiated products demand systems from a combination of micro and macro data: The new car market. *Journal of Political Economy 112*(1), 68–105.

Berry, S. T., J. Levinsohn, and A. Pakes (1995). Automobile Prices in Market Equilibrium. *Econometrica 63*(4), 841–890.

Bhattacharya, J. and W. B. Vogt (2003). A simple model of pharmaceutical price dynamics. *Journal of Law and Economics 46*(2), 599–626.

Blume-Kohout, M. E. and N. Sood (2013). Market Size and Innovation: Effects of Medicare Part D on Pharmaceutical Research and Development. *Journal of Public Economics 97*, 327–336.

Buxbaum, J. D., M. E. Chernew, A. M. Fendrick, and D. M. Cutler (2020). Contributions of public health, pharmaceuticals, and other medical care to us life expectancy changes, 1990–2015. *Health Affairs (Millwood) 39*(9), 1546–1556.

Chernew, M., D. M. Cutler, K. Ghosh, and M. B. Landrum (2016). Understanding the improvement in disability-free life expectancy in the us elderly population. In D. A. Wise (Ed.), *Insights in the Economics of Aging*, pp. 161–201. National Bureau of Economic Research, Inc.

Ching, A. T., R. Clark, I. Horstmann, and H. Lim (2016). The effects of publicity on demand: The case of anti-cholesterol drugs. *Marketing Science 35*(1), 158–181.

Ching, A. T. and M. Ishihara (2010). The effects of detailing on prescribing decisions under quality uncertainty. *Quantitative Marketing and Economics 8*(2), 123–165.

Dave, D. M. (2013). Effects of Pharmaceutical Promotion: A Review and Assessment. *NBER Working Paper 18830*.

David, G., S. Markowitz, and S. Richards-Shubik (2010). The Effects of Pharmaceutical Marketing and Promotion on Adverse Drug Events and Regulation. *American Economic Journal: Economic Policy 2*(4), 1–25.

de Frutos, M. A., C. Ornaghi, and G. Siotis (2013). Competition in the pharmaceutical industry: How do quality differences shape advertising strategies? *Journal of Health Economics 32*(1), 268–285.

Donohue, J. M., M. Cevasco, and M. B. Rosenthal (2007). A decade of direct-to-consumer advertising of prescription drugs. *New England Journal of Medicine 357*(7), 673–681.

Doraszelski, U. and K. L. Judd (2012). Avoiding the curse of dimensionality in dynamic stochastic games. *Quantitative Economics 3*(1), 53–93.

Doraszelski, U. and S. Markovich (2007). Advertising dynamics and competitive advantage. *The RAND Journal of Economics 38*(3), 557–592.

Dubé, J., G. Hitsch, and P. Manchanda (2005). An empirical model of advertising dynamics. *Quantitative Marketing and Economics 3*, 107–144.

Dubois, P. (2025). Pharmaceutical regulation and incentives for innovation in an international perspective. In *Econometric Society Monographs: 2025 World Congress*, Volume 3. Cambridge University Press. Forthcoming.

Dubois, P., O. de Mouzon, F. Scott Morton, and P. Seabright (2015). Market size and pharmaceutical innovation. *The RAND Journal of Economics 46*(4), 844–871.

Ericson, R. and A. Pakes (1995). Markov-Perfect Industry Dynamics: A Framework for Empirical Work. *Review of Economic Studies 62*(1), 53–82.

Esponda, I. and D. Pouzo (2016). Berk–nash equilibrium: A framework for modeling agents with misspecified models. *Econometrica 84*(3), 1093–1130.

Fershtman, C. and A. Pakes (2012). Dynamic Games with Asymmetric Informa- tion: A Framework for Empirical Work. *Quarterly Journal of Economics 127*, 1611–1661.

Filson, D. (2012). A Markov-perfect equilibrium model of the impacts of price controls on the performance of the pharmaceutical industry. *RAND Journal of Economics 43*(1), 110–138.

Fudenberg, D. and D. K. Levine (1993). Self-confirming equilibrium. *Econometrica 61*(3), 523–545.

Gormsen, N. J. and K. Huber (2025). Corporate discount rates. *American Economic Review 115*(6), 2001–2049.

Grennan, M., K. R. Myers, A. Swanson, and A. Chatterji (2024). No free lunch? welfare analysis of firms selling through expert intermediaries. *The Review of Economic Studies*, rdae090.

Ho, K. and A. Pakes (2025). Policy options for the drug pricing conundrum. *Proceedings of the National Academy of Sciences Vol 122 # 9*.

Hoxha, M., G. E. Rovati, and A. B. Cavanillas (2017). The leukotriene receptor antagonist montelukast and its possible role in the cardiovascular field. *European Journal of Clinical Pharmacology 73*(7), 799–809. Epub 2017 Apr 4. Review.

Iizuka, T. (2004). What Explains the Use of Direct-to-Consumer Advertising of Prescription Drugs? *Journal of Industrial Economics 52*(3), 349–379.

Iizuka, T. and G. Z. Jin (2005). The Effect of Prescription Drug Advertising on Doctor Visits. *Journal of Economics & Management Strategy 14*(3), 701–727.

Iizuka, T. and G. Z. Jin (2007). Direct to consumer advertising and prescription choice. *The Journal of Industrial Economics 55*(4), 771–771.

Jehiel, P. (2005). Analogy-based expectation equilibrium. *Journal of Economic Theory 123*(2), 81–104.

Jehiel, P. and D. Samet (2007). Valuation equilibrium. *Theoretical Economics 2*(2), 163–185.

Lakdawalla, D. and T. Philipson (2011). Does intellectual property restrict output? an analysis of pharmaceutical markets. *Journal of Law and Economics 55*(1), 151–187.

Lakdawalla, D., N. Sood, and Q. Gu (2013). Pharmaceutical advertising and Medicare Part D. *Journal of Health Economics 32*(6), 1356–1367.

Lakdawalla, D. N. (2018). Economics of the Pharmaceutical Industry. *Journal of Economic Literature 56*(2), 397–449.

Ling, D. C., E. R. Berndt, and M. K. Kyle (2002). Deregulating direct-to-consumer marketing of prescription drugs: Effects on prescription and over-the-counter product sales. *The Journal of Law and Economics 45*(S2), 691–723.

Liu, H., Q. Liu, and P. Chintagunta (2017). Promotion spillovers: Drug detailing in combination therapy. *Marketing Science 36*.

Maskin, E. and J. Tirole (1988a). A theory of dynamic oligopoly, I: Overview and quantity competition with large fixed costs. *Econometrica 56*, 549–569.

Maskin, E. and J. Tirole (1988b). A theory of dynamic oligopoly, ii: Price competition, kinked demand curves, and edgeworth cycles. *Econometrica 56*(3), 571–599.

Osborne, M. J. and A. Rubinstein (1998). Games with procedurally rational players. *American Economic Review 88*(4), 834–847.

Pakes, A. and McGuire (1994). Computing markov-perfect nash equilibria: Numerical implications of a dynamic differentiated product model. *The RAND Journal of Economics 25*, 555–589.

Pakes, A. and P. McGuire (2001). Stochastic Algorithms, Symmetric Markov Perfect Equilibria, and the Curse of Dimensionality. *Econometrica 69*(5), 1261–1281.

Ridley, D. B. (2015). Payments, promotion, and the purple pill. *Health Economics 24*(1), 86–103.

Rosenthal, M. B., E. R. Berndt, J. M. Donohue, R. G. Frank, and A. M. Epstein (2002). Promotion of prescription drugs to consumers. *New England Journal of Medicine 346*(7), 498–505.

Rust, J. (1996). Numerical dynamic programming in economics. *Handbook of Computational Economics 1*, 619–729.

Rust, J. (1997). Using randomization to break the curse of dimensionality. *Econometrica 65*, 487–516.

Ryan, S. (2012). The costs of environmental regulation in a concentrated industry. *Econometrica 80*, 1019–1062.

Shapiro, B. (2018). Positive Spillovers and Free Riding in Advertising of Prescription Pharmaceuticals: The Case of Antidepressants. *Journal of Political Economy 126*(1), 381–437.

Shapiro, B. T. (2022). Promoting wellness or waste? evidence from antidepressant advertising. *American Economic Journal: Microeconomics 14*(2), 439–77.

Sinkinson, M. and A. Starc (2019). Ask your doctor? direct-to-consumer advertising of pharmaceuticals. *Review of Economic Studies 86*(2), 836–881.

Sweeting, A. (2013). Dynamic product positioning in differentiated product markets: The effect of fees for musical performance rights on the commercial radio industry. *Econometrica 81*(5), 1763–1803.

Villas-Boas, M. (1993). Predicting advertising pulsing policies in an oligopoly: A model and empirical test. *Marketing Science 12*, 88–102.

Wosinska, M. (2005). Direct-to-consumer advertising and drug therapy compliance. *Journal of Marketing Research 42*, 323–332.

A Appendix

A.1 Disease Indication Markets

The Anticholesterol Market: Several medications are used to lower blood cholesterol levels. Statins are recommended for most patients. Statins are the only cholesterol-lowering drugs that have been directly associated with a reduction in the risk of heart attack or stroke. They work in the liver to prevent cholesterol from forming. The other drugs in this market include PCSK9 inhibitors which bind to and inactivate a protein in liver in order to lower LDL (bad) cholesterol, selective cholesterol absorption inhibitors which work by preventing cholesterol from being absorbed in the intestine, resins which work in the intestines by promoting increased disposal of cholesterol, lipid-lowering therapies which are not very effective in lowering LDL cholesterol, Niacin which works in the liver by affecting the production of blood fats, and Omega-3 Fatty Acid Ethyl Esters are used in tandem with dietary changes, to help people with very high triglyceride levels (over 500 mg/dL) lower their levels.

The Antiulcer Market: Ulcers are sores on the lining of the stomach or small intestine. They also could be on the esophagus (throat). Most ulcers are located in the small intestine. These ulcers are called duodenal ulcers. Stomach ulcers are called gastric ulcers. The antiulcer drugs used are mostly proton pump inhibitors or selective histamine Type 2 receptor Antagonists/Blockers.

The Antiasthma Market: Asthma is a condition in which a person's airways become inflamed, narrow and swell, and produce extra mucus, which makes it difficult to breathe. Three types of drugs are used in treating and preventing asthma attacks: i) Bronchodilators relax the smooth muscles that line the airway. They may be taken by mouth, injected, or inhaled. ii) Corticosteroids block the inflammation that narrows the airways. Used regularly, these drugs help prevent asthma attacks, however, corticosteroids cannot stop an attack that is already underway. They may be taken by mouth, injected, or inhaled. iii) Cromolyn also is taken regularly to prevent asthma attacks but cannot stop an attack that already has started. The drug works by preventing certain cells in the body from releasing substances that cause asthma symptoms.

The Antidepressant Market: There are five main types of antidepressants. Serotonin and noradrenaline reuptake inhibitors (SNRIs; like Cymbalta, and Effexor) raise levels of serotonin and norepinephrine, two neurotransmitters in the brain that play a key role in stabilizing mood. Selective serotonin reuptake inhibitors (SSRIs) block the reuptake, or absorption, of serotonin in the brain and are the most widely used antidepressant as they are quite effective and have fewer side effects than the other antidepressants. Tricyclic antidepressants (TCAs) are used to treat depression, fibromyalgia, some types of anxiety, and they can help

control chronic pain. Monoamine oxidase inhibitors (MAOIs) were commonly prescribed before the introduction of SSRIs and SNRIs, and inhibits the action of monoamine oxidase, a brain enzyme. Monoamine oxidase helps break down neurotransmitters, such as serotonin. Noradrenaline and specific serotoninergic antidepressants (NASSAs).

A.2 Advertising Data

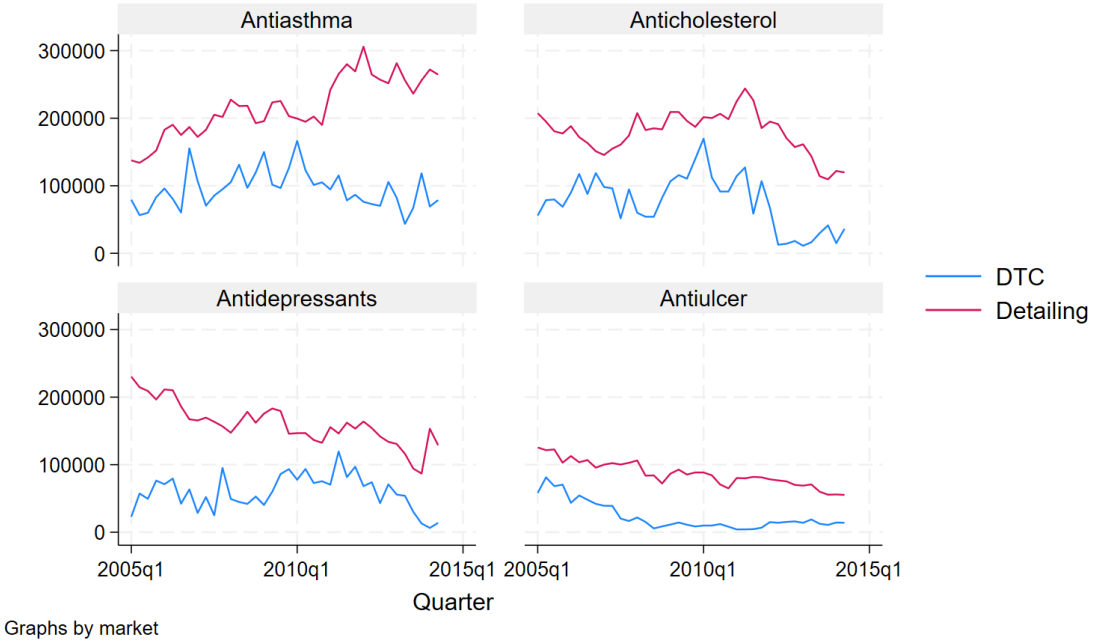
This appendix provides more detail on different aspects of the data we use. Table 10 provides total advertising expenditures for detailing and DTCA in the U.S. by year for our four markets and for the pharmaceutical industry as a whole. The figure 7 graphs the time series on these advertising variables for the four markets.

Table 10: Advertising Expenses on Antiulcer, Anticholesterol, Antiasthma and overall in the US

	DTC Advertising			Detailing		
Year	(Four markets)	(All markets)	Ratio	(Four markets)	(All markets)	Ratio
2005	1,196	4,251	0.281	2,764	11,297	0.245
2006	1,463	4,925	0.297	2,748	11,587	0.237
2007	1,211	4,944	0.245	2,643	11,370	0.232
2008	1,170	4,426	0.264	2,839	12,255	0.232
2009	1,498	4,383	0.342	2,935	12,103	0.243
2010	1,533	4,061	0.378	2,697	11,340	0.238
2011	1,467	3,952	0.371	3,160	13,001	0.243
2012	1,196	3,374	0.354	3,007	12,755	0.236
2013	1,167	3,726	0.313	2,549	12,713	0.200
2014	654	2,730	0.239	1,975	9,476	0.208

Note: Promotion expenditures in millions of US$ per year (3 quarters only for 2014). First column is DTCA spending on the four markets considered and second column is overall expenses. The third is the ratio. Fourth column is Detailing promotion spending on the four markets considered and fifth column is overall expenses. The last column is the ratio.

Figure 7: Total Advertising Expenses by Market



Note: Total by market in 1000US$ per quarter.

A.3 Medical Expenditure Panel Survey (MEPS)

MEPS is a nationally representative survey of the US civilian non-institutionalized population, with a construction of a new panel of sample households each year. The counts are nationally representative, as they are being calculated using the MEPS survey weights. These survey weights account for the complexity of the sampling design, thus potential selection, non-response and post-stratification (for details on the sampling design I refer to the next paragraph). Appropriate standard errors are then calculated using the Taylor-series linearization method. Data entails information on health services the American population uses, the frequency and cost of used services as well as payment schemes and insurance data. We use both the publicly available Household Component as well as the prescribed medicine files. The prescribed medicines files contain information from survey participants about their prescribed drugs, as well as (if subjects give permission) more detailed information from pharmacies about the type, dosage and payment for each prescription. This enables us to compute the patient populations in the US that are "diagnosed with each of the four illnesses Asthma, High Cholesterol, Ulcer and Depression but without any prescription for the illness".

Table 11 contains the list of drugs identified as potentially indicated for each of the four illnesses and

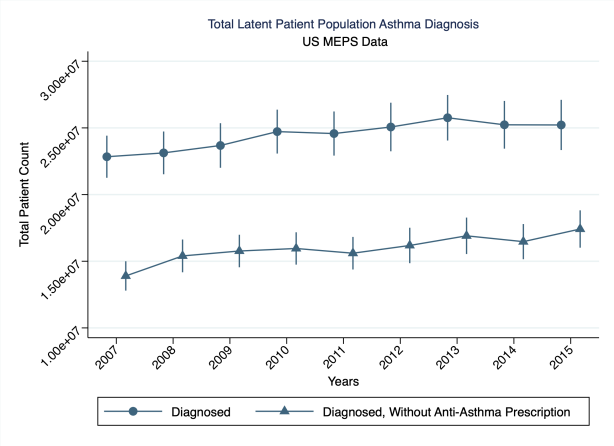
whose quantities prescribed are observed in the IMS(IQVIA) data.

Table 11: List of Medications for specific illnesses

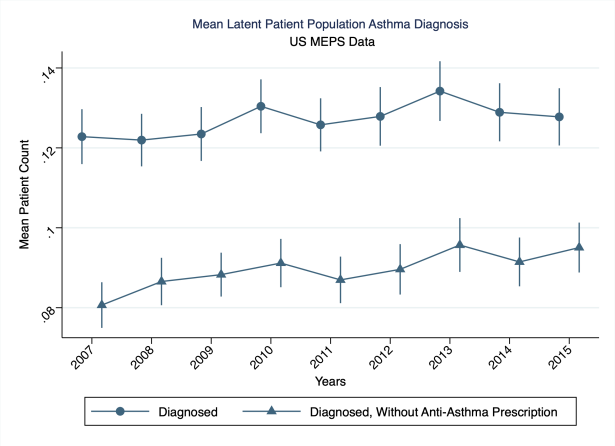
Anti Asthma	Accolate, Accuneb, Advair, Aerobid, Albuterol, Albutsulfipratrop, Alupent, Alvesco, Anoroellipta, Asmanex, Atrovent, Azmacort, Breoellipta, Brethine, Bronchial, Brovana, Budesonide, Combivent, Daliresp, Dulera, Duoneb, Ephedrine, Flovent, Foradil, Intal, Ipratropium, Levalbuterol, Maxair, Montelukast, Perforomist, Primatene, Proair, Proventil, Pulmicort, Qvar, Serevent, Singulair, Spiriva, Symbicort, Terbutaline, Theophylline, Tudorza, Uniphyl, Ventolin, Vospire, Xopenex, Zafirlukast, Zyflo
Anti hypercholesterolemia	Altoprev, Antara, Atorvastatin, Cholestyramine, Colestid, Colestipol, Crestor, Fenofibrate, Fenofibric, Fluvastatin, Gemfibrozil, Lescol, Lipitor, Livalo, Lofibra, Lopid, Lovastatin, Mevacor, Niacin, Niaspan, Pravachol, Pravastatin, Pravigard, Prevalite, Simvastatin, Tricor, Triglide, Trilipix, Welchol, Zetia, Zocor
Anti Depression	Amitriptyline, Aplenzin, Brintellix, Budeprion, Bupropion, Celexa, Citalopram, Clomipramine, Cymbalta, Deplin, Desipramine, Desyrel, Doxepin, Duloxetine, Effexor, Emsam, Escitalopram, Fetzima, Fluoxetine, Fluvoxamine, Imipramine, Lexapro, Luvox, Mirtazapine, Nefazodone, Nortriptyline, Pamelor, Paroxetine, Paxil, Pexeva, Pristiq, Prozac, Remeron, Sarafem, Sertraline, Serzone, Tofranil, Tranylcypro, Trazodone, Venlafaxine, Viibryd, Wellbutrin, Zoloft
Anti (Peptic) Ulcer	Aciphex, Axid, Carafate, Cimetidine, Cytotec, Dexilant, Famotidine, Helida, Kapidex, Lansopamoxclarit, Lansoprazole, Misoprostol, Nexium, Nizatidine, Omeprazole, Pantoprazole, Pepcid, Prevacid, Prevpac, Prilosec, Protonix, Pylera, Rabeprazole, Ranitidine, Sucralfate, Zantac, Zantacotc, Zegerid

The following graphs and tables represent different latent patient populations in the US during the years 1996 - 2020 for each of the four illnesses considered Asthma, High Cholesterol, (Peptic) Ulcer and Depression. Yearly counts of the latent patient population are defined as (for example) "Being diagnosed with asthma but without any prescription for antiasthma medication".

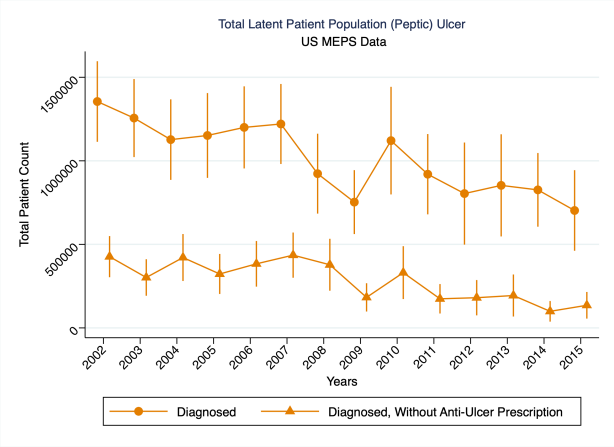
Figure 8: Latent Patient Population Counts By Illness Over Time, US MEPS Data - I



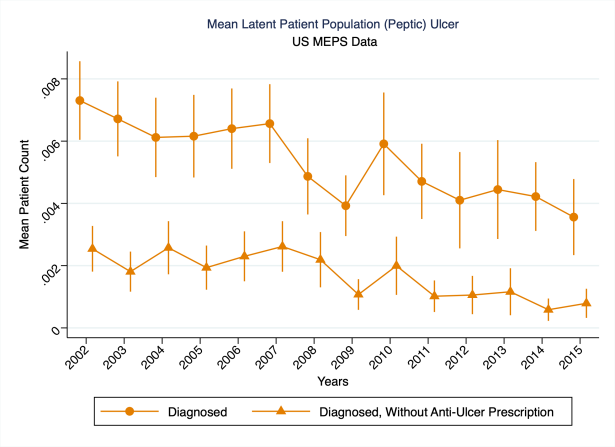
(a) Asthma Diagnosis & No Antiasthma Prescription, Total Counts



(b) Asthma Diagnosis & No Antiasthma Prescription, Mean Counts



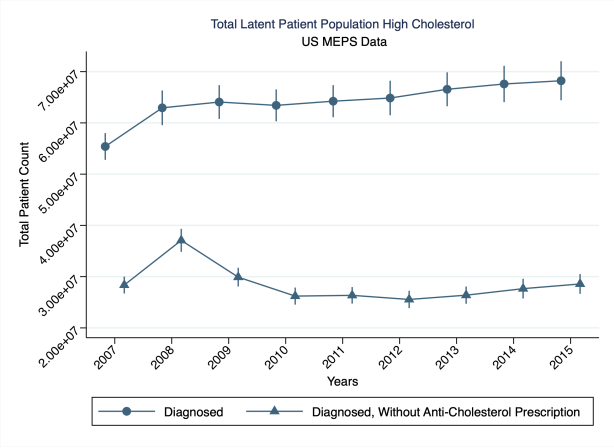
(c) Ulcer Diagnosis & No Antiulcer Prescription, Total Counts



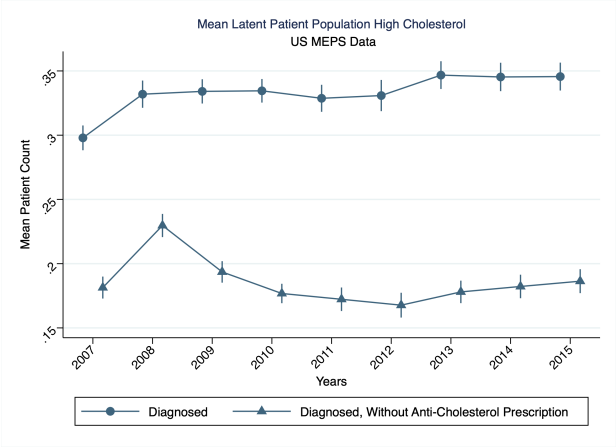
(d) Ulcer Diagnosis & No Antiulcer Prescription, Mean Counts

Notes: Vertical bars indicate confidence intervals.

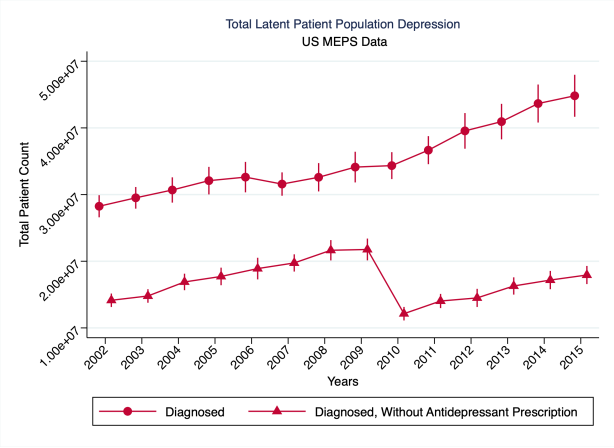
Figure 9: Latent Patient Population Counts By Illness Over Time, US MEPS Data - II



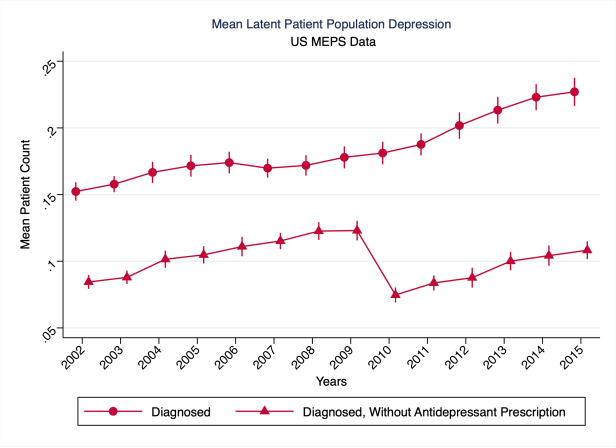
(a) Cholesterol Diagnosis & No Anticholesterol Prescription, Total Counts



(b) Cholesterol Diagnosis & No Anticholesterol Prescription, Mean Counts



(c) Depression Diagnosis & No Antidepressant Prescription, Total Counts



(d) Depression Diagnosis & No Antidepressant Prescription, Mean Counts

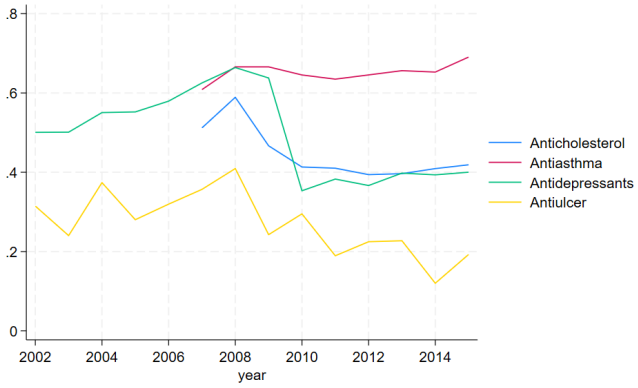
Notes: Vertical bars indicate confidence intervals.

The following figures and tables display the potential patient population in terms of diagnosis, thus nationally representative counts of those diagnosed with asthma, (peptic) ulcer, high cholesterol or depression. Diagnosis for a specific illness was determined using ICD-9 and ICD-10 codes for (peptic) ulcer and depression (see Table 12) and for asthma and high cholesterol using the prespecified definitions of the MEPS data providers.

Table 12: List of ICD codes for specific illnesses

Disease	ICD-9/10 Codes
Depression	296, 300, 309, F43, 298, 311, F06, F31, F32, F33, F34
(Peptic) Ulcer	531, 532, 533, 534, K25, K26, K27, K28, 139

Figure 10: Estimated outside good market share s_{0t} using the MEPS data



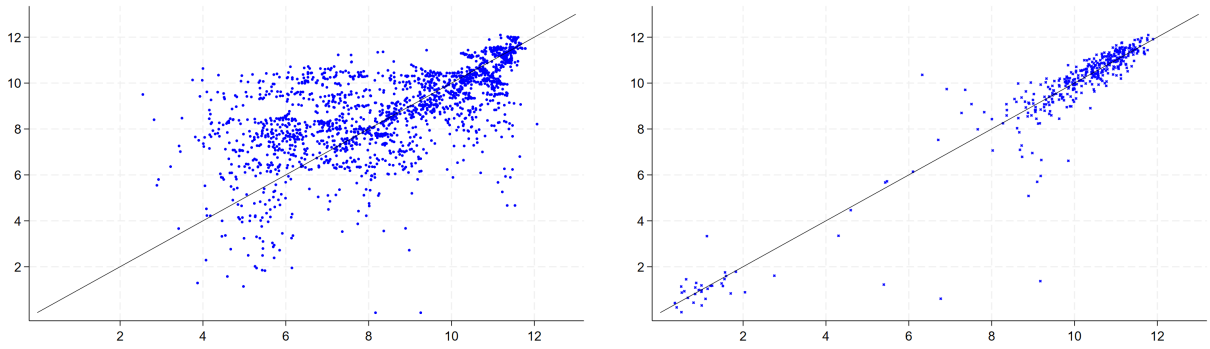
A.4 Details on EBE results

Table 13: Experience Based Equilibrium parameters

	Anticholesterol	Antiasthma	Antidepressants	Antiulcer
Quarterly discount rate (β)	.97	.914	.922	.97
$\theta_{1,d}$	.9312	.9110	1.030	.7788
$\beta_{w,d}$ (Other detail)	-.095	.0892	-.026	-.083
$\beta_{w,d2}$ (time to loe)	-.0104	.00660	-.003	.00760
$\theta_{0,d}$	-.634	-1.49	-2.23	.1041
$\theta_{1,D}$	.8384	.4780	-.0817	.7840
$\beta_{w,D}$ (Other DTC)	.0953	.0033	.2380	-.0916
$\beta_{w,D2}$ (time to loe)	.0723	-.0149	-.0149	.0195
$\theta_{0,D}$	-2.0145	4.9375	9.6518	.1386

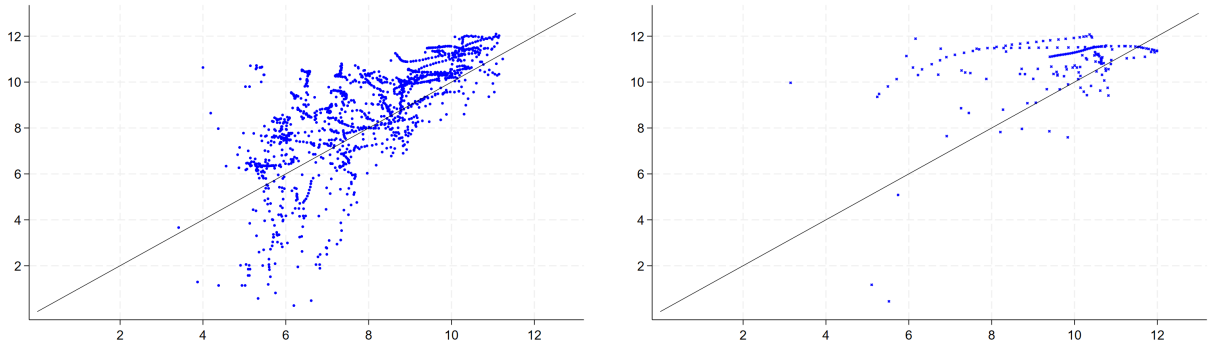
Note: EBE parameters when prices are allowed to change in the counterfactual.

Figure 11: Fit of the EBE and Empirical Model using expectation equation (15) from $t-1$ to predict t



Note: Scatter plot of observed log advertising on vertical axis against predicted log advertising on horizontal axis. Detailing on the left, DTCA on the right. All four markets together.

Figure 12: Fit of the EBE and Empirical Model using expectation equation (15) from $t=0$ to predict t

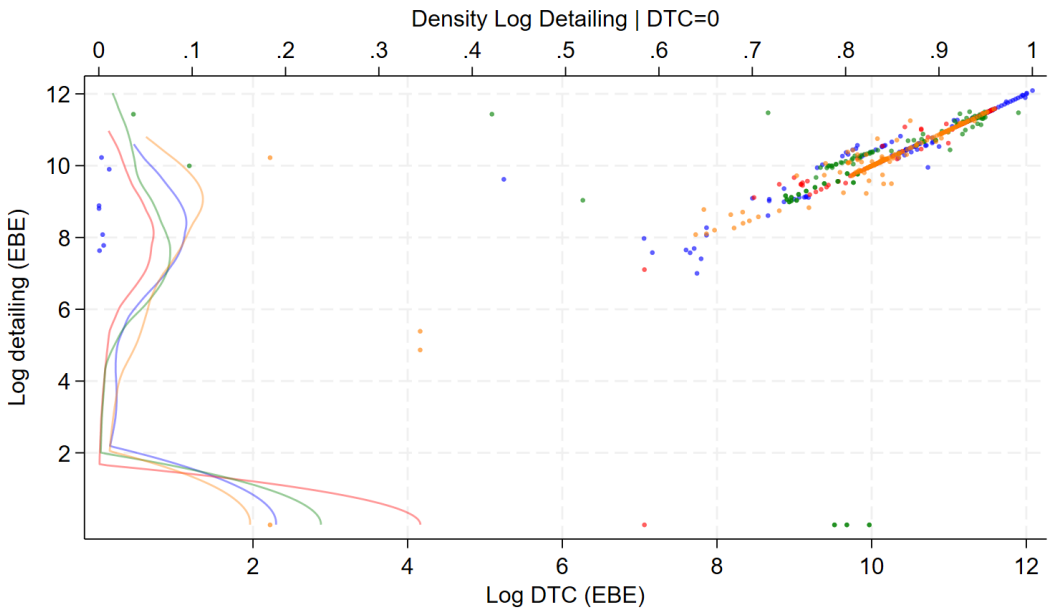


Note: Scatter plot of observed log advertising on vertical axis against predicted log advertising on horizontal axis. Detailing on the left, DTCA on the right. All four markets together.

Table 14: Experience Based Equilibrium parameters when DTC Advertising is banned

	Anticholesterol	Antiasthma	Antidepressants	Antiulcer
$\theta_{1,d}$	.9310	.9121	1.028	.6526
$\beta_{w,d}$ (Other detail)	-.101	.1093	-.033	-.169
$\beta_{w,d2}$ (time to loe)	-.0093	.00519	-.007	.00457
$\theta_{0,d}$	-.614	-1.66	-2.50	1.894

Figure 13: Detailing versus DTCA for dynamic EBE equilibrium in each market



Note: Detailing on the vertical axis versus DTCA on horizontal. Colors for each market in left graph (Blue for Anticholesterol, Red for Antiulcer, Green for Antidepressants, Orange for Antiasthma). The density distribution of Detailing when DTCA is zero is plotted horizontally on the left vertical axis and the associated probabilities can be read off the top of the table.

Table 15: Total quarterly market level detailing, DTCA and net profit with and without DTCA (θ_0 fixed)

	Anticholesterol		Antiasthma		Antidepressants		Antiulcer	
All	Total	Change	Total	Change	Total	Change	Total	Change
Quarterly discount rate (β)	.97		.914		.922		.97	
A. Net Profit								
Data	3,405,022		3,387,940		1,786,078		2,685,483	
EBE	3,440,861	1%	3,494,841	3.1%	1,926,155	7.8%	2,691,954	.2%
EBE with price change	3,406,281	0%	3,471,253	2.4%	1,921,553	7.5%	2,694,432	.3%
EBE no DTC	3,016,037	-11.3%	2,425,417	-27.5%	1,643,767	-7.3%	2,557,059	-4.7%
EBE branded drugs	3,149,304	1.1%	3,223,429	3.5%	1,474,854	10.4%	2,473,362	.3%
EBE no DTC branded drugs	2,699,888	-14.2%	2,138,145	-33.6%	1,179,421	-20%	2,327,273	-5.9%
EBE generic drugs	291,557	-.5%	271,412	-1.4%	451,301	.1%	218,593	-.9%
EBE no DTC generic drugs	316,149	8.4%	287,272	5.8%	464,346	2.8%	229,786	5.1%
B. Detailing								
Detailing data	179,036		217,271		160,495		86,708	
Detailing EBE	149,883	-16.2%	165,357	-23.8%	78,405	-51.1%	84,616	-2.4%
Detailing EBE no DTC	134,480	-10.2%	151,969	-8%	66,546	-15.1%	23,435	-72.3%

Note: Totals in 1,000 US$ per quarter. All % changes of EBE allowing DTCA are compared to the data values. All % changes of EBE without DTCA are compared to the EBE allowing DTCA values.

Table 16: EBE prediction of zero detailing

	Anticholesterol	Antiasthma	Antidepressants	Antiulcer
No price change				
P00	0.953	0.965	0.908	0.977
P10	0.047	0.035	0.092	0.023
P01	0.000	0.000	0.000	0.000
P11	1.000	1.000	1.000	1.000
No price change, θ_0 fixed				
P00	0.949	0.954	0.958	0.979
P10	0.051	0.046	0.042	0.021
P01	0.000	0.000	0.000	0.000
P11	1.000	1.000	1.000	1.000
With price change				
P00	0.966	0.965	0.904	0.836
P10	0.034	0.035	0.096	0.164
P01	0.000	0.000	0.000	0.000
P11	1.000	1.000	1.000	1.000
With price change, θ_0 fixed				
P00	0.966	0.954	0.952	0.979
P10	0.034	0.046	0.048	0.021
P01	0.000	0.000	0.000	0.000
P11	1.000	1.000	1.000	1.000

Note: PXY means the probability that predicted advertising is $X \in \{0, 1\}$ given it is $Y \in \{0, 1\}$ in the data, where 0 means no advertising and 1 means positive advertising. The fit is better when P00 and P11 are closer to 1. Generics are not included as they do not advertise and the fit is perfect for them.

A.5 Robustness: All ξ regressions by market

Table 17: Regression of ξ_{jt} for Anticholesterol

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$\xi_{j,t-2}$	0.883*	0.883*	0.882*	0.883*	0.885*	0.900*	0.929*	0.917*
	(0.015)	(0.015)	(0.015)	(0.015)	(0.015)	(0.016)	(0.014)	(0.021)
$\log(1 + a_{dj,t-1})$	0.012	0.013			0.010	0.006		
	(0.007)	(0.007)			(0.009)	(0.009)		
$\log(1 + a_{Dj,t-1})$	0.020*	-0.111			0.018*	-0.628		
	(0.007)	(0.109)			(0.008)	(1.155)		
$\log((1 + a_{dj,t-1})(1 + a_{Dj,t-1}))$			0.016*	0.012			0.007*	0.006
			(0.003)	(0.007)			(0.003)	(0.006)
$\log(1 + a_{dj,t-1}) \log(1 + a_{Dj,t-1})$		0.012		0.001		0.058		0.001
		(0.010)		(0.001)		(0.103)		(0.001)
Generic	0.254*	0.254*	0.273*	0.251*	0.233*	0.207*	0.000	0.120*
	(0.063)	(0.063)	(0.055)	(0.063)	(0.073)	(0.084)	(.)	(0.055)
Constant	0.178*	0.178*	0.196*	0.174*	0.156*	0.116	0.101*	0.150*
	(0.050)	(0.050)	(0.041)	(0.050)	(0.062)	(0.073)	(0.027)	(0.044)
Adj. R-Square	0.8351	0.8352	0.8352	0.8352	0.8350	0.8389	0.8691	0.8706
N	833	833	833	833	833	811	645	645

Note: OLS for first four columns, 2SLS for next four columns. * for $p < .05$, ** for $p < .01$, and *** for $p < .001$

Table 18: Regression of ξ_{jt} for Antiasthma

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$\xi_{j,t-2}$	0.937*	0.937*	0.945*	0.937*	0.937*	0.955*	0.963*	0.951*
	(0.012)	(0.012)	(0.012)	(0.012)	(0.013)	(0.014)	(0.023)	(0.024)
$\log(1+a_{dj,t-1})$	0.028*	0.028*			0.027*	0.022*		
	(0.006)	(0.006)			(0.008)	(0.009)		
$\log(1+a_{Dj,t-1})$	0.000	0.017			0.001	-0.560		
	(0.006)	(0.076)			(0.007)	(2.642)		
$\log((1+a_{dj,t-1})(1+a_{Dj,t-1}))$			0.013*	0.028*			0.013*	0.034*
			(0.003)	(0.006)			(0.002)	(0.010)
$\log(1+a_{dj,t-1})\log(1+a_{Dj,t-1})$		-0.002		-0.002*		0.051		-0.003*
		(0.007)		(0.001)		(0.238)		(0.001)
Generic	0.206*	0.207*	0.137*	0.206*	0.201*	0.171	0.153*	0.262*
	(0.049)	(0.049)	(0.042)	(0.049)	(0.058)	(0.090)	(0.055)	(0.083)
OTC	0.111	0.111	0.073	0.111	0.108	0.109		
	(0.071)	(0.071)	(0.070)	(0.071)	(0.073)	(0.083)		
Constant	0.210*	0.210*	0.136*	0.210*	0.204*	0.178*	0.141*	0.251*
	(0.040)	(0.040)	(0.030)	(0.040)	(0.051)	(0.089)	(0.035)	(0.067)
Adj. R-Square	0.8376	0.8375	0.8368	0.8376	0.8376	0.8322	0.8574	0.8583
N	1289	1289	1289	1289	1289	1248	1005	1005

Note: OLS for first four columns, 2SLS for next four columns. * for $p < .05$, ** for $p < .01$, and *** for $p < .001$

Table 19: Regression of ξ_{jt} for Antidepressants

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$\xi_{j,t-2}$	0.882*	0.884*	0.889*	0.882*	0.872*	0.877*	0.929*	0.922*
	(0.010)	(0.010)	(0.009)	(0.010)	(0.010)	(0.033)	(0.015)	(0.016)
$\log(1+a_{dj,t-1})$	0.032*	0.031*			0.045*	0.048*		
	(0.004)	(0.004)			(0.006)	(0.017)		
$\log(1+a_{Dj,t-1})$	0.004	-0.098*			-0.003	0.838		
	(0.004)	(0.036)			(0.005)	(0.778)		
$\log((1+a_{dj,t-1})(1+a_{Dj,t-1}))$			0.018*	0.031*			0.014*	0.020*
			(0.002)	(0.004)			(0.002)	(0.006)
$\log(1+a_{dj,t-1})\log(1+a_{Dj,t-1})$		0.009*		-0.002*		-0.075		-0.001
		(0.003)		(0.001)		(0.070)		(0.001)
Generic	0.271*	0.260*	0.205*	0.264*	0.336*	0.429*	0.177*	0.207*
	(0.028)	(0.028)	(0.023)	(0.028)	(0.036)	(0.143)	(0.025)	(0.040)
Constant	0.250*	0.239*	0.185*	0.243*	0.316*	0.409*	0.155*	0.185*
	(0.025)	(0.025)	(0.020)	(0.025)	(0.034)	(0.144)	(0.023)	(0.040)
Adj. R-Square	0.9021	0.9027	0.9008	0.9018	0.9013	0.8522	0.9239	0.9236
N	1227	1227	1227	1227	1227	1193	948	911

Note: OLS for first four columns, 2SLS for next four columns. * for $p < .05$, ** for $p < .01$, and *** for $p < .001$

Table 20: Regression of ξ_{jt} for Antiulcer

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$\xi_{j,t-2}$	0.912*	0.913*	0.914*	0.912*	0.912*	0.944*	0.934*	0.935*
	(0.011)	(0.011)	(0.011)	(0.011)	(0.011)	(0.041)	(0.019)	(0.019)
$\log(1 + a_{dj,t-1})$	-0.005	-0.005			-0.008	0.007		
	(0.005)	(0.005)			(0.006)	(0.020)		
$\log(1 + a_{Dj,t-1})$	0.016*	-0.123			0.022*	-6.216		
	(0.007)	(0.095)			(0.008)	(7.806)		
$\log((1 + a_{dj,t-1})(1 + a_{Dj,t-1}))$			0.003	-0.006			0.006*	0.014
			(0.003)	(0.005)			(0.003)	(0.013)
$\log(1 + a_{dj,t-1}) \log(1 + a_{Dj,t-1})$		0.013		0.002*		0.585		-0.002
		(0.009)		(0.001)		(0.733)		(0.003)
Generic	0.033	0.032	0.048	0.031	0.029	-0.017	0.073*	0.087*
	(0.030)	(0.030)	(0.029)	(0.030)	(0.032)	(0.090)	(0.034)	(0.044)
OTC	0.019	0.018	0.011	0.020	0.023	-0.047		
	(0.031)	(0.031)	(0.031)	(0.031)	(0.031)	(0.112)		
Constant	0.041	0.040	0.054*	0.040	0.039	-0.018	0.084*	0.098*
	(0.022)	(0.022)	(0.022)	(0.022)	(0.024)	(0.101)	(0.029)	(0.040)
Adj. R-Square	0.8472	0.8473	0.8468	0.8473	0.8471	0.2969	0.8395	0.8379
N	1222	1222	1222	1222	1222	1194	930	930

Note: OLS for first four columns, 2SLS for next four columns. * for $p < .05$, ** for $p < .01$, and *** for $p < .001$

A.6 Details on the Empirical Model estimation method

The method of moments estimation of the empirical model relies on the moments provided by equations (16), (14) and (15).

The inner loop consists in finding for each j the $c_j(\theta)$ that satisfies:

$$\sum_t \left(1_{\{a_{h,jt}=0\}} \left(\log(1+a_{h,jt}) - \underline{v}_{h,jt}(\theta, c_j) - E[\nu_{h,jt}|a_{h,jt}>0]\right)\right) = 0$$

Then, the outer loop does the following minimizations:

$$\min_{\sigma_h} \sum_{j,t} \left(1_{\{a_{h,jt}=0\}} - \left(1 - \Phi\left(-\frac{\underline{v}_{h,jt}(\theta, c_j(\theta))}{\sigma_h}\right)\right)\right)^2 \tag{21}$$

$$\min_{\theta} \sum_{j,t} (1_{\{a_{h,jt}=0\}} \tilde{\nu}_{h,jt}(\theta) W_{jt}) \tag{22}$$

where

$$\tilde{\nu}_{h,jt}(\theta) \equiv \log(1+a_{h,jt}) - \underline{v}_{h,jt}(\theta, c_j(\theta)) - E\left[\nu_{h,jt}|\nu_{h,jt} > -\underline{v}_{h,jt}(\theta, c_j(\theta)), J_{jt}\right]$$

is assumed mean zero or uncorrelated with instruments W_{jt} (we either assume $\tilde{\nu}_{h,jt}(\theta)$ is mean zero in which case we use Non Linear Least Squares with $W_{jt} \equiv \tilde{\nu}_{h,jt}(\theta)$ or we use GMM using $W_{jt} \equiv a_{h,jt-1}$. Whether we need instruments depends on the timing of advertising decisions relative to when profits become known.

The outer loops starts by estimating equation (15) using a control function for $E(\nu_{h,jt}|J_{jt}, a_{h,jt}>0, \theta)$. The estimates of θ and c that this generates are used to compute an initial value of $\underline{v}(\theta, c_j)$. These are then substituted into equation (14) to estimate the variance parameters σ_d, σ_D. These estimates are then used as starting values of the full algorithm.

Then, the outer loop proceeds in sequence between the minimization of (21) given θ and the minimization of (22) given σ_h for $h = d, D$, until convergence.

A.7 EBE algorithm with Prices Changes

In the case where we allow prices to change in the Experience Based Equilibrium, we compute new prices solution at each beginning of iteration l of the EBE algorithm, following a price model that accounts for the change in ξ compared to the data. Using the previous iteration $l-1$ values for ξ_{jt}^{l-1} we use a price model to obtain new prices p_{jt}^l that are used for iteration l when computing profits. As the Bertrand Nash equilibrium in prices clearly does not fit the data, we prefer to use a reduced form pricing model consistent with what is observed in the data and assume it would follow in the EBE calculation.

We use the observed relationship between prices and ξ we estimate (we tested several specifications until getting this preferred one) a partially linear model including product fixed effects, a well as market and drug characteristics (time to loss of exclusivity and number of competitors) w_{jt}:

$$p_{jt} = p_{0j} + f(\xi_{jt}) + \alpha_w w_{jt} + \epsilon_{jt} \tag{23}$$

For any iteration value ξ_{jt}^{l-1}, we use the following counterfactual price p_{jt}^l at iteration l: $p_{jt}^l \equiv p_{jt} - \hat{f}(\xi_{jt}) + \hat{f}(\xi_{jt}^{l-1})$. For the four markets, we obtain $\hat{f}$ is an increasing function.

A.8 Aggregation of generics and the pricing equation

The Bertrand Nash equilibrium in prices is estimated to show that it cannot provide meaningful marginal cost estimates. The reason is that average prices for the US need not satisfy an aggregate price first order condition for the aggregate demand because demand of different buyers in the US is heterogeneous. However, we still estimate these first order conditions for the sake of comparing marginal cost estimates. To do so, we take into account the aggregation of generics as follows.

For generics, demand is modeled using aggregate generics as a single product j, but prices are chosen by each generic company with smaller market share than the aggregate share of generics of the same molecule. We thus need to consider that prices are chosen by each individual generic company and not by the "aggregate" generic entity.

Denoting s_{jkt} the market share of the generic firm k of generic product j, if generics are identical in price and characteristics for the consumers, we have the aggregate share s_{jt} of generics of product j with same molecule as the sum of market shares s_{jkt} of generics k of product j: $s_{jt} = \sum_{k=1}^{K} s_{jkt}$ if there are K generics.

With the random coefficient logit model, we have

$$s_{jkt} = \int s_{ijkt} dF\left(\beta_p^i\right) = \int \frac{1}{K} s_{ijt} dF\left(\beta_p^i\right) = \frac{1}{K} s_{jt}$$

because we assume generics are identical and thus have equal market shares ($p_{jkt} = p_{jk't} = p_{jt}$, $s_{ijkt} = s_{ijk't}$).

Then,

$$\frac{\partial s_{jkt}}{\partial p_{jkt}} = \int \beta_p^i s_{ijkt} \left(1 - s_{ijkt}\right) dF\left(\beta_p^i\right) = \frac{1}{K} \int \beta_p^i s_{ijt} \left(1 - \frac{1}{K} s_{ijt}\right) dF\left(\beta_p^i\right)$$

and

$$\frac{\partial s_{jkt}}{\partial p_{jrt}} = -\int \beta_p^i s_{ijkt} s_{ijrt} dF\left(\beta_p^i\right) = -\frac{1}{K} \int \beta_p^i s_{ijt} s_{ijrt} dF\left(\beta_p^i\right)$$

implying that for generics j

$$\frac{\partial s_{jt}}{\partial p_{jt}} = \sum_{k=1}^{K} \frac{\partial s_{jkt}}{\partial p_{jkt}} = \int \beta_p^i s_{ijt} \left(1 - \frac{1}{K} s_{ijt}\right) dF\left(\beta_p^i\right) = K \frac{\partial s_{jkt}}{\partial p_{jkt}}$$

and

$$\frac{\partial s_{jt}}{\partial p_{j't}} = \sum_{k=1}^{K} \frac{\partial s_{jkt}}{\partial p_{j'kt}} = -\sum_{k=1}^{K} \frac{1}{K} \int \beta_p^i s_{ijt} s_{ij't} dF\left(\beta_p^i\right) = -\int \beta_p^i s_{ijt} s_{ij't} dF\left(\beta_p^i\right) = K \frac{\partial s_{jkt}}{\partial p_{j'kt}}$$

A.9 Model Variant

We present here the model estimated to which the empirical results correspond to and which differs from the model presented in the definition of the advertising lags that affect ξ.

A.9.1 Empirical model

The starting difference of the model presented in the main part of the paper concerns the ξ equation which empirically proves to be depending on two quarterly lags of advertising expenses as follows[33]

$$\xi_{jt} = \rho_\xi \xi_{jt-2} + f(a_{D,jt} + a_{D,jt-1}, a_{d,jt} + a_{d,jt-1}) + z_{jt}\beta_z + \mu_{jt} + \mu_{jt-1} \tag{24}$$

where

$$\begin{aligned} f(a_{D,jt} + a_{D,jt-1}, a_{d,jt} + a_{d,jt-1}) = \beta_D \log(1 + a_{D,jt} + a_{D,jt-1}) + \beta_d \log(1 + a_{d,jt} + a_{d,jt-1}) \\ + \beta_{D,d} \left(\log(1 + a_{D,jt} + a_{D,jt-1}) \log(1 + a_{d,jt} + a_{d,jt-1})\right) \end{aligned} \tag{25}$$

As we find that $\beta_{D,d} = 0$ and $\beta_d = \beta_D \equiv \beta_a$, we have thus that

$$\frac{\partial f(a_{D,jt} + a_{D,jt-1}, a_{d,jt} + a_{d,jt-1})}{\partial a_{h,jt}} = \frac{\beta_a}{1 + a_{h,jt} + a_{h,jt-1}}$$

Then, the expected discounted sum of profits writes:

$$W(a|\xi_{jt}, J_{jt}) \equiv \mathcal{E}\left[\sum_{\tau=0}^{\infty} \beta^\tau \left(\pi(\cdot)_{t+\tau} - a_{d,jt+\tau} - a_{D,jt+\tau}\right) |\xi_{jt}, J_{jt}, a\right]$$

[33] We account for these two lags in the estimation as shown in Appendix A.10.

Moreover, we have

$$\mathcal{E}[\frac{\partial \pi_{jt}}{\partial \xi_{jt}} \times \frac{\partial \xi_{jt}}{\partial a_{h,jt}} | J_{jt}, \xi_{jt}] = \mathcal{E}[\frac{\partial \pi_{jt}}{\partial \xi_{jt}} | J_{jt}, \xi_{jt}] \ \times \ \frac{\beta_a}{1 + a_{h,jt} + a_{h,jt-1}}$$

Then, we use the approximation of the perception as:

$$\mathcal{E}\Big[\sum_{\tau=1}^{\infty} (\beta \rho_\xi)^\tau \frac{\partial \pi(\cdot)_{jt+\tau}}{\partial a_{h,jt}} | J_{jt}, \xi_{jt}\Big] = \tilde{\theta}_{0,h} \left(\mathcal{E}[\frac{\partial \pi(\cdot)_{jt}}{\partial \xi_{jt}} | J_{jt}] \right)^{\theta_{1,h}} \frac{\beta_a}{1 + a_{h,jt} + a_{h,jt-1}} \exp[w_{h,jt-1}\beta_{w,h} + \omega_{h,jt}] \quad (26)$$

Then, taking logs we obtain:

$$\begin{aligned} (i) \quad P(a_{h,jt} = 0) \quad &= \quad Pr\left\{ 0 \geq -\log(1 + a_{h,jt-1}) + \theta_{0,h} + \theta_{1,h} \log \frac{\partial \pi(\cdot)_{jt-1}}{\partial \xi_{jt-1}} + w_{jt-1}\beta_{w,h} + \omega_{h,jt} \right\}, \text{ and} \\ (ii) \quad a_{h,jt} > 0 \quad &\Rightarrow \quad \log(1 + a_{h,jt} + a_{h,jt-1}) = \theta_{0,h} + \theta_{1,h} \log \frac{\partial \pi(\cdot)_{jt-1}}{\partial \xi_{jt-1}} + w_{jt-1}\beta_{w,h} + \omega_{h,jt}. \end{aligned} \quad (27)$$

Defining

$$\begin{aligned} \underline{v}_{h,jt}(\theta, c_j) = (1 - \rho_h)\theta_{0,h} + \theta_{1,h} &\left(\log \frac{\partial \pi(\cdot, c)_{jt-1}}{\partial \xi_{jt-1}} - \rho_h \log \frac{\partial \pi(\cdot, c)_{jt-2}}{\partial \xi_{jt-2}} \right) \\ &+ (w_{jt-1} - \rho_h w_{h,jt-2})\beta_{w,h} + \rho_h \log(1 + a_{h,jt-1} + a_{h,jt-2}) \end{aligned} \quad (28)$$

so if $a_{h,jt} > 0$ is obtained as (conditionally on $a_{h,jt-1} > 0$) [34]:

$$\log(1 + a_{h,jt} + a_{h,jt-1}) = \underline{v}_{h,jt}(\theta, c_j) + \nu_{h,jt}$$

and

$$P(a_{h,jt} = 0) = \Phi\left(\frac{\log(1 + a_{h,jt-1}) - \underline{v}_{h,jt}(\theta, c_j)}{\sigma_h} \right)$$

[34] If $a_{h,jt-1} = 0$, then conditionally on last positive advertising $a_{h,jt-n} > 0$, we have

$(\log(1 + a_{h,jt} | a_{h,jt} > 0, a_{h,jt-1} = 0, .., a_{h,jt-n+1} = 0, a_{h,jt-n} > 0) = \underline{v}^n_{h,jt}(\theta, c_j) + \nu_{h,jt} + \rho_h \nu_{h,jt-1} + \rho_h^2 \nu_{h,jt-2} + .. + \rho_h^n \nu_{h,jt-n}$

with

$$\begin{aligned} \underline{v}^n_{h,jt}(\theta, c_j) = (1 - \rho_h^n)\theta_{0,h} + \theta_{1,h} &\left(\log \frac{\partial \pi(\cdot, c)_{jt-1}}{\partial \xi_{jt-1}} - \rho_h^n \log \frac{\partial \pi(\cdot, c)_{jt-(n+1)}}{\partial \xi_{jt-(n+1)}} \right) \\ &+ (w_{jt-1} - \rho_h^n w_{jt-(n+1)})\beta_{w,h} + \rho_h^n \log(1 + a_{h,jt-n} + a_{h,jt-(n+1)}) \end{aligned}$$

and the the selection term is

$$E[\nu_{h,jt}|J_{jt}, a_{h,jt} > 0, a_{h,jt-1} > 0, \theta] = E\left[\nu_{h,jt}|\nu_{h,jt} \geq \frac{\log(1+a_{h,jt-1}) - \underline{v}_{h,jt}(\theta, c_j)}{\sigma_h}\right]$$
$$= \sigma_h \frac{\phi\left(\frac{\log(1+a_{h,jt-1})-\underline{v}_{h,jt}(\theta,c_j)}{\sigma_h}\right)}{1-\Phi\left(\frac{\log(1+a_{h,jt-1})-\underline{v}_{h,jt}(\theta,c_j)}{\sigma_h}\right)}$$

Hence

$$E[\log(1+a_{h,jt}+a_{h,jt-1})|a_{h,jt} > 0, a_{h,jt-1} > 0, J_{jt}] = \underline{v}_{h,jt}(\theta, c_j) + \sigma_h \frac{\phi\left(\frac{\log(1+a_{h,jt-1})-\underline{v}_{h,jt}(\theta,c_j)}{\sigma_h}\right)}{1-\Phi\left(\frac{\log(1+a_{h,jt-1})-\underline{v}_{h,jt}(\theta,c_j)}{\sigma_h}\right)} \tag{29}$$

Concerning the inner loop of the estimation algorithm of the empirical model, we recover marginal costs as solution of:

$$0 = \sum_t \Big(\log(1+a_{h,jt}+a_{h,jt-1}) - \rho_h \log(1+a_{h,jt-1}+a_{h,jt-2}) - \theta_{0,h}(1-\rho_h)\Big)$$
$$- \sum_t \Bigg[\theta_{1,h} \log\left(\frac{\partial D(\cdot)_{jt}}{\partial \xi_{jt}}(p_{jt}-c_j)\right) - \rho_h \theta_{1,h} \log\left(\frac{\partial D(\cdot)_{jt-1}}{\partial \xi_{jt-1}}(p_{jt-1}-c_j)\right)$$
$$+ \beta_{w,h}(w_{jt} - \rho_h w_{jt-1}) + E[\nu_{h,jt}|J_{jt}, a_{h,jt} > 0, a_{h,jt-1} > 0, \theta]\Bigg] \tag{30}$$

And for the outer loop the error term $\tilde{\nu}_{h,jt}$ is defined as:

$$\tilde{\nu}_{h,jt} \equiv \log(1+a_{h,jt}+a_{h,jt-1}) - \rho_h \log(1+a_{h,jt-1}+a_{h,jt-2}) - \theta_{0,h}(1-\rho_h)$$
$$-\theta_{1,h}\left[\log\left(\frac{\partial D(\cdot)_{jt}}{\partial \xi_{jt}}(p_{jt}-c_j(\theta))\right) - \rho_h \log\left(\frac{\partial D(\cdot)_{jt-1}}{\partial \xi_{jt-1}}(p_{jt-1}-c_j(\theta))\right)\right]$$
$$+\beta_{w,h}(w_{jt} - \rho_h w_{jt-1}) + E[\nu_{h,jt}|J_{jt}, a_{h,jt} > 0, \theta]\Big]$$

A.9.2 EBE Algorithm

Simulating K sample paths from iteration l's policy functions. For each $k = 1, .., K$ we do the following steps to simulate a sample path:

- Simulate $\nu^k_{h,jt} \sim \mathcal{N}(0, \sigma^2_h)$ and $\mu^k_{jt} \sim \mathcal{N}(0, \sigma^2_\mu)$, for all products j and $t = 1, .., T$

- Compute all $\omega^k_{h,jt}$ for all j,t using: $\omega^k_{h,jt} = \rho_h \omega^k_{h,jt-1} + \nu^k_{h,jt}$

- Iterate from $t = t_0 + 2$ to T (initializing with data) to compute:

  – Compute $a^k_{h,jt}$:

$$a^k_{h,jt} = 0 \text{ if } 0 \geq -\log(1 + a^k_{h,jt-1}) + \theta^l_{0,h} + \theta^l_{1,h} \log \frac{\partial \pi^k_{jt-1}}{\partial \xi'_{jt-1}} + w_{jt}\beta_{w,h} + \omega^k_{h,jt})$$

  otherwise $a^k_{h,jt} > 0$ such that

$$\log(1 + a^k_{h,jt-1} + a^k_{h,jt}) = \theta^l_{0,h} + \theta^l_{1,h} \log \frac{\partial \pi^k_{jt-1}}{\partial \xi_{jt-1}} + w_{jt}\beta_{w,h} + \omega^k_{h,jt} \tag{31}$$

  – Compute

$$\xi^k_{j,t} = \rho_\xi \xi^k_{j,t-2} + \beta_a \log((1 + a^k_{Dj,t} + a^k_{Dj,t-1})(1 + a^k_{dj,t} + a^k_{dj,t-1})) + z_{j,t}\beta_z + \mu^k_{j,t} + \mu^k_{j,t-1}$$

  – Calculate net profit and profit derivative as

$$\pi^k_{jt}\left(p_t, \xi^k_t\right) = (p_{jt} - c_j)M_t s^k_{jt} - a^k_{dj,t} - a^k_{Dj,t} \quad \textit{and} \quad \frac{\partial \pi^k_{jt}(p_t, \xi^k_t)}{\partial \xi_{jt}} = (p_{jt} - c_j)M_t \frac{\partial s^k_{jt}}{\partial \xi_{jt}}$$

  using market shares

$$s^k_{jt}\left(p_t, \xi^k_t\right) = \int \frac{\exp(\beta^i_p(X_j)p_{jt} + \beta_{m(j)} + \beta_x X_{jt} + \xi^k_{jt})}{1 + \sum_{j'} \exp(\beta^i_p(X_{j'})p_{j't} + \beta_{m(j')} + \beta_x X_{j't} + \xi^k_{j't})} dF(\beta^i_p(X))$$

Computing the solutions of the Kuhn-Tucker condition for advertising $a^{*l}_{h,jt}$.

- We solve for $a^{*l}_{h,jt}$:

$$a^{*l}_{h,jt} \frac{\partial W^l_{j,t}}{\partial a^l_{h,jt}} \approx 0$$

  where

$$W^l_{j,t} \equiv \pi^l_{j,t+\tau} - a^l_{d,jt} - a^l_{D,jt} + \beta \mathcal{E}_k \left[\sum\nolimits_{\tau=1}^{\infty} \beta^\tau \left(\pi^{kl}_{j,t+\tau} - a^{kl}_{d,jt+\tau} - a^{kl}_{D,jt+\tau}\right)\right]$$

Since

$$\xi_{jt}^{kl} = \rho_\xi \xi_{jt-2}^{kl} + \beta_a \log((1 + a_{d,jt}^{kl} + a_{d,jt-1}^{kl})(1 + a_{D,jt}^{kl} + a_{D,jt-1}^{kl})) + z_{j,t}^{kl}\beta_z + \mu_{jt}^{kl} + \mu_{jt-1}^{kl}$$

we have

$$\frac{\partial W_{j,t}^{l}}{\partial a_{h,jt}^{l}} = \frac{\partial \pi_{j,t}^{l}}{\partial \xi_{jt}} \frac{\beta_a}{1 + a_{h,jt}^{l} + a_{h,jt-1}^{l}} + \beta \mathcal{E}_k \left[\frac{\partial \pi_{j,t+1}^{kl}}{\partial \xi_{jt+1}} \frac{\beta_a}{1 + a_{h,jt+1}^{kl} + a_{h,jt}^{kl}} \right]$$
$$+\beta^2 \mathcal{E}_k \left[\frac{\partial \pi_{j,t+2}^{kl}}{\partial \xi_{jt+2}} \rho_\xi \frac{\beta_a}{1 + a_{h,jt}^{kl} + a_{h,jt-1}^{kl}} \right] + \beta^3 \mathcal{E}_k \left[\frac{\partial \pi_{j,t+3}^{kl}}{\partial \xi_{jt+3}} \rho_\xi \frac{\beta_a}{1 + a_{h,jt+1}^{kl} + a_{h,jt}^{kl}} \right] + ... - 1$$

that is

$$\frac{\partial W_{j,t}^{l}}{\partial a_{h,jt}^{l}} = -1 + \frac{\partial \pi_{j,t}^{l}}{\partial \xi_{jt}} \frac{\beta_a}{1 + a_{h,jt}^{l} + a_{h,jt-1}^{l}}$$
$$+\mathcal{E}_k \left[\sum_{\tau=1}^{\infty} 1_{\{\tau \text{ even}\}} \frac{\partial \pi_{j,t+\tau}^{kl}}{\partial \xi_{jt+\tau}} \frac{\beta_a}{1 + a_{h,jt}^{l} + a_{h,jt-1}^{kl}} \beta^\tau \rho_\xi^{int(\tau/2)} + 1_{\{\tau \text{ odd}\}} \frac{\partial \pi_{j,t+\tau}^{kl}}{\partial \xi_{jt+\tau}} \frac{\beta_a}{1 + a_{h,jt+1}^{kl} + a_{h,jt}^{l}} \beta^\tau \rho_\xi^{int(\tau/2)} \right]$$

Denoting

$$\partial W_{h,jt}^{even,k} \equiv \sum_{\tau=2}^{\infty} 1_{\{\tau \text{ even}\}} \frac{\partial \pi_{j,t+\tau}^{kl}}{\partial \xi_{jt+\tau}} \beta^\tau \rho_\xi^{int(\tau/2)}$$

and

$$\partial W_{h,jt}^{odd,k} \equiv \sum_{\tau=0}^{\infty} 1_{\{\tau \text{ odd}\}} \frac{\partial \pi_{j,t+\tau}^{kl}}{\partial \xi_{jt+\tau}} \beta^\tau \rho_\xi^{int(\tau/2)}$$

where $int(.)$ means the integer part.

Thus, $a_{h,jt}^{*l} = 0$ if $\frac{\beta_a}{1 + a_{h,jt-1}^{*l}} \left(\frac{\partial \pi_{j,t}^{l}}{\partial \xi_{jt}} + \mathcal{E}_k \left[\partial W_{h,jt}^{even,k} \right] \right) + \mathcal{E}_k \left[\frac{\beta_a}{1 + a_{h,jt+1}^{kl}} \partial W_{h,jt}^{odd,k} \right] < 1$ and otherwise is solution of:

$$\frac{\beta_a}{1 + a_{h,jt}^{*l} + a_{h,jt-1}^{*l}} \left(\frac{\partial \pi_{j,t}^{l}}{\partial \xi_{jt}} + \mathcal{E}_k \left[\partial W_{h,jt}^{even,k} \right] \right) + \mathcal{E}_k \left[\frac{\beta_a}{1 + a_{h,jt+1}^{kl} + a_{h,jt}^{*l}} \partial W_{h,jt}^{odd,k} \right] = 1 \qquad (32)$$

Perceptions of post sample profits. We have

$$\frac{\partial W^l_{j,t}}{\partial a^{*l}_{h,jt}} = -1 + \frac{\partial \pi^l_{j,t}}{\partial \xi_{jt}} \frac{\beta_a}{1 + a^{*l}_{h,jt} + a^{*l}_{h,jt-1}}$$
$$+\mathcal{E}_k \left[\sum_{\tau=1}^{T_{h,j}-2} 1_{\{\tau \text{ even}\}} \frac{\partial \pi^{kl}_{j,t+\tau}}{\partial \xi_{jt+\tau}} \frac{\beta_a}{1 + a^{*l}_{h,jt} + a^{*l}_{h,jt-1}} \beta^\tau \rho_\xi^{int(\tau/2)} + 1_{\{\tau \text{ odd}\}} \frac{\partial \pi^{kl}_{j,t+\tau}}{\partial \xi_{jt+\tau}} \frac{\beta_a}{1 + a^{kl}_{h,jt+1} + a^{*l}_{h,jt}} \beta^\tau \rho_\xi^{int(\tau/2)} \right]$$
$$+\mathcal{E}_k \left[\sum_{\tau=T_{h,j}-1}^{\infty} 1_{\{\tau \text{ even}\}} \frac{\partial \pi^{kl}_{j,t+\tau}}{\partial \xi_{jt+\tau}} \frac{\beta_a}{1 + a^{*l}_{h,jt} + a^{*l}_{h,jt-1}} \beta^\tau \rho_\xi^{int(\tau/2)} + 1_{\{\tau \text{ odd}\}} \frac{\partial \pi^{kl}_{j,t+\tau}}{\partial \xi_{jt+\tau}} \frac{\beta_a}{1 + a^{kl}_{h,jt+1} + a^{*l}_{h,jt}} \beta^\tau \rho_\xi^{int(\tau/2)} \right]$$

As

$$\frac{\partial W^l_{j,T_{h,j}}}{\partial a^{*l}_{h,jT_{h,j}}} = -1 + \frac{\partial \pi^l_{j,T_{h,j}}}{\partial \xi_{jT_{h,j}}} \frac{\beta_a}{1 + a^{*l}_{h,jT_{h,j}} + a^{*l}_{h,jT_{h,j}-1}}$$
$$+\mathcal{E}_k \left[\sum_{\tau=1}^{\infty} 1_{\{\tau \text{ even}\}} \frac{\partial \pi^{kl}_{j,T_{h,j}+\tau}}{\partial \xi_{jT_{h,j}+\tau}} \frac{\beta_a}{1 + a^{*l}_{h,jT_{h,j}} + a^{*l}_{h,jT_{h,j}-1}} \beta^\tau \rho_\xi^{int(\tau/2)} + 1_{\{\tau \text{ odd}\}} \frac{\partial \pi^{kl}_{j,T_{h,j}+\tau}}{\partial \xi_{jT_{h,j}+\tau}} \frac{\beta_a}{1 + a^{kl}_{h,jT_{h,j}+1} + a^{*l}_{h,jT_{h,j}}} \beta^\tau \rho_\xi^{int(\tau/2)} \right]$$

we use the following approximation

$$\mathcal{E}_k \left[\sum_{\tau=T_{h,j}-1}^{\infty} 1_{\{\tau \text{ even}\}} \frac{\partial \pi^{kl}_{j,t+\tau}}{\partial \xi_{jt+\tau}} \frac{\beta_a}{1 + a^{*l}_{h,jt} + a^{*l}_{h,jt-1}} \beta^\tau \rho_\xi^{int(\tau/2)} + 1_{\{\tau \text{ odd}\}} \frac{\partial \pi^{kl}_{j,t+\tau}}{\partial \xi_{jt+\tau}} \frac{\beta_a}{1 + a^{kl}_{h,jt+1} + a^{*l}_{h,jt}} \beta^\tau \rho_\xi^{int(\tau/2)} \right]$$
$$\approx \beta^{T_{h,j}} \rho_\xi^{int((T_{h,j}+1)/2)} \left(1 - \frac{\partial \pi^l_{j,T_{h,j}}}{\partial \xi_{jT_{h,j}}} \frac{\beta_a}{1 + a^{*l}_{h,jT_{h,j}} + a^{*l}_{h,jT_{h,j}-1}} \right)$$

Hence

$$\frac{\partial W^l_{j,t}}{\partial a^{*l}_{h,jt}} = -1 + \frac{\partial \pi^l_{j,t}}{\partial \xi_{jt}} \frac{\beta_a}{1 + a^{*l}_{h,jt} + a^{*l}_{h,jt-1}}$$
$$+\mathcal{E}_k \left[\sum_{\tau=1}^{T_{h,j}-2} 1_{\{\tau \text{ even}\}} \frac{\partial \pi^{kl}_{j,t+\tau}}{\partial \xi_{jt+\tau}} \frac{\beta_a}{1 + a^{*l}_{h,jt} + a^{*l}_{h,jt-1}} \beta^\tau \rho_\xi^{int(\tau/2)} + 1_{\{\tau \text{ odd}\}} \frac{\partial \pi^{kl}_{j,t+\tau}}{\partial \xi_{jt+\tau}} \frac{\beta_a}{1 + a^{kl}_{h,jt+1} + a^{*l}_{h,jt}} \beta^\tau \rho_\xi^{int(\tau/2)} \right]$$
$$+\beta^{T_{h,j}} \rho_\xi^{int((T_{h,j}+1)/2)} \left(1 - \frac{\partial \pi^l_{j,T_{h,j}}}{\partial \xi_{jT_{h,j}}} \frac{\beta_a}{1 + a^{*l}_{h,jT_{h,j}} + a^{*l}_{h,jT_{h,j}-1}} \right)$$

and we solve for $a^{*l}_{h,jt}$ solution of $\dfrac{\partial W^l_{j,t}}{\partial a^{*l}_{h,jt}} = 0$.

Updating the policy function.

- The policy function update $\theta^{l+1}, \beta^{l+1}$ is obtained using $\nu_{h,jt} \to \mathcal{N}(0, \sigma_h^2)$ and the maximum likelihood estimation of:

$$\log(1 + a_{h,jt}^{*l} + a_{h,jt-1}^{*l}) = \max \left\{ \log(1 + a_{h,jt-1}^{*l}), \theta_{0,h}^{l+1} + \theta_{1,h}^{l+1} \left(\frac{1}{K} \sum_k \log[\frac{\partial \pi^{k,l}(\cdot)_{h,jt-1}}{\partial \xi_{jt-1}}] \right) \right.$$
$$\left. + \beta_{w,h}^{l+1} \left(\frac{1}{K} \sum_k w_{h,jt-1}^{k.l} \right) + \rho_h \frac{1}{K} \sum_k \omega_{h,jt-1}^{k,l} + \nu_{h,jt} \right\}$$

- We then compute $W_{j,t}^l \equiv \frac{1}{K} \sum_k W_{j,t}^{k,l}$ and iterate until $Y^l \equiv \frac{1}{\sum_{j,t} 1} \sum_{j,t} \left(\frac{W_{j,t}^l - W_{j,t}^{l-1}}{W_{j,t}^{l-1}} \right)^2 < 10^{-5}$ to get equilibrium parameters $\theta_{0,h}^L, \theta_{1,h}^L, \beta_{w,h}^L$ at this last iteration L.

A.10 Econometric details for equation (3)

The exact specification of equation (3) is such that we estimate the following equation:

$$\xi_{jt} = \rho_\xi \xi_{jt-2} + \beta_a \log((1 + a_{D,jt} + a_{D,jt-1})(1 + a_{d,jt} + a_{d,jt-1})) + z_{jt}\beta_z + \mu_{jt} + \mu_{jt-1}$$

We thus need to correct for standard errors in estimating this linear regression because of the serial correlation between error terms $\epsilon_{jt} \equiv \mu_{jt} + \mu_{jt-1}$.

Correcting the standard errors in the ξ equation. Serial correlation implies that for OLS, the true variance covariance is obtained as

$$E[(X'X)^{-1}X'\epsilon\epsilon'X(X'X)^{-1}] = (X'X)^{-1}\Omega(X'X)^{-1}$$

where if $t = 1, ...T$, and Ω_t is the t^{th} row of Ω then

- $\Omega_T = [\sigma^2, .5\sigma^2, 0, ..., 0]$,
- and for $1 < t < T$, $\Omega_t = [0, 0, ...0, \omega_{t-1} = .5\sigma^2, \omega_t = \sigma^2, \omega_{t+1} = .5\sigma^2, 0, ...0]$,
- $\Omega_1 = [0, 0, 0, ...0, \omega_2 = .5\sigma^2, \omega_1 = \sigma^2]$

For the IV regression if there is the same number of instruments as the right hand side variable, the correct value of the covariance matrix is $(Z'X)^{-1}Z'\Omega Z(Z'X)^{-1}$. If there are more instruments then right hand side variables replace Z with the first stage prediction of X (however it would be wrong to replace both X and Z with the first stage prediction, which is an error many make).

As $var(\epsilon_{jt}) = \sigma^2$ and $cov(\epsilon_{jt}, \epsilon_{jt-1}) = \frac{\sigma^2}{2}$, the variance covariance matrix of residuals is

$$\Omega = \begin{bmatrix} \sigma^2 & \frac{\sigma^2}{2} & 0 & 0 & 0 & .. & 0 \\ \frac{\sigma^2}{2} & \sigma^2 & \frac{\sigma^2}{2} & 0 & 0 & .. & 0 \\ 0 & \frac{\sigma^2}{2} & \sigma^2 & \frac{\sigma^2}{2} & 0 & .. & 0 \\ . & . & . & . & . & . & . \\ . & . & . & . & . & . & . \\ 0 & .. & 0 & 0 & \frac{\sigma^2}{2} & \sigma^2 & \frac{\sigma^2}{2} \\ 0 & .. & .. & .. & 0 & \frac{\sigma^2}{2} & \sigma^2 \end{bmatrix}$$

We also correct standard errors using the following transformation with T such that $T'T = \Omega^{-1}$ which amounts to use the T operator as follows:

$$Ty_{jt} = y_{jt} - (y_{jt-1} + y_{jt+1})$$

since

$$\begin{aligned} cov(T\epsilon_{jt}, T\epsilon_{jt-1}) &= cov(\epsilon_{jt} - (\epsilon_{jt-1} + \epsilon_{jt+1}), \epsilon_{jt-1} - (\epsilon_{jt-2} + \epsilon_{jt})) \\ &= \sigma^2(\frac{1}{2} - 1 - 1 + \frac{1}{2} + \frac{1}{2} + \frac{1}{2}) = 0 \end{aligned}$$

Thus T diagonalize Ω, that is $T\Omega T' = I$ and $T'T = \Omega^{-1}$, which allows to compute correct standard errors by usual formulas.