Demand Inertia and the Hidden Impact of Pharmacy Benefit Managers*

Josh Feng, Luca Maini†

February 28, 2023

Abstract

Do Pharmacy Benefit Managers (PBMs) reduce spending on prescription drugs? Reduced-form evidence suggests that PBMs enforce a tradeoff between net-of-rebate prices and access to drugs within each market. However, net-of-rebate prices grow consistently over time and appear unresponsive to competitor entry. We argue that inertia in drug demand can reconcile these facts. To formally analyze the roles played by PBMs and demand inertia, we build a dynamic structural model of drug pricing and estimate it using net-of-rebate prices of three major statins from 1996–2013. Counterfactuals suggest that, relative to a market with price-setting by drug manufacturers and patients who face coinsurance, PBMs reduce overall spending by 28 percent, without greatly limiting patient access. Without demand inertia, the presence of PBMs would cause prices to fall significantly as competitors enter.

*Feng would like to acknowledge his PhD advisers Ariel Pakes, David Cutler, and Josh Lerner. We would also like to thank Ernie Berndt, Jen Brown, Ramon Caminal, Lauren Cohen, Brigham Frandsen, Edward Glaeser, Bart Hamilton, Matt Higgins, Robin Lee, Alex Mackay, Chris Metcalf, Anders Munk-Nielsen, David Ridley, David Schmidt, Brett Wendling, and Heidi Williams for providing detailed comments. Next, we thank several industry contacts for providing institutional knowledge and data: Ryan Baum, Richard Evans, and Scott Hinds at SSR Health for providing data and suggestions on estimating negotiated prices, and Bob Nease and Edmund Pezalla for insights into PBMs. We also thank conference and seminar participants at ASHEcon, Bates White, Duke Fuqua, the Federal Trade Commission, Harvard University, Harvard Business School, IIOC, NBER Productivity Lunch, and UBC Sauder. We benefited from the computing resources at the Research Computing Environment, supported by the Institute for Quantitative Social Science in the Faculty of Arts and Sciences at Harvard University. This research was supported by the National Institute on Aging, Grant Number T32-AG000186 and Grant Number R24-AG048059, and by the National Institute for Heath Care Management.

†Feng: Department of Entrepreneurship and Strategy, David Eccles School of Business, University of Utah; email: josh.feng@eccles.utah.edu. Maini: Department of Health Care Policy, Harvard Medical School; email: maini@hcp.med.harvard.edu

What is the best approach to controlling prescription drug prices? Branded drug manufacturers enjoy significant pricing power in U.S. markets, thanks to patent protection, price-insensitive demand, and consumer inertia. To compensate for this imbalance in market power, US drug prices are often negotiated by intermediaries called Pharmacy Benefit Managers (PBMs), large entities that design drug benefits on behalf of almost 80 percent of all insurance enrollees across the commercial market (e.g., employers, unions, and large insurers), Medicare Part D, and the ACA health exchanges.[1] PBMs claim that they help control spending and improve patient welfare, but critics argue that they have done very little to slow the rapid price growth of branded drugs.[2] To date, empirical research on PBMs is scarce, and the role these intermediaries play in moderating drug prices remains poorly understood.

In this article, we shed light on the role of PBMs and argue that their effect on branded drug prices is masked by the presence of inertia in drug demand. Inertia has two theoretical effects. First, it means that manufacturers want to increase prices over time, as their focus gradually shifts from building up market share to extracting higher markups from existing customers. Second, it limits the impact of later entrants by putting them at a disadvantage when competing for existing patients. As a result, inertia reduces competition and conceals the true impact of PBMs, making them appear ineffective at controlling drug prices.

To carry out our analysis of PBMs, we bring together data on net prices, formulary coverage, and individual patient claims. Net prices, unlike the list prices typically cited in the media, incorporate price concessions that manufacturers make to PBMs. Our net price data comes from SSR Health and covers over 1000 branded drugs owned by publicly traded companies between 2007–2020. The data is at the market level, and reflects an average across all PBMs. Formulary coverage determines the out-of-pocket cost patients would need to pay for each drug. Formularies are designed as tiered menus. Most formularies contain at least two tiers for branded drugs: a low-cost "preferred" tier, and a higher-cost "non-preferred" tier. Drugs can also occasionally be excluded from a formulary altogether, in which case patients must pay full price in order to purchase a prescription. Our formulary data comes from MMIT Analytics and contains information on the formulary placement of over 600 branded drugs from 2011–2020. Net price and formulary placement are the two key outcomes that arise from the negotiations between

[1] Numbers are based on authors' own calculations using MMIT data and are discussed in greater detail in Section 1.

[2] See Financial Times article, "Trump is losing the war on drug prices" (July 31, 2018), for an overview of the debate, including the opinions of former HHS Secretary Alex Azar: https://www.ft.com/content/2b57448e-94a4-11e8-b747-fb1e803ee64e (retrieved February 17, 2023).

manufacturers and PBMs. PBMs extract lower prices by letting manufacturers compete for better formulary placement. Finally, for our structural analysis, we use individual-level prescription claims data from MarketScan and the Medical Expenditure Panel Survey (MEPS) from 1996–2013. MarketScan tracks patients' choices over time and includes plan identifiers, allowing us to estimate a switching cost demand system, while MEPS collects data from a representative sample of consumers and therefore better reflects aggregate market shares.

We begin our analysis by documenting an apparent contradiction at the core of the debate over the impact of PBMs. On the one hand, drugs with higher prices receive worse coverage, suggesting that PBMs enforce a tradeoff between price and access. On the other hand, net prices and coverage generosity (as measured by placement on preferred formulary tiers) do not fall significantly after the entry of close therapeutic substitutes. These two findings are hard to reconcile. If PBMs are effective at enforcing a coverage-price tradeoff, why are they unable to leverage an increase in competition into greater discounts?

We argue that inertia in drug demand can produce equilibrium outcomes consistent with these patterns. The vast majority of drugs are prescribed for an extended period of time to treat chronic conditions, and several papers show that patients rarely switch between therapeutic substitutes (see e.g., Crawford and Shum, 2005; Sinkinson and Starc, 2018; Lee, 2016; Feng, 2023).[3] Demand inertia leads to market segmentation and increasing equilibrium price paths (e.g., Klemperer, 1987b,a; Dube et al., 2009).[4] The market segmentation effect dampens the ability of PBMs to leverage new market entrants to extract higher discounts. Moreover, although net prices rarely fall over time, PBMs could potentially be preventing higher rates of price growth.

To formally account for PBMs and demand inertia, we develop a dynamic structural model of drug prices. In the model, drug manufacturers and a representative PBM play a dynamic pricing game where each stage takes the form of a Stackelberg-Nash game (Sudhir, 2001; Besanko et al., 2003). Drug manufacturers submit prices to the PBM, which in turn assigns drugs to either one of two cost sharing tiers or excludes the drug from coverage. We model PBM behavior using a flexible control function that incorporates

[3] Alpert (2016) uses a data driven approach to estimate that approximately 88 percent of prescriptions are for maintenance use in the treatment of chronic conditions. Substitution between therapeutic substitutes refers to switches between different molecules. This definition excludes switches from branded to generic versions of the same product, which generally do not suffer from inertia because generic substitution laws allow pharmacists to substitute branded drugs with their generic version at point of sale.

[4] In markets with very similar products, low switching costs can actually lead to lower prices (Doganoglu, 2010; Cabral, 2016). These conditions are unlikely to be satisfied in the pharmaceutical market, where products are heterogeneous and the switching costs we estimate are quite large relative to product quality.

coverage generosity and expected spending. Consumers then make drug choice decisions given the insurance design. The game unfolds over a finite number of periods, ending when all drugs lose patent protection.

We estimate the model using claims data and net prices on the market for medium-intensity statins from 1996 to 2013. We capture inertia through a switching cost model of demand, using quasi-exogenous variation in choice sets to separate switching cost from unobserved heterogeneity in patient-drug match quality. We then estimate the parameters of the PBM control function by matching predicted prices to observed average net prices.

After estimating the model, we run a series of simulations to show how the current equilibrium compares to a counterfactual world where manufacturers set prices unilaterally. In our main counterfactual, we assume health plans cover drugs at a uniform flat coinsurance rate, which we set at 33 percent to match the ratio of net prices to copays we observe in the data. We find that the total cost of statins would increase by almost 50 percent. Even after accounting for potential PBM fees, we find that total spending is almost 40 percent higher.[5] PBMs not only lower spending, but they also generate better insurance contracts from a patient's perspective. Cost sharing for drugs increases significantly in our counterfactual, leading to worse access to medication and lower patient welfare.

We also simulate counterfactual outcomes under alternative formulary structures to highlight the mechanisms through which PBMs achieve spending reductions. We find that altering the tier structure (e.g., single-tier vs. multi-tier formularies) has a small impact on equilibrium outcomes, likely because consumers are not very sensitive to cost sharing. Instead, the most effective tool available to PBMs is the threat of exclusion. Eliminating the ability of PBMs to exclude multiple drugs leads to a substantial spending increase. Crucially, the impact on patient access is negligible, because drugs are rarely excluded in equilibrium even when exclusion is possible. However, the threat of exclusion creates downward pressure on price even when not acted upon.

Finally, we simulate equilibrium prices in a market with no demand inertia. To do so, we eliminate switching costs from the demand system. Under this hypothetical scenario, drug manufacturers price higher at the beginning of the life-cycle, but reduce prices significantly whenever a (branded or generic) competitor enters the market.

Our article contributes to two strands of literature on drug pricing. First, it is part of a small but growing literature that engages with the role of PBMs and proposes more so-

[5] We use PBM financial filings and a back-of-the-envelope calculation to estimate PBM fees. While the figures we obtain are likely to be noisy, they have a small impact on our overall results. PBMs would need to generate a 47 percent profit margin (versus the 6.5 percent margin we estimate using financial filings) in order to reverse the result that they reduce overall spending.

phisticated models of drug price-setting.[6] Papers in this space include Conti et al. (2021), which proposes a theoretical model of drug pricing that incorporates PBMs, and Olssen and Demirer (2023), which studies how Part D plans cover statins. Our paper stands out in this space by being the first to empirically estimate a model that explicitly incorporates PBMs.[7] Second, we contribute to the literature on strategic formulary setting (Lavetti and Simon, 2018; Geruso et al., 2019). Instead of analyzing within-drug variation in coverage across market segments to speak to the effects of policy or payer incentives on coverage, we analyze and model the role of net prices in driving average formulary outcomes within drug markets.

More broadly, this article contributes to the literatures on vertical relationships in health care markets (see, e.g., Ho and Lee, 2017, 2019), by focusing on the repeated interaction nature of these relationships and incorporating dynamics in demand. Dynamic firm incentives could be important in a number of health care bargaining settings given evidence on inertia in hospital demand (Raval and Rosenbaum, 2018) and demand for insurance (Handel, 2013; Ericson, 2014; Fleitas, 2016; Ho et al., 2017). Our results also speak to the literature on pricing in markets with demand inertia (Klemperer, 1987b,a; Dube et al., 2009; Honka, 2014; Shcherbakov, 2016). Our analysis evaluates outcomes in a market with a more complex pricing structure, and may be relevant in other settings, such as retail, where supermarkets and wholesale clubs leverage shelf space and their ability to exclude items to extract discounts from suppliers.

The article is organized as follows. Section 1 discusses drug demand, the PBM industry, and data sources. Section 2 documents stylized facts on equilibrium outcomes across drug markets. Section 3 presents a structural model of drug pricing. Section 4 presents model estimates. Section 5 provides counterfactuals. Section 6 concludes.

1 The Role of Pharmacy Benefit Managers in Prescription Drug Markets

Pharmacy Benefit Managers (PBMs) are intermediaries that negotiate the price of most drugs on behalf of large payers, organizations such as insurance companies and employ-

[6] Most paper that model US drug prices assume manufacturers set prices directly (see, e.g., Dunn, 2012; Arcidiacono et al., 2013; Bokhari and Fournier, 2013; Dubois et al., 2019). A few papers have used bargaining models between manufacturers and governments to describe price-setting for drugs outside the US (see, e.g., Dubois and Saethre, 2016; Dubois et al., 2019)

[7] Brot-Goldberg et al. (2022) provides a thorough overview of the state of the literature on PBMs. A related paper that does not mention PBMs explicitly is Dafny et al. (2022), which uses a bargaining model between a manufacturer and an insurer to model price negotiation in the market for MS drugs.

ers that pay for health care services . PBMs emerged in the late 1960s as the FDA approved a growing number of drugs, and pharmaceuticals quickly became a sizable component of healthcare spending. Unfamiliar with prescription drugs, many health insurers opted to outsource drug benefits management to PBMs. Today, PBMs manage prescription drugs for about 80 percent of insured individuals across the commercial market (large insurers, employers, unions), Medicare Part D, Managed Medicaid, and Affordable Care Act health exchanges.[8] PBMs negotiate prices by letting manufacturers compete for favorable placement on drug formularies.

Formularies are tiered menus of drugs that determine the out-of-pocket cost patients pay at the pharmacy. The number of tiers can vary widely across formularies, but they usually contain at least three: the lowest and cheapest tier, usually reserved for generic drugs; a middle tier for preferred brand drugs; and the highest tier for non-preferred brand drugs. PBMs can also exclude a drug from a formulary. When that happens, patients generally need to pay the full price of the drug.

Prices and formulary tiers are determined jointly through a bidding process typically conducted separately by therapeutic class. [9] For each therapeutic class, manufacturers send grids of tier-contingent bids for their products. A bid consists of a discount rate off of a list price that can depend on a drug's own formulary tier and those of its therapeutic substitutes. After receiving the bids, PBMs combine them to construct formularies that they then offer to payers.[10]

PBMs then submit the formularies they create to payers, which in turn evaluate them based on cost (as determined by the negotiated prices), and generosity of formulary coverage (including considerations such as member disruptions from drug exclusions.[11]A few other factors beyond cost and formulary coverage can influence the evaluation of a proposal. PBMs provide some additional services, such as processing pharmaceutical claims for reimbursement. Many PBMs also own and operate mail-order pharmacies that ship medications to patients. Mail-order prescriptions usually save money for the payer because they are cheaper than prescriptions supplied through retail pharmacies. Conversely, unlike intermediaries in other markets, PBMs do not play much of an informational, search, or liquidity role.

[8] Appendix A.1 provides more details on these calculations.

[9] The definitions of therapeutic classes can differ slightly across PBMs, but they are generally confined to products that are substitutable for one another.

[10] We provide a more detailed discussion of this process in Appendix A.2.

[11] Disclosures from several state health plans describe this process in detail. See, e.g., the 2016 North Carolina State Health Plan PBM procurement process document: https://files.nc.gov/ncshp/documents/board-of-trustees/Contract-Approval-PBM-Services-BOT-Presentation-Redacted_Redacted.pdf, retrieved January 10, 2023. Documents from other state health plans are also available online.

While PBMs can extract lower prices by aggregating the demand of many payers, they retain some of the savings they generate, which makes it hard to assess their overall effectiveness. Self-reported numbers from a few large PBMs suggest that they retain 10–15 percent of savings off the list price, though this number is hard to verify.[12] Concerns over PBM margins have increased following a wave of recent mergers.[13] However, while the concentration in the market is very high, it has remained relatively stable over the past decade.[14] More concerning is that throughout the past twenty years, both drug manufacturers and pharmacy chains have periodically owned large PBMs. Of particular relevance to our structural analysis is the integration of Medco and Merck that lasted until 2003, which we discuss in greater detail in Section 4.

2 Descriptive Evidence on Prescription Drug Market Outcomes

We begin our analysis by looking at the correlation between drug prices and formulary coverage, and the impact of competition on prices.

We use data from two sources. Our data on net prices comes from SSR Health, a company that collects drug-level net-of-rebates revenues from financial filings of publicly-traded manufacturers. The SSR data contains over 1000 branded drugs from 2007 to 2019. We also obtain data on formulary coverage from MMIT Analytics, which collects formularies from all US health plans. MMIT data covers the years from 2011 through 2021. MMIT records the tier of each drug for each plan (preferred tier, non-preferred tier, excluded from the formulary). We use this disaggregated data to compute average coverage generosity statistics at the drug-year level, weighting by the number of enrollees for each plan. The MMIT sample includes 543 drugs from the top 29 disease areas by average annual sales.

Table 1 provides summary statistics on price growth, formulary coverage generosity measures, and the competitive environment faced by drugs in our sample. In the full SSR

[12]This figure comes from an Express Scripts executive interviewed for a Wall Street Journal article, "Drugmakers Point Finger at Middlemen for Rising Drug Prices," available at https://www.wsj.com/articles/drugmakers-point-finger-at-middlemen-for-rising-drug-prices-1475443336, retrieved February 17, 2023.

[13]The three largest PBMs have all acquired a rival PBMs in the recent past: Express Scripts merged with Medco in 2012, OptumRx merged with Catamaran in 2015, and CVS merged with Aetna in 2018.

[14]In 2020, the three largest PBMs (CVS Health, Express Scripts, and Optum Rx) managed 77% of all prescription claims (see the numbers at https://www.drugchannels.net/2021/04/the-top-pharmacy-benefit-managers-pbms.html, retrieved Jan. 9th, 2023). Appendix A.1 provides some numbers on PBM market shares and concentration measures.

Table 1: Summary Statistics for Price and Formulary Data

	Mean	Median	SD	N
Annual WAC growth	0.09	0.07	0.20	7,542
Annual net price growth	0.04	0.03	0.61	5,913
Percent Covered	0.87	0.89	0.11	3,421
Percent Preferred	0.23	0.15	0.21	3,421
Number of Same ATC-4 Competitors	2.34	2.00	2.86	6,783
Same ATC-4 Entry	0.12	0.00	0.32	6,783
Same ATC-4 LOE	0.07	0.00	0.25	6,783

Notes: drug-by-year level summary statistics for the core dataset. WAC refers to wholesale acquisition cost. Competitors are defined as the number of other manufacturers that own drugs in the same four-digit Anatomical Therapeutic Chemical classification (ATC-4) class. LOE refers to competitors losing exclusivity (generic versions of therapeutic substitutes enter). WAC and net price statistics are based on the full SSR Health sample, whereas the coverage statistics are based on the SSR-MMIT merged sample. The reduced data points for net price and competition are due to missing data and incomplete ATC classifications.

sample, we find that list price (measured by Wholesale Acquisition Cost, or WAC) grows at about 9 percent on average over the 2007–2020 period, while net price (which incorporate rebates) grows at 4 percent. Both measures confirm the overarching impression that drug prices increase over time.

We measure the coverage rate of each drug as the fraction of patients whose formulary covers it. A drug is, on average, included in formularies for about 90 percent of insured patients. However, PBMs give a drug, on average, preferred coverage status for less than a quarter of the insured population. The presence of exclusion and the relative narrowness of preferred tiers suggests that PBMs are willing to reduce or even deny coverage of certain prescription drugs.

Finally, we find that the average drug faces two branded competitors, faces entry of a new branded competitor once every eight years, and entry of a (non-own) generic competitor every fourteen years. We define classes of competitors using the first four digits of the Anatomical Therapeutic Chemical classification (ATC-4).[15] The classification roughly corresponds to the equivalence classes used by PBMs in their bidding process.

[15]The ATC system classifies molecules according to the organ on which they act and their therapeutic and chemical properties.

2.1 Cross-Sectional Patterns in Prices and Formulary Coverage

To provide suggestive evidence on whether PBMs use formulary position to extract larger discounts, we begin by analyzing the relationship between net prices and formulary coverage within drug classes. Our sample for this analysis consists of the 2011–2019 panel of 543 drugs for which we have both formulary and price data.

Formally, we run the following regression

$$y_{jt} = \psi_1 \log\left(WAC_{jt}\right) + \psi_2 \log\left(net_{jt}\right) + \upsilon_{k(j)t} + \epsilon_{jt} \tag{1}$$

where j indexes drugs, t indexes the year, and $k(j)$ indicates the therapeutic class of drug j. υ_{kt} are ATC-4-by-year fixed effects.[16] As outcomes of interest y_{jt}, we use the fraction of covered lives with any access and with preferred access. These come from the MMIT Analytics data. The predictors on the right-hand side are list price (measured by Wholesale Acquisition Cost, or WAC) and net price. We include the list price for two reasons. First, if the cross-sectional differences in per-unit price are driven by differences in how manufacturers define the size of units (e.g., number of milligrams) across drugs, adding list price would account for scale issues. Second, we want to separately test whether PBMs offer better coverage to drugs with higher list prices, all else equal. We restrict our analysis to branded drugs in the years when they have market exclusivity.

Results suggest that PBMs enforce a net price and formulary tradeoff within drug markets. Table 2 presents estimates showing that lower net prices lead to better formulary coverage. All formulary coverage generosity measures are negatively correlated with net price, with preferred coverage exhibiting the strongest effects. After adding list price as a predictor, we find that net price remains negatively correlated with coverage generosity, while list prices only play a small and insignificant role. The effect sizes are the largest for preferred coverage, although the correlation is small in absolute terms. A 20 percent difference in net price is associated with just a 3 percent increase in preferred coverage relative to the average rate.

While suggestive, these estimates should not be interpreted causally and are likely to be biased downward by unobserved factors, such as drug quality. Manufacturers of higher-quality drugs within a given class can charge higher prices while demanding better coverage. Any differentiation in strategy across PBMs would also dampen the estimated correlation. Each PBM could create a sharp tradeoff between net price and formulary generosity. Still, if different drugs occupy favorable tiers on different formularies, these relationships would be weak at the overall market level. More generally, we observe equi-

[16] We verify robustness of our results to drug class definition (Table A1 in Appendix B.1).

Table 2: Prices vs. Formulary Correlations

	Frac. Preferred		Frac. Covered	
Log(net price)	-0.0228**	-0.0294*	-0.00656*	-0.0103
	(0.00672)	(0.0138)	(0.00315)	(0.00742)
Log(WAC)		0.0104		0.00586
		(0.0159)		(0.00979)
Fixed Effects	Class-by-Year			
Observations	2520	2520	2520	2520

Notes: Estimates of Equation 1, comparing within class-by-year cells. Competitors are defined as sharing the same ATC-4 classification. "Net price" refers to annual net-of-rebate prices from SSR Health. "WAC" is the list price, also tracked by SSR Health. "Frac. Preferred" and "Frac. Covered" refers to the fraction of insured lives who face preferred tier coverage of a given drug and have the drug covered by insurance, respectively. Both measures are calculated using MMIT data. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

librium outcomes, which might not reflect the price–formulary tradeoff facing each drug manufacturer.

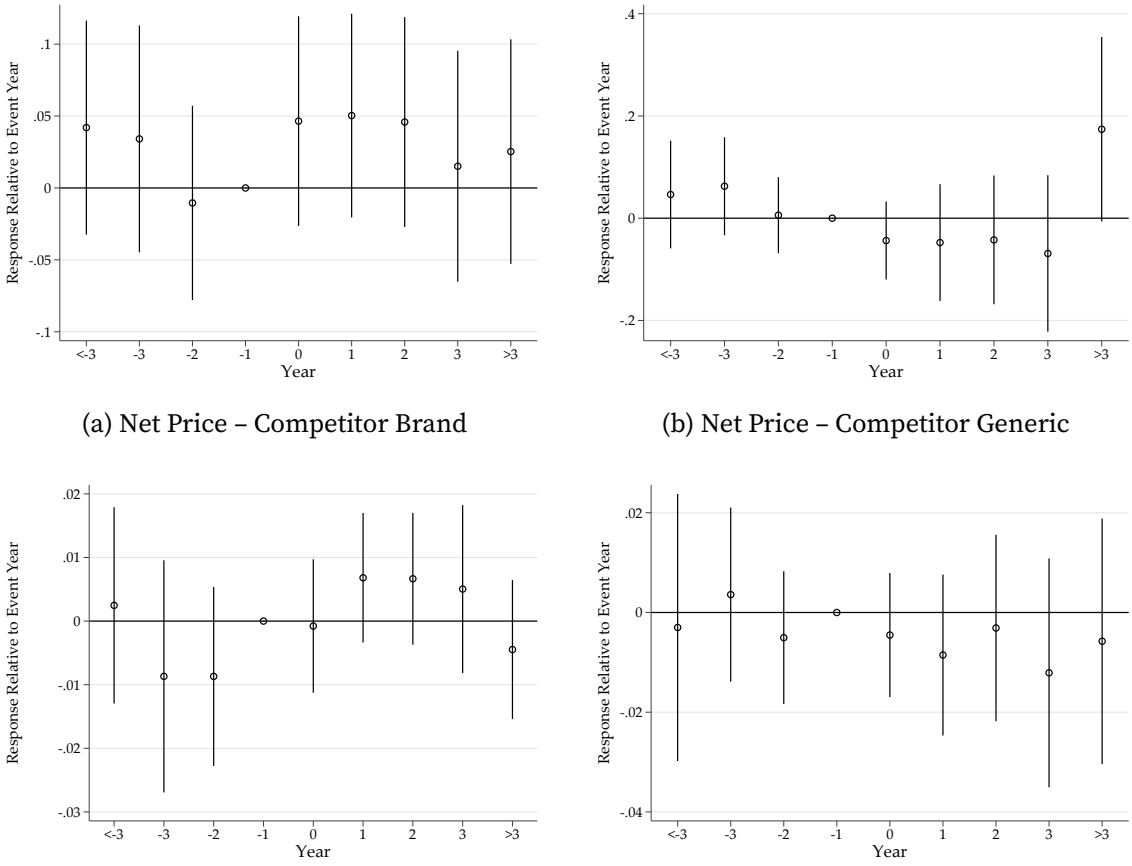
2.2 Response to Competitor Entry

If PBMs use formulary position to extract discounts, then the entry of competitors into a given market should lead incumbents to lower net prices.

To investigate this hypothesis, we conduct event studies around drug entry events. We consider two types of events: the entry of a new branded competitor and the entry of a generic version of a branded competitor.[17] For this analysis we use the largest sample available. When studying prices, we use the full 2007–2019 panel of over 1000 drugs from SSR Health. When studying formulary coverage, we use the 2011–2019 panel of 543 drugs from MMIT.

Following Sun and Abraham (2020), we use a two-way fixed-effects framework with time-varying responses to measure the response of branded incumbents to the entry of branded substitutes and the entry of generic versions of branded competitors. We exclude cases in which the same manufacturer owns the substitute. Formally, we estimate the following regression:

[17] We do not consider entry of direct generic competitors because brand drugs generally acquiesce to generic competition by suspending discounts and focusing on customers with strong brand preferences only (Frank and Salkever, 1992; Grabowski et al., 2016).



(a) Net Price – Competitor Brand

(b) Net Price – Competitor Generic

(c) Frac. Preferred – Competitor Brand

(d) Frac. Preferred – Competitor Generic

Figure 1: Time-Varying Estimates of Responses to Entry Events

Notes: Time-varying estimates of responses of market outcomes to entry events. "Net price" refers to annual net-of-rebate prices from SSR Health. "Frac. Preferred" refers to the fraction of insured lives who face preferred tier coverage of a given drug, based on MMIT data. "Competitor Brand" refers to the entry of a branded competitor in the same ATC-4 therapeutic area. "Competitor Generic" refers to a competing branded drug in the same ATC-4 losing exclusivity, as tracked by SSR Health.

$$y_{jt} = \zeta_j + \lambda_t + \sum_{k \in B_j} \left[\tau_1 \sum_{l<-3} D^l_{jkt} + \sum_{\substack{i=-3 \\ i \neq -1}}^{3} \xi_l D^l_{jkt} + \tau_2 \sum_{l>3} D^l_{jkt} \right] + \sum_{k \in G_j} \left[\tau_3 \sum_{l<-3} D^l_{jkt} + \sum_{\substack{i=-3 \\ i \neq -1}}^{3} \eta_l D^l_{jkt} + \tau_4 \sum_{l>3} D^l_{jkt} \right] + \varepsilon_{jt} \quad (2)$$

where y_{jt} represents price and formulary outcomes, ζ_j and λ_t are drug and year fixed-effects, respectively, and the remaining terms allow for flexible, time-varying responses to each event. Formally, $D^l_{jkt} = \mathcal{I}\{t - E_{jk} = l\}$, and indicates whether the current period t is in year l relative to treatment event k experienced by drug j. B_j is the set of branded competitor entry events faced by drug j, and G_j is the set of competitor generic entry events. The specification allows for different responses to each type of competitor entry event but constrains the response to be the same across all events of the same type. We again include only years in which a branded drug has exclusivity.

Figure 1 presents the time-varying estimates using ATC-4 drug class definitions. We find no clear patterns in net price response for branded competitor entry events. We again find little evidence of a large price response for competitor generic entry events, with some suggestive evidence that prices drop slightly soon after competitor loss of exclusivity followed by long-run increases. Formulary outcomes also exhibit minimal responses to both types of entry events. As a robustness check, Table A2 in Appendix B.1 provides results using alternative definitions of therapeutic class.

2.3 Discussion and Motivation for Structural Analysis

The two empirical facts documented in the previous section are seemingly at odds with each other. On the one hand, the cross-sectional patterns suggest that PBMs enforce a tradeoff between net prices and formularies within each drug class. On the other hand, the event study estimates do not show clear evidence that competition leads to lower prices or changes in formulary status. This contradiction appears elsewhere in the literature. For example, both Hwang et al. (2019) and Kakani et al. (2020) find that drugs in Medicare Part D protected classes—which PBMs cannot exclude from coverage—exhibit significantly higher price growth than drugs in other classes. At the same time, papers focusing on prices tend to find that product proliferation in a specific therapeutic area has little or no effect on prices (Lu and Comanor, 1998; Hartung et al., 2015; Howard et al., 2015).

We argue that demand inertia can potentially reconcile these facts. Inertia is pervasive in markets where consumers make repeated choices (e.g., Dube et al., 2009; Handel, 2013; Honka, 2014; Shcherbakov, 2016; Raval and Rosenbaum, 2018), and, as discussed earlier, the majority of prescription drugs treat chronic conditions, which, by definition, involve repeated choices. Feng (2023) provides quasi-experimental evidence of significant inertia in chronic drug demand.

The theoretical literature suggests that demand inertia can dampen price responses to

competitor entry. Theory predicts that inertia will lead to higher and increasing equilibrium prices by segmenting markets (see e.g. Klemperer, 1987a,b; Dube et al., 2009). Applying the logic of the segmented market to our event study analysis, by the time competitors enter, existing drugs likely would have established a loyal consumer base, dampening the effects of competition on prices. Furthermore, as discussed earlier, payers dislike the exclusion of popular drugs because of disruption to patients, so PBMs are limited in their ability to reduce coverage generosity for incumbent drugs significantly.

Because products will build up a loyal patient base over time, the effect of inertia should grow over time. Hence, if our hypothesis is correct, we should see incumbents reacting more strongly to an entrant earlier in the life-cycle. To test this hypothesis, we run a slightly modified version of the regression in Equation 2 that includes an indicator for branded competitor entry and an interaction term that allows the response to vary linearly with the age of the incumbent at the time of competitor entry. Table A4 in Appendix B.3 reports the results. Entry of a competitor leads to a 6 percent decline in net price at age zero, but the magnitude of the effect diminishes by about 1 percent each year that the incumbent drug has been on the market. These price concessions lead to a 3 percent increase in the fraction of lives with preferred access to the incumbent drug, as the manufacturer races to build a loyal patient base more quickly. The magnitude of this effect also declines by about 0.4 percent each year the incumbent has been on the market.

While these analyses provide suggestive evidence on the pricing and coverage dynamics of the market, it is hard to gauge how PBMs use formularies to extract price concessions using only reduced-form tests. Doing so would require us to observe data from drug markets where PBMs do not operate or where PBMs use alternative formulary designs (e.g., formularies without tiers or exclusion). Unfortunately, this data is not available. Therefore, in the rest of the article, we turn to structural analysis, which allows us to conduct counterfactuals where we alter the structure of PBM-manufacturer relationships or eliminate PBMs altogether.

3 A Structural Model of Drug Pricing in Markets with Demand Inertia

We consider a finite-period dynamic game. Manufacturers of branded therapeutic substitutes within a class compete over prices until their drugs lose exclusivity. At this point, generic manufacturers enter the market, and branded manufacturers receive a zero terminal payoff.

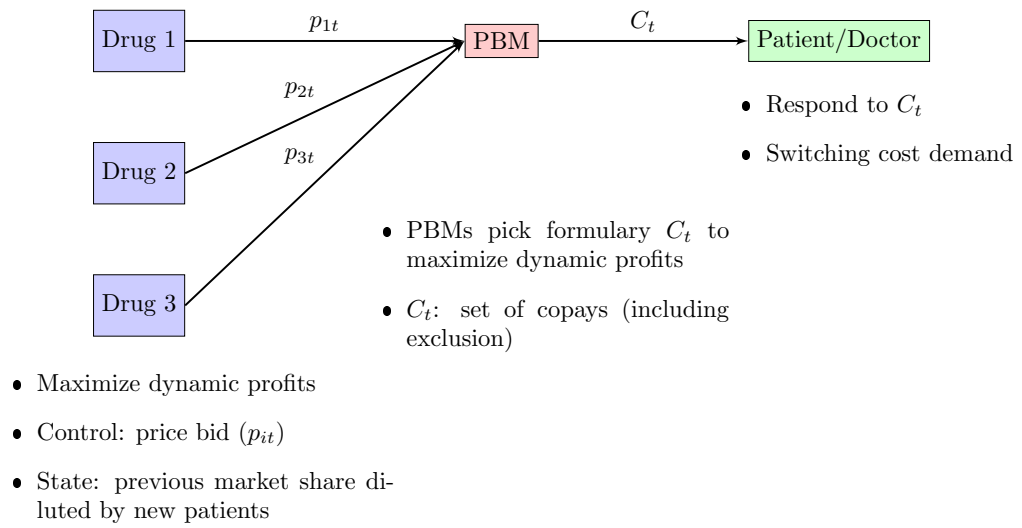


Figure 2: Model Summary

Notes: A summary of the period game structure in our model. Each "drug" box represents a drug manufacturer. p_{it} are the price bids and C_t the formulary chosen by the PBM.

In each period, manufacturers play a Stackelberg-Nash pricing game by submitting a per-unit price to a representative PBM (see Figure 2 for a visual representation of this game).[18] The PBM then designs a formulary that maximizes its objective function by assigning each drug to either the low-copay tier or high-copay tier or excluding it from coverage altogether. We model the objective function of the PBM as a flexible control function that includes patient welfare, drug spending, and a penalty for exclusion (which can generate additional disruptions for existing patients beyond switching medications). Finally, consumers pick a drug based on the formulary.[19]

Dynamic considerations arise in this game because demand in the current period influences future demand through inertia. Hence, manufacturers, all else equal, will value

[18] As discussed in Section 1, manufacturers actually submit multiple prices, one for each formulary arrangement. We make the simplifying assumption that firms send a single price to simplify computation of the equilibrium. In theory, our model could accommodate the submission of a formulary-specific price vector. In practice, the computational time required to solve the model makes such an extension unfeasible.

[19] In reality, the choice of a prescription drug is a joint process that involves up to three agents: patients, physicians, and pharmacists. Upon reaching a diagnosis, a physician decides which drug to prescribe with the help of the patient. Some drugs are administered by physicians in offices. For all other drugs, patients fill prescriptions at a pharmacy. In the latter case, pharmacists must fill the prescription for the same compound (e.g., the active ingredient, strength, form, and days supply) prescribed by the physician but may substitute a branded prescription with an identical generic one if available. Because manufacturers ultimately only care about overall demand and not how demand is generated, we treat all agents as a single, joint decision-maker.

preferred status earlier on in the life-cycle of their drug—and bid accordingly. Over time, the benefit of preferred status declines, as consumers already using the drug are more likely to keep using it even if it is moved to the non-preferred tier. At the same time, accumulating a large market share also makes it hard for the PBM to restrict coverage. As a result, manufacturers will bid less aggressively as their drugs age.

To account for demand inertia, we include the previous period's market shares as the state variable of the dynamic game. Manufacturers observe the state variable at the beginning of each period before choosing price bids. We assume that firms play a Markov Perfect Equilibrium, with perfect foresight and information, which helps ease the computational burden.

We now present each element of the model in reverse order, starting with the demand system, followed by the choice problem of the pharmacy benefit manager, and concluding with the manufacturer bidding game.

3.1 A Model of Demand for Prescription Drugs with Inertia

We present a dynamic model for prescription drugs that accounts for inertia.

We assume consumers choose from a set $\mathcal{J}$ of molecules within a therapeutic class.[20] The set $\mathcal{J}$ includes the outside option $j = 0$, which indicates the option of seeking treatment outside the therapeutic class. We assume that a single manufacturer—either a branded manufacturer or a representative generic manufacturer—produces each molecule. This assumption implies that branded manufacturers exit the market after generic entry. We make this assumption for simplicity and because, empirically, generic manufacturers capture 93 percent of a molecule's market share within a year of patent expiration (Grabowski et al., 2016).

We let N_t denote the number of consumers in period t. In each period, some new consumers arrive.[21] Consumers who were already in the market make repeated choices. Consumers only differ ex-ante in their past period choice and are otherwise identical.

Formally, we model consumer choice utility as

$$u_{ijt} = \delta_{jt} - \alpha c_{jt} + \gamma I_{s_{it-1}=j, j\neq 0} + \nu I_{s_{it-1}\neq 0, j\neq 0} + \varepsilon_{ijt} \tag{3}$$

[20] A molecule is the chemical structure of the active ingredient of a drug. Branded drugs and their generic versions share the same molecule.

[21] We do not consider consumer exit, though extending the model to include it is not hard. In the specific setting of our empirical analysis (the market for cholesterol-lowering drugs) exit is unlikely to be important, as high-cholesterol is a chronic condition, but also has a negligible mortality rate. In other settings, consumer exit may be more important.

where i indexes the consumer, j the molecule, and t the year. δ_{jt} reflects period-specific molecule quality. The time-varying component in quality accounts for variation in advertising, changes in medical evidence, and evolution in the quality of the outside option, whose value (δ_{0t}) is normalized to zero in each period. α measures sensitivity to the out-of-pocket cost c_{jt}, which is determined by the formulary.[22] We model inertia by including two switching cost terms. γ is the cost of switching between molecules: s_{it-1} indexes the option that patient i picked in period $t-1$, while $I_{s_{it-1}=j,j\neq 0}$ is an indicator for whether a patient i chose a non-outside-option molecule j in period $t-1$. ν is the cost of switching from any molecule to the outside option. This cost reflects the fact that once a patient has started on a maintenance medication, they are very unlikely to stop (but are more likely to switch medications instead, especially if their current medication is excluded from the formulary).[23] Intuitively, the first switching cost term allows for an extra advantage for the incumbent brand over other brands, while the second switching cost term gives all drug brands an advantage over the outside option if the patient is already taking medication. Finally, ε_{ijt} represents a consumer- and molecule-specific taste shock, which is distributed according to type-I extreme value distribution. We do not explicitly account for persistent taste differences in the demand system, because firms in our dynamic game would have to track preference-dependent previous market shares, which would increase the dimensionality of the state space beyond what is computationally feasible. [24]

Under these assumptions, we can express choice probabilities as a function of previous choices and the current formulary. To do so, we define the formulary of consumer i as a set of copays $C_t = \{c_{jt}\}_{j\in\mathcal{J}}$ for each drug in $\mathcal{J}$. We then denote the probability that patient i chooses molecule j as

$$ms_{ijt}(s_{it-1}, C_t) = \frac{\exp\left(\delta_{jt} - \alpha c_{jt} + \gamma I_{s_{it-1}=j,j\neq 0} + \nu I_{s_{it-1}\neq 0,j\neq 0}\right)}{\sum_{k\in\mathcal{J}} \exp\left(\delta_{kt} - \alpha c_{kt} + \gamma I_{s_{it-1}=k,k\neq 0} + \nu I_{s_{it-1}\neq 0,k\neq 0}\right)} \tag{4}$$

From individual choice probabilities, we can aggregate up to obtain a formula for market shares. To do so, let $x_t = \{x_{jt}\}_{j\in\mathcal{J}}$ denote the fraction of current patients who chose prod-

[22] The out-of-pocket cost for the outside option is also normalized to zero. In addition, we are assuming all patients face the same representative formulary to simplify exposition. In our estimation, we use individual-level data and relax this assumption.

[23] The parametrization of switching costs is a simplified version of the model in Feng (2023), which allows each drug to have a different switching cost. The findings of that paper suggest brand-specific switching costs are very similar to one another, so we impose this additional constraint to ease the computational burden.

[24] Not explicitly modeling unobserved heterogeneity will lead us to overstate the copay elasticity for new patients. However, Feng (2023) finds unobserved heterogeneity to be quantitatively small relative to switching costs, using a simulated maximum likelihood approach, which suggests that our demand system will be fairly accurate in capturing the behavior of patients already taking medication.

uct j in period $t-1$. Notice that this is not the same as the market share in the previous period because, in each period, new patients may come in—although, for a given number of new patients, the two are mechanically related by the recursive formula

$$x_{jt} = ms_{jt-1}\left(\mathbf{x_{t-1}}, C_{t-1}\right) \times \frac{N_{t-1}}{N_t} + \mathbb{I}\{j = 0\} \times \frac{N_t - N_{t-1}}{N_t} \tag{5}$$

where N_t denotes the number of consumers at time t, the second term accounts for the entry of new patients, which we treat as if they chose the outside option in the previous period, and $ms_{jt-1}\left(\mathbf{x_{t-1}}, C_{t-1}\right)$ is the market share of product j in period $t-1$. In turn, the market share is given by

$$ms_{jt}\left(\mathbf{x_t}, C_t\right) = \sum_{k \in \mathcal{J}} the x_{kt} ms_{ijt}\left(k, C_t\right) \tag{6}$$

The utility specification also gives us a closed-form solution for the total welfare generated by a given formulary configuration, which we refer to as "formulary surplus" and denote as $W\left(\mathbf{x_{t-1}}, C_t\right)$. Using the properties of the extreme value distribution, we get that

$$W\left(\mathbf{x_t}, C_t\right) = \sum_{k \in \mathcal{J}} x_{kt} \cdot \log\left[\sum_{l \in \mathcal{J}} \exp\left(\delta_{lt} - \alpha c_{lt} + \gamma I_{k=l, l\neq 0} + \nu I_{l\neq 0}\right)\right] \tag{7}$$

3.2 Formulary choice model

A representative PBM chooses the formulary arrangement to maximize an objective function that includes both drug cost and formulary surplus. Essentially, our approach is to use a flexible parametric control function to capture the PBM's role as an agent for the payer and its profit incentives that arise from vertical integration with manufacturers (as discussed in Section 1).

Formally, given a vector p_t of bids from branded drug manufacturers, the PBM chooses a formulary arrangement $C_t = \left\{c_{jt}\right\}_{j \in \mathcal{J}}$, where $c_{jt} \in \{c_g, \underline{c}_t, \bar{c}_t, \infty\}$, with $c_g < \underline{c}_t < \bar{c}_t$. The values of c_g, $\underline{c}_t$, and $\bar{c}_t$, the copays associated with the generic, preferred, and nonpreferred tiers, are assumed to be exogenous.[25] Copay c_g is only available to generic drugs. We assume that PBMs automatically place generics in a separate generic copay tier, with a lower out-of-pocket cost than either branded-drug copay tier.[26] Crucially, the presence of

[25] This reflects the fact that insurers rather than PBMs pick cost sharing levels.

[26] Empirical evidence (Grabowski et al., 2016) shows that brand manufacturers essentially acquiesce to generic entry by abstaining from competing on price and focusing on a small subset of very loyal patients. Moreover, based on our formulary data, virtually all formularies have separate generic tiers with copays

generics influences PBM formulary decisions because they present a cheap and efficient option that would be available even if all branded drugs were excluded. We denote $\mathcal{C}_t$ as the set of all possible formularies.

In each period, given price bids $\mathbf{p_t}$, the PBM chooses $C_t \in \mathcal{C}_t$ to maximize:

$$\max_{C_t} \frac{W(\mathbf{x_t}, C_t)}{\alpha} - \theta_s \sum_{j \in \mathcal{J}} p_{jt} ms_{jt}(\mathbf{x_t}, C_t) - \theta_e \sum_{j \in \mathcal{J}} \mathbb{I}(c_{jt} = \infty) x_{jt} + \sum_{j \in \mathcal{J}} \theta_j \times ms_{jt}(\mathbf{x_t}, C_t) + \theta_\omega \omega_{Ct} \tag{8}$$

The expression has four components. The first component is "formulary surplus" $W(\mathbf{x_t}, C_t)$. We divide formulary surplus by our estimated copay-elasticity to transform utils into dollars. The second is expected spending per member. The third is an additional penalty—increasing in the share of existing customers for each drug—for excluding molecules from the formulary. This term helps capture disruption costs and any additional preference heterogeneity not captured by the demand system. We also add a fourth component that captures any potential incentives for the PBM to favor specific drugs. This term accounts for the fact that drug manufacturers owned a few PBMs in the 1990s and 2000s. Integrated PBMs might favor drugs owned by their parent company in their formularies. Finally, the error term ω_{Ct} captures PBM preferences for specific formulary configurations unobserved by manufacturers before making price bids. We assume the error term is distributed according to type-I extreme value and allow for a flexible variance through θ_w. The error term has a dual function. First, it allows the model to fit the data. Second, it facilitates estimation by smoothing out expected manufacturer revenue as a function of the price bid.[27]

Using once again the properties of the extreme value distribution, we can write the probability that a given formulary arrangement is picked as

$$\mathcal{P}(C_t | \mathbf{x_t}, \mathbf{p_t}) = \frac{\exp\left(\frac{1}{\theta_\omega} \text{ObjFun}(C_t, \mathbf{x_t}, \mathbf{p_t})\right)}{\sum_{C \in \mathcal{C}_t} \exp\left(\frac{1}{\theta_\omega} \text{ObjFun}(C, \mathbf{x_t}, \mathbf{p_t})\right)} \tag{9}$$

where $\text{ObjFun}(C_t, \mathbf{x_t}, \mathbf{p_t})$ is the function in the maximization in Equation 8.

below even those of preferred branded drugs.

[27] Without the error term, formulary selection is deterministic, and firm profits can present discontinuities. This does not rule out the existence of an equilibrium, but it does make the computation of an equilibrium extremely difficult.

3.3 Bidding for formulary placement

Finally, we model competition between multiple branded drug manufacturers, each producing a single molecule, who compete for formulary placement by submitting price bids to the representative PBM. Because each manufacturer produces a single product, we index manufacturers using j as well. Manufacturer enter and exit the market at exogenously set times T_j^{en} and T_j^{ex}, which are known to all competitors. Each manufacturer also has perfect foresight about the evolution of the market size variable N_t.

The components of the dynamic game are the state variable, which consists of $\mathbf{x_t} = \{x_{jt}\}_{j \in \mathcal{J}}$ (i.e., the fraction of current patients that chose each product, as defined in Equation 5), and the control, which is the vector of price bids submitted by manufacturers. The state variable captures the history-dependent nature of demand. Manufacturers trade off lower prices and higher market share in the present against higher profits from a more dominant market position in the future.

Each drug manufacturer j chooses a price bid p_{jt} in period t to maximize dynamic profits, represented by the following value function:

$$V_{jt}(\mathbf{x_t}) = \max_{p_{jt}} \sum_{C_t \in \mathcal{C}_t} \left[p_{jt} N_t ms_{jt}(\mathbf{x_t}, C_t) \mathcal{P}(C_t | \mathbf{x_t}, \mathbf{p_t}) + \beta V_{jt+1}(\mathbf{x_{t+1}}(C_t, N_{t+1})) \right] \tag{10}$$

where β is the discount factor, which we set equal to 0.88, following the cost of capital calculation in DiMasi et al. (2003). $\mathbf{x_{t+1}}(C_t, N_{t+1})$—the value of $\mathbf{x_{t+1}}$ as a function of the chosen formulary—is deterministically set by the expressions in Equations 5 and 9. Also, note that we are assuming firms face no fixed or marginal costs. In practice, manufacturing and distribution costs of branded prescription drugs represent a small fraction of their price, so this approximation is unlikely to have a large impact on our estimation.

Manufacturers stop making strategic decisions after their terminal period T_j^{ex} (i.e., when their product loses exclusivity). When that happens, generic manufacturers enter the market. While manufacturers take into account the presence of generic competitors when bidding, we do not model generic manufacturers as strategic bidders, because generics are always placed in a lower-cost tier relative to branded products. In general, PBMs play a much more passive role for generic drugs, whose prices are instead determined through price competition among many homogeneous products. As previously stated, we assume that firms play a Markov Perfect Equilibrium over a finite period.

3.4 Discussion of assumptions

Before moving on to the estimation we discuss two assumptions we make in the model and how they affect estimation. First, we abstract from competition between PBMs and consider a single representative PBM. This assumption is undesirable, but necessary because of data limitations and to make estimation feasible. Data limitations arise because we do not observe PBM-specific prices, but only average prices at the US market level. Hence, we have no variation on the supply side that could help us identify PBM-specific parameters.[28] Even if we did have this data, estimating an equilibrium with multiple competing PBMs would require a lot of computing time. The threat of multiple equilibria would also be much greater. The main downside of considering only a representative PBM is that it makes the threat of formulary exclusion worse than it might be in practice. In a world with multiple PBMs, manufacturers retain some leverage if they can be included on some, but not all, formularies. This option does not exist in our setting. However, the relative importance of exclusion versus being placed in the non-preferred tier—the other tool at the PBM's disposal—should not change.

Second, we rely on a relatively simple Nash-Stackelberg price-setting game, rather than a more sophisticated bargaining model. While bargaining might approximate the actual price-setting process more closely, we do not believe this assumption has a meaningful impact on our results. Our model captures the key tradeoff faced by manufacturers: lower prices for better formulary coverage. Manufacturer response to key market structure changes will also be qualitatively identical: entry of competitors leads to lower bids, all else equal, and changes in the formulary structure such as removal of the threat of exclusion also have the expected effect on equilibrium outcomes.[29]

The use of a simple price-setting model also comes with advantages. The presence of inertia would introduce complicated dynamic considerations in a bargaining game, which would compromise computational feasibility. Moreover, drug manufacturers and PBMs negotiate both price and formulary position. This dual negotiation does not fit within the framework of Nash bargaining models, which typically only account for a single bargaining outcome (see e.g. Gowrisankaran et al., 2015; Ho and Lee, 2017, 2019).

[28] This limitation is shared by all papers on the pharmaceutical industry as far as we know.

[29] Additionally, we note that our model closely follows a descriptions of the process used by Express Scripts to set formularies. Based on conversations with former executives and regulators, Express Scripts runs annual auctions in each drug market. It lays out a set of formulary arrangements for each market; each arrangement places each drug in either the low-copay tier, high-copay tier, or excluded tier. Companies submit a price bid for each arrangement, and Express Scripts then picks the arrangement that maximizes its profits.

4 A Structural Analysis of the Medium-Intensity Statin Market

4.1 Background and data

Because the estimation of the model is very computationally intensive, we choose a single therapeutic area for our structural exercise. We focus on medium-intensity statins, a popular class of cholesterol-lowering drugs, for three reasons. First, high-cholesterol is usually a chronic condition, so patients that use statins usually take them for a very long time. Hence, the demand for statins is likely to have a strong degree of inertia, which is the motivation behind our exercise. Second, because statins are a relatively old class of drugs, we are able to observe data on sales of statins up until the point when all statins lose patent exclusivity. This is a necessary requirement for estimating our model. Third, our data spans a period during which we see competition among up to three branded statins and two generic entry events, giving us enough variation to estimate the parameters of our model. Finally, we also note that the statin market is interesting in and of itself. A recent report by the Center for Disease Control and Prevention (CDC) estimates that 27.9 percent of Americans over the age of 40 are currently taking anti-cholesterol medication.[30]

To conduct the estimation, we collect additional data on prescription claims and the prevalence of high-cholesterol from Truven MarketScan and the Medical Expenditure Panel Survey (MEPS). The MarketScan database consists of a large panel of prescription drug claims from a set of large employers—nearly 60 million individuals in 2012. Our version of the data covers the years from 1996–2013. We use Marketscan claims to estimate our demand system. MEPS also contains prescription drug claims but from a much smaller sample of individuals. However, the sample of patients surveyed through MEPS is nationally representative, which we use to scale up our demand estimates to the US-market level.[31]

We also manually collect data on net prices of statins up to 1996. This yields an additional 11 years of data compared to the SSR Health data, which starts in 2007. To construct estimated net prices, we collect net revenues from the financial filings of statin manufacturers, and total quantities from MEPS data, which we aggregate at the product level,

[30] National Center for Health Statistics Data Brief No. 177, December 2014.

[31] We cannot estimate demand using MEPS for two reasons. First, MEPS uses an overlapping panel, meaning we cannot follow the same patient for longer than 18 months. This is a major drawback, as it makes it very hard to know if a patient was already on statins prior to the start of the panel. Second, MEPS does not contain plan identifiers, which makes it impossible to determine the out-of-pocket cost of all the options available to each patient.

across pills of different strengths. Our net price estimate is the ratio of the two, under the assumption that each strength carries the same rebate percentage. To ensure that our estimates are robust to this assumption, we calculate net prices under two additional sets of assumptions (constant dollar rebates, and a more sophisticated calculation that tries to exclude rebates to institutional buyers that don't use PBMs like Medicaid and the Department of Veteran Affairs) and repeat the estimation using these alternative price series.[32]

Figure 3 summarizes the evolution of (inflation-adjusted) per-pill net prices, average per-pill copays, and market shares in the statin market from 1996 to 2013. The market consists of specific dosages for three branded drugs, Zocor (simvastatin, first launched in 1992), Lipitor (atorvastatin, launched in 1997), and Crestor (rosuvastatin, launched in late 2003).[33]

Zocor has a higher net price than Lipitor and Crestor, for two reasons. First, it is the incumbent in the class. Second, it is marketed by Merck, which also owned a large PBM (Medco) until 2003. Zocor's price starts falling after Merck separated from Medco.

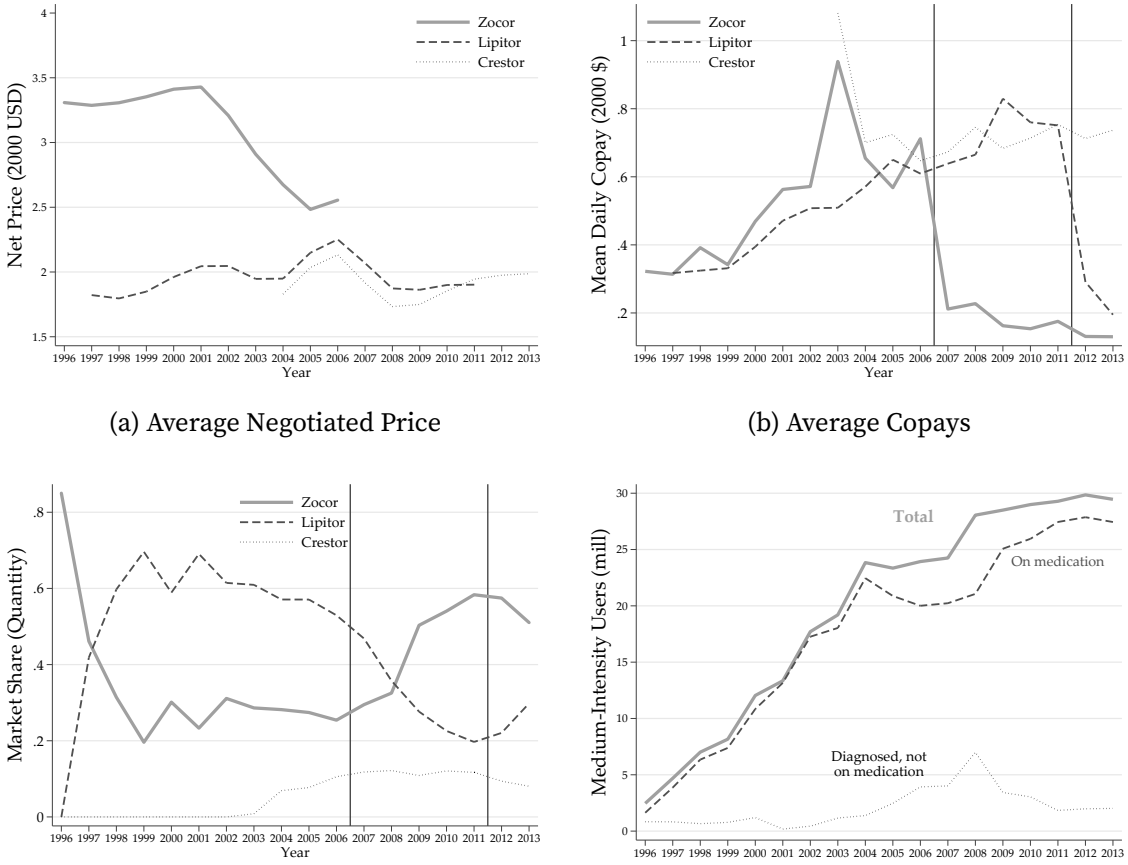
Average out-of-pocket cost (calculated across all filled scripts tracked by MEPS) are fairly similar across drugs, up to the point when a drug becomes generic. Generic entry occurs in 2007 for Zocor and in 2012 for Lipitor.

In spite of the similarity in out-of-pocket cost, Lipitor has a dominant market share (again calculated using claims from MEPS) until the entry of generic Zocor, while Crestor consistently has lower market share than its two competitors. Some of these patterns are likely due to unobserved heterogeneity in drug quality, but inertia probably also plays a role in relegating Crestor (the latest entrant) to a much lower market share.

We also provide our calculations for market size, again based on MEPS. We first compute the total number of individuals taking medium-intensity statins in a given year, using weights provided by MEPS to arrive at a nationally representative total. Next, we find all individuals not taking medication who are diagnosed with high cholesterol. For these individuals, we check whether they take medication in the following year, as a proxy of whether they were at risk of taking cholesterol medication in the focal year. We mark these individuals as choosing the "outside option" and add it to the number of individu-

[32] See Appendix C.1 for a detailed description of each methodology.

[33] A few additional statin products are also available on the market, but are generally used to treat low-intensity patients (Mevacor, Lescol, Pravachol), or high-intensity patients (Vytorin). Vytorin's lowest dosage is also indicated for moderate-intensity use, but its market share is below 1 percent, so we exclude it from our analysis for simplicity. We treat these alternatives as the joint outside option of our demand model. Appendix C.1 describes the market and provides details on the mapping from dosages to treatment intensity. It also summarizes the volume sold of each drug and dosage. For all three drugs we consider, the medium-intensity market is the largest source of sales.



(a) Average Negotiated Price

(b) Average Copays

(c) Market Shares

(d) Market Size

Figure 3: Market Summary

Notes: Graphs summarizing key statistics for the three major statins we analyze. Negotiated prices are based on assuming a constant percentage rebate amount across all dosages of each drug. Copays represent the average across all recorded scripts for that drug in MEPS. Both prices and copays are normalized to a per-pill number (equivalent to per-day). Market shares are calculated based on quantities from MEPS. In Panels (b) and (c), the two vertical lines represent the entry of generic Zocor in 2006 and generic Lipitor in 2011. In Panel (b), the cost sharing for Zocor and Lipitor after loss-of-exclusivity are based on cost sharing associated with their generic versions.

als taking medication to arrive at the total market size in a given year. The market grows significantly in the early part of the period, and becomes more stable starting in 2004.

4.2 Demand Estimates

We start by estimating the demand system, using individual-level claims data from MarketScan. The main challenge is separating switching costs from unobserved heterogeneity. If persistent choices were solely driven by unobserved heterogeneity in patient preferences, then demand would not create dynamic incentives for drug manufacturers.

Our solution for identifying switching costs is to use the choices of other individuals who started treatment in the same year as an instrument for the focal patient's previous choice. The approach is based on the idea that time-specific factors unrelated to a given patient's unobserved preferences will affect choices: entry of new drugs, changes in medical evidence and prescription habits, and advertising intensity across drugs. Patients starting treatment are influenced by these factors to choose a drug over others. We can then use the average choices of people in an individual's starting cohort as an exogenous driver of the brand to which a patient is loyal.[34] Intuitively, our approach compares two individuals with similar tastes, who happen to start treatment at different times with time-specific factors captured by the average choices of others in their respective cohorts. Subsequent differences in their choices at any given point in time help us identify the degree of inertia.

Formally, given a patient i who started treatment in year y, our instrument for the previous choice indicator, $I_{s_{it-1}=j}$, is the leave-one-out cohort mean $\bar{I}_{y,i,j,t-1}$ of $I_{s_{k,t-1}=j}$ for other patients who also started treatment in year y, for each molecule j and time period t:

$$\bar{I}_{y,i,j,t} = \frac{1}{|Z_y| - 1} \times \sum_{k \in \{Z_y \setminus i\}} I_{s_{k,t-1}=j} \tag{11}$$

where Z_y is the set of patients who started treatment in year y, and $|Z_y|$ indicates the cardinality of set Z_y. In other words, the leave-one-out cohort mean for choice j and patient i is the fraction of patients other than i whose previous choice was j.

[34] One concern would be that patients' starting times are related to their unobserved preferences (e.g., they delay treatment because they anticipate the future entry of a drug they unobservably prefer). We assume that this does not happen. Feng (2023) provides evidence that the flow of new anti-cholesterol patients has similar characteristics over time.

We use $\bar{I}_{y,i,j,t}$ to predict the choice of patient i as:

$$I_{s_{k,t-1}=j} = \tau_1 + \tau_2 \bar{I}_{y,i,j,t} + \upsilon_{ijt} \tag{12}$$

We use the same approach to instrument for $I_{s_{i,t-1}\neq 0}$.

Because we are estimating a non-linear model, we cannot take a simple two-stage least squares approach. Instead, we follow two approaches outlined in the literature. The first approach comes from Newey (1987) and consists of substituting a predicted value from the instrument for the endogenous regressor (two-stage predictor substitution, or 2SPS). The second approach comes from Terza et al. (2008) and consists of adding the "first stage" residual as an additional regressor (two-stage residual inclusion, or 2SRI). The 2SPS provides an analogous approach to two-stage least squares for linear regressions. The equivalent here is to substitute in $\bar{I}_{j=s_{k,t-1}}$ for $I_{j=s_{k,t-1}}$ in the logit model. The 2SRI approach involves keeping the basic specification but adding in the residual υ_{ijt} to proxy for the serially correlated error terms.

We estimate a conditional logistic model via maximum likelihood using our predicted choice probabilities from Equation 4 and yearly-level panel data on individuals' anti-cholesterol drug choices over time constructed from MarketScan. We determine choices by assigning the most frequently prescribed drug to each individual and year. One key challenge is calculating copays c_{jt} for options not chosen by the individual. To do this, we make use of plan identifiers in MarketScan, and calculate the median monthly copay for each plan-drug-year across individuals who chose a given drug.

We do not use an instrument for c_{jt}. While instrumenting for copay level is desirable, the unobserved correlation between copays and unobserved patient-level or drug-level factors does not have a clear direction. Patients could select into plans that have a low copay associated with the drugs they like, leading to an upward bias in the estimated copay elasticity. On the other hand, copays might be higher for higher quality drugs, or drugs with more loyal patients because manufacturers might be more willing to accept worse coverage in exchange for charging higher net prices. Hence, our estimates might be biased upward or downward, and, in practice, these two effects might cancel each other out. To confirm that our estimates are reliable, we compare them to estimates from Einav et al. (2016), which estimates copay elasticities using quasi-random variation in copays arising from the Medicare Part D donut hole. Our estimates are very similar (we report the results of this comparison in Appendix D.1).

Table 3 reports the estimates using the two approaches, and for the simple regres-

sions without instruments.[35] The first column reports results without instrumenting for the previous choice indicators, which serves as a baseline. Instrumenting for previous choices leads to a decline in the estimated switching cost relative to the baseline case, consistent with the idea that some choices are driven by unobserved heterogeneity in match quality. For the IV specifications, we cluster standard errors at the starting cohort level, in order to reflect the level at which the instrument is constructed.

The two IV methodologies return quantitatively different but qualitatively similar results. For our dynamic analysis, we use the estimates from the 2SPS approach in the fourth column: $(\alpha, \gamma, \nu) = (0.668, 1.46, 1.385)$, but also provide model estimates using the 2SRI parameter values in Appendix D.2 (the estimates are very similar). The implied monthly switching costs γ and ν are \$65 and \$62, which is slightly more than double the average monthly copay.

Because MarketScan only includes patients on employer-sponsored health insurance, we use the nationally representative MEPS data to adjust drug quality parameters (δ_{jt}). To perform the adjustment we assume that copay elasticity and switching costs for the whole population are identical to those estimated from MarketScan. We then recover period-specific drug quality by matching the predictions of the demand model to MEPS market shares. Formally, let $\hat{\alpha}, \hat{\gamma}, \hat{\nu}$ denote our estimated parameters. Then, we find δ_{jt} that solves the equation:

$$ms_{jt}^{MEPS} - \sum_{k \in \mathcal{J}} x_{kt} ms_{ijt}\left(k, C_t; \hat{\alpha}, \hat{\gamma}, \hat{\nu}, \delta_{jt}\right) = 0, \forall j, t \tag{13}$$

There are two main reasons we use MEPS data to adjust quality. First, MarketScan is not meant to be a representative sample of the US market. MEPS is representative and covers all segments of the market, including Medicare. Second, the sample of patients in MarketScan can change significantly from year-to-year, which leads to fluctuations in market shares that add noise to quality estimates.

Figure 4 shows the implied quality estimates over time. We find that Lipitor has a significantly higher quality at the start of the period relative to Zocor and that Crestor has lower implied quality than the other two. These patterns reflect the reduced form statistics from Figure 3 (Panels B and C). Lipitor has higher quality because it has a higher market share than Zocor, but a similar out-of-pocket cost. Similarly, Crestor has lower implied quality because it also has similar out-of-pocket cost to Lipitor, but lower market share. All drugs exhibit declining quality. This makes sense, as drug quality is measured relative to the outside option, whose quality increases over time as additional alternative

[35] We report first stage estimates in Appendix D.1.

Table 3: Demand System Estimates

	Basic CL	2SPS	2SRI	2SPS	2SRI
Cost Sharing (α)	0.719***	0.633***	0.724***	0.668***	0.739***
	(0.00253)	(0.101)	(0.141)	(0.111)	(0.149)
Same Drug (γ)	3.341***	2.674***	2.502**	1.460***	2.070***
	(0.00235)	(0.567)	(0.778)	(0.204)	(0.277)
Any Drug (ν)				1.385**	0.330
				(0.494)	(0.494)
Drug-by-Year FE	x	x	x	x	x
N	49,413,573	49,413,446	49,413,446	49,413,446	49,413,446

Notes: Estimates of Equation 3 using maximum likelihood. Basic CL refers to specifications where we do not instrument for previous choice. 2SPS (two-stage predictor substitution) and 2SRI (two-stage residual inclusion) refer to the two instrumental variables approaches discussed in the main text. The results show an inelastic response to cost sharing and significant inertia. Incorporating the instrument and the term for previous intensive margin choice diminishes some of the raw estimates. The instrumental variables specifications (2SPS and 2SRI) have standard errors clustered at the cohort level to reflect the source of the exogenous variation used in the estimates. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

treatment options to statins are launched.

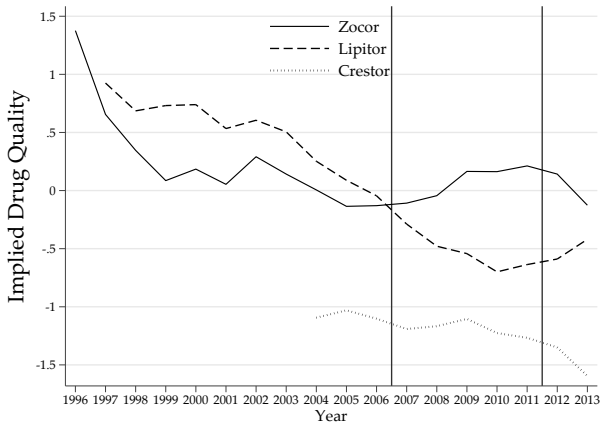


Figure 4: Implied Quality Estimates

Notes: A plot of the δ_{jt} coefficients from Equation 3, with generic Zocor and generic Lipitor qualities replacing their branded counterparts in the series once they enter. The two vertical lines represent the entry of generic Zocor in 2006 and and generic Lipitor in 2011.

4.3 Estimating the Dynamic Model

The final step of the estimation is the recovery of the parameters in the PBM's objective function. We estimate these parameters by matching model predictions for equilibrium bids to observed prices. Because we do not have a closed-form solution for the optimal bidding strategy, our computation of equilibria relies on an iterative optimization routine, where we set an initial guess for a set of prices, and then allow firms to best-respond one-by-one until an equilibrium is reached.

To calculate the PBM's objective function from Equation 8 we use our demand estimates to calculate welfare and spending for any given formulary arrangement. Our chosen parametrization also includes a boost for favoring Zocor in the formulary arrangement from 1996 to 2003. This term accounts for the fact that Merck, the manufacturer of Zocor, owned Medco, one of the largest PBMs, from 1993 until 2003. The term is linear in the market share of Zocor and included only in the years during which Merck owned Medco.[36]

Because our demand data is estimated over a distribution of patients with access to different formularies, we calculate firm revenue under the assumption that the PBM does

[36]Appendix D.4 provides parameter estimates for alternative parametrizations of the PBM's objective function. These include allowing PBMs to have dynamic incentives, allowing PBMs to value generic market share, and allowing previous market share to affect some fraction of new patients (e.g., through doctor behavior). We find that these factors are not quantitatively important for model fit.

not pick a single formulary. Instead, the PBM randomly assigns each possible formulary to patients according to the probabilities defined in equation 9.[37]

To generate moment conditions, we assume that net prices are observed with measurement error ι_{jt}. This is almost certainly true, since our data consists of estimated yearly averages. We allow for heteroskedasticity in the error term, but assume that the error is otherwise independently distributed across drugs and years. We do not include any source of structural error.[38]

To recover estimates for the vector of parameters $\grave{} = (\theta_s, \theta_e, \theta_j, \theta_\omega)$ we minimize the sum of squared errors:

$$\min_{\theta} \sum_{t} \sum_{j \notin G_t} \left(\hat{P}_{jt}(\grave{}; \hat{\alpha}, \hat{\gamma}, \hat{\nu}) - P^o_{jt} \right)^2 \tag{14}$$

where G_t is the set of drugs that has lost exclusivity, $\hat{P}_{jt}(\cdot)$ are the equilibrium prices of the finite period game under parameters θ and demand estimates $\hat{\alpha}, \hat{\gamma}, \hat{\nu}$ based on the first order conditions associated with Equation 10, and P^o_{jt} is data.

The estimation is computationally intensive because it requires computing the value function of each manufacturer over a large state space and number of periods for each guess of the model parameters. We describe the computation in detail in Appendix C.2.

A potential threat to our estimation approach is the presence of multiple equilibria. Multiplicity may occur in the stage game if manufacturers play separating strategies where one firm responds to a low price with a high price to focus on the inelastic part of the demand curve. While we do not have a formal proof of uniqueness, we do not believe that multiplicity is an issue in our setting, for three reasons. First, our game has characteristics that limit the scope of multiple equilibria. It has a finite horizon, which reduces the space of potential strategies, and it involves direct competition in a limited number of periods—the three manufacturers compete all together in only three periods, from 2004 to 2006. Second, even though we do not have a formal proof of uniqueness, we analyze the best-response functions of each manufacturer to check that, for our estimated parameters, they satisfy stagewise uniqueness in at least some of the stages. In our setting, stagewise uniqueness is a sufficient condition for a unique equilibrium because our game

[37] This assumption implies that there is no correlation between a patient's current choice of drug and their formulary. This is undesirable but necessary for estimation. In general, we expect patients taking a specific medication to be on formularies that cover their medication more generously. This exacerbates the effect of inertia, which in turn should make the role of PBMs even more important than we estimate it to be.

[38] The main concern in our estimation would be if PBMs had some preference for specific drugs in certain years in a way that is unobservable to the econometrician, but observable to manufacturers. This would lead manufacturers to send higher bids and still retain good formulary placement, which could bias our other coefficients. However, we believe this type of preference is unlikely, other than in the case of Zocor, which we account for in the estimation.

has a finite horizon. Hence, backward induction can be used to establish uniqueness (Doraszelski and Pakes, 2007; Besanko et al., 2010). Figure A8 in Appendix C.3 shows that the best response functions only intersect once in a few select states during the 2007–2011 period. More generally, we do not find multiple intersections at any points in state space during this duopoly period. Finally, in our estimation, we use a variety of different initial guesses for the vector of price bids, including guesses where each firm starts at a high or a low price, and guesses where firms start from different price benchmarks. In all instances, the recovered equilibrium is the same.

Because our demand parameters are estimated with error, we compute standard errors using a two-step approach (Murphy and Topel, 1985; Newey and McFadden, 1994). Our approach is as follows. Let n be the total number of periods for which we have branded drug price data, α represent the vector of demand parameters (α, γ, ν), and θ represent the vector of PBM control parameters. Furthermore, let k indexing both dimensions of our data (j and t). Our minimum distance estimator can be rewritten as

$$\hat{\theta} = \arg\min_{\theta} \frac{1}{n} \sum_{k=1}^{n} \left(P_k\left(\theta, \hat{\alpha}\right) - \tilde{P}_k \right)^2$$

where $P_k(\theta, \hat{\alpha})$ are the equilibrium prices of the finite period game under parameters θ and $\hat{\alpha}$ (the demand estimates). The expression is minimized at the true α_0 and θ_0. Let $Q_n(\theta, \alpha) = \frac{1}{n} \sum_{k=1}^{n} (P_k(\theta, \alpha) - \tilde{P}_k)^2$. Under the assumption that Q is twice differentiable, the asymptotic distribution of the second stage estimator is

$$\sqrt{n}\left(\hat{\theta} - \theta_0\right) \to N\left(0, B_0^{-1}\left(\Lambda_0 V_\alpha \Lambda_0' + \Gamma_0 \Omega_0 \Gamma_0'\right) B_0^{-1}\right)$$

where B_0 is the Hessian of the objective function $B(\theta) = \frac{\partial^2}{\partial\theta\partial\theta'} Q(\theta, \alpha_0)$ evaluated at the true parameter value θ_0, $\Lambda_0 = \frac{\partial^2}{\partial\theta\partial\alpha'} Q(\theta, \alpha)$ evaluated at $\theta = \theta_0, \alpha = \alpha_0$, V_α comes from the asymptotic distribution of the demand estimates, $\Gamma_0 = \frac{\partial P(\theta,\alpha)}{\partial\theta}$, and $\Omega_0 = \text{Diag}[(P_k(\theta, \hat{\alpha}) - \tilde{P}_k)^2]$. Intuitively, the first component in the variance expression accounts for the impact of demand errors through a delta method approach, and the second component is the standard minimum distance asymptotic variance expression.

We run our estimation routine to match each of the three negotiated prices series we described in Section 4.1.[39] We also use data on market size from MEPS, tier-specific copay amounts $\underline{c}_t, \bar{c}_t$ from MarketScan, and generic prices and copays from MEPS (set at the median values $p_g = 1$ and $c_g = 0.33$).

[39] We also run a robustness check where we use SSR Health data instead of our own calculations in the 2007–2013 period. This check yields virtually identical coefficients to the percentage discount specification (Column 1 of Table 4).

Table 4: Parameter Estimates from Dynamic Game

	Estimates		
Parameter	Pct Discount	Dollar Discount	Private
Sensitivity to Total Cost	1.70	1.67	1.68
θ_s	(0.13)	(0.13)	(0.41)
Aversion to Excluding Popular Drugs	1.50	1.55	1.54
θ_e	(0.05)	(0.05)	(0.25)
Zocor preference (Merck-Medco Integration)	1.11	1.01	0.89
θ_j	(0.37)	(1.61)	(0.02)
Standard deviation of unobservable factors	0.063	0.064	0.044
θ_ω	(0.005)	(0.040)	(0.022)

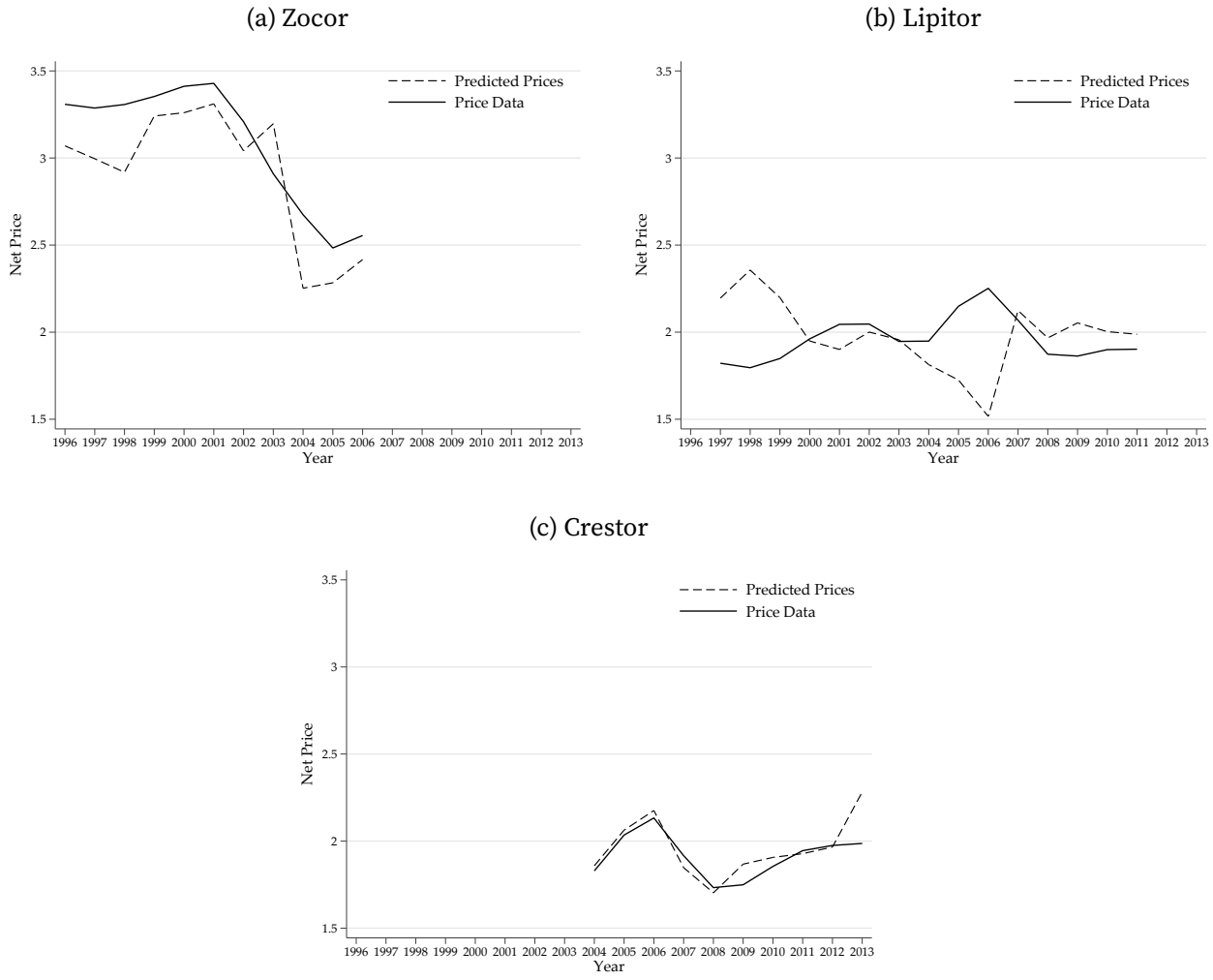
Notes: Parameter estimates based on different price series: "Pct Discount" reflects prices based on a constant percentage discount across dosages, "Basic" refers to the basic construction of negotiated prices, and "Private" reflects prices purged of sales and quantity in other market segments not directly affected by PBMs. Standard errors are computed using the two-step approach described in the main text.

Table 4 shows the parameter estimates and standard errors under different price series. Column 1, our preferred specification, reports the estimates using price data that assumes a constant percentage discount across dosages. The importance of cost reduction relative to formulary surplus is estimated to be 1.7. θ_e is large and significant, suggesting that PBMs are less likely to exclude drugs with many existing users. This is consistent with the idea that disruption is a criterion payers consider when choosing PBMs. θ_j is also quantitatively large ($1.11) and statistically significant, which suggests that Medco had strong incentives to increase the market share of Zocor.

Columns 2 and 3 provide evidence of the robustness of our estimates to data construction methodology. One concern with our core estimates is that the data do not reflect the actual prices drug manufacturers submit to PBMs. The second price series assumes a constant dollar discount across dosages, while the third attempts to isolate average prices across the commercial and Part D segments. The point estimates for the three series are quite similar, although there are some inconsistencies in the sizes of standard errors.

In terms of model fit, we are able to capture three key patterns in the data (Figure 5). First, the dynamics in the model help justify Lipitor setting lower prices despite having higher implied quality in the late 1990s and early 2000s. Lipitor was launched five years after Zocor and had market exclusivity that lasted until 2011 versus 2006 for Zocor. Therefore, the model predicts that, in the late 1990s and early 2000s, Lipitor had a much stronger incentive to invest in market share. However, it could only do so by setting significantly lower prices than Zocor, since Zocor was still favored because of its relationship with

Figure 5: Model Fit – Predicted Prices



Notes: Prices predicted by the best-fit model versus the prices observed in the data.

Medco. Second, even though it incorporates inertia, the model can justify the relatively flat prices of anti-cholesterol drugs thanks to the PBM's ability to prevent more extreme dynamic strategies.[40] Third, the model matches the small increase in Zocor's net price at the end of its life-cycle. This increase is motivated by Zocor's manufacturer's desire to extract more surplus from its loyal patient base.

Our model also returns formulary predictions—which can provide a better sense of model fit since we did not use formulary data in the estimation. While we do not have systematic coverage information for statins, we compare the model's predictions, shown in Figure A10 in Appendix on page A20, do qualitatively line up with the data from Medi-

[40] Figure A13 in Appendix on page A25 shows that equilibrium prices in a market without PBMs would be higher and increasing at a faster rate. Section 5 provides more details.

care Part D (Figure 3, available only from 2008 onwards), as well as with nonprofit and media reports in the early 2000s. In the pre-Crestor period, the model predicts that Zocor is more likely to be in the higher-copay tier while formulary exclusions are rare.[41] After Crestor enters, it is excluded with some probability.[42] After generic Zocor enters, both Lipitor and Crestor see some formulary exclusions, with Lipitor experiencing a lower exclusion rate.

5 Counterfactuals: PBM Impact and Policy

We use the estimated model to understand whether PBMs reduce prices and overall spending and, if so, how they achieve these reductions. We divide this section into three parts. First, we assess the impact of PBMs on prescription drug spending by comparing the current equilibrium to an equilibrium where manufacturers set prices unilaterally, and patients face an exogenously set coinsurance rate. Second, we compute equilibrium outcomes in counterfactual scenarios where the PBM can commit to choosing some subset of the possible formulary arrangements to highlight the mechanisms PBMs use to reduce prices. Finally, we examine the role of inertia in masking the impact of PBMs by presenting counterfactual price paths absent switching costs.

5.1 Do PBMs Reduce Drug Expenditures?

We start by exploring how PBMs affect drug prices and spending. We compute equilibrium outcomes assuming drug manufacturers set prices and patients face an exogenously set coinsurance rate. We choose this structure for two reasons. First, we assume manufacturers set prices directly because there are too many payers for one-on-one negotiations to make sense. Second, we use coinsurance rates because copays imply perfectly inelastic demand with respect to price, which leads to infinite prices under profit maximization. We fix the coinsurance rate at 33 percent, matching the observed ratio of cost sharing to net drug prices in our data.

The counterfactual results suggest that, relative to the current equilibrium, removing PBMs and imposing a 33 percent coinsurance would increase payments to drug manu-

[41] Table 9 of a 2001 report by the California HealthCare Foundation (http://www.chcf.org/publications/2001/06/prescription-drug-coverage-and-formulary-use-in-california-different-approaches-and-emerging-trends, retrieved February 9, 2023). The table shows that all the major PBMs put Lipitor in the preferred tier, whereas Zocor was only in the preferred tier of the formularies of Medco and AdvancePCS.

[42] Crestor initially was excluded from some formularies. See The Guardian article "US blow to AstraZeneca": https://www.theguardian.com/business/2003/oct/04/4, retrieved February 9, 2023.

facturers by 47.5 percent, or approximately $78.2 billion (in 2000 dollars) over the 18-year period of our data. Panel A of Table 5 shows that net revenue earned by drug manufacturers from 1996–2013 would be $242.5 billion under the coinsurance structure instead of $164.3 billion under the baseline model. Figure A13 in Appendix D.5 also shows that absent the moderating effect of PBMs, prices are higher and grow faster in the coinsurance counterfactual.

Accounting for payments to PBMs does not affect the results. Using gross revenue and variable costs reported in the financial filings of eight large, publicly traded PBMs, we estimate that PBMs earn a margin of about 6.5 percent, which is equal to $10.7 billion in net revenue from the statin market.[43] Therefore, the total cost to payers is $175.0 billion in the baseline model, which is 28 percent less than the total cost under the "33% Coinsurance" counterfactual. $10.7 billion is equivalent to about 13.7 percent of the savings generated relative to the 33 percent coinsurance counterfactual. The result that PBMs save money holds over a vast range of reasonable profit margins. In order to capture all the savings that they generate, PBMs would have to have a profit margin in excess of 47 percent.

Despite generating lower spending, PBMs are able to generate greater formulary surplus. In our model, formulary surplus (defined in equation 7) represents the revealed-preference welfare generated by a given formulary arrangement net of other costs paid by the patient (i.e., copays and premiums). Because patient costs are mechanically equivalent to manufacturer and PBM revenue, formulary surplus can be loosely interpreted as overall societal welfare within the limits of our model.[44] Formulary surplus under the 33 percent coinsurance counterfactual falls by close to 20 percent, or $46 billion (in 2000 dollars). The decline is caused by an increase in cost sharing that in turn reduces access to statins.

Figure 6 provides a year-by-year comparison between the baseline model and the 33 percent coinsurance counterfactual, plotting the percentage reduction in payments to drug manufacturers. The savings rate increases rapidly as Zocor faces branded competitors in Lipitor and Crestor, and it reaches its peak right before the entry of generic Zocor. It then falls after each generic entry event. The results reflect the notion that PBMs are most

[43] Appendix A.2 provides a detailed discussion of our methodology, which requires making assumptions about the average discount rate off of list price, and PBM margins from sources other than drug savings. Our manually collected data on PBM financials and market shares can be found here: https://sites.google.com/view/jfeng/data. However, our estimate is not necessarily accurate, because PBMs have multiple lines of business beyond price negotiation and formulary setting (such as the aforementioned claims processing and mail order pharmacies). The accuracy of our estimates relies either on margins being similar across activities or price negotiations dominating PBM activities.

[44] This interpretation treats payments to manufacturers and PBMs as pure transfers. It also ignores consumption and expenditures outside of the statin market.

Table 5: Summary of Counterfactual Market Outcomes

	Average Unit Price ($)			Payments ($bill)			Formulary
Scenario	Zocor	Lipitor	Crestor	Drug Comp	PBM	Total	Surplus
Baseline	2.79	1.91	1.94	164.3	10.7	175.0	252.4
Panel A – No PBM							
33% Coinsurance	4.96	5.17	1.74	242.5	0	242.5	206.4
49% Coinsurance	3.33	3.47	1.19	177.8	0	177.8	202.1
Full margin pricing	2.60	2.67	1.00	157.5	0	157.5	226.9
Partial margin pricing	3.41	3.53	1.27	188.3	0	188.3	217.4
Panel B – Committing to Formulary Structures							
Exclude $\leqslant 1$	4.77	3.82	2.49	269.0	0	269.0	253.0
Commit one exclusive	1.85	1.01	1.39	97.2	20.3	117.6	237.2
Commit one preferred	2.33	1.89	1.77	154.1	12.4	166.4	244.9
Only Preferred Tier	2.82	1.90	1.79	166.2	10.7	176.9	256.7

Notes: Statistics summarizing the results from the counterfactual exercises. Average prices are computed in a weighted manner across the 18 years from 1996–2013, are scaled to daily prices, and only include branded prescriptions. Total payments is computed by multiplying equilibrium negotiated prices by market share and by market size (assuming people who choose a drug take it 90 percent of the days in a year), and summing over all years. Counterfactual PBM payments are set to 16 percent of savings generated, reflecting the retention rate of savings in the baseline scenario.

effective in markets that contain multiple branded competitors and no generics. In the absence of PBMs, manufacturers pursue more pronounced "invest-then-harvest" pricing strategies. The presence of generics makes PBMs more effective negotiators (given that generics offer a cheap and always available alternative for formularies) but also limits the scope of potential savings because generics increase competition under most market structures.[45]

We also simulate two additional scenarios. First, we calculate how high coinsurance rates would have to be to match spending in the current equilibrium. We find that a 49% coinsurance rate would lead to roughly the same amount of spending. Interestingly, this increase in coinsurance only leads to a small decline in formulary surplus relative to the

[45]Savings rates are negative in the first few years for two reasons. First, under the coinsurance model, a lower price always results in a lower out-of-pocket cost, which can increase market share. Conversely, under the formulary model, the lowest possible copay is the copay associated with the preferred tier. Once a drug has achieved the preferred tier, there is no way to lower the out-of-pocket cost. This rules out some of the more extreme pricing strategies as manufacturers try to build a loyal patient base early on. Second, building a base is more profitable under the coinsurance model, where the firm can charge a very high price with only a small effect on out-of-pocket cost, rather than in the PBM model, where the PBM can exclude a drug from the formulary if the price is too high.

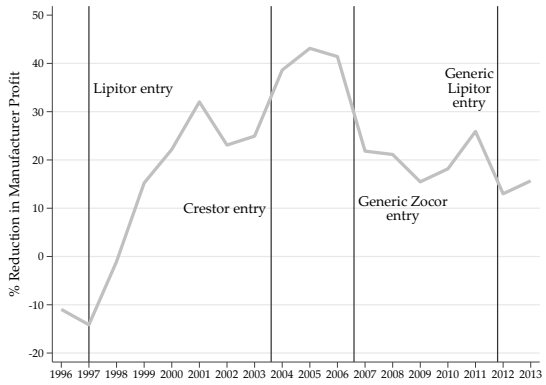


Figure 6: Percentage Reduction in Drug Manufacturer Profits by Year

Notes: A plot of the reduction in payments to drug manufacturers when moving from the 33 percent coinsurance counterfactual to the current baseline.

33% coinsurance scenario, as we find that, in equilibrium, manufacturers would reduce prices so that pocket costs are almost unchanged.

Second, we simulate the introduction of margin pricing (Einav et al., 2016). Under margin pricing, payers cover the entire cost of the drug with the lowest price, but patients must pay the difference in price to access other drugs. Implementing margin pricing leads to a 10 percent reduction in spending relative to the baseline scenario, but results in lower patient welfare due to higher cost sharing associated with some options. Even a partial version of margin pricing, in which patients only pay 50 percent of the difference, achieves slightly lower patient welfare.[46]

5.2 Formulary Structure and Market Outcomes

In our second set of counterfactuals, we vary the set of possible formularies that PBMs can choose, in order to study the underlying mechanisms that make PBMs effective. We report results in Panel B of Table 5. We assume throughout that PBMs are able to capture 13.7 percent of the savings they generate relative to the 33 percent coinsurance counterfactual discussed in Section 5.1, just as they do in the current equilibrium.

We simulate three alternative scenarios. First, we assume that the PBM can exclude at most one drug from coverage ("Exclude $\leqslant 1$"). Under this structure, spending increases by over 60 percent relative to the baseline scenario. At the same time, the impact on formulary surplus is barely noticeable (a 0.2 percent increase) because PBMs almost never

[46] See Table 4.5 of the WHO report "Medicines Reimbursement Policies in Europe." (https://www.euro.who.int/__data/assets/pdf_file/0011/376625/pharmaceutical-reimbursement-eng.pdf, retrieved February 9, 2023).

exclude two or more drugs under the baseline model.[47] The results under this counterfactual are consistent recent empirical evidence on the inflationary impact of Medicare Part D protected class rules, which mandate coverage of "substantially all" drugs in certain classes, leading to higher prices (Hwang et al., 2019; Kakani et al., 2020). Our analysis suggests that this increase in spending might not come with significant gains in patient welfare because very few drugs would be excluded in equilibrium even when exclusion is possible. In equilibrium, exclusion simply serves as an off-equilibrium threat that prevents manufacturers from increasing their bids.

Second, we consider a case in which PBMs can commit to excluding all but one branded drug in advance. This counterfactual mimics the approach taken by Louisiana's Medicaid program, which negotiated an exclusive contract with one Hepatitis-C drug. The results ("Commit one exclusive") imply a 33 percent reduction in spending relative to baseline. Formulary surplus falls, but only by about 6 percent. This suggests that PBMs could be even more aggressive in their efforts to restrict formularies. However, we note that this result has limited external validity—statins are considered close substitutes, whereas drugs in other classes might exhibit greater horizontal differentiation.

Third, we show that leveraging formulary tiers can also generate savings, but on a much smaller scale relative to exclusion. We run counterfactuals in which PBMs commit to either i) only including one branded drug in the preferred tier ("Commit one preferred"), or ii) removing the non-preferred tier altogether ("Only Preferred Tier"). The results of the first counterfactual suggest that committing to only including one drug in the preferred tier leads to about a five percent reduction in overall spending. The results of the second show that removing the threat of assignment to the non-preferred tier would only marginally increase spending. Both of these numbers represent marginal changes relative to the effects of increasing or decreasing the threat of exclusion. We also note that while altering the exclusion affects all drugs in the same way, changing the tier structure has a very heterogeneous impact on different drugs. Both tiering counterfactuals result in significantly lower prices for Crestor, the last drug in the class, because the changes to the tiering structure make it more likely that Crestor will be excluded. However, removing the non-preferred tier increases Zocor's price because it essentially forces the PBM to cover Zocor at low copays in the initial periods, when Zocor is the only drug in the class. Conversely, Lipitor's prices are essentially unaffected.

Overall, our estimates point to the importance of the exclusion threat as a key tool for PBMs to extract savings, albeit with two caveats. First, our model does not factor in

[47] Appendix D.3 provides evidence that PBMs choose to exclude multiple drugs less than 2 percent of the time, except in the year before generic entry events.

competition between PBMs, which will generally weaken competition between drug manufacturers in the case where all PBMs commit to covering one branded drug. If a drug is excluded by one PBM, a different PBM might have stronger incentives to cover it in order to differentiate itself, tempering the effect we estimate in the model. Second, we assume that formulary exclusion completely deters consumers from choosing the drug. In practice, health plans occasionally cover excluded drugs for existing patients. Some patients may also choose to pay the full price for a drug.

5.3 Demand Inertia and Market Outcomes

In our third counterfactual, we investigate the role played by demand inertia in shaping equilibrium prices. To do so, we set the switching cost parameters in the demand system to zero and recompute the equilibrium. We find that overall price levels are lower, and that incumbent prices fall in response to branded and generic competitor entry, from a high of $1.60 when Zocor is the only drug in class, to about $1.30 when all three drugs are available. Figure 7 presents the computed equilibrium prices. We find significant drops in incumbent price levels after the entry of both Lipitor and Crestor. Crestor's price increases towards the end of the period, when both generic Zocor and generic Lipitor are on the market. This pricing strategy is a result of competition with two generic products, who are automatically assigned to a low copay tier. In a model without inertia, this implies that only patients with strong idiosyncratic preferences for Crestor will choose it. Therefore, Crestor's profit maximizing strategy is to extract higher surplus from this smaller set of patients.

The qualitative price pattern from 1996 to 2006 (before the entry of generic Zocor) is similar to that observed in the Hepatitis C market, suggesting that in the absence of inertia, the hidden impact of Pharmacy Benefit Managers would be easier to recognize in the data. The recent treatments in Hepatitis C are cures rather than maintenance drugs, and each drug cures almost all patients without severe side effects, making access to multiple options less valuable. The data from SSR Health, presented in Figure A3 of Appendix B.4, shows that net-of-rebate prices have fallen significantly as branded competitors have entered, unlike the average net price trends documented earlier.

6 Conclusion

This paper uses data on prices and formulary coverage and a novel structural model to analyze the role PBMs play in drug markets. Reduced-form evidence does not provide a clear

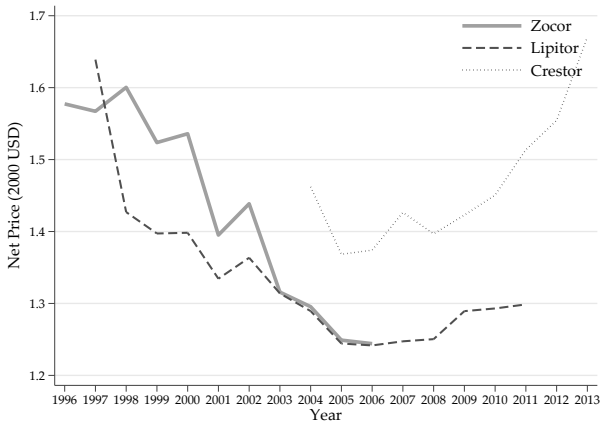


Figure 7: Pricing Counterfactual, No Inertia

Notes: A plot of predicted equilibrium prices if demand system had no inertia terms, holding everything else constant.

answer as to whether PBMs help reduce prices and spending on prescription drugs. However, counterfactuals based on a structural model of the anti-cholesterol market show that PBMs reduce payments to drug manufacturers by 32 percent relative to a direct pricing system with coinsurance rates. Even though we estimate that PBMs capture a fraction of these savings, they still reduce overall spending by 28 percent relative to a counterfactual world where manufacturers set prices and patients pay a 33% coinsurance rate. Further analysis suggests that the threat of exclusion is a key tool for PBMs, even though this threat is rarely realized in equilibrium. As a result, policies limiting the ability of PBMs to exclude drugs impose a large burden on spending, without substantially improving patient welfare.

It is important to note that our analysis focuses on the role played by PBMs in reducing spending, and is not meant to provide a complete picture of the role of PBMs in the pharmaceutical market. For example, many PBMs influence the distribution of prescription drugs both directly, through in-house mail-delivery pharmacies, and indirectly, by forming preferred retail pharmacy networks (Starc and Swanson, 2021). Most importantly, our paper does not model list prices and rebates separately, but rather focuses on net prices. While this simplification is irrelevant for total spending, list prices can affect how the cost of drugs is passed through to consumers. Recent media narratives suggest that PBM contracts are structured in a way that incentivizes manufacturers to offer high list prices and high rebates, which can increase out of pocket costs for patients and reduce access

to treatment.[48]

We note two other minor caveats. First, spending reduction, while welfare-enhancing in the short-run, may adversely affect innovation incentives. Recent evidence shows that the use of formulary exclusion is tied to a reduction in R&D investment, but that the shift is focused among drug candidates in therapeutic areas with more available therapies (Agha et al., 2022). More generally, there is no universally accepted estimate for the elasticity of pharmaceutical innovation with respect to revenue, and other frictions in the pharmaceutical markets—like the presence of insurance and monopoly power granted by the patent system—could actually imply that innovation incentives are already higher than optimal. Second, our welfare estimates are based on a revealed preference framework rather than direct health outcomes. Patients may suffer from formulary exclusions that force them to switch drugs in ways that our model cannot capture.

Overall, we view our paper as a stepping stone toward the analysis of broader issues in the pharmaceutical market and beyond. Our model and estimation framework could be adapted to study dynamic issues in a variety of settings such as: (i) competition between biologic and biosimilar drugs; (ii) demand for antibiotics; and (iii) competition in markets where the introduction of new technology (like a cure, or a vaccine) leads to a shrinking patient population over time. Outside of the pharmaceutical market, our model can be used to study how other purchasing organizations leverage demand structure to negotiate discounts. The literature has documented demand inertia and rebate negotiations in many industries, such as hospital services and supermarkets, making the model's structure generally applicable. Going one step further, the model could be applied to study the incentives for buyers to manipulate demand structure, such as the number of copay tiers in insurance plans or a supermarket's layout, in order to achieve lower prices through bargaining with suppliers.

[48] See, e.g., discussion of Amgen's pricing strategy for its Humira biosimilar: https://www.drugchannels.net/2023/02/the-warped-incentives-behind-amgens.html, retrieved February 9, 2023.

References

Agha, Leila, Soomi Kim, and Danielle Li, "Insurance Design and Pharmaceutical Innovation," *American Economic Review: Insights*, 2022, *4* (2), 191–208.

Aksoy-Pierson, Margaret, Gad Allon, and Awi Federgruen, "Price competition under mixed multinomial logit demand functions," *Management Science*, 2013, *59* (8), 1817–1835.

Alpert, Abby, "The anticipatory effects of Medicare Part D on drug utilization," *Journal of Health Economics*, 2016, *49*, 28–45.

Arcidiacono, Peter, Paul B. Ellickson, Peter Landry, and David B. Ridley, "Pharmaceutical followers," *International Journal of Industrial Organization*, 2013, *31*, 538–553.

Besanko, David, Jean Pierre Dube, and Sachin Gupta, "Competitive Price Discrimination Strategies in a Vertical Channel Using Aggregate Retail Data," *Management Science*, 2003, *49* (9), 1121–1138.

_ , Ulrich Doraszelski, Yaroslav Kryukov, and Mark Satterthwaite, "Learning-by-Doing, Organizational Forgetting, and Industry Dynamics," *Econometrica*, 2010, *78* (2), 453–508.

Bokhari, Farasat A S and Gary M. Fournier, "Entry in the ADHD drugs market: Welfare impact of generics and me-too's," *Journal of Industrial Economics*, 2013, *61*, 339–392.

Brot-Goldberg, Zarek C, Catherine Che, and Benjamin R Handel, "Pharmacy Benefit Managers and Vertical Relationships in Drug Supply: State of Current Research Pharmacy Benefit Managers and Vertical Relationships in Drug Supply: State of Current Research," 2022.

Cabral, Luís, "Dynamic pricing in customer markets with switching costs," *Review of Economic Dynamics*, 2016, *20*, 43–62.

Conti, Rena M, Brigham Frandsen, Michael L Powell, and James B Rebitzer, "Common Agent or Double Agent? Pharmacy Benefit Managers in the Prescription Drug Market," 2021.

Crawford, Gregory S and Matthew Shum, "Uncertainty and Learning in Pharmaceutical Demand," *Econometrica*, 2005, *73*, 1137–1173.

Dafny, Leemore, Kate Ho, and Edward Kong, "How Do Copayment Coupons Affect Branded Drug Prices and Quantities Purchased?," 2022.

DiMasi, Joseph A., Ronald W. Hansen, and Henry G. Grabowski, "The price of innovation: New estimates of drug development costs," *Journal of Health Economics*, 2003, *22*, 151–185.

Doganoglu, Toker, "Switching costs, experience goods and dynamic price competition," *Quantitative Marketing and Economics*, 2010, *8* (2), 167–205.

Doraszelski, Ulrich and Ariel Pakes, "A Framework for Applied Dynamic Analysis in IO," in "Handbook of Industrial Organization," Vol. 3 2007, pp. 1887–1966.

Dube, Jean-Pierre, GÃŒnter J Hitsch, and Peter E Rossi, "Do Switching Costs Make Markets Less Competitive?," *Journal of Marketing Research*, 2009, *46*, 435–445.

Dubois, Pierre and Morten Saethre, "On the Role of Parallel Trade on Manufacturers and Retailers Profits in the Pharmaceutical Sector," 2016, (1), 1–45.

_ , Ashvin Gandhi, and Shoshana Vasserman, "Bargaining and International Reference Pricing in the Pharmaceutical Industry," 2019, pp. 1–75.

Dunn, Abe, "Drug Innovations and Welfare Measures Computed from Market Demand: The Case of Anti-Cholesterol Drugs," *American Economic Journal: Applied Economics*, 2012, *4*, 167–189.

Einav, Liran, Amy Finkelstein, and Heidi Williams, "Paying on the margin for medical care: Evidence from breast cancer treatments," *American Economic Journal: Economic Policy*, 2016, *8*, 52–79.

Ericson, Keith M., "Consumer inertia and firm pricing in the medicare part d prescription drug insurance exchange," *American Economic Journal: Economic Policy*, 2014, *6*, 38–64.

Feng, Josh, "History-Dependence in Drug Demand: Identification and Implications for Entry Incentives," *Review of Economics and Statistics*, 2023.

Fleitas, Sebastian, "Dynamic Competition and Price Regulation When Consumers Have Inertia : Evidence from Medicare Part D," 2016, pp. 1–70.

Frank, Richard G. and David S. Salkever, "Pricing, Patent Loss and the Market for Pharmaceuticals," *Southern Economic Journal*, 1992, *59* (2), 165.

Geruso, Michael, Timothy Layton, and Daniel Prinz, "Screening in contract design: Evidence from the ACA health insurance exchanges," *American Economic Journal: Microeconomics*, 2019, *11*, 64–107.

Gowrisankaran, Gautam, Aviv Nevo, and Robert Town, "Mergers when prices are negotiated: Evidence from the hospital industry," *American Economic Review*, 2015, *105*, 172–203.

Grabowski, Henry, Genia Long, Richard Mortimer, and Ani Boyo, "Updated trends in US brand-name and generic drug competition," *Journal of Medical Economics*, 2016, *19* (9), 836–844.

Handel, Benjamin R., "Adverse Selection and Inertia in Health Insurance Markets :," *American Economic Review*, 2013, *103*, 2643–2682.

Hartung, D. M., Dennis N. Bourdette, Sharia M. Ahmed, and Ruth H. Whitham, "The cost of multiple sclerosis drugs in the US and the pharmaceutical industry: Too big to fail?," *Neurology*, 2015, *85* (19), 1728.

Ho, Kate and Robin S. Lee, "Insurer Competition in Health Care Markets," *Econometrica*, 2017, *85*, 379–417.

_ **and** _ , "Equilibrium Provider Networks: Bargaining and Exclusion in Health Care Markets," *American Economic Review*, 2019, *109* (2), 473–522.

_ **, Joseph Hogan, and Fiona Scott Morton**, "The Impact of Consumer Inattention on Insurer Pricing in the Medicare Part D Program," 2017.

Honka, Elisabeth, "Quantifying search and switching costs in the US auto insurance industry," *RAND Journal of Economics*, 2014, *45*, 847–884.

Howard, David H., Peter B. Bach, Ernst R. Berndt, and Rena M. Conti, "Pricing in the market for anticancer drugs," *Journal of Economic Perspectives*, 2015, *29* (1), 139–162.

Hwang, Thomas J., Stacie B. Dusetzina, Josh Feng, Luca Maini, and Aaron S. Kesselheim, "Price Increases of Protected-Class Drugs in Medicare Part D, Relative to Inflation, 2012-2017," *JAMA - Journal of the American Medical Association*, 2019, *322*, 267–269.

Kakani, Pragya, Michael Chernew, and Amitabh Chandra, "Rebates in the Pharmaceutical Industry: Evidence From Medicines Sold in Retail Pharmacies in the U.S.," 2020.

Klemperer, Paul, "The Competitiveness of Markets with Switching Costs," *RAND Journal of Economics*, 1987, *18*, 138–150.

_ , "Markets with Consumer Switching Costs," *The Quarterly Journal of Economics*, 1987, *102* (2), 375–394.

Lavetti, Kurt and Kosali Simon, "Strategic formulary design in Medicare Part D plans," *American Economic Journal: Economic Policy*, 2018, *10*, 154–192.

Lee, Chung-Ying, "Brand Loyalty and Learning in Pharmaceutical Demand," *Academia Economic Papers*, 2016, *44*, 333–359.

Lipton, Helene L, David H Kreling, Ted Collins, and Karen C Hertz, "Pharmacy Benefit Management Companies: Dimensions of Performance," Technical Report 1999.

Lu, Z John and William S Comanor, "Strategic Pricing of New Pharmaceuticals," *Review of Economics and Statistics*, 1998, *80*, 108–118.

Murphy, Kevin M. and Robert H. Topel, "Estimation and inference in two-step econometric models," *Journal of Business and Economic Statistics*, 1985, *3* (4), 88–97.

Newey, Whitney K, "Efficient Estimation of Limited Dependent Variable Models with Endogenous Explanatory Variables," *Journal of Econometrics*, 1987, *36*, 231–250.

Newey, Whitney K. and Daniel McFadden, "Large sample estimation and hypothesis testing," *Handbook of Econometrics*, 1994, *4*, 2111–2245.

Olssen, Alexander L and Mert Demirer, "Drug Rebates and Formulary Design: Evidence from Statins in Medicare Part D," 2023.

Raval, Devesh and Ted Rosenbaum, "Why Do Previous Choices Matter for Hospital Demand? Decomposing Switching Costs from Unobserved Preferences," *Review of Economics and Statistics*, 2018, *100* (5), 906–915.

Shcherbakov, Oleksandr, "Measuring consumer switching costs in the wireless industry," *RAND Journal of Economics*, 2016, *47*, 366–393.

Sinkinson, Michael and Amanda Starc, "Ask Your Doctor? Direct-to-Consumer Advertising of Pharmaceuticals," *The Review of Economic Studies*, 2018.

Starc, Amanda and Ashley Swanson, "Preferred Pharmacy Networks and Drug Costs," *American Economic Journal: Economic Policy*, aug 2021, *13* (3), 406–446.

Sudhir, K., "Structural analysis of manufacturer pricing in the presence of a strategic retailer," *Marketing Science*, 2001, *20* (3), 244–264.

Sun, Liyang and Sarah Abraham, "Estimating dynamic treatment effects in event studies with heterogeneous treatment effects," *Journal of Econometrics*, 2020, (xxxx), 1–25.

Terza, Joseph V., Anirban Basu, and Paul J. Rathouz, "Two-stage residual inclusion estimation: Addressing endogeneity in health econometric modeling," *Journal of Health Economics*, 2008, *27*, 531–543.

Thomadsen, Raphael, "Existence of pure strategy equilibria among geographically dispersed firms," *Economics Bulletin*, 2005, *12* (1).

Appendix to "Demand Inertia and the Hidden Impact of Pharmacy Benefit Managers"

Josh Feng, *University of Utah – David Eccles School of Business*
Luca Maini, *Harvard Medical School*

August 2023

A Additional Background on Pharmacy Benefit Managers

A.1 Evidence on concentration and prevalence of PBMs

The PBM industry is highly concentrated. As of 2019, the three largest PBMs (CVS-Aetna, Express Scripts, and Optum Rx) control approximately 80% of the market. During the time span covered by our analysis, the market was less concentrated, but not dramatically so. Exact data on market share is hard to come by, but Lipton et al. (1999) report that in 1997 the top four PBMs combined to reach approximately a 67% market share.[49] We also calculated a revenue-based measure of HHI using the PBM financial filings we describe in the main text and in the next section. The index fluctuates between 0.45 and 0.3 between 2000 and 2015 without a clear trend (Figure A1).

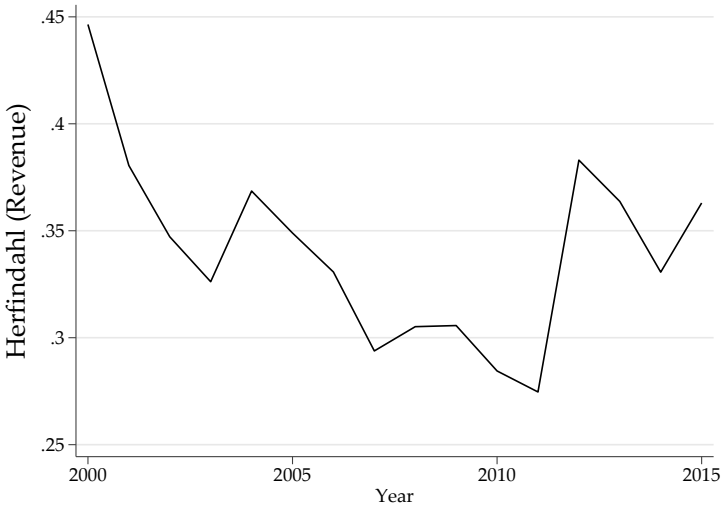


Figure A1: Revenue-based HHI in PBM market

[49] Authors' own calculations based on numbers in Table 1.

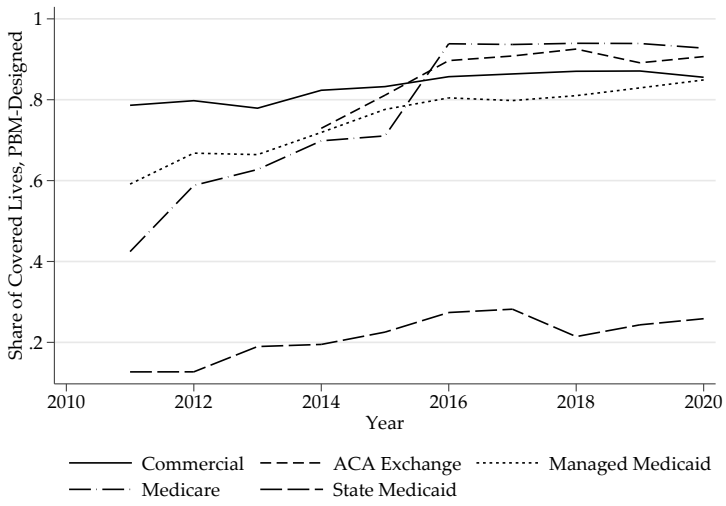


Figure A2: Fraction of lives with PBM-designed (custom or national) formularies

Notes: A plot of the share of lives in each insurance segment that is covered by a plan where the formulary is designed by a PBM, based on MMIT data.

The function of PBMs may differ slightly based on the market they serve. The majority of self-insured employers rely on PBMs to handle all aspects of prescription drug insurance, including negotiating prices, determining drug coverage through national formularies, and processing prescription drug claims. On the other hand, most health insurance companies only use PBMs for price negotiation, and design their own custom formularies. In the case of Medicare Part D, PBMs may also offer plans alongside traditional insurance providers. Lastly, some payers use PBMs solely for claim management, and make independent decisions regarding prices and coverage. Figure A2 shows the fraction of covered lives for whom PBMs negotiate drug prices across the commercial, Medicare, Health Exchanges, and Medicaid channels.

A.2 Description of the Bidding Structure

To guide our modeling decisions, we spoke with several former PBM industry executives and regulators and reviewed the recent Senate investigation report on insulin manufacturers.[50] In this section, we summarize what we learned.

The process of setting prices occurs separately for each drug class.[51] PBMs list several potential formulary arrangements to each drug manufacturer. Each formulary ar-

[50]The report is available at https://www.finance.senate.gov/download/grassley-wyden-insulin-report, retrieved February 13, 2023.

[51]Cross-market agreements are possible but are not systematically used during the bidding process.

rangement specifies the coverage tier for the drug owned by the manufacturer and for its therapeutic alternatives.[52] Manufacturers then send bids indicating the rebate they are willing to grant for each arrangement. In general, manufacturer surrender higher rebates for formularies that cover their drugs more generously and limit coverage for their competitors.

Based on these bids, PBMs create a set of "national" formularies, which they offer to payers. Payers can either choose a formulary offered by the PBM or create their own custom formulary by selecting arrangements from different drug classes. The price of the custom formulary is determined by the bids already submitted by each manufacturer.

PBMs face a trade-off in choosing national formularies, as they prefer lower prices but face costs when setting restrictive formularies. In particular, setting restrictions on popular drugs in chronic markets can result in payer dissatisfaction and may not lead to volume reduction due to consumer inertia and limited sensitivity to cost sharing.PBM Financial Filings

We use PBM financial filing information to estimate their gross profit margins.[53] PBM financial filings are complex. Some PBMs (e.g., Caremark) are divisions within larger companies. In addition, PBMs do not break down their revenues by drug market. However, they generally charge payers on a per member or per script basis. Finally, PBM revenues and margins are hard to measure, because the accounting measures incorporate payments for drugs that come from payers and are passed on to drug companies.[54]

We use the following process to build a list of PBMs and associated financial numbers:

1. We start off with a list of the largest PBMs in the 2011–2020 period based on MMIT data: Express Scripts, Medco, CVS/Caremark, Optum, Catamaran, Humana

2. For each year between 2000–2019 we look through the financial filings of each of these five PBMs, and check which competitors are mentioned. Using this approach, we add Catalyst (active 2000–2011) and NextRx (active 2000–2009). Others competitors mentioned include Prime Therapeutics and MedImpact, but neither is publicly traded, so we cannot collect information on them.

[52]The contracts we have seen do not mention specific competitors. However, they specify the number of other competing drugs in the same tier of coverage as the drug owned by the manufacturer in question. These contracts are found in the Appendices to the insulin report, which can be accessed through links available at the end of the report.

[53]The data used in our calculations is available online at https://sites.google.com/view/jfeng/data.

[54]A 2017 article from Bloomberg provides a more detailed discussion of this issue: bloomberg.com/news/articles/2017-03-31/drug-middlemen-have-slim-profit-margins-just-ask-them, retrieved February 17, 2023.

3. For each of the seven PBMs we collect revenue, gross profit, SG&A (which approximates overhead costs), net income, and operating income from financial filings

4. We then compute variable cost as the difference between revenue and gross profit

The revenues number include both rebates from drug companies and fees payed by clients, whereas profit is what the PBM retains. Hence, variable cost should primarily reflects payments to drug companies, plus some small processing costs for prescription management. If we assume that other costs are small enough, we can approximate the PBMs gross profit margin (i.e., not including fixed costs) as

$$\text{PBM Margin} = \frac{\text{PBM Gross Profit}}{\text{PBM Revenues} - \text{PBM Gross Profit}} = \frac{\text{PBM Revenues} - \text{PBM Variable Cost}}{\text{PBM Variable Cost}}$$

Our estimated margin is approximately 6.5 percent, and is fairly consistent over time for the largest PBMs. This margin comes in part from retaining a fraction of the rebates, and in part from fees charged to payers. Express Scripts claims that it keeps about 10 percent of all rebates and passes the rest onto payers.[55] If we normalize list prices to \$1 and assume an average discount of 30 percent, Express Scripts would keep \$0.03 of the \$0.30 discount, which would equal 4.3 percent of variable cost (\$0.70).

B Additional Reduced-Form Evidence

B.1 Alternative Market Definitions

We present evidence that the empirical facts documented in Section 2 are robust to definition of drug class. The results in the main text use ATC-4. Drugs with the same ATC-4 generally have very similar chemical properties and treat the same disease. ATC-4 is commonly used in the pharmaceutical literature to define classes (see, e.g., Dubois et al., 2019).

Here, we show results using three alternative class measures. The first is the ATC-3, which is the next higher tier in the ATC classification. The next two are proprietary classifications that are included in our MMIT and SSR Health data.

[55] Wall Street Journal article, "Drugmakers Point Finger at Middlemen for Rising Drug Prices," available at https://www.wsj.com/articles/drugmakers-point-finger-at-middlemen-for-rising-drug-prices-1475443336, retrieved February 17, 2023.

Correlation Between Net Price and Formulary Coverage We run the cross-sectional regressions using the three classifications not reported in the main text. Table A1 reports the results. We consistently find that net price has an impact on formulary coverage generosity, particularly the rate of preferred coverage. List price, on the other hand, has a noisily-measured correlation with formulary generosity in all specifications without any consistency in the magnitude or sign of the coefficient.

Table A1: Prices vs. Formulary Correlations

(a) ATC-3

	Frac. Preferred		Frac. Covered	
Log(net price)	-0.0323**	-0.0344*	-0.00888**	-0.00868
	(0.00616)	(0.0134)	(0.00246)	(0.00625)
Log(WAC)		0.00303		-0.000293
		(0.0150)		(0.00742)
Fixed Effects	Class-by-Year			
Observations	2520	2520	2520	2520

(b) MMIT Class

	Frac. Preferred		Frac. Covered	
Log(net price)	-0.0329*	-0.0582*	-0.0126*	-0.0233*
	(0.0140)	(0.0245)	(0.00631)	(0.0107)
Log(WAC)		0.0397		0.0168
		(0.0348)		(0.0119)
Fixed Effects	Class-by-Year			
Observations	2740	2740	2740	2740

(c) SSR Health Class

	Frac. Preferred		Frac. Covered	
Log(net price)	-0.0226**	-0.0286**	-0.00837**	-0.00861
	(0.00514)	(0.00857)	(0.00296)	(0.00607)
Log(WAC)		0.00889		0.000362
		(0.0119)		(0.00665)
Fixed Effects	Class-by-Year			
Observations	2425	2425	2425	2425

Notes: Estimates of Equation 1, comparing within class-by-year cells. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Response to Branded and Generic Competitor Entry Next, we present results on response to entry in Table A2. We report results from a simplified version of Equation 2

$$y_{it} = \alpha_i + \gamma_t + \beta_1 \times \mathtt{BrandedEntryCounter}_{it} + \beta_2 \times \mathtt{GenericEntryCounter}_{it} + \varepsilon_{it} \quad \text{(A1)}$$

We consistently find across all the class definitions that entry events have a minimal effect on net prices. There is weak evidence in some cases of a small drop in net price in response to competitor loss-of-exclusivity (generic versions of competitors enter), but the estimates rule out any large effects. The only consistent and significant result we find is a units response to competitor loss-of-exclusivity. The units sold drops somewhere between 10–20 percent for the incumbent branded drug.

Table A2: Response to Competitor Entry

(a) ATC-4

	log(WAC)	log(net)	log(units)	Preferred	Covered
Competitor	-0.047**	0.007	0.097**	0.002	0.013**
Launch Counter	(0.015)	(0.023)	(0.035)	(0.004)	(0.003)
Competitor	0.0079	-0.014	-0.118+	-0.007	0.003
LOE Counter	(0.023)	(0.040)	(0.069)	(0.010)	(0.005)
Drug FE	x	x	x	x	x
Year FE	x	x		x	x
Age FE			x		
Observations	6706	5392	6701	2941	2941

(b) ATC-3

	log(WAC)	log(net)	log(units)	Preferred	Covered
Competitor	-0.044**	0.016	0.094**	0.003	0.009**
Launch Counter	(0.011)	(0.018)	(0.030)	(0.004)	(0.002)
Competitor	0.0082	-0.041	-0.110**	-0.001	-0.001
LOE Counter	(0.023)	(0.025)	(0.035)	(0.006)	(0.003)
Drug FE	x	x	x	x	x
Year FE	x	x		x	x
Age FE			x		
Observations	6706	5392	6701	2941	2941

(c) MMIT Class

	log(WAC)	log(net)	log(units)	Preferred	Covered
Competitor	0.005	0.001	0.093	-0.003	-0.003
Launch Counter	(0.028)	(0.041)	(0.080)	(0.012)	(0.007)
Competitor	0.030	0.019	-0.448**	-0.015	-0.012*
LOE Counter	(0.031)	(0.039)	(0.088)	(0.012)	(0.005)
Drug FE	x	x	x	x	x
Year FE	x	x		x	x
Age FE			x		
Observations	7061	5796	7055	3372	3372

(d) SSR Health Class

	log(WAC)	log(net)	log(units)	Preferred	Covered
Competitor	-0.027*	0.008	-0.006	-0.0002	0.008**
Launch Counter	(0.012)	(0.019)	(0.033)	(0.004)	(0.003)
Competitor	0.053**	0.035	-0.149**	0.006	-0.002
LOE Counter	(0.017)	(0.034)	(0.052)	(0.007)	(0.004)
Drug FE	x	x	x	x	x
Year FE	x	x		x	x
Age FE			x		
Observations	5883	5211	5881	2609	2609

Notes: Estimates of Equation A1, using different class definitions to identify competitors. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Table A3: Prices vs. Formulary Correlations with Manufacturer Fixed Effects

	Frac. Preferred		Frac. Covered	
Log(net price)	-0.0266**	-0.0512**	-0.0127	-0.0230
	(0.00666)	(0.0149)	(0.00806)	(0.0160)
Log(WAC)		0.0322+		0.0136
		(0.0169)		(0.0184)
Fixed Effects	Class-by-Year, Manufacturer			
Observations	2520	2520	2520	2520

Notes: Estimates of Equation 1, comparing within class-by-year cells. The results here include a manufacturer fixed effect. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

B.2 Accounting for Manufacturer Fixed Effects

We provide additional specifications related to the cross-sectional tradeoff within class between net prices and formulary coverage. As noted in the main text, our evidence should not be interpreted causally. Here, we included manufacturer fixed effects to try to account for systematic differences across companies in drug quality and negotiation experience, which they could apply across all of their drugs in different disease areas. Table A3 presents the estimates.

B.3 Suggestive Evidence of the Strategic Impact of Inertia

To provide suggestive evidence on how inertia affects incumbents' response to entry, we run a regression where we allow the price response to vary linearly by incumbent age at time of competitor entry. The idea is that incumbents that have been on the market for longer will have a larger patient base, and therefore are less likely respond to competitor entry by lowering prices and trying to obtain better formulary coverage. Formally, we run the following specification

$$y_{it} = \alpha_i + \gamma_t + \beta_1 \times \texttt{BrandedEntryCounter}_{it} + \beta_2 \times \texttt{BrandedEntryCounter}_{it} \times \texttt{Age}_i + \varepsilon_{it} \tag{A2}$$

The first term allows for a response size of $\beta_1$at age zero. The second term, computed by incrementing the counter by the number of years an incumbent has been on the market ("age") at the time of competitor entry, allows for the response size to vary linearly with

the incumbent's age at the time of entry. Table A4 reports the results. Confirming our intuition, we find a negative, albeit statistically insignificant, net price response at age zero of about six percent. This negative effect is then offset by a positive interaction term. The coefficients combine to suggest that the negative price response goes away by age seven. We find the opposite result for preferred coverage. Younger incumbents respond to entry by obtaining more favorable formulary position. This is then offset by a negative interaction term, so older incumbents appear to have weaker incentives to obtain better formulary position in response to competitor entry.

Table A4: Incumbent Response Variation by Age at Entry

	log(net)		Preferred	
Competitor Launch Counter	0.00614	-0.0599	0.00143	0.0319**
	(0.0242)	(0.0431)	(0.00339)	(0.00664)
Age Scaled Competitor Launch Counter		0.00833+		-0.00363**
		(0.00473)		(0.000721)
Drug FE	x	x	x	x
Year FE	x	x	x	x
Observations	5392	5392	2941	2941

Notes: Estimates of Equation A2. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

B.4 Hepatitis C Case Study

Finally, we present evidence from SSR data showcasing price trends in the Hepatitis C market. In this market, demand inertia is not relevant thanks to the recent introduction of novel therapies that can cure the disease in a few weeks. The first cure for Hepatitis C, Sovaldi, entered the market in 2013, and was quickly followed by a series of competitors. Figure A3 shows that net prices have declined each year since 2013.[56]

[56]The declining prices are exacerbated by the presence of a different kind of dynamic effect: market shrinking due to the exhaustion of the stock of patients. Prior to 2013, Hepatitis C treatments were much less effective and had significant side effects. As a result, many patients simply chose to remain untreated. When the new cures arrived on the market, the stock of potential patients was very large, but it quickly shrank because the number of treated patients was much greater than the flow of new patients. This environment likely exacerbated price competition as firms raced to treat as many patients as possible.

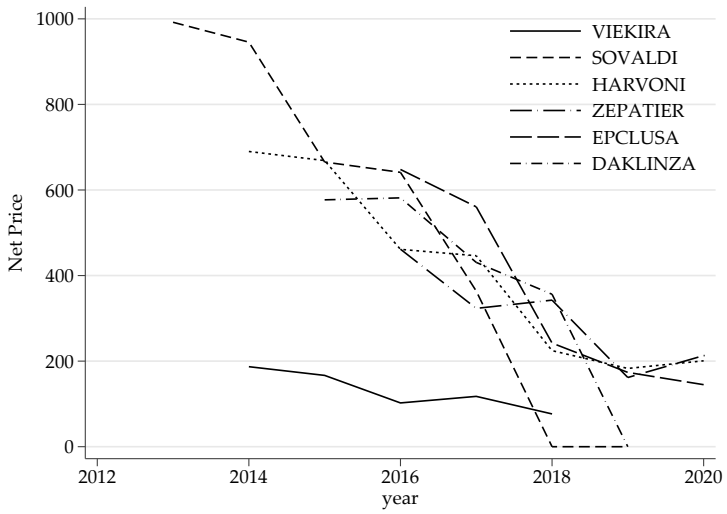


Figure A3: Hepatitis C Drug Prices

C Data Construction and Structural Estimation Details

C.1 Construction of pre-2007 net price series

Our structural analysis covers 1996–2013. Since SSR Health only tracks net prices starting in 2007, we collect additional data on US net revenues and volume sold by drug and year going back to 1996 to arrive at average negotiated prices. This requires some additional assumptions, which we detail in this section.

Calculation of US volume sold We start by calculating volume sold. We use MEPS, which is available throughout the entire 1996–2013 span. We begin by collecting scripts data at the drug-dosage level in each year:

1. We extract all scripts in MEPS related to Zocor, Crestor, Lipitor, Mevacor, and Vytorin.[57]

2. We take the RXQUANTY variable (days supplied) and multiply it by the weight given to each respondent

3. We aggregate total quantity by product, strength, and year

We break out quantity by dosage, because we want to focus on the medium intensity market. Table A5 summarizes the dosages and treatment intensities associated with different

[57] We only include Mevacor because Merck reports aggregate sales for Mevacor and Zocor in the early years of our sample.

medications, along with total quantity of brand sales (number of pills) from 1996–2013 calculated based on MEPS data. The medium intensity market is by far the largest, and the majority of sales of Zocor, Lipitor, and Crestor are in the medium intensity segment. This supports our modeling assumption that they are mainly competing against each other rather than against Vytorin and Zetia.

Table A5: Drug and Dosage vs. Treatment Intensity

Drug	Dosage	Treatment Intensity	Number of Units (billion)	Share Within Molecule	Share within the moderate intensity market
Lipitor	10 MG	Moderate	15.6	0.47	39.0%
Lipitor	20 MG	Moderate	10.8	0.32	27.0%
Lipitor	40 MG	High	5.2	0.16	
Lipitor	80 MG	High	1.7	0.05	
Vytorin	10 MG-10 MG	Moderate	0.3	0.07	0.8%
Vytorin	10 MG-20 MG	High	1.9	0.42	
Vytorin	10 MG-40 MG	High	1.7	0.38	
Vytorin	10 MG-80 MG	High	0.6	0.13	
Crestor	5 MG	Moderate	1.2	0.14	3.0%
Crestor	10 MG	Moderate	4.2	0.49	10.5%
Crestor	20 MG	High	2.4	0.28	
Crestor	40 MG	High	0.8	0.09	
Zocor	5 MG	Low	0.5	0.04	
Zocor	10 MG	Low	2.1	0.18	
Zocor	20 MG	Moderate	4.9	0.42	12.3%
Zocor	40 MG	Moderate	3.0	0.25	7.5%
Zocor	80 MG	High	1.3	0.11	

Notes: Categorization of different strengths of each drug into treatment classes and the associated units sold for the branded version of each dosage from 1996–2013. The focus of the analysis is on Zocor 20mg, Lipitor 10mg, and Crestor 10mg, by far the most popular dosages for starting patients. Vytorin strengths report Zetia first, Zocor second. 10 MG-10MG Vytorin has negligible market share. Sources: Prescriber's Letters (2011), available at: https://www.epocrates.com/sites/default/files/res/dacc/2017/StatinDoseTablePL1703.pdf, retrieved July 15th, 2021.

Calculation of net US revenue Drug manufacturers report data as follows over the 1996–2007 period (all data comes from 10-K filings unless otherwise noted):

1. 1996–1999: Merck reports global net sales for all its cholesterol drugs (Mevacor and Zocor); Pfizer reports Lipitor global net sales

2. 2000–2003: Merck reports Zocor global net sales separately; Pfizer reports Lipitor global net sales

3. 2004–2006: Merck reports US net sales for Zocor in its conference calls

4. 2003–2013: all companies report net US sales in its 10-K filings.

To supplement the data above, we use IMS Top-Line Market Data, which tracks US and global gross revenues by drug, and is available starting in 2000.[58]

Using these data sources, we use the following procedure to calculate net US revenues:

1. In years with reported net U.S. revenue, we just use the number available.

2. In years where only net global revenue are reported, we arrive at a net U.S. revenue number by making the assumption that invoice sales are equal to net sales in non-US markets. Our exact formula is

$$\begin{aligned}\text{Net US Revenue}=&\text{Net Global Revenue}\\ &-(\text{Gross Global Revenue, IMS}-\text{Gross US Revenue, IMS})\end{aligned}$$

3. We adjust Zocor sales in 2006 to account for the fact that generics entered in the middle of the year. We use quarterly revenue from the first two quarters and project out to the whole year.

4. In the 1996–1999 period, we multiply world net sales by the U.S. to world net sales ratio from 2000 to arrive at an estimate for U.S. net sales. We further multiply U.S. net sales reported by Merck for all anti-cholesterol drugs by the ratio of Zocor scripts to total scripts (Zocor plus Mevacor) to arrive at a net sales number for Zocor.

To validate the assumption in (2) we plot the difference in US gross-to-net sales and Global gross-to-net sales in Figure A4. The figure shows that the two quantities are roughly equivalent.

Finally, we collect data on gross prices. To do so, we collect average wholesale price (AWP) and median price paid (labeled as "pay" in the MarketScan data) for each drug by dosage by seller from the MarketScan data. Each entry in the data contains these two quantities. The AWP is constant at the national level at a given point in time, while the "pay" variable partially reflects pharmacy markups.

[58] Based on business news articles and Top-Line Market Data: http://www.imshealth.com/en/about-us/news/top-line-market-data, retrieved January 14, 2017. The Top-Line data is available on Wayback Machine, and we have also included all IMS revenue data as an additional tab in the PBM financials excel file.

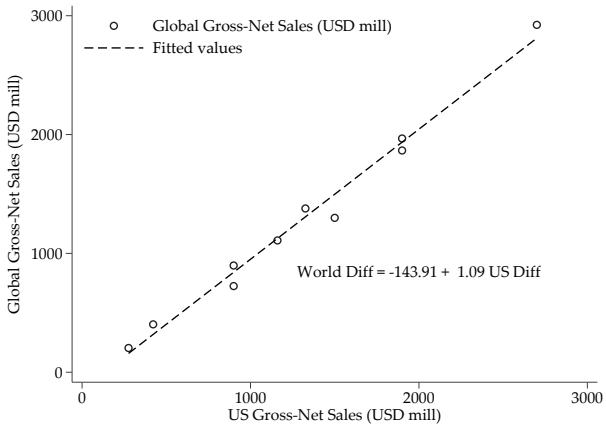


Figure A4: Global vs. US Rebates ($bill)

Notes: A binscatter showing a comparison of discounts and rebates between the US market and global numbers, covering the three main statins in the years 2007–2011. The calculations are based on US and global gross revenue numbers from IMS Topline and company-reported US and global net revenue.

Construction of Net Price Series Under Alternative Assumptions A major challenge in terms of calculating net prices is that net sales estimates are at the drug level, but different dosages of the same drug are sometimes priced differently. Otherwise, we could just divide net sales by quantity. For example, a 40MG Lipitor tablet has a gross price that is $1 more expensive than the more commonly used 10MG tablet. Crestor, on the other hand, sets the same price for all dosage levels. Without additional knowledge, we cannot pin down the negotiated price for medium intensity dosages of each drug, which is the focus of our analysis.

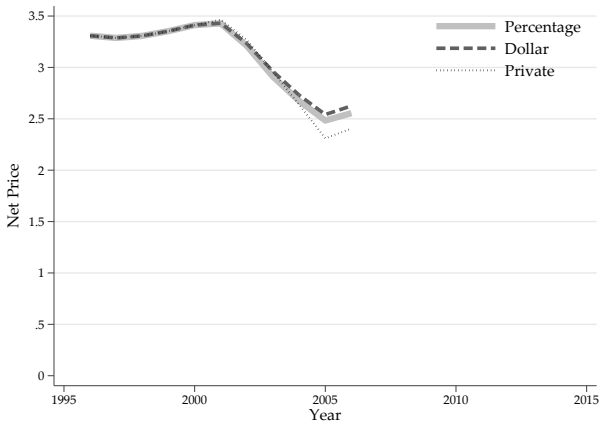
To solve this problem, we assume that percentage discounts per pill are equal across dosages for the same drug. To do so, we compare the difference in overall gross and net revenue for each drug. We then take the percentage discount off the list price of the medium intensity dosage. In cases where there are multiple dosages and different prices associated with each, we take a weighted average by volume. In the end, we arrive at average negotiated price in the U.S. market for each of the three drugs in each year.

We also compute two additional prices series. Figure A5 presents a comparison of the results from the three approaches.

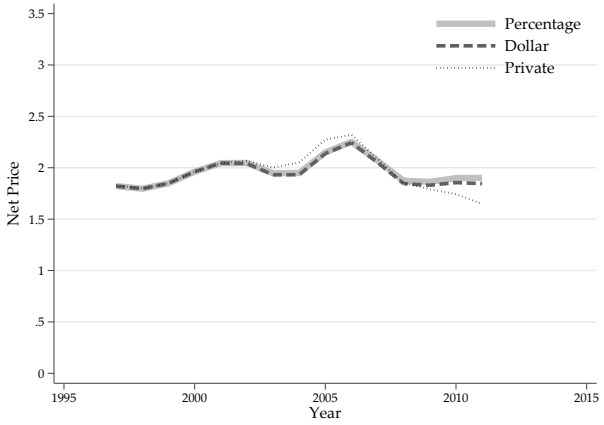
As a first alternative, we derive net prices assuming constant dollar discounts across dosages. In the core derivation, the percentage discount per pill is assumed to be the same across dosages. An alternative assumption is that dollar discounts are the same across dosages. There are minimal differences between the two approaches. The biggest deviations are about 5 cents, and Crestor is barely affected, as its dosages have similar prices.

Figure A5: Price Comparison – Three Approaches

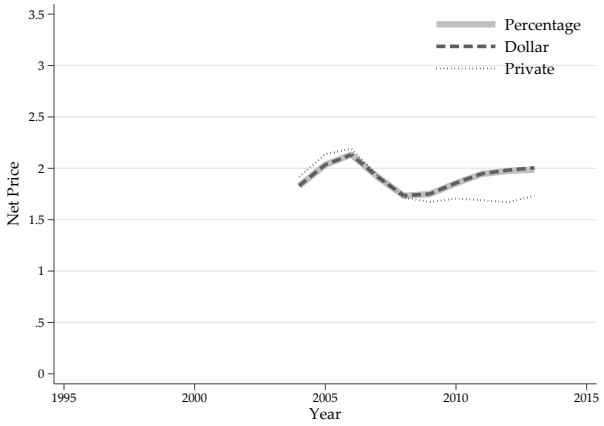
(a) Zocor



(b) Lipitor



(c) Crestor



Notes: Net prices under the three different methodological approaches.

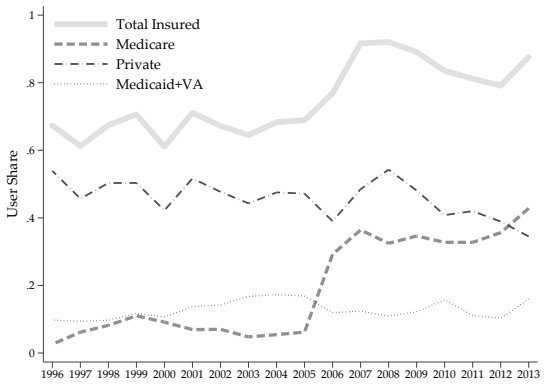


Figure A6: Insurance Coverage Type for Cholesterol Patients

Notes: A plot of the percentage of patients taking anti-cholesterol medication who are covered by each type of insurance (a small percentage are covered by multiple). Calculations based on MEPS data.

The differences that are present depend on how the prices of low-intensity and high-intensity dosages relate to those of the medium-intensity dosages. When high-intensity dosages are priced higher, they absorb more of the discount under the percentage model, which results in higher prices.

A more substantive adjustment is to isolate the negotiated prices in markets covered by PBMs, which generally results in lower derived prices. PBMs serve Part D plans and private insurance, which does not include Medicaid, Veterans Administration (VA), and uninsured patients. Uninsured patients pay an amount close to list price, and make up about 15–30 percent of the individuals taking anti-cholesterol medication. Medicaid and VA patients pay less than or equal to the private market, and account for 10–15 percent of people. Figure A6 documents these numbers. Our approach is to first purge the net revenue data of uninsured patients paying list price, which pushes down the effective net price in private markets. Then, to create as high a potential negotiated price as possible, we assume that the VA pays \$0 and that Medicaid receives a 30 percent discount off the net price in the private market. We then obtain the net price by solving out the following accounting identity for p_{PBM}:

$$p_{PBM}(s_{Medicare} + s_{Private} + 0.7s_{Medicaid}) + s_{uninsured}p_{list} = p_{net}$$

where s represents the share of individuals taking medication covered under each type of insurance, p_{net} is the overall net price derived earlier, and p_{list} is the list price. Even this adjustment does not lead to substantial changes in the price series, with the exception of the price of Crestor in 2009—2013. Column 3 of Table 4 reports estimates that use this

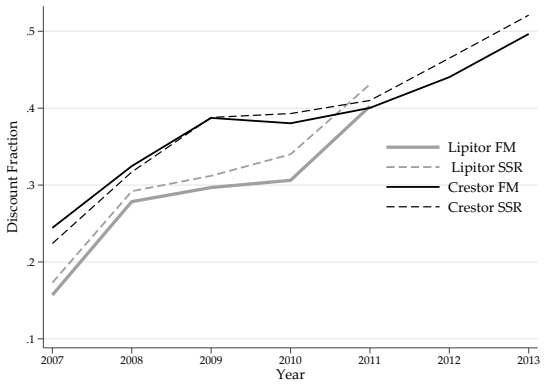


Figure A7: Comparison of Percentage Discounts

Notes: FM refers to our discount series calculated using the methodology described in this Appendix. SSR refers to percentage discounts calculated by SSR Health using proprietary data.

price series. The main difference is that the penalty of excluding a drug altogether is lower in this specification, but the overall implications of the model are quite similar. In our counterfactuals we stick with the constant percentage discount price series because it requires fewer assumptions.

Comparisons to SSR Health Estimates Figure A7 provides a comparison of the percentage discounts calculated using our approach versus the SSR Health data. Our methodology approximates the SSR Health data very accurately for both Lipitor and Crestor.

C.2 Computational Step for the Estimation of the Bidding Model

Matching estimated net prices series to the model's prediction is the core exercise of our paper. In this section we describe how we perform the estimation algorithm:

1. Start from a guess for the parameters in the PBM's objective function: $\theta' = \left(\theta'_s, \theta'_e, \theta'_j, \theta'_\omega\right)$
2. Compute the value function for each firm, starting from the last period—when only Crestor is active—and solving the model using backwards induction:
   (a) In period t, fix a value x_t for the state space, and compute optimal prices and payoffs for each firm as follows:
      i. Guess an initial vector of prices $\left(p^0_{\text{ZOCOR}}, p^0_{\text{LIPITOR}}, p^0_{\text{CRESTOR}}\right)$
      ii. For each drug j, starting with Zocor and ending with Crestor, find $p^\star_j\left(p^0_{-j}\right) = \arg V_{jt}\left(x_t\right)$, where $V_{jt}\left(x_t\right)$ is the value function of drug j defined as the max-

imization problem in equation 10. Intuitively, $p_j^\star\left(p'_{-j}\right)$ is the best response of manufacturer j holding fixed the prices of the other two drugs

iii. Replace p_j^0 with $p_j^1 = p_j^\star\left(p_{-j}^0\right)$ in the initial vector of guesses. Repeat step ii. for all drugs that have not lost exclusivity yet (prices for generics are fixed).

iv. Repeat steps ii. and iii. until $\sum_j \left(p_j^n - p_{-j}^{n-1}\right)^2 < \varepsilon$ for some pre-established threshold ε (we use $\varepsilon = 10^{-8}$). This is the equilibrium.

v. Calculate the value function for all drugs j at the estimated values $\left(p^n_{\text{ZOCOR}}, p^n_{\text{LIPITOR}}, p^n_{\text{CRESTOR}}\right)$

(b) Repeat step (a) for all values $x_t \in [0, 0.05, 0.12, 0.2, 0.3, 0.5, 0.7]^3$ (excluding combinations where the sum of the market shares sums to more than one).

(c) To find the value function for values of x_t that are not in $[0, 0.05, 0.12, 0.2, 0.3, 0.5, 0.7]^3$ we use the scattered interpolant function in MATLAB, with the natural neighbor option, which yields a smoother function than a simple linear interpolation.[59]

(d) Move to period $t-1$, and repeat steps (a)–(c)

3. Once we have a value function for each period and state, we play out the game, giving Zocor a 0.4 market share coming into 1996. This yields a set of predicted equilibrium prices associated with the parameter vector θ', which we can then compare to the observed net prices in the data.

Given the routine for computing equilibrium prices given a parameter vector, we then run a search for the parameter vector that minimizes the distance between predicted and observed net prices. The routine is computationally intensive. Each evaluation takes 5–10 minutes using parallelized code on a cluster with 15 workers. The main bottleneck is the set of periods during which there are three active players. In these periods, the state space is very large.

C.3 Evidence of uniqueness of equilibrium

The algorithm described above is based on a minimum distance estimator, which is only valid if the game has a unique equilibrium. While we do not have a formal proof of

[59] We use a sparse vector to reduce computational time. The vector is more densely populated at lower market shares because those states are more likely to be observed in the data.

uniqueness, we offer some heuristic analysis that suggest multiplicity is not an issue in our model.

Because our game has a finite horizon, a sufficient condition for uniqueness is that each stage game has a unique equilibrium. Results in the literature have provided conditions for existence uniqueness in price setting games under certain classes of demand systems, focusing on price and market share bounds (Thomadsen, 2005; Aksoy-Pierson et al., 2013). While our demand system fits under the conditions, our model also contains an intermediary, which will alter the effective demand curve facing firms. This makes it difficult to directly apply the conditions derived in the existing literature.

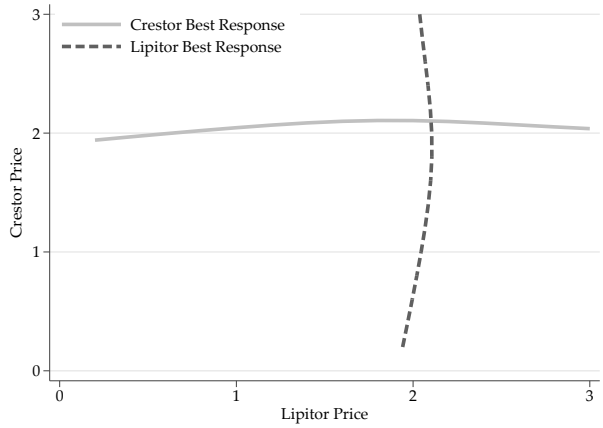
To show that our stage games are likely to have a unique equilibrium, we calculate the best response functions in the years during which only two manufacturers compete (1997–2003 for Zocor and Lipitor, and 2007–2011 for Lipitor and Crestor). This allows for a simple view of best response functions in a two-player game. While the best response functions are non-monotonic, we do find that they only intersect once for the values of the parameter θ that we estimate and for the values of the state space that we observe. Figure A8 plots the best response functions for the five years in which Lipitor and Crestor competed. One thing to notice is that prices are not particularly responsive to one another. This is largely due to the fact that generic Zocor has a large market share, so undercutting the branded competitor has a relatively small impact on market share. We conducted additional simulations (not shown) confirming that if generic Zocor had lower market shares, the best response functions would be more responsive. In these simulations too, we find that the functions satisfy the single-crossing property.

D Additional Structural Model Discussion and Results

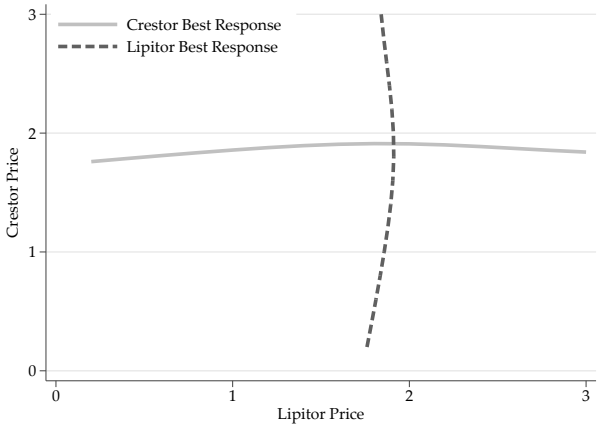
D.1 Copay elasticities and first-stage estimates

One potential issue in the demand estimation is that we do not have a suitable instrument for copays. While instrumenting for copays is desirable, it is not clear what bias the absence of an instrument introduces. Patients could select into plans that have a low copay associated with the drugs they like, leading to an upward bias in the estimated copay elasticity. On the other hand, copays might be higher for higher quality drugs, or drugs with more loyal patients because manufacturers might be more willing to accept worse coverage in exchange for charging higher net prices. Hence, our estimates might be biased upward or downward, and, in practice, these two effects might cancel each other out.

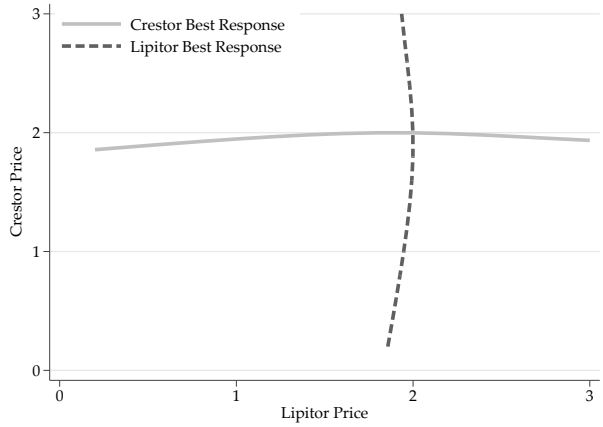
To validate our estimates, we compare them to the estimates from Einav et al. (2016),



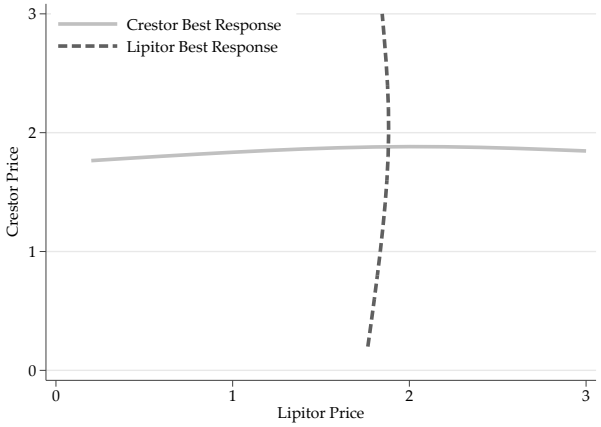
(a) 2007



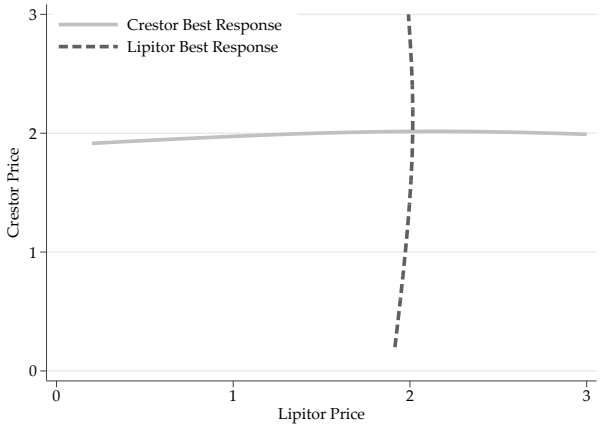
(b) 2008



(c) 2009



(d) 2010



(e) 2011

Figure A8: Best Response Function Plots

Notes: A plot of best response functions for Lipitor and Crestor after the entry of generic Zocor and before the entry of generic Lipitor (2007–2011). In each year, we show the state that is closest to the observed state in equilibrium play: 2007, [0.5, 0.3, 0.05]; 2008, [0.5, 0.2, 0.05]; 2009, [0.5; 0.2; 0.05]; 2010, [0.7; 0.12; 0.05]; 2011, [0.7;0.12;0.05].

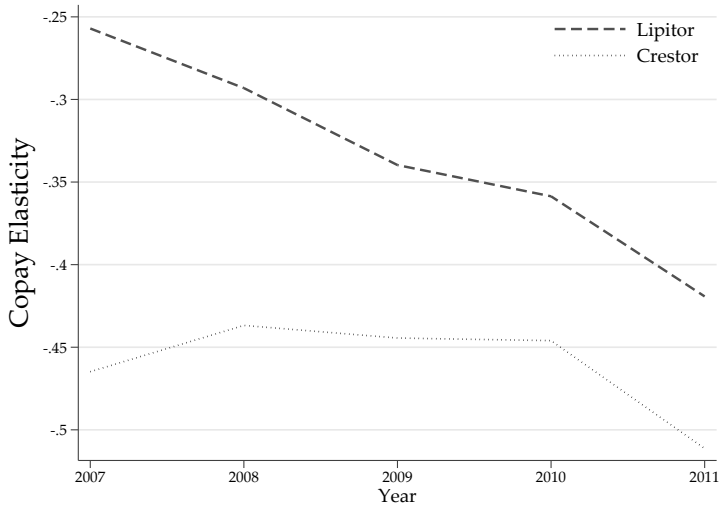


Figure A9: Predicted Copay Elasticities 2007–2011

which exploits the Medicare Part D donut hole as an exogenous source of variation in copays. Einav et al. (2016) use data from 2007–2011, and estimate copay elasticities for the most popular branded drugs, including Lipitor and Crestor. They estimate that Lipitor and Crestor have a copay elasticity of -0.33 and -0.31 respectively.

When we compute the copay elasticities of Lipitor and Crestor using our demand system in the same 2007–2011 period, we find that our demand system generates similar implied elasticities. Figure A9 plots the elasticities for Lipitor and Crestor by year. Our estimates for Lipitor range from -0.26 to -0.42 and our estimates for Crestor range from -0.43 to -0.51. These are quantitatively similar to the findings in Einav et al. (2016). The differences likely arise from our imposition of a functional form and differences in sample. Implicitly, Einav et al. (2016) are measuring the response of all patients who reach the donut hole, potentially a slightly selected sample of older individuals who take more drugs. These patients are likely to be more inelastic than the overall population.

In addition, we provide statistics in Table A6 on the distribution of year-to-year copay changes at the person-drug level in our MarketScan sample. Copay changes only occur in 44 percent of person-years. For reference, the median copay for branded statins is \$15 per month and the average is \$21.70.

Finally, we report the first stage regressions associated with the instrumental variables approach, which is used in both the 2SPS and 2SRI estimation approaches. Table A7 reports the results. As shown in the table, the leave-one-out cohort mean is strongly predictive of a given indivdual's previous choice indicators. The coefficients are close to one, which makes sense given that we are predicting shares across all individual-drug pairs,

Table A6: Distribution of Copay Changes

	5th	10th	25th	Median	75th	90th	95th
Change in Monthly Copay	-$10	-$2	$0	$0	$0	$5	$11

Notes: Distribution of year-to-year copay changes at the person-drug level, using the MarketScan sample used in demand estimation.

Table A7: First Stage Instrumental Variables Estimates

	2SPS and 2SRI	
	Incumbent	Non-Outside
Leave-one-out measure (Incumbent)	0.99988	
	(0.00024)	
Leave-one-out measure (Non-Outside)		0.99981
		(0.00009)
R^2	0.2564	0.7311
F	>99,999	>99,999
N	49,696,288	49,696,292

Notes: Incumbent refers to the term $I_{s_{it-1}=j,j\neq 0}$ and non-outside refers to the term $I_{s_{it-1}\neq 0,j\neq 0}$. We construct the cohort means ("leave-one-out measure") based on Equation 11, and run first-stage regressions based on Equation 12.

and each cohort has a large number of individuals.

D.2 Estimation results under different demand parameters and discount rates

For our core specification, we make several choices related to the offline parameters that are then inputted into the estimation of the dynamic game. This includes which demand parameters to use (2SPS in the core results) and picking a discount factor for manufacturers (0.88 in the core results). Table A8 provides estimates using 2SRI demand parameters and using different discount rates. The model fit is similar across the specifications, and the parameter estimates are also similar.

D.3 Formulary Predictions

The dynamic estimation of our model targets the net prices observed in the data. However, our model can also generate predictions for formulary coverage, given those prices.

Table A8: Parameter estimates using different "offline" parameters

	Estimates			
Parameter	Baseline	2SRI	$\beta = 0.91$	$\beta = 0.85$
θ_s	1.70	1.74	1.66	1.65
θ_e	1.50	1.39	1.44	1.48
θ_j	1.11	0.86	0.94	0.94
θ_ω	0.063	0.050	0.055	0.064
Total error	0.46	0.43	0.48	0.47

Notes: All parameter estimates are based on the "Pct Discount" price series. Total error refers to the sum of squared errors across all branded drug-years between 1996 to 2013 (excluding periods after generic entry for each drug).

Figure A10 presents the predictions of the estimated model for formulary coverage of each drug. The model creates a probabilistic distribution over all possible formulary arrangements. For each drug, we plot the average expected copay facing patients (conditional on formulary inclusion) and the probability the drug is excluded altogether.

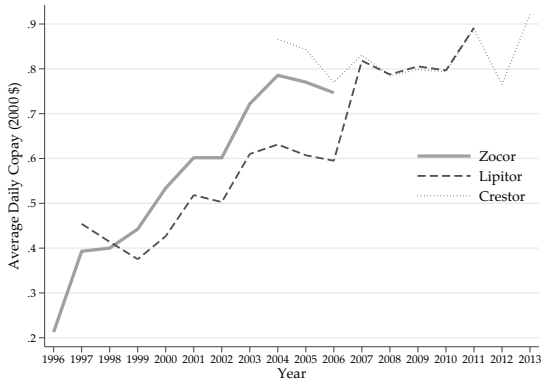
We find that our model's predictions match qualitative and quantitative evidence on formulary structure in the statin market. The model predicts that Zocor is more likely to be placed in the non-preferred tier in earlier years, in line with both anecdotal evidence and the mean daily copay series presented in Figure 3 of the main text.[60] We note that copays are increasing for all drugs over time. This does not come from the model itself—which predicts formulary tier—but from the MarketScan data, which shows increasing copays at all tier levels over time.

The model also predicts low exclusion rates in earlier years, with some exclusion for Crestor.[61] Figure A11 plots the formulary exclusion rates for Lipitor and Crestor in Medicare Part D formularies starting in 2008, when data becomes available. The two drugs also face softer restrictions such as step therapy and prior authorization, outcomes not included in the model, which make a direct quantitative comparison infeasible.

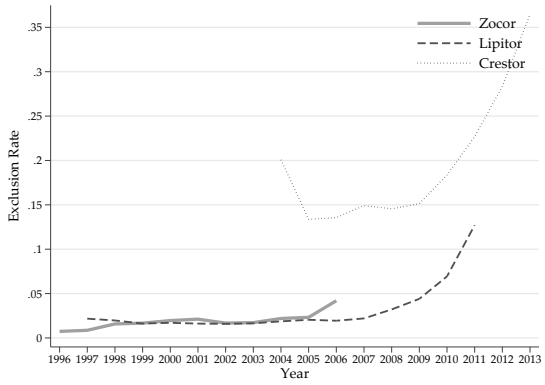
Finally, we also plot the probability that multiple drugs are excluded in equilibrium. As discussed in the counterfactual analysis in Section 5.1, if PBMs cannot exclude more

[60] For anecdotal evidence, see Table 9 of a 2001 report by the California HealthCare Foundation. The table shows that all the major PBMs put Lipitor in the preferred tier, whereas Zocor was only in the preferred tier of the formularies of Medco and AdvancePCS: "Prescription Drug Coverage and Formulary Use in California: Difference Approaches and Emerging Trends" available at: http://www.chcf.org/publications/2001/06/prescription-drug-coverage-and-formulary-use-in-california-different-approaches-and-emerging-trends. Note that the MEPS series is plotting copays conditional on choosing a drug, whereas our model is predicting the unconditional average copay facing patients.

[61] Crestor initially was excluded from some formularies. See The Guardian article "US blow to AstraZeneca": https://www.theguardian.com/business/2003/oct/04/4.



(a) Average Copay (Conditional on Inclusion)



(b) Exclusion Rate

Figure A10: Additional Model Predictions – Formulary Outcomes

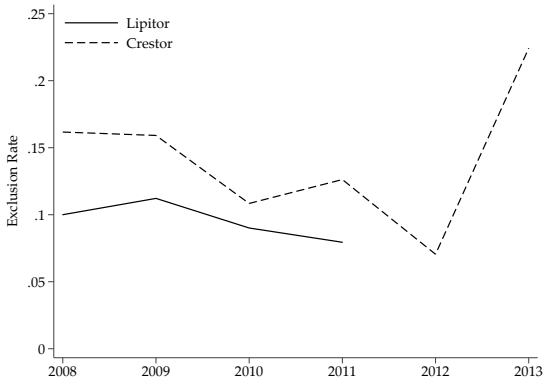


Figure A11: Formulary Exclusions for Anti-Cholesterol Drugs (Part D)

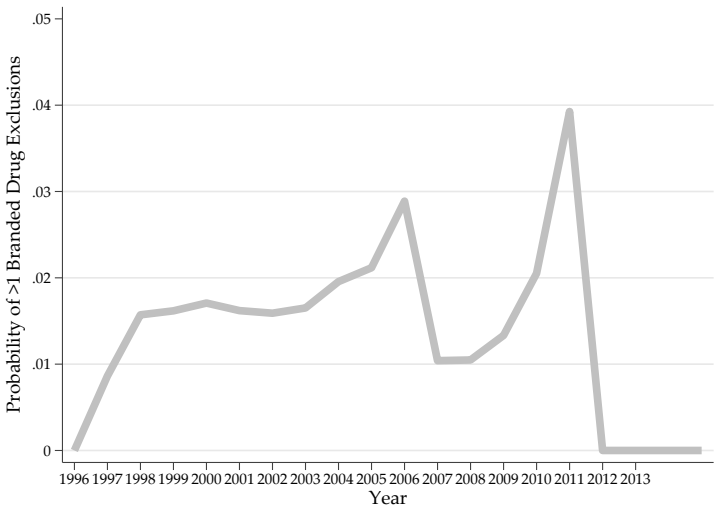


Figure A12: Double Exclusion Rates in Equilibrium

Notes: A plot of the rate at which the representative PBM excludes multiple brand drugs from the formulary, based on predicted play under the core model estimates.

than one drug, the equilibrium prices and spending are significantly higher. However, Figure A12 shows that this is a rare outcome in equilibrium—below 2 percent in all years except the years before generic entry.

D.4 Estimation results under alternative specifications of the PBM's objective function

We use a flexible control function approach to model the PBM's objective function. To test the robustness of this approach, we re-estimate the model using three alternative specifications. The results are qualitatively identical. We describe our additional analysis here.

Allowing for dynamic considerations Just like firms are forward-looking, PBMs could also incorporate dynamic considerations in their decision making. For example, PBMs might prefer starting new patients onto drugs who are about to experience loss-of-exclusivity, in order to reduce their future costs. At the same time, since patients frequently switch insurance, PBMs may only realize a small fraction of these potential future savings. We test for PBM dynamic incentives by introducing an additional term:

$$\beta^{PBM} V_{t+1}^{PBM}(x_{t+1})$$

The term $V_{t+1}^{PBM}(x_{t+1})$ represents the PBM's value function, and it's calculated dynamically through the same procedure used to compute the manufacturers' value functions we describe in Appendix C.2.

Incentives to boost generic market share Many PBMs also run mail-order pharmacies, which most often ship generic drugs. Therefore, PBMs might have additional incentives to switch patients onto generic drugs. To account for this effect we include an additional term for the total market share of generic drugs:

$$\theta_{generic} \times ms_{generic}(x_t, C_t)$$

The importance of this term depends on the mail order share of prescriptions drugs, which empirically is not very high.[62]

Carryover of inertia through doctors to new patients Finally, we modify Equation 5 to allow for a fraction κ of patients to inherit last period's choices: $x_{jt} = ms_{jt-1} \times \frac{N_{t-1}+\kappa(N_t-N_{t-1})}{N_t}$. This type of carryover could be due to doctors or social peers. A high value of κ would affect dynamic incentives for firms, especially when market size is growing.

Table A9 shows the results and compares them to the baseline analysis. We find that including a dynamic term for PBM revenue and a boost for generic market share lead to a small increase in model fit. The coefficients are also significant. However, the coefficients are also very small and do not affect other estimated coefficients. Because adding even one coefficient significantly increases computational time, we ultimately decided against including them in our main specification.

[62] According to the Kaiser Family Foundation (using 2019 data from IQVIA), patients fill about 11.6 prescriptions at retail pharmacies on average, with only 0.52 fills through mail order. Sources: https://www.kff.org/health-costs/state-indicator/number-of-mail-order-prescription-drugs-per-capita/ and https://www.kff.org/health-costs/state-indicator/retail-rx-drugs-per-capita/ (both retrieved February 17, 2023). Because mail order utilization has increased over time, the rate should be even smaller during our period of analysis.

Table A9: Parameter Estimates from Additional Specifications

		Estimates		
Parameter	Baseline	(1)	(2)	(3)
θ_s	1.70	1.71	1.72	1.81
	(0.13)	(0.26)	(0.45)	(0.26)
θ_e	1.50	1.77	1.61	1.59
	(0.05)	(0.22)	(0.35)	(0.17)
θ_j	1.11	1.24	1.13	1.29
	(0.37)	(0.17)	(0.27)	(0.17)
θ_ω	0.063	0.061	0.064	0.063
	(0.005)	(0.006)	(0.005)	(0.002)
β^{PBM}	–	0.10	–	–
	–	(0.02)	–	–
$\theta_{generic}$	–	–	0.11	–
	–	–	(0.02)	–
κ	–	–	–	0.08
	–	–	–	(0.05)
Total Error	0.458	0.435	0.439	0.456

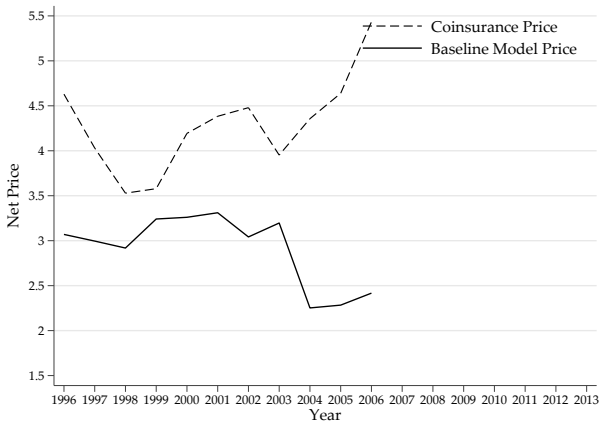
Notes: All parameter estimates are based on the "Pct Discount" price series. Total error refers to the sum of squared errors across all branded drug-years between 1996 to 2013 (excluding periods after generic entry for each drug). Standard errors are computed using the two-step procedure outlined in the main text.

D.5 Predicted prices under 33% coinsurance counterfactual

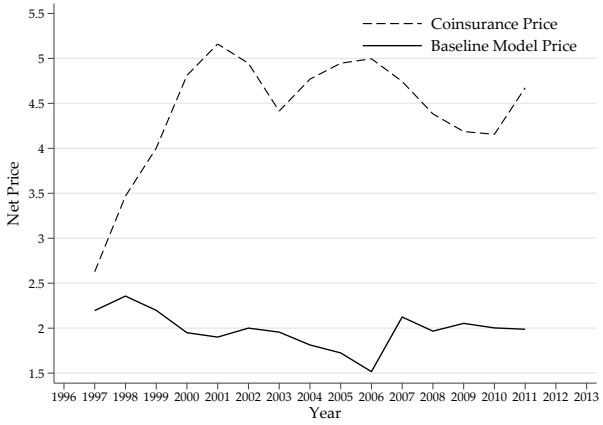
Figure A13 presents a comparison of the equilibrium prices under the 33 percent coinsurance counterfactual to the equilibrium prices under the baseline estimated model. The coinsurance prices are higher, and exhibit greater growth over time. This is particularly obvious in the case of Lipitor, which, absent PBMs, would experience dramatic price growth in the first five years after launch.

Figure A13: Price Comparison – Baseline vs. Coinsurance

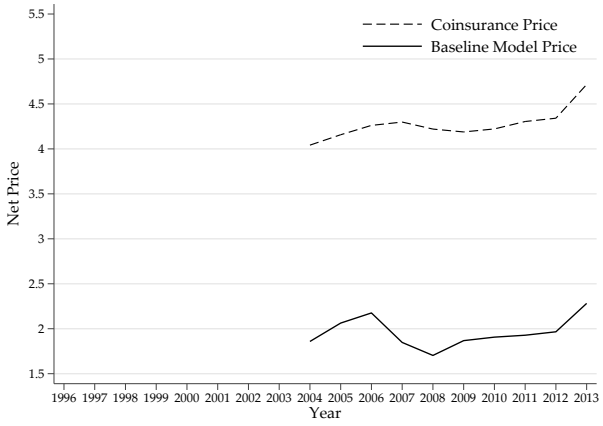
(a) Zocor



(b) Lipitor



(c) Crestor



Notes: Prices predicted by the best-fit model versus the prices observed in the data.