Pharmaceutical Advertising in Dynamic Equilibrium.

Pierre Dubois and Ariel Pakes

(Toulouse School of Economics and Harvard University)

Toulouse School of Economics

August 5, 2024

Background

• U.S. is the largest pharmaceutical market in the world in both revenue (recently $\approx$ \$600B per annum) and promotional spending: $\approx$ \$ 7B on DTC (3/4 on TV), & $\approx$ \$20B on detailing (most on free samples). Our data is from 2005-14 when the quantities were about 2/3 of current quantities.

• Only two developed countries allow DTC of prescription drugs: New Zealand & U.S. (since 1985 in print, but with restrictions that did not allow TV until 1997).

• Markets we analyze: treatments for Asthma, Cholesterol, Depression, & Ulcer: $\approx$ 25% of DTC, and 22% of all detailing.

• Not all drugs are advertised. Depending on the market 50-70% do detailing & 6-17% do DTC.

Background

• However average DTC expenditures (conditional on DTC) as large or larger than average detailing expenditures (conditional on detailing).

• *Is DTC socially useful?* Arguments:

- *Against*; (i) incentives for excessive use, & (ii) returns largely a result of business stealing (no net benefit to society).
- *For*; make; (i) consumers aware that they can treat a condition before it becomes serious (particularly those that do not regularly see doctors) &/or adherence to regimen, (ii) providers aware of possible treatments.

• Keep track of DTC's impact on profits.

- The welfare impact of pharmaceuticals is in trillions of dollars a decade, and R&D costs (private + public) are less than a third of that .
- The worry about pharma profits is accentuated by current policy proposals which if implemented, would likely have dramatic negative effects on pharma R& D (see Ho and Pakes, 2023).

Framework for the Analysis.

• Markets have more than 25 competitors, are divided into therapeutic groups with an average of 3 to 11 products appearing between 2005-14, and face considerable uncertainty from research and regulatory processes.

• Cognitive constraints make Markov Perfect or Bayesian MP questionable.

• Both the theory and the empirical literature are well aware of this

- Theory considers weaker notions of equilibrium; Fudenberg and Levine (1993), Osborne and Rubinstein (1998), Esponda and Pouzo (2016)
- Empirical/Computational literature uses approximations; Benkard et. al. 2008, and the review of recent work by Aguirreberia, et al 2024)

- Start with empirical model.
  - Assumes agents maximize given their perceptions.
  - **Does not** assume their perceptions are "correct" or are "equilibrium perceptions".
- Empirical model
  - Uses data to determines which variables observable to the researcher advertising responds to.
  - Allows for **serially correlated** unobserved state variables to account for impact of variables that the firm knows but we do not observe.

• Introduces a notion of equilibrium that

- conditions on the same variables as the empirical model (our observable, and serially correlated unobserved, states).
- and imposes certain equilibrium conditions on the.relationship between perceptions and actual outcomes (see below).

• Provide a novel algorithm that is relatively quick and easy to use to compute equilibrium policies.

• These policy functions differ from those of the empirical model. So we compare the *in-sample* equilibrium policies to;

- those from the empirical model, and
- to the data.

• Then we recompute equilibrium for our counterfactuals (which are *out of sample*).

Experience Based Equilibrium.

- Equilibrium conditions:
  1. firms chose policies to maximize their perceptions of EDVs conditional on the variables they use to determine their expenditures,
  2. perceptions are consistent with outcomes at states that are visited repeatedly (but not all states).
- Characteristics of equilibrium:
  - Has *asymmetric information*: the two serially correlated "unobserved" states (for DTC & for detailing) not directly observed by competitors.
  - Generates a Markov process for the vector formed from combining the state variables of all firms.
  - *However firms only condition on (& formulate policies for) the states in their own information set (eases cognitive burden).*

• The Markov process wanders into a subset of points that are visited repeatedly (the recurrent class).

• The firm can learn the correct values of their policies from past behavior in (and only in) the recurrent class.

• They can do this by

- keeping averages of past profits and average of past perceived continuation values in memory, and
- updating the averages with current profits and the perceived continuation value at the state reached.

• There is still a question of realism but:

- we compare equilibrium to empirical policies in-sample,
- have shown elsewhere that firms can learn equilibrium policies (though it can take a lot of time).

Literature.

- On DTC: Find positive effects on aspects of health care.
  - Sinkinson and Starc (2019) on demand for drugs (statins), and
  - Shapiro (2018) effects on labor supply (anti-depressants).
  - Jin and Iizuka (2005) on doctor visits (a large set of drugs).
  - Wozinska M. (2005), DTC and compliance.
  - Large literature published in health journals.

- Frameworks used for empirical/computational work on market dynamics.
  - Full Info Markov Perfect (FIMP): Ericson and Pakes (1995).
  - Approximations to FIMP: Benkard et al. (2008) & following literature.
  - Experience Based Equilibrium: Fershtman and Pakes (2012).
  - Anonymous sequential games in related markets: Brancaccio et al. (2020) and others on transportation markets.
  - See Aguirregabaria et. al. (forthcoming) for a recent detailed review.

Steps in the analysis

- Begin with
  - Estimate of a BLP demand system, and recover the "quality" (ξ) terms for each drug in each period (quarter).
  - Estimate a controlled Markov process for how quality evolves over time, where the controls are the two types of advertising.

- With this we can compute;
  - the impact of advertising on demand ($\Rightarrow$ advertising incentives), &
  - if we add marginal costs (see below) business stealing incentives: defined as the profit loss were the firm to stop advertising when its competitors did not (similar to Sinkinson and Starc, but with a demand model).

Steps in the analysis

- Estimate an empirical advertising model. Needs to
  - estimate marginal cost and allow for both
  - serially correlated unobserved states, and for intervals with
  - advertising=0 for several periods between positive advertising sequences (advertising "pulsates", Dubé et al. 2018)

- Introduce the equilibrium framework and a novel (and easy to compute) algorithm for equilibrium policies.

- Compare equilibrium policies; (i) to the policy functions from the empirical model, and (ii) to the data.

- Counterfactuals
  - no DTC (still doing robustness checks).
  - tax advertising and use the proceeds to advertise without identifying actual drugs (not done).

Demand.

- An individual with the given health condition maximizes utility among $j \in \{0, 1, \ldots, J\}$ where $j = 0$ is the choice not to be medicated

$$U_{ijt} = \beta_p^i p_{jt} + \beta_g g_{jt} + \beta_{m(j)} + \beta_x X_{jt} + \xi_{jt} + \varepsilon_{ijt}$$

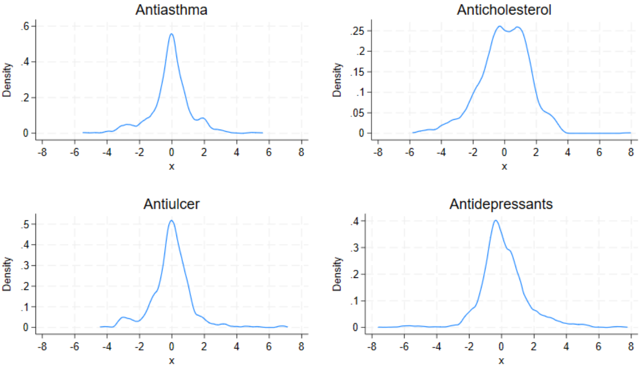
  - p_{jt} is the price of drug j at time t, & $\beta_p^i \sim N(\beta_p, \sigma_p^2)$,
  - β_m is a molecule fixed effect,
  - g_{jt} is a dummy variable indicating patent expiration,
  - X_{jt} are controls which vary by market,
  - $\xi_{j,t}$ is the quality term,
  - ε_{ijt} is an i.i.d. extreme value error.

BLP demand models.

	Anticholesterol	Antiasthma	Antidepressants	Antiulcer
Mean Price Coefficient				
Price	-2.167***	-4.292***	-1.547***	-1.860***
	(0.2440)	(0.7405)	(0.4419)	(0.2105)
Price × Generic	-0.740**	-7.477*	-4.287	
	(0.2418)	(3.2038)	(2.6022)	
Price × OTC				-7.231***
				(0.8210)
Standard Deviation Price Coefficient				
Price	1.015***	1.625***	0.550***	0.671***
	(0.1457)	(0.2790)	(0.1624)	(0.1120)
Price × Generic		4.018***	3.021*	
		(0.9874)	(1.5108)	
Controls				
Generic		1.611	0.415	2.115***
		(2.6141)	(1.0504)	(0.2590)
OTC				-0.461*
				(0.2254)
Drug fixed effects	Molecule ×Patent status	Molecule	Molecule	Product
Year dummies	✓		✓	
ATC4×Year dummies		✓		✓
N	1088	1733	1592	1602
Parameters	20	231	44	100
ATC4 categories	4	14	3	5
# products	32	58	46	47

- Need: on average 100 parameters per market: they are
  - Both μ & σ of price coefficient highly significant in all markets.
  - Price & s.e. of price also interacts with generic and OTC.
  - Dummies for; time, Generic, OTC., molecules, and
  - ATC4, or ATC4 $\otimes$ time dummies.

- Kernels $\Rightarrow$ distribution of $\xi \approx$ normal which we use in what follows.



Advertising and the Evolution of Quality.

- $\{\xi_{j,t}\}_{j,t}$ insure market shares match at t given observable determinants.

- Advertising does not enter demand directly, but only through ξ which is a controlled Markov process

$$\xi_{j,t} = \rho_\xi \xi_{j,t-1} + f(a_{Dj,t}, a_{d,j,t}) + z_{j,t}\beta_z + d_t\beta_{dt} + \mu_{j,t}$$

- The mean of the increment is a function of detailing (a_d) and DTC (a_D).

$$f(a_{dj,t}, a_{Dj,t}) = \beta_{a_d} \log(a_{dj,t}) + \beta_{a_D} \log(a_{Dj,t}) + \beta_{a_{d,D}} \log(a_{dj,t}) log(a_{Dj,t}).$$

- Tests indicate in all markets:
  - We accept $d_t \equiv 0$ (we have time dummies in demand),
  - $\beta_{a_d} \approx \beta_{a_D} \equiv \beta_a$ and highly significant,
  - $\beta_{a_{d,D}} = 0$. This implies that the cross partial between $a_{d,j,t}$ and $a_{D,j,t}$ in the profit function is positive. So, not surprisingly, when both detailing and DTC are positive, they are strategic complements.

Advertising and the Evolution of Quality.

Table: Regression of ξ for all markets

	Anticholesterol		Antiasthma		Antidepressants		Antiulcer	
$\xi_{j,t-2}$	0.915***	0.915***	0.993***	0.994***	0.979***	0.978***	0.991***	0.992***
	(0.012)	(0.013)	(0.009)	(0.009)	(0.007)	(0.007)	(0.006)	(0.006)
$\log((1+a_{dj,t-1})(1+a_{Dj,t-1}))$	0.033***	0.033***	0.015***	0.012	0.013***	0.019***	0.011**	0.008
	(0.005)	(0.009)	(0.004)	(0.007)	(0.003)	(0.005)	(0.004)	(0.006)
$\log(1+a_{dj,t-1})\log(1+a_{Dj,t-1})$		0.000		0.000		-0.001		0.001
		(0.001)		(0.001)		(0.001)		(0.001)
Generic	0.271***	0.270***	-0.089	-0.090	0.024	0.037	0.018	0.018
	(0.049)	(0.053)	(0.048)	(0.048)	(0.027)	(0.028)	(0.029)	(0.029)
OTC			0.163	0.153			-0.007	-0.008
			(0.084)	(0.088)			(0.032)	(0.032)
Constant	0.314***	0.313***	0.122**	0.110*	0.062*	0.088**	0.059*	0.054
	(0.051)	(0.063)	(0.043)	(0.051)	(0.025)	(0.030)	(0.026)	(0.028)
Adj. R-Square	0.9606	0.9605	0.9626	0.9626	0.9733	0.9732	0.9795	0.9795
N	816	816	1259	1259	1198	1198	1194	1194

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

Business Stealing Incentives.

- What happens to profits if a firm stops DTC & its competitors do not?

Set DTC=0 for only one firm and look at profit change.

	Anticholesterol		Antiasthma		Antidepressants		Antiulcer	
Net profit	Total	Δ	Total	Δ	Total	Δ	Total	Δ
Data	3,161,049		2,536,463		1,690,988		2,647,500	
Unilateral deviation	1,400,555	-55.6%	1,543,792	-39.1%	1,418,212	-16.1%	1,884,444	-28.8%

Note: 1,000 US$ per quarter.

- Profit losses are huge: on average 16-55% (using our cost estimates).
- *One Goal*: compare this to the losses that would be incurred if we banned DTC for all participants. For this we need an advertising model; *one that includes the response of detailing*.

Empirical Model of Advertising.

• $W(a|\xi_{j,t}, J_{j,t})$ is the decision maker's perception of the EDV of returns conditional on $a = (a_D, a_d)$, where $D =$DTC & $d =$detailing, so

$$W(a|\xi_{j,t}, J_{j,t}) \equiv \mathcal{E}\left[\sum_{\tau=1}^{\infty} \beta^{\tau} \pi(\cdot)_{t+\tau} \Big| \xi_{j,t}, J_{j,t}\right],$$

- $\mathcal{E}(\cdot|\cdot)$ is the agent's expectations operator (which bares no necessary relationship to actual outcomes).
- $(J_{j,t}, \xi_{j,t})$ are the set of variables the firm conditions on when making its advertising decisions ($\xi \in J$ but we want to keep track of it).
- $J_{j,t}$ Includes variables observed by the researcher ($\{w_{h,j,t}\}_{h \in \{d,D\}}$) &
- Separate serially correlated unobservables for (a_d, a_D), $\{\omega_{h,j,t}\}_{\{h \in d,D\}}$.

• Management maximizes given its perceptions: $\Rightarrow$ marginal return for $h \in \{d, D\}$ (which we set equal to one) is

$$\mathcal{E}\Big[\sum_{\tau=0}^{\infty}(\beta\rho_{\xi})^{\tau}\frac{\partial\pi(\cdot)_{t+\tau}}{\partial a_{h,t}}|J_{j,t},\xi_{j,t}\Big].$$

• Assume management knows $\partial\pi_{t-1}/\partial\xi_{t-1}$ and $\partial\xi_t/\partial a_h = \beta_{a_h}/a_h$, but *future terms depends on competitors' response* which can differ by "h"

$$\approx \left(\mathcal{E}[\frac{\partial\pi(\cdot)_t}{\partial\xi_t}|J_{j,t},\xi_{j,t}]\right)^{\theta_{1,h}}\frac{\beta_{a,h}}{a_{h,t}}\exp[\theta_{0,h}+w_{h,j,t}\beta_{w,h}+\omega_{h,j,t}],$$

- $\{w_{h,t}\}_{h\in\{d,D\}}$ observables which we can condition on &
- $\{\omega_{h,t}\}_{h\in d,D}$ unobservables which we assume are serially correlated or

$$\omega_{h,t}=\rho_{\omega,h}\omega_{h,t-1}+\nu_{h,t}, \qquad \nu_{h,t}\sim\mathcal{N}(0,\sigma_h^2).$$

Results: Variables that advertising responds to.

- Observables (i.e. in data).
  - The derivative of log profits w.r.t. advertising. The value function is an iterate of the profit function, but far too complex to compute. So using the impact of advertising on profits as a proxy makes sense.
  - Time to loss of patent exclusivity. Advertising goes down as we approach the end of the patent life.
  - Advertising of competitors matters through its impact on profits. In addition advertising of competitors in the same therapeutic (ATC4) class typically had an independent effect.

- Disturbances are highly serially correlated. $\Rightarrow$ the variables that the firm conditions on & we do not observe are highly serially correlated.

Accounting for Properties of the Data.

• Assume the firm knows the demand system, production costs, and the process generating ξ. Still have to overcome three issues.

1. We need estimates of

$$\frac{\partial \pi(\cdot)_{t-1}}{\partial \xi_{t-1}} \approx \frac{\partial D(\cdot)_{t-1}}{\partial \xi_{t-1}}(p_{t-1} - c)$$

where

- p is list prices which are known to us, and
- c is *marginal cost minus rebates* & is not known.

• We estimate c from the advertising equation & compare to estimates from the more traditional static Nash pricing assumption.

2. After starting advertising they often stop advertising for an interim period and then start up again.

- Table: Between the first (t^0) & the last (t^1) advertising period;
  - $a_h = 0$ periods; 7 to 21% for a_d, & 16 to 32% for a_D.
  - $a_h = 0$ followed by $a_h > 0$; 3 to 4% for a_d, and 2 to 4% for a_D.
  - $\Rightarrow$ there are long sequences of zero's in between positive advertising observations.

Table: Properties of the Data.

	(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	0.746	0.143	0.121	0.717	0.058	0.027	0.253	0.030
Antiulcer	1314	0.377	0.078	0.062	0.343	0.229	0.045	0.187	0.028
Anticholesterol	907	0.643	0.190	0.162	0.610	0.115	0.028	0.177	0.048
Antidepressants	1325	0.537	0.105	0.087	0.500	0.133	0.039	0.185	0.040

• To the extent that the decision to stop advertising is a function of factors we can not condition on (our $\{\omega_{h,j,t})\}_{h\in\{d,D\}}$), we need to correct for them or there will be a selection problem. We use Kuhn-Tucker conditions which allow for zero advertising.

3. The disturbance term (our $\{\omega_{h,j,t}\}_{h\in d,D}$) is serially correlated so the derivative of profits, which appears as a determinant of advertising, is correlated with it. We need an estimation algorithm which accounts for this as well as the selection correction.

Correcting for serial correlation and selection.

- Go back to the f.o.c. and take logs. The fixed costs then imply

$$(i)\ a_{h,t} = 0 \quad \text{when} \quad 0 \geq \theta_{0,h} + \theta_{1,h} \ln[\frac{\partial \pi(\cdot)_t}{\partial \xi_t}]_{|a_{h,t}=0} + w_t \beta_{w,h} + \omega_{h,t},$$

$$(ii)\ \text{if } a_{h,t} > 0 \quad \text{then} \quad \ln[a_{h,t}] = \theta_{0,h} + \theta_{1,h} \ln[\frac{\partial \pi(\cdot)_t}{\partial \xi_t}]_{|a_{h,t}} + w_t \beta_{w,h} + \omega_{h,t}.$$

- $\omega_{h,t}$ correlated with rhs variables. If we compute $a_{h,t} - \rho_h a_{h,t-1}$ the disturbance on the rhs becomes

$$\omega_{h,t} - \rho_h \omega_{h,t-1} \equiv \nu_{h,t}$$

which is not correlated with anything in the past. Then If $a_{h,t} > 0$

$$a_{h,t} = \underline{v}_{h,t}(\theta, c) + \nu_{h,t}, \quad \text{where}$$

$$\underline{v}_{h,t}(\theta, c) \equiv \theta_{0,h}(1 - \rho_h) + \theta_{1,h} \left[\ln[\frac{\partial \pi(\cdot, c)_t}{\partial \xi_t}] - \rho_h \ln[\frac{\partial \pi(\cdot, c)_{t-1}}{\partial \xi_{t-1}}] \right]$$
$$+ \beta_{w,h}[w_t - \rho_h w_{t-1}] - \rho_h a_{h,t-1}.$$

Estimation algorithm.

- **Parameter heterogeneity**
  - $\Theta \equiv \{\rho_{\omega_h}, \sigma_{\nu_h}, \theta_{0,h}, \theta_{1,h}, \beta_{w,h}\}_{h\in\{d,D\}}$ differ across markets.
  - $c \equiv \{c_j\}_j$; we estimate separate marginal costs for each product in each market.
- **Nested algorithm.**
  - *Inner loop*: for each θ find the $\{c_j\}_j$ that satisfy the average of the dynamic f.o.c. for each drug $\frac{1}{T_j}\sum_t \Big(\log[a_{h,j,t}] - \underline{v}_{h,j,t}(\theta, c_j) - E[\nu_{h,j,t}|a_{h,j,t} > 0, J_t]\Big) = 0.$
  - *Outer Loop*: NLLS searches over Θ parameters in
    - $Pr(a_{h,t=0}|J_t, \Theta, c) = P(\nu_t \leq \underline{v}_{h,t}(\theta, c)|c = c(\Theta))$, &
    - $E[a_{h,t}|a_{h,t} > 0, J_t, \Theta, c] = \underline{v}_{h,t}(\theta, c) + E[\nu_{h,t}|\nu_t \geq \underline{v}_{h,t}(\theta, c)|c = c(\Theta)]$

Results: Marginal Cost Estimates

- We did this in two ways
  - From the dynamic f.o.c. for the two types of advertising
  - From the static equilibrium condition for price.

- Results for branded drugs:
  - Dynamic f.o.c.; market average margins varied between 70 and 86 % for patented drugs.
  - Static f.o.c.: market average margins 30 to 46% for patented drugs.

- We also did this separately for generics:
  - Static f.o.c: markups are higher for generics than for branded drugs (48 to 66%), with average marginal costs sometimes negative. This is highly unlikely and convinced us to use the dynamic results.
  - There is very little advertising for generics (essentially none for DTC), so we do not use generic advertising in what follows.

Marginal Costs & Margins

Costs and Margins from Price and Detailing First Order Conditions

		Advertising Eq.		Price FOC	
	Price	Cost	Margin	Cost	Margin
Anticholesterol	2.95	0.45	0.80	6.60	0.43
Antiasthma	1.85	0.55	0.70	1.06	0.46
Antidepressants	5.49	0.91	0.86	4.00	0.30
Antiulcer	3.01	0.32	0.78	2.25	0.30

Results: Advertising Parameters (12 per market).

- Coefficients of observables.
  - Profit Derivative: highly significant and positive in all 8 equations. Larger for DTC than detailing (except Ulcer market).
  - Time to LOE: always positive; larger and more significant for DTC than detailing (except Ulcer market).
  - Coefficients of rival advertising from same ATC4. Is negative in six sometimes significant; positive near zero in two.

- Parameters of Unobservable Processes
  - Serially correlated unobserved states: significant in all 8 equations. Values above .9 except for DTC in the ulcer market (.61).
  - σ_d: between .27 and .37 & all significant.
  - σ_D vary from .09 to .55 & all significant but ulcer.

	Anticholesterol	Antiasthma	Antidepressants	Antiulcer
ρ_d	0.955**	0.962**	0.927**	0.903**
	(0.030)	(0.010)	(0.014)	(0.014)
$\theta_{1,d}$	0.499**	0.211**	0.474**	0.479**
	(0.161)	(0.057)	(0.000)	(0.096)
β_d	-0.017	-0.025*	-0.009	0.013
	(0.024)	(0.013)	(0.016)	(0.013)
$\beta_{2,d}$	0.005	0.027	0.024	0.076**
	(0.074)	(0.018)	(0.029)	(0.019)
$\theta_{0,d}$	10.413**	9.373**	9.382**	9.453**
	(1.524)	(0.532)	(0.405)	(0.504)
ρ_D	0.915**	0.971**	0.977**	0.605**
	(0.024)	(0.029)	(0.013)	(0.027)
$\theta_{1,D}$	2.350**	0.909**	0.677**	0.298*
	(0.166)	(0.364)	(0.186)	(0.174)
β_D	-0.060**	-0.043	0.011	-0.007
	(0.027)	(0.029)	(0.026)	(0.025)
$\beta_{2,D}$	0.196**	0.051	0.440**	0.043**
	(0.084)	(0.135)	(0.199)	(0.012)
$\theta_{0,D}$	12.140**	11.548**	8.637	9.465**
	(0.860)	(0.775)	(6.131)	(0.388)
σ_d	0.284**	0.268**	0.302**	0.368**
	(0.029)	(0.024)	(0.030)	(0.038)
σ_D	0.549**	0.185**	0.319**	0.089
	(0.109)	(0.059)	(0.072)	(0.057)

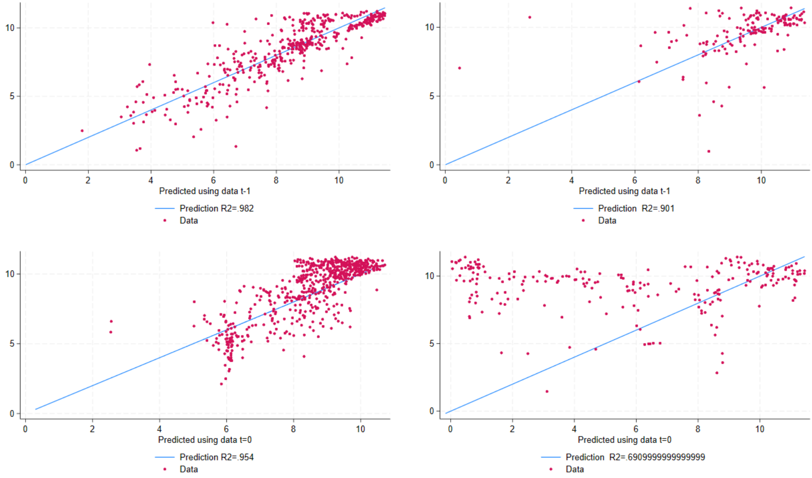
Fit of the Empirical Model.

• Four panels plot actual advertising (the y-axis) when advertising is positive against the model's predictions for positive advertising. The difference is due to the impact of randomness in the advertising data & estimation error in the empirical model. Fit measures below.

• Top two panels: condition on the actual value of the variables in prior period: i.e. condition on what our model predicts the firm knows when it makes its advertising decisions other than the advertising disturbance. The data should be randomly distributed around the 45 degree line and given the serial correlation the fit should be good.

• Bottom two panels: predict advertising using estimated model that conditions only on what is known in initial period. The data in later periods is a non-linear function of the randomness in past advertising, and our estimates need not be symmetrically distributed around 45 degree line, Also the fit should deteriorate as we are predicting up to 38 periods out.

Results: Fit of the Empirical Model.



Note: Detailing on the left, DTC on the right. Using data t-1 for top graphs and only data at t=0 for bottom graphs.

EBE With Estimated Primitives.

- *Experience Based Equilibrium (Fershtman and Pakes, 2012):*
  1. As in empirical model: firms chose policies that maximize their perceptions of EDVs conditional on the variables they use to determine their advertising expenditures, and
  2. **New condition**: perceptions are consistent with outcomes at states visited repeatedly by the Markov process generated by the policies (on the "recurrent" class).

- *In Sample Analysis of Equilibrium:*
  - Using estimated demand function, costs, and stochastic processes, develop a computational algorithm to compute fixed point for EBE policies for both DTC and detailing.
  - These will differ from the empirical model's policies. Compare the EBE policies to the empirical model's policies both (i) visually and (ii) in terms of fit.

Multiplicity.

• Markov Perfect Bayes is a special case of the equilibrium used here & has multiplicity, so there will be multiplicity here. The additional equilibria generated by EBE is because there can be different perceptions of values outside the recurrent class that may never be corrected, and can lead to different recurrent classes (interior and boundary points).

• We use two conditions to limit the equilibria that the algorithm can generate.

- We start the computation at the initial state in the data (history limits the possible equilibria).
- We use the boundary consistency refinement in Asker (et. al. 2023). Boundary perceptions forced to be consistent with current perceptions of the value of future streams of profits outside of the recurrent class.

Algorithm: Overview

- Problems in computing fixed point for dynamic market equilibria.

1. Dimension of state space (how many points do we need policies for).
   - Do not compute policies for all possible states. Only compute separate policies for each product in each time period (small number of points). This is "pointwise": i.e. does not discretize the states or use approximations.
2. Computing continuation values (dimension of summation for integral).
   - Substitute simulation for integration. Average of simulated sequences of continuation values. 100 sequences, small number of draws on on random variables that must be held constant across iterations. Can use last period's investment measures post sample incentives (so post sample is firm's perceptions).

- Algorithm starts at initial point in the data. It assumes each point visited is in the recurrent class of a boundary consistent equilibrium thereafter.

Algorithm Details.

• Algorithm is iterative. Iterations are indexed by l & associated with policy functions. Initial estimate is the empirical model's policy function.

• Use policy functions in memory to simulate K sample paths for $\{\pi^{k,l}(\cdot)_{j,t}\}$ and $\{\omega^{k,l}_{h,j,t}\}$ for each firm, keeping the underlying random draws of (μ, u, ν) the same over iterations (otherwise it will not converge).

• Compute the averages over the sample paths of $\{\pi^{k,l}(\cdot)_{j,t}\}$ and $\{\omega^{k,l}_{h,j,t}\}$ say $\{\pi^{l}(\cdot)\}_{j,t}$, & $\{\omega^{l}_{h,j,t}\}_{h,j,t}$. Compute the average discounted value $W^{l}_{j,t} \equiv \sum_{\tau=1}^{\infty} \beta^{\tau} \pi^{l}(\cdot)_{j,t+\tau}$.

• Use these averages to estimate the marginal returns to a_h as

$$\sum_{\tau=t}^{\infty} (\beta\rho)^{\tau} \frac{\partial \pi^{l}(\cdot)_{t+\tau}}{\partial \xi_{t+1}} \frac{\beta_{a,h}}{a_{h,t}} \approx \mathcal{E}\Big[\sum_{\tau=t}^{\infty} (\beta\rho)^{\tau} \frac{\partial \pi(\cdot)_{t+\tau}}{\partial \xi_{t+1}} \frac{\beta_{a,h}}{a_{h,t}} | J_t, \xi_t\Big].$$

- Terminal period's contribution: Use the last observed advertising, say T_j, to approximate for $\sum_{\tau=T_j}^{\infty}(\beta\rho)^{\tau}\frac{\partial\pi(\cdot)_{t+\tau}}{\partial\xi_{t+1}}\beta_{a,h}$ (eliminates need to assume firm's perceptions about post sample process is the same as in sample).

- Update the policy function. Regress the dependent variable

$$\log\left[\sum_{\tau=t}^{\infty}(\beta\rho)^{\tau}\frac{\partial\pi^{l}(\cdot)_{t+\tau}}{\partial\xi_{t+1}}\beta_{a,h}\right]-\omega_{h,t}^{l}$$

on

$$\approx\theta_{0,h}^{l+1}+\theta_{1,h}^{l+1}\log\left(\frac{\partial\pi^{l}(\xi',J')}{\partial\xi'}\right)+w_t\beta_{w,h}^{l+1}$$

- Finally compute

$$Y^{l}\equiv\frac{1}{\sum_{j,t}1}\sum_{j,t}\left(\frac{W_{j,t}^{l}-W_{j,t}^{l-1}}{W_{j,t}^{l-1}}\right)^{2}.$$

- If $Y^{l}<10^{-5}$ stop. If $Y^{l}>10^{-5}$ proceed to iteration $l+1$.

EBE Parameter Estimates (s.e. not corrected)

	Asthma	Cholesterol	Depressants	Ulcer
$\theta^{l}_{1,d}$	.8119	.8780	.5874	.7460
	(.0007)	(.0014)	(.0008)	(.0009)
$\beta^{l}_{w,d}$ (Other ad)	.1396	.0076	.2149	-.042
	(.0005)	(.0029)	(.0006)	(.0030)
$\beta^{l}_{w,d2}$ (time to loe)	.0486	.0790	.0737	.0738
	(.0000)	(.0000)	(.0000)	(.0000)
$\theta^{l}_{0,d}$	-4.24	-4.19	-3.57	-3.20
	(.0612)	(.2986)	(.0466)	(.1865)
$\theta^{l}_{1,D}$	.4169	.4739	.3812	.3300
	(.0010)	(.0019)	(.0006)	(.0006)
$\beta^{l}_{w,D}$ (Other ad)	-.010	.0026	-.034	-.199
	(.0004)	(.0008)	(.0004)	(.0008)
$\beta^{l}_{w,D2}$ (time to loe)	.0229	.0531	.0711	.0500
	(.0000)	(.0000)	(.0000)	(.0000)
$\theta^{l}_{0,D}$	-2.95	-3.26	-2.68	-1.52
	(.0830)	(.1699)	(.0403)	(.0665)

$$\log(a)^{l}_{h,j,t} - \omega^{l}_{h,j,t} = \theta^{l+1}_{0,h} + \theta^{l+1}_{1,h} log\left(\frac{\partial \pi^{l}(\xi', J')_t}{\partial \xi'}\right) + w_t \beta^{l+1}_{w,h} + e_{h,j,t}.$$

Results: EBE Parameter Estimates.

• In equilibrium in both detailing and DTC and in all 4 industries (we don't have standard errors corrected for first stage estimates yet):

- the derivative of log profits positive and larger in detailing than DTC (opposite of empirical model),
- the time to l.o.e. is positive everywhere, and larger in detailing than DTC (also opposite empirical model), and
- the sign of the impact of advertising of others in the same therapeutic class differs by market and advertising type.

• The magnitudes of the coefficients differ quite a bit from those in the empirical models.

Compare: EBE to Empirical Model & to the Data.

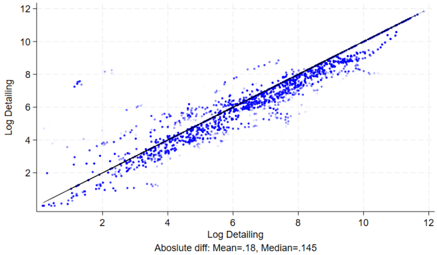
- First visual results then measures of fit.
  - EBE vs Empirical model
  - EBE vs Data.

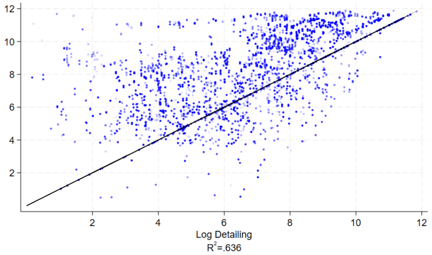
- All figures are for all markets. We condition on initial advertising values and simulate sequences of 37 periods thereafter.
  - We insured the random draws used to compute the empirical and the EBE policies are identical; $\Rightarrow$ differences due solely to policy functions.
  - Color intensity decreases for farther out predictions.

- Keep in mind that
  - The empirical model is derived solely from the data.
  - The EBE imposes additional restrictions from theory that the data do not necessarily abide by.

Results: EBE vs (i) Empirical and (ii) vs Data

Detailing

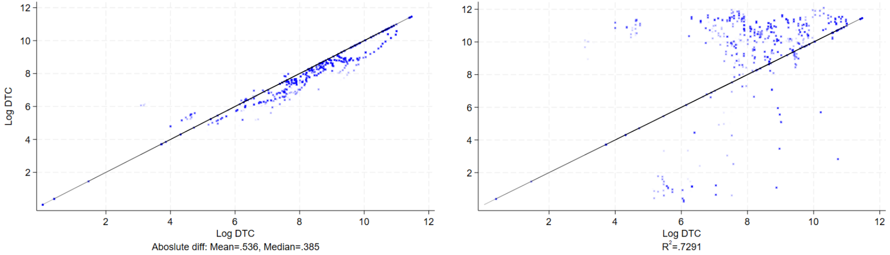




- Left panel: empirical model vs EBE. Right panel: Data vs EBE.

- EBE is noticeably different from empirical model, but mimics the data about as well.

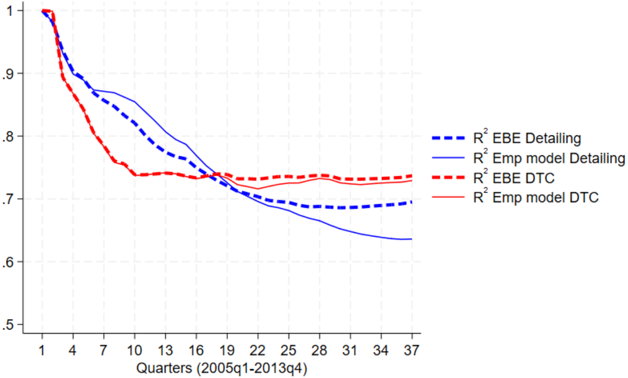
DTC.



- Left panel: Empirical model vs EBE. Right panel: Data vs EBE.

- EBE is noticeably different from empirical model, but mimics the data about as well.

Results: R^2 for different prediction horizons.



- Necessarily start similarly, but the restrictions imposed by theory help as we move to longer run predictions (especially in detailing).
- These are quite good fits for micro data; in the 37th period the EBE prediction for DTC has an $R^2 \approx .75$ and for detailing $\approx .7$.

Counterfactuals: Additional Assumptions.

- Assumptions used when evaluating counterfactuals.
  - We are assuming the counterfactual world is in equilibrium, and the underlying processes for unobservables do not change in the counterfactual world.
  - We present a selected equilibria. The equilibrium we selected
    - uses the states of the firm in the initial period as an initial condition. So we are looking at what would have happened in the in-sample period, were the counterfactual institutions in place, and
    - considers only equilibria that are boundary consistent.
  - Alternative not used: assume a learning theory (as in Fershtman and Pakes, 2012) and use it to evaluate counterfactuals; this would enable us to simulate many equilibria.

Remaining Counterfactuals Issues.

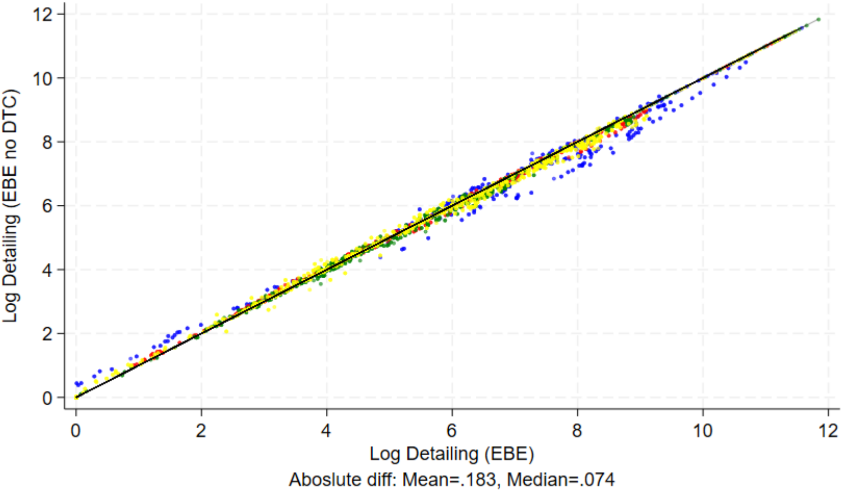
- Just started work here and have only done the no DTC counterfactual.

- These runs use current prices. We have the programs running for a counterfactual which sets prices to maximize the sum of current profits over the time period in the data (but the French take a holiday in August).
  - This requires adding a module to the algorithm which computes a fixed point for prices at every iteration and then feeds it back into the algorithm already described.
  - So we are computing a fixed point when there are three controls for each drug when we allows for DTC and two when we omit DTC.

- We would also like to compute a counterfactual
  - for a regulation which bans advertising for a named brand, but directs the FDA to advertise the availability of treatments for a given disease.
  - This would be funded by a sales tax on the relevant drugs.

What happens to detailing when we ban DTC?

Equilibrium Detailing if DTC banned (y-axis) versus when DTC allowed.



Detailing & Profits; with & without DTC.

Detailing, DTC, net profit (in 1000's quarterly)

	Anticholesterol		Antiasthma		Antidepressants		Antiulcer	
All products	Total	% Δ	Total	% Δ	Total	% Δ	Total	% Δ
Detailing	68,827		66,028		59,118		33,927	
DTC	37,523		24,874		14,008		16,429	
Net profit	1,948,735		2,460,545		1,559,511		2,003,420	
Detailing (no DTC)	59,638	-13.3%	62,844	-4.8%	57,852	-2.1%	32,412	-4.4%
Net profit (no DTC)	1,604,859	-17.6%	2,203,779	-10.4%	1,435,168	-7.9%	1,804,020	-9.9%

- Profits fall in every market, but
  - by more in anti-cholesterol and less in anti-depressants (perhaps because high cholesterol is a less salient condition),
  - this mimics the ordering across markets for business stealing (unilateral deviations) except now the losses are about half as large.

• Detailing falls in all markets. However not in all firms (look to quantiles of the distribution).

• Firms who did little or no detailing when DTC was allowed often increase their detailing once DTC is not allowed.

• If DTC no longer induces patients to ask for a particular product, products that did not find it profitable to send agents to doctors when there was DTC, might find it profitable to do so when DTC is not allowed.

• We are checking whether the firms that own the drugs that lose disproportionately are also the firms that do disproportionate R&D. This is likely, and if pharma research is as welfare improving as other results indicate it is, "leveling the profit field" may not be what we want to do; at least not without a mitigating pro research policy.

Detailing & Profits; with & without DTC.

Anticholesterol	10%	25%	50%	75%	90%
Log detailing	0	0	6.05	7.79	8.88
Log DTC	0	0	0	0	7.94
Net profit	365	2,983	22,071	71,810	215,626
Log detailing (no DTC)	0	0	6.06	7.42	8.64
Net profit (no DTC)	448	3,145	20,109	53,009	149,947
Antiasthma					
Log detailing	0	1.05	5.02	7.18	8.16
Log DTC	0	0	0	0	7.14
Net profit	375	1,582	10,466	66,868	220,442
Log detailing (no DTC)	0	1.18	5.11	7.05	8.01
Net profit (no DTC)	356	1,642	11,221	64,837	184,298
Antidepressants					
Log detailing	0	0	3.59	6.11	7.53
Log DTC	0	0	0	0	5.24
Net profit	226	1,399	5,684	29,155	86,924
Log detailing (no DTC)	0	0	3.61	6	7.5
Net profit (no DTC)	226	1,450	5,747	28,412	84,972
Antiulcer					
Log detailing	0	0	0	5.25	7.1
Log DTC	0	0	0	0	0
Net profit	176	952	5,022	26,214	121,707
Log detailing (no DTC)	0	0	0	5.23	7.08
Net profit (no DTC)	153	1,043	5,293	25,554	118,201

Changes in consumer take-up.

• A direct indicator of the impact on the fraction accessing the drugs to alleviate their conditions is the “inside good share”. We find that it falls noticeably for all markets but Asthma (where it goes down, but only by about 2%, perhaps because Asthma is so salient).

• We are looking to buttress this with additional data because;

- data on the outside good share is indirect (based on an NIH survey)
- and some of the fall in inside share may be a fall in those using it but not benefiting (the ”excessive use” argument). However
    - there is evidence that too few Americans are taking these drugs (especially for cholesterol).
- We are trying to access data that would enable us to analyze the impact of DTCA by geography, income, and educational group. This should shed more direct light on which consumers DTC impacts.

• The loss in profit and the loss in the inside goods share indicate that if we do ban DTC we might want to replace it with some other system to:

1. Maintain inside good share. This is what we hope our second counterfactual will do (advertising without brand names).
2. To induce more R&D expenditures (sales tax to support non-branded advertising would likely decreases profits). How to do this is a topic of ongoing research (by many others, as well as by us).

That is all we have for now.

Thanx for listening!