Demand Inertia and the Hidden Impact of Pharmacy Benefit Managers

Josh Feng† Luca Maini‡
(University of Utah) (Harvard Medical School)

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.

†josh.feng@eccles.utah.edu
‡maini@hcp.med.harvard.edu

*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. 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. This paper is based on the second part of a 2017 manuscript titled “History-Dependent Demand and Intermediaries: Explaining Prescription Drug Pricing Dynamics.”
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. Formulary coverage determines the out-of-pocket cost patients pay for drugs. 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 nego-

\(^1\)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).
tiations between 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 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). Demand inertia leads to market segmentation and increasing equilibrium price paths (e.g., Klemperer, 1987b,a; Dube et al., 2009). 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

---

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.
the drug from coverage. We model PBM behavior using a flexible control function that incorporates 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 results 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

\(^5\) 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.
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 sophisticated 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.

---

\(^6\) 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.
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 like insurance companies and employers. 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 eval-

\(^8\) Appendix ?? 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 ???. Another option would be to model this process using Nash bargaining. While Nash bargaining effectively nests our approach, we believe that our bidding model captures the essential tradeoffs faced by PBMs and manufacturers. Our bidding model also has the advantage of being simple enough to accommodate dynamic estimation, which is necessary to capture inertia in drug demand.

\(^11\) Disclosures from several state health plans describe this process in detail. See,
uation 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.

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. Concerns over PBM margins have increased following a wave of recent mergers. However, while the concentration in the market is very high, it has remained relatively stable over the past decade. 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 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.


13The 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.

14In 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 ?? provides some numbers on PBM market shares and concentration measures.
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. 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 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). The classification roughly corresponds to the equivalence classes used by PBMs in their bidding process.

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

$$ y_{jt} = \psi_1 \log (WAC_{jt}) + \psi_2 \log (net_{jt}) + \nu_{k(j)t} + \epsilon_{jt} $$

The ATC system classifies molecules according to the organ on which they act and their therapeutic and chemical properties.
Table 1: Summary Statistics for Price and Formulary Data

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

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).

where j indexes drugs, t indexes the year, and k(j) indicates the therapeutic class of drug j. v_{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 are driven by differences in unit definitions 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.

\(^{16}\)We verify robustness of our results to drug class definition (Table ?? in Appendix ??).
Table 2: Prices vs. Formulary Correlations

<table>
<thead>
<tr>
<th></th>
<th>Frac. Preferred</th>
<th>Frac. Covered</th>
</tr>
</thead>
<tbody>
<tr>
<td>Log(net price)</td>
<td>-0.0228**</td>
<td>-0.00656*</td>
</tr>
<tr>
<td></td>
<td>(0.00672)</td>
<td>(0.00315)</td>
</tr>
<tr>
<td>Log(WAC)</td>
<td>0.0104</td>
<td>0.00586</td>
</tr>
<tr>
<td></td>
<td>(0.0159)</td>
<td>(0.00979)</td>
</tr>
</tbody>
</table>

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.

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 equilibrium 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. 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.

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).
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:

\[ y_{jt} = \zeta_j + \lambda_t + \sum_{k \in B_j} \left[ \tau_1 \sum_{l < -3} D_{jkt}^l + \sum_{l = -3}^{-2} \xi_l D_{jkt}^l + \sum_{l = 0}^{3} \xi_l D_{jkt}^l + \tau_2 \sum_{l > 3} D_{jkt}^l \right] + \sum_{k \in G_j} \left[ \tau_3 \sum_{l < -3} D_{jkt}^l + \sum_{l = -3}^{-2} \eta_l D_{jkt}^l + \sum_{l = 0}^{3} \eta_l D_{jkt}^l + \tau_4 \sum_{l > 3} D_{jkt}^l \right] + \varepsilon_{jt} \]  

(2)

where \( y_{jt} \) represents price and formulary outcomes, \( \zeta_j \) and \( \lambda_t \) are drug and year fixed-effects, respectively, and the remaining terms allow for flexible, time-varying responses to each event. Formally, \( D_{jkt}^l = I\{t - E_{jk} = l\} \), 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 ?? in Appendix ?? provides results using alternative definitions of therapeutic class.
Response Relative to Event Year

(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.

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 ?? in Appendix ?? 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.

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). 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.

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

---

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.
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.

Figure 2: Model Summary

Notes: A summary of the period game structure in our model. Each “drug” box represents a drug manufacturer. \( p_{1t}, p_{2t}, p_{3t} \) are the price bids and \( C_t \) the formulary chosen by the PBM.

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 \( J \) of molecules within a therapeutic class. The set \( J \) includes the outside option \( j = 0 \), which indicates the option of seeking treat-
ment 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. 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

$$u_{ijt} = \delta_{jt} - \alpha c_{jt} + \gamma I_{s_{it-1}=j,j\neq0} + \nu I_{s_{it-1}\neq0,j\neq0} + \varepsilon_{ijt} \quad (3)$$

where $i$ indexes the consumer, $j$ the molecule, and $t$ the year. $\delta_{jt}$ 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 ($\delta_{0t}$) is normalized to zero in each period. $\alpha$ measures sensitivity to the out-of-pocket cost $c_{jt}$, which is determined by the formulary. We model inertia by including two switching cost terms. $\gamma$ is the cost of switching between molecules: $s_{it-1}$ indexes the option that patient $i$ picked in period $t - 1$, while $I_{s_{it-1}=j,j\neq0}$ is an indicator for whether a patient $i$ chose a non-outside-option molecule $j$ in period $t - 1$. $\nu$ is the cost of switching cost from any molecule to the outside option: $I_{s_{it}\neq0,j\neq0}$ is an indicator for whether patient $i$ picked any non-outside-option molecule in the previous period. 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). Finally, $\varepsilon_{ijt}$ represents a consumer- and molecule-specific taste shock, which is distributed according to

---

20 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.

21 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.

22 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 drug-specific switching cost are very similar to one another, so we impose this additional constraint to ease the computational burden.
type-I extreme value distribution.

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 $C_t = \{c_{jt}\}_{j \in J}$ for each drug in $J$. We then denote the probability that patient $i$ chooses molecule $j$ as

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

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

$$x_{jt} = ms_{jt-1}(x_{t-1}, C_{t-1}) \times \frac{N_{t-1}}{N_t} + I\{j = 0\} \times \frac{N_t - N_{t-1}}{N_t}$$ (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 $ms_{jt-1}(x_{t-1}, C_{t-1})$ is the market share of product $j$ in period $t-1$. In turn, the market share is given by

$$ms_{jt}(x_t, C_t) = \sum_{k \in J} x_{kt} ms_{ijt}(k, C_t)$$ (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 $W(x_{t-1}, C_t)$. Using the properties of the extreme value distribution, we get that

$$W(x_t, C_t) = \sum_{k \in J} x_{kt} \cdot \log \left[ \sum_{l \in J} \exp(\delta_{lt} - \alpha c_{lt} + \gamma I_{k=1, l \neq 0} + \nu I_{l \neq 0}) \right]$$ (7)

### 3.2 Formulary choice model

A representative PBM chooses the formulary arrangement to maximize an objective function, which we model as a flexible parametric control function to capture the many
different ways in which PBMs earn revenue.

Formally, given a vector $p_t$ of bids from branded drug manufacturers, the PBM chooses a formulary arrangement $C_t = \{c_{jt}\}_{j \in J}$, where $c_{jt} \in \{c_g, c_t, c_t, \infty\}$, with $c_g < c_t < \infty$. The values of $c_g$, $c_t$, and $\infty$ are assumed to be exogenous.\(^{23}\) 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.\(^{24}\) Crucially, the presence of 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 $p_t$, the PBM chooses $C_t \in \mathcal{C}_t$ to maximize:

\[
\max_{C_t} \frac{W(x_t, C_t)}{\alpha} - \theta_s \sum_{j \in J} p_t m_{sjt}(x_t, C_t) - \theta_e \sum_{j \in J} I(c_{jt} = \infty) x_{jt} + \sum_{j \in J} \theta_j \times m_{sjt}(x_t, C_t) + \theta_\omega \omega_{Ct} \tag{8}
\]

The expression has four components. The first component is “formulary surplus” $W(x_t, C_t)$. We divide formulary surplus by our estimated copay-elasticity to transform utils into dollars. The second is expected spending per member. The coefficient $\theta_s$ translates dollars spent on prescription drugs into the utility metric. 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 $\omega_{Ct}$ 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 $\theta_\omega$. The error term has a dual function. First, it allows the model to fit the data. Second, it facilitates estimation by smoothing out

\(^{23}\)This reflects the fact that insurers rather than PBMs pick cost sharing levels.

\(^{24}\)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 below even those of preferred branded drugs.
expected manufacturer revenue as a function of the price bid.\textsuperscript{25}

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

\[ P(C_t|x_t, p_t) = \frac{\exp\left(\frac{1}{\theta}\text{ObjFun}(C_t, x_t, p_t)\right)}{\sum_{C \in \mathcal{C}_t} \exp\left(\frac{1}{\theta}\text{ObjFun}(C, x_t, p_t)\right)} \quad (9) \]

where \text{ObjFun}(C_t, x_t, p_t) is argument of the maximization in Equation 8.

### 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 \( T^\text{en}_j \) and \( T^\text{ex}_j \), 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 \( x_t = \{x_{jt}\}_{j \in J} \) (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:

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

where \( \beta \) is the discount factor, which we set equal to 0.88, following the cost of capital calculation in \textit{DiMasi et al. (2003)}. \( x_{t+1}(C_t, N_{t+1}) \) — the value of \( x_{t+1} \) as a function of the chosen formulary—is deterministically set by the expressions in Equations 5 and

\textsuperscript{25}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.
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 \( T^e \) (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.\(^{26}\) 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 man-

\(^{26}\)This limitation is shared by all papers on the pharmaceutical industry as far as we know.
ufacturers: 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.\textsuperscript{27}

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).

\section{A Structural Analysis of the Medium-Intensity Statin Market}

\subsection{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.

\textsuperscript{27}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.
taking anti-cholesterol medication.\textsuperscript{28}

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.\textsuperscript{29}

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, 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.\textsuperscript{30}

Figure 3 summarizes the evolution of (inflation-adjusted) net prices, average co-pays, 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).\textsuperscript{31}

\textsuperscript{28}National Center for Health Statistics Data Brief No. 177, December 2014.
\textsuperscript{29}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.
\textsuperscript{30}See Appendix \ref{app:meth} for a detailed description of each methodology.
\textsuperscript{31}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 \ref{app:meth} 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
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 individuals 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, then demand would not create dynamic incentives for drug manufacturers and PBMs.

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 inertia 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. Assuming that the choice to
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. 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.
start treatment is unrelated to these factors, 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.\textsuperscript{32}

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

$$I_{y,i,j,t} = \frac{1}{|Z_y| - 1} \times \sum_{k \in [Z_y \setminus i]} I_{s_{k,t-1}=j}$$  \hspace{1cm} \text{(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 $j$.

We use $I_{y,i,j,t}$ to predict the choice of patient $i$ as:

$$I_{s_{k,t-1}=j} = \tau_1 + \tau_2 I_{y,i,j,t} + \nu_{ijt}$$  \hspace{1cm} \text{(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 $I_{j=s_{k,t-1}}$ for $I_{j=s_{k,t-1}}$ in the logit model. The 2SRI approach involves keeping the basic specification but adding in the residual $\nu_{ijt}$ 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

\hspace{1cm} \text{Footnote:} 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.
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 ??).

Table 3 reports the estimates using the two approaches, and for the simple regressions without instruments. 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. Both IV methodologies return similar results. 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. 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 conduct additional tests under different demand parameters. The implied monthly switching costs $\gamma$ and $\nu$ 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 ($\delta_{jt}$). 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,
### Table 3: Demand System Estimates

<table>
<thead>
<tr>
<th></th>
<th>Basic CL</th>
<th>2SPS</th>
<th>2SRI</th>
<th>2SPS</th>
<th>2SRI</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Cost Sharing ($\alpha$)</strong></td>
<td>0.719***</td>
<td>0.633***</td>
<td>0.724***</td>
<td>0.668***</td>
<td>0.739***</td>
</tr>
<tr>
<td></td>
<td>(0.00253)</td>
<td>(0.101)</td>
<td>(0.141)</td>
<td>(0.111)</td>
<td>(0.149)</td>
</tr>
<tr>
<td><strong>Same Drug ($\gamma$)</strong></td>
<td>3.341***</td>
<td>2.674***</td>
<td>2.502**</td>
<td>1.460***</td>
<td>2.070***</td>
</tr>
<tr>
<td></td>
<td>(0.00235)</td>
<td>(0.567)</td>
<td>(0.778)</td>
<td>(0.204)</td>
<td>(0.277)</td>
</tr>
<tr>
<td><strong>Any Drug ($\nu$)</strong></td>
<td>1.385**</td>
<td>0.330</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.494)</td>
<td>(0.494)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Drug-by-Year FE</strong></td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td><strong>N</strong></td>
<td>49,413,573</td>
<td>49,413,446</td>
<td>49,413,446</td>
<td>49,413,446</td>
<td>49,413,446</td>
</tr>
</tbody>
</table>

**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.
we find $\delta_{jt}$ that solves the equation:

$$m_{jt}^{\text{MEPS}} - \sum_{k \in J} x_{kt} m_{ijt} \left( k, \nu, \hat{\alpha}, \hat{\gamma}, \hat{\nu}, \hat{\beta} \right) = 0, \forall j, t \quad (13)$$

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 treatment options to statins are launched.

![Figure 4: Implied Quality Estimates](image)

**Notes:** A plot of the $\delta_{jt}$ 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.33

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 not pick a single formulary. Instead, the PBM randomly assigns each possible formulary to patients according to the probabilities defined in equation 9.34

To generate moment conditions, we assume that net prices are observed with measurement error $\eta_{jt}$. 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.35

To recover estimates for the vector of parameters $\hat{\theta} = (\theta_s, \theta_e, \theta_j, \theta_\omega)$ we minimize

33Appendix ?? 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.

34This 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.

35The 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.
the sum of squared errors:

$$\min_\theta \sum_t \sum_{j \notin G_t} (\hat{P}_{jt} (\cdot ; \hat{\alpha}, \hat{\gamma}, \hat{\nu}) - p_{jt}^o)^2$$

(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 $\theta$ and demand estimates $\hat{\alpha}, \hat{\gamma}, \hat{\nu}$ based on the first order conditions associated with Equation 10, and $p_{jt}^o$ 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 ??.

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 has a finite horizon. Hence, backward induction can be used to establish uniqueness (Doraszelski and Pakes, 2007; Besanko et al., 2010). Figure ?? in Appendix ?? 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, $\alpha$ represent the vector of demand parameters $(\alpha, \gamma, \nu)$, 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

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

where $P_k(\theta, \hat{\alpha})$ are the equilibrium prices of the finite period game under parameters
$\theta$ and $\hat{\alpha}$ (the demand estimates). The expression is minimized at the true $\alpha_0$ and $\theta_0$.

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 differen-
tiable, the asymptotic distribution of the second stage estimator is

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

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 $\theta_0$, $\Lambda_0 = \frac{\partial^2}{\partial \theta \partial \alpha} Q(\theta, \alpha)$ evaluated at $\theta = \theta_0, \alpha = \alpha_0$, $V_\alpha$
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 expres-
sion accounts for the impact of demand errors through a delta method approach, and
the second component is the standard minimum distance asymptotic variance expres-
sion.

We run our estimation routine to match each of the three negotiated prices series
we described in Section 4.1.\textsuperscript{36} We also use data on market size from MEPS, tier-specific
copay amounts $c_t$, $\bar{c}_t$ from MarketScan, and generic prices and copays from MEPS (set
at the median values $p_g = 1$ and $c_g = 0.33$).

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. $\theta_c$ 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. $\theta_j$ 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 con-
struction methodology. One concern with our core estimates is that the data do not

\textsuperscript{36}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

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Pct Discount</th>
<th>Dollar Discount</th>
<th>Private</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\theta_s$</td>
<td>1.70</td>
<td>1.67</td>
<td>1.68</td>
</tr>
<tr>
<td></td>
<td>(0.13)</td>
<td>(0.13)</td>
<td>(0.41)</td>
</tr>
<tr>
<td>$\theta_e$</td>
<td>1.50</td>
<td>1.55</td>
<td>1.54</td>
</tr>
<tr>
<td></td>
<td>(0.05)</td>
<td>(0.05)</td>
<td>(0.25)</td>
</tr>
<tr>
<td>$\theta_j$</td>
<td>1.11</td>
<td>1.01</td>
<td>0.89</td>
</tr>
<tr>
<td></td>
<td>(0.37)</td>
<td>(1.61)</td>
<td>(0.02)</td>
</tr>
<tr>
<td>$\theta_w$</td>
<td>0.063</td>
<td>0.064</td>
<td>0.044</td>
</tr>
<tr>
<td></td>
<td>(0.005)</td>
<td>(0.040)</td>
<td>(0.022)</td>
</tr>
</tbody>
</table>

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.

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 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. 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.

---

37 Figure ?? in Appendix on page ?? shows that equilibrium prices in a market without PBMs would be higher and increasing at a faster rate. Section 5 provides more details.
Figure 5: Model Fit – Predicted Prices

(a) Zocor

(b) Lipitor

(c) Crestor

Notes: Prices predicted by the best-fit model versus the prices observed in the data.
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 ?? in Appendix on page ??, do qualitatively line up with the data from Medicare 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.\(^{38}\) After Crestor enters, it is excluded with some probability.\(^{39}\) 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

\(^{38}\)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.

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 manufacturers 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 ?? in Appendix ?? 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.\textsuperscript{40} 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.\textsuperscript{41} 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

\textsuperscript{40}Appendix ?? 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.

\textsuperscript{41}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

<table>
<thead>
<tr>
<th>Scenario</th>
<th>Average Unit Price ($)</th>
<th>Payments ($bill)</th>
<th>Formulary Surplus</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Zocor</td>
<td>Lipitor</td>
<td>Crestor</td>
</tr>
<tr>
<td>Baseline</td>
<td>2.79</td>
<td>1.91</td>
<td>1.94</td>
</tr>
<tr>
<td><strong>Panel A – No PBM</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>33% Coinsurance</td>
<td>4.96</td>
<td>5.17</td>
<td>1.74</td>
</tr>
<tr>
<td>49% Coinsurance</td>
<td>3.33</td>
<td>3.47</td>
<td>1.19</td>
</tr>
<tr>
<td>Full margin pricing</td>
<td>2.60</td>
<td>2.67</td>
<td>1.00</td>
</tr>
<tr>
<td>Partial margin pricing</td>
<td>3.41</td>
<td>3.53</td>
<td>1.27</td>
</tr>
<tr>
<td><strong>Panel B – Committing to Formulary Structures</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Exclude (\leq 1)</td>
<td>4.77</td>
<td>3.82</td>
<td>2.49</td>
</tr>
<tr>
<td>Commit one exclusive</td>
<td>1.85</td>
<td>1.01</td>
<td>1.39</td>
</tr>
<tr>
<td>Commit one preferred</td>
<td>2.33</td>
<td>1.89</td>
<td>1.77</td>
</tr>
<tr>
<td>Only Preferred Tier</td>
<td>2.82</td>
<td>1.90</td>
<td>1.79</td>
</tr>
</tbody>
</table>

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.

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 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.\(^{42}\)

\(^{42}\)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 negative savings rates.
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.

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 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.43

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 $\leq 1$”). 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 exclude two or more drugs under the baseline model. 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

---

44Appendix ?? provides evidence that PBMs choose to exclude multiple drugs less than 2 percent of the time, except in the year before generic entry events.
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 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 ?? of Appendix ??, 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 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 re-
ducing 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.\textsuperscript{45}

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 struc-

ture, 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.
References


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


