Patient Costs and Physicians' Information*

Michael J. Dickstein[1], Jihye Jeon[2], and Eduardo Morales[3]

[1]New York University

[2]Boston University

[3]Princeton University

November 8, 2025

Abstract

Health plans in the U.S. increasingly use cost-sharing to steer demand for prescription drugs. However, the effectiveness of these incentives depends both on physicians' price sensitivity and their knowledge of patient costs. We employ a moment inequality model to identify physician preferences without fully specifying their information. Applying this model to diabetes care, we find that physicians generally lack detailed price information and are more price sensitive than full-information models imply. We also identify substantial heterogeneity in preferences and information by physician training, suggesting a benefit from targeted information interventions.

JEL Classifications: C13, I13.

Keywords: moment inequalities, pharmaceutical markets, physician decision-making.

*We thank Stéphane Bonhomme, Chris Conlon, Liran Einav, Haoran Pan, Ariel Pakes, Marc Rysman, Xiaoxia Shi, and Amanda Starc for insightful conversations, and seminar participants at ASSA Annual Meeting, CEPR Virtual IO Seminar, IIOC, Georgetown University, KAEA virtual seminar, Harvard University, MIT, New York University, Northwestern University, Pennsylvania State University, SITE IO of Healthcare and Credit Markets Conference, and Tinos Industrial Organization Conference for helpful comments. We especially thank Alon Eizenberg and Eilidh Geddes for very useful discussions. We also thank the staff at the Oregon Health Authority for assistance using the APAC dataset. All errors are our own. Email: michael.dickstein@nyu.edu, jjeon@bu.edu, ecmorale@princeton.edu.

1 Introduction

Real per capita annual expenditure on prescription drugs increased from $140 to more than $1,000 in the US between 1980 and 2018 (CBO, 2022). In response, as a way to change drug consumption patterns, private insurance plans have embedded price incentives in formularies that map drugs to tiers and require patients to pay more for higher-tier drugs. The adoption of such formularies is now widespread. In 2022, for example, 84% of workers with employer-sponsored health insurance had plans with at least three tiers (KFF, 2022).

The success of price incentives in shifting demand for prescription drugs depends on preferences—how physicians and patients value the efficacy of a treatment against its cost—and awareness of the monetary incentives. For policymakers seeking to steer prescription drug demand toward cheaper alternatives, it is critical to distinguish information from preferences. If, for example, usage of an expensive drug is high relative to lower-cost options, is it because the drug is more effective, because physicians and patients are price-inelastic, or because physicians are unaware of the differences in out-of-pocket costs?

We develop a model that allows us to estimate physicians' sensitivities to out-of-pocket costs and the value they place on a drug's efficacy. Importantly, we do so by allowing physicians to vary unobservably in the information they use to form expectations about drug- and patient-specific out-of-pocket costs. Combining estimates of physicians' preferences with inferences about their information, we evaluate the effect of policies to inform physicians about out-of-pocket costs at the point of prescribing. Our results suggest substantial benefits from targeting information interventions, in our setting by prioritizing physicians with general practice training over specialist physicians.

We focus our analysis on the study of prescription drug choices for patients with type 2 diabetes. We choose diabetes care as our market of interest due to both the size of the affected population and the rapid growth in treatment costs.[1] Using Oregon's All Payer All Claims database (Oregon Health Authority, 2011-2016) for the years 2011 to 2016, we form a sample of prescription drug insurance claims for diabetic patients covered by private insurance. We collect treatment choices, patient prices, and patient and physician characteristics. Using these data, we start by documenting two facts. First, there is significant dispersion in out-of-pocket costs, both across insurance plans for a given drug and across drugs and plan types for a given insurer. For example, for the class of treatments known as DPP-4 inhibitors, the mean out-of-pocket cost for a 30-day supply lies between $42 and $46, depending on the drug. However, the interquartile range computed across insurers and plans for a given

[1] In the U.S., for example, 37 million people lived with diabetes in 2019 (CDC, 2022) and the medical costs of diabetes totaled $237 billion in 2017, rising 26% above 2012 inflation-adjusted levels (ADA, 2017).

drug-year pair is large relative to the mean, ranging between $22 and $28, depending on the drug. Second, we observe that physicians often choose drugs that are not the patient's cheapest option. For example, although there are only three drugs in the class of DPP-4 inhibitors, physicians choose the lowest-cost option only 38% of the time, roughly the same as if they had chosen randomly. If physicians had selected the cheapest alternative at each visit, the average patient would have saved $17 per month.

The observed relationship between drug choices and out-of-pocket costs may reflect preferences or could indicate that physicians lack information on the price incentives the patient faces. To separate these two mechanisms, we develop a model of prescription drug choice in which the physician selects a treatment based on the effectiveness of each drug and her sensitivity to *expected* out-of-pocket costs. Although we assume physicians' expectations are rational, we allow the information set to vary flexibly between physicians and patient visits. In doing so, physicians' expectations function like unobserved covariates in our model.

We estimate the model using a moment inequality procedure that combines two sets of moments. The first set, labeled "odds-based" moments, generalizes the approach in Dickstein and Morales (2018) to settings with more than two choices. The second set, labeled "bounding" moments, follows Porcher et al. (2025). When researchers combine these moment inequalities with instrument functions that depend on variables that belong to the physician's information set, the resulting identified set includes the true value of the preference parameters. Thus, our moment inequalities provide bounds on preference parameters even when the researcher only partly observes the agent's information set.

Before applying our model to the study of diabetes care, we perform a simulation to illustrate the properties of both our moment inequalities and alternative full-information estimation approaches. We have four conclusions. First, following Manski (1991), we show the maximum likelihood estimates of the preference parameters are inconsistent when the researcher incorrectly specifies the agent's information set. Second, consistent with our theoretical analysis, the moment inequalities yield an identified set that contains the true parameter value when our instruments—i.e. the variables we assume agents know—form a *subset* of agents' true information sets. Third, if our instruments coincide with the agent's complete information set, the identified set defined by the moment inequalities is a singleton; i.e., it includes the true parameter only. Fourth, and finally, we show that a researcher can use specification tests of moment inequality models to test whether a vector of covariates belongs to agents' information sets.

In our study of diabetes care, we use our insurance claims data to quantify the determinants of the physician's treatment choice. We recover and compare estimates from both a traditional maximum likelihood approach and an inequality approach. We then use the

inequality framework to test several assumptions on the content of physicians' information sets with the goal of learning how physicians form their price expectations. Finally, we use the estimated model to predict the effect of an intervention that provides patient-specific price information to physicians at the time of prescribing.

In our application, we find that maximum likelihood estimates of preference parameters vary significantly with the specification of the physician's information set. For example, for the own-price elasticity of a product in the choice set, our estimates imply an elasticity equal to -0.58 when we assume that providers know the patient-specific costs for each drug. This estimate is equal to -1.69 when we assume instead that providers only know the average prices by drug and plan type, and it equals -3.57 when we assume that providers only know the average price of each drug in the previous year.

In addition, we find that the different informational assumptions and the corresponding parameter estimates also imply different predictions for how demand reacts to counterfactual changes in out-of-pocket costs. As an illustration, we consider a policy change in which three insurers, which collectively represent about half of the patients in our sample, decide to cut out-of-pocket costs for a drug by 50%. Depending on which information set the researcher assumes, the empirical model predicts the now cheaper drug's prescription share will increase anywhere from 3.8 to 26.1 percentage points. Here, the models that assume physicians know the most about prices predict the smallest increase, as these models, when combined with the data, find relatively inelastic demand.

This sensitivity to informational assumptions, both in parameter estimates and counterfactual predictions, motivates our shift to a moment inequality framework. With our inequalities, we first recover confidence sets on preference parameters and test assumptions on the shared content of physicians' information sets, treating the preferences of all physicians as homogeneous. In our setting, we reject the null hypothesis that all physicians have perfect information on out-of-pocket costs. Instead, our data and model suggest that physicians' information sets, when tested collectively, consist only of coarse price measures, such as the prior year's average price for each drug by insurance plan type. By concluding that physicians lack perfect information about patient costs, we depart from the common approach in the literature that uses realized patient costs in prescription drug choice models (see literature reviews by Goldman et al., 2007; Baicker and Goldman, 2011).

Next, we use our moment inequalities to study whether physicians are heterogeneous in their preferences and in the information they use to form their price predictions. To perform this analysis, we first divide physicians into groups based on five criteria: medical specialty, graduation year, sex, rank of the medical school attended, and concentration of insurance plan types among the physician's patient population. We then test whether physicians have

different price information between groups, while also allowing each group to differ in their price sensitivity and their perception of the efficacy of each drug. We find significant heterogeneity in information and preferences by physician training. In particular, endocrinologists appear to have more detailed price information than primary care physicians. In addition, we estimate that endocrinologists are less price sensitive than primary care physicians, while both types of physicians broadly agree on the relative efficacy of each drug.

Finally, given that we find heterogeneity in information when comparing groups of physicians, we evaluate the effect of targeting an information intervention. In particular, we analyze the effects of providing information to different physician specialties. While we predict endocrinologists are only 1.4% to 3.8% more likely to prescribe the cheapest drug in the choice set when given full information on patient prices, primary care physicians are 6.8% to 32.5% more likely to do so. This difference arises because endocrinologists, in our analysis, have both better price information in the status quo and put less weight on patient costs in their prescription decisions. Thus, if providing physicians with patient-specific prices at the point of prescribing is costly, our results suggest a value of targeting this provision toward general practitioners.[2]

Our paper relates to several research areas. First, we contribute to a literature that studies physicians' drug treatment choices. This literature studies the influence of several factors, including financial incentives (Iizuka, 2012; Dickstein, 2018), advertising and detailing (Ching and Ishihara, 2012; Grennan et al., 2024), as well as the interplay with secondary markets (Schnell, 2025) and competitive forces (Currie et al., 2025). Our results emphasize the role that physician information plays in prescribing behavior: we show that price elasticity estimates are biased if researchers misspecify physicians' information sets, and we present evidence that indicates physicians' price information is imperfect. That physicians form price expectations using only aggregate price information is consistent with Shrank et al. (2005), who use survey data to illustrate that physicians have limited information on prices, and with Carrera et al. (2018) and Desai et al. (2022), who show that prescribing behavior reacts to informational shocks about out-of-pocket costs.

Second, our research adds to a literature studying the role of information frictions in healthcare markets, summarized in Handel and Schwartzstein (2018). These frictions arise in health insurance choice (Handel and Kolstad, 2015; Handel et al., 2019; Brown and Jeon, 2024), and also in the process through which physicians determine the quality of treatments (Crawford and Shum, 2005; Chintagunta et al., 2009; Ching, 2010). We depart from this

[2]The difference between endocrinologists and primary care physicians in information and price sensitivity may reflect differential training or a different patient population; e.g., if endocrinologists treat more severely ill patients, they may value efficacy over price more.

literature in that we do not micro-found physicians' information sets, but rather apply moment inequalities to infer their content. In this way, our paper relates to Ito (2014), who provides a framework to uncover a consumer's perceived price, finding evidence that consumers behave suboptimally.

Finally, we contribute to a literature that uses moment inequalities to estimate agents' preferences. This literature, reviewed in Kline et al. (2021), Kline and Tamer (2023), and Canay et al. (2023), has early examples in Pakes (2010), Holmes (2011), and Pakes et al. (2015). Prior applications of moment inequalities in the healthcare context include Ho (2009), Ho and Pakes (2014), and Maini and Pammolli (2023). Applications in other contexts include Eizenberg (2014), Illanes (2017), Wollmann (2018), Morales et al. (2019), and Houde et al. (2023). Our contribution is to generalize the odds-based moment inequality introduced in Dickstein and Morales (2018), and subsequently applied in Bombardini et al. (2023), to discrete choice settings with more than two choices. In our analysis, we apply our odds-based inequalities as well as the bounding inequalities introduced in Porcher et al. (2025) to the multinomial logit model. In the context of drug choice, we demonstrate that these inequalities allow us to recover informative bounds on physicians' preferences. We also show how to derive analogous inequalities for nested logit and multinomial probit models.

The rest of the paper proceeds as follows. In Section 2, we describe our setting and data, and present statistics that motivate our analysis. In sections 3 to 5, we present a model of prescription drug choice, introduce the inequalities we use for estimation, and describe how we implement them in our setting. In Section 6, we present simulation results comparing the properties of the moment inequality and maximum likelihood estimators. In sections 7 and 8, we present the estimates and counterfactual results we obtain when applying both maximum likelihood and moment inequality estimators to the diabetes setting. In Section 9, we derive moment inequalities for alternative discrete choice models. Section 10 concludes.

2 Empirical Setting and Data

Our analysis focuses on care for type 2 diabetes patients. In Section 2.1, we describe the typical treatment protocol. In Section 2.2, we present our data along with descriptive statistics on out-of-pocket costs and physicians' treatment choices.

2.1 Diabetes Care

Treatment of type 2 diabetes often begins after abnormal results for a fasting plasma glucose test or a hemoglobin A1c test. Following the guidelines from American Diabetes Association

(2017), physicians generally start patients on metformin, an inexpensive oral medication. If the patient does not achieve the desired average blood glucose level, the physician may add a second therapeutic treatment. At this dual-therapy stage, the physician can choose among several classes of diabetes drugs, including DPP-4 inhibitors and SGLT2 inhibitors, among others. Drugs in these classes differ in their efficacy and side effects. For example, DPP-4 inhibitors and SGLT2 inhibitors have similar efficacy in controlling blood glucose. However, relative to patients on DPP-4 inhibitors, patients on SGLT2 inhibitors see greater weight loss but also a higher risk of bone fractures.

In the baseline, we focus on patients in the dual-therapy stage and study the choice of treatment within the class of DPP-4 inhibitors. As a robustness check, we later consider choices both across and within the classes of DPP-4 and SGLT2 inhibitors.

2.2 Data

We use data from Oregon's All-Payer All-Claims (APAC) database for the period 2011-2016. Our sample includes medical and prescription drug claims for patients with private insurance through the individual insurance market and through group insurance. For each medical claim, we observe the patient's diagnosis as well as patient demographics, insurance coverage, and the identity of the patient's healthcare provider. We link these medical claims to the patient's drug claims, where we observe the treatment prescribed and the patient's out-of-pocket cost. Finally, we validate the information on physician characteristics in the APAC data with data from two public registries that contain information on physicians' specialty, sex, and medical school graduation year.[3]

Sample creation. We focus our analysis on a sample of claims where providers, patients, and treatments satisfy certain conditions. First, we only consider claims linked to physician specialties that typically provide care and medication management for patients with diabetes, including family medicine, internal medicine, and endocrinology. Second, for the main analysis, we focus on patients who receive a diagnosis of type 2 diabetes and who are prescribed a drug in the class of DPP-4 inhibitors; in robustness analyses, we add patients to the sample who receive drugs in other dual-therapy classes. In addition, we restrict our sample in several minor ways, such as excluding drug claims that reflect refills rather than active physician choices; see Appendix A.1 for more details.

The resulting sample includes 10,956 claims prescribed by 1,513 providers for a set of 3,729 patients. Among providers, 95% are primary care physicians with specialties in in-

[3]Specifically, we use the National Plan and Provider Enumeration System registry (CMS, 2011-2016b) and the Doctors and Clinicians National Downloadable File (CMS, 2011-2016a). Throughout the paper, we use "physician" to refer to medical doctors and other providers who prescribe treatments in our sample.

ternal medicine or family medicine, while the remaining 5% are endocrinologists. Although endocrinologists make up a small fraction of providers in our sample, they represent a larger share of claims: on average, endocrinologists record 67 claims with a diagnosis of type 2 diabetes per quarter, compared to 19 per quarter for primary care physicians.

The average patient is 57 years old. More than half of patients enroll in a preferred provider organization (PPO) plan, 19% choose a health maintenance organization (HMO) or a point of service (POS) plan, and the remaining 24% have self-insured plans. The largest carrier enrolls 18% of the sampled patients, and the four largest carriers by patient volume jointly account for 40% of the patients.

Drug choice. The class of DPP-4 inhibitors includes three drug treatments during our sample period. The most popular treatment accounts for approximately 70% of non-refill drug claims, with the remaining 30% of claims distributed roughly equally between the other two treatments. We use copayments as our measure of the out-of-pocket costs patients face when filling a prescription.[4] When specifying the copayment levels for all drugs and patients, however, we face a missing data problem. We do not observe the full drug formulary each patient faces; using our claims data, we can only infer copayments using observations from patients who filled a prescription. To generate the complete cost list for a patient, we employ a random forest model that uses our observed data to impute missing drug prices for all plans and years. We provide more details on this imputation in Appendix A.2.

In Table 1, we document significant heterogeneity in monthly copayments. We report statistics of the distribution of copayments across plans for each drug in panel A, and across plans and drugs for each carrier in panel B. For example, for the drug Janumet, the mean monthly copayment is close to $42 and the standard deviation is close to $28. The coefficient of variation is thus 0.68, only slightly larger than the coefficient of variation for Januvia and Tradjenta, the other two DPP-4 inhibitor drugs available in our sample. Comparing both panels in Table 1, we find that the dispersion in copayments within a carrier is generally smaller than the dispersion within a drug, with coefficients of variation for within-carrier dispersion equal to roughly 0.35. Carrier B is an exception, with a larger dispersion.

Given the significant heterogeneity in copayments, both across insurance plans for a given drug and across drugs and plans for a given carrier, physicians may find it difficult to predict the specific copayment an insured patient would face for each drug. In Figure 1, we show that, conditional on prescribing one of the three drugs in the class of DPP-4 inhibitors, physicians prescribe the cheapest option for each patient at roughly the same rate as if they had chosen the drug randomly. Patients who do not receive the cheapest drug pay an average of $17 more per month in out-of-pocket costs.

[4]Deductibles and coinsurance are non-zero for only 3% and 4% of the patients in our sample, respectively.

Table 1: Distribution of Monthly Copayments for DPP-4 Inhibitors

Panel A: By Drug

Drug	Mean	St. Dev.	IQR
Janumet	41.62	28.30	22.21
Januvia	43.87	27.95	24.12
Tradjenta	46.24	22.10	27.76

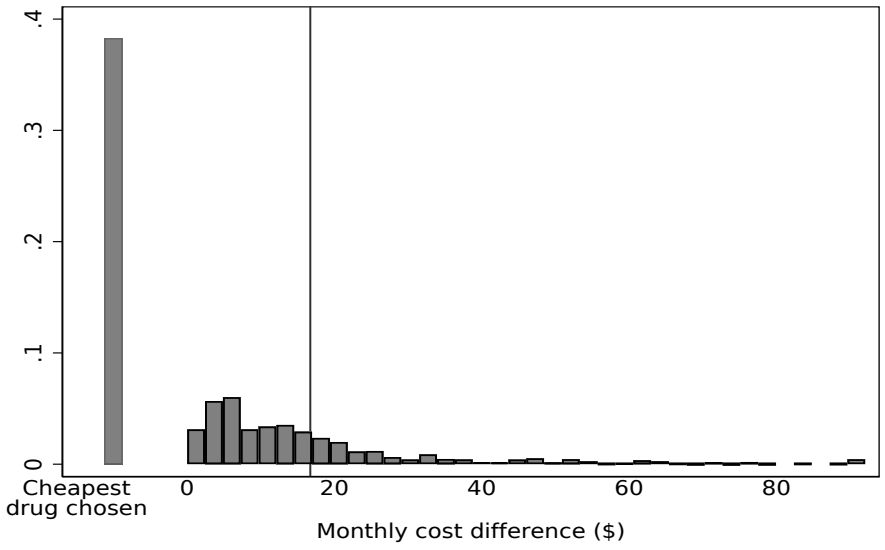
Panel B: By Carrier

Carrier	Mean	St. Dev.	IQR
A	42.22	16.15	13.38
B	36.74	22.56	25.03
C	80.34	28.60	35.38
D	31.23	11.24	15.85

Note: We report summary statistics of the distribution of out-of-pocket costs (in $ per month) in our sample. In each panel, we report the mean, standard deviation (*St. Dev.*), and interquartile range (*IQR*). *St. Dev.* and *IQR* are computed after residualizing costs to take out drug-year fixed effects (in panel A) or carrier-year fixed effects (in panel B). Panel A reflects variation in prices across plans for each drug, while panel B reflects variation across plans and drugs for each carrier.

The choice pattern in Figure 1 may (a) reflect physicians' preferences for drug characteristics other than price; (b) indicate that physicians lack information about their patients' out-of-pocket costs for drugs in the choice set; or (c) reflect estimation noise due to our need to predict copayments for drug-plan pairs not observed in the data. To better understand physicians' information and preferences, we next present a model and estimation approach that accounts both for quality differences across drugs and for the possibility that physicians have imperfect information about patient prices. Our estimation approach can also account for classical measurement error in measured prices.

Figure 1: Out-of-pocket Costs and Physicians' Choices for DPP-4 Inhibitors



Notes: We report the observed probability of choosing the drug with the lowest out-of-pocket cost in the class as well as the distribution of differences in monthly copayments between the observed drug chosen for a patient and the cheapest drug available to that patient in the class, conditional on the difference being greater than zero. We winsorize the cost differences at the 99th percentile. The vertical line represents the mean monthly difference in costs conditional on the difference being positive.

3 Model of Prescription Choice

We model a physician's choice of prescription drug at each patient visit. We index visits by i and drugs by j. At each visit i, we assume the physician's utility from choosing drug j is

$$\mathcal{U}_{ij} = u_{ij} + \varepsilon_{ij}, \tag{1a}$$

$$u_{ij} = u(x_{ij}, \mu), \tag{1b}$$

where x_{ij} is a vector of characteristics of drug j at visit i, μ is a vector of preference parameters, and ε_{ij} is the physician's idiosyncratic preference for drug j. Defining a variable d_{ij} that equals one if the physician prescribes drug j at visit i and zero otherwise, we assume

$$d_{ij} \equiv \mathbb{1}\{\mathbb{E}[\mathcal{U}_{ij}|\mathcal{J}_i] \geqslant \max_{j'=1,\ldots,J} \mathbb{E}[\mathcal{U}_{ij'}|\mathcal{J}_i]\}, \qquad \text{for } j = 1, \ldots, J, \tag{2}$$

where J is the cardinality of the set of drugs that the physician may prescribe at visit i, $\mathcal{J}_i$ is the physician's information set, and $\mathbb{E}[\cdot|\mathcal{J}_i]$ is a conditional expectation operator reflecting the physician's beliefs. We assume physicians' expectations are rational. Thus, for any random vector $\mathcal{A}_i$, $\mathbb{E}[\mathcal{A}_i|\mathcal{J}_i]$ denotes the expectation with respect to the distribution of $\mathcal{A}_i$ conditional on $\mathcal{J}_i$ in the population of interest. As a result, physicians' expectational errors are mean zero conditional on their information set.

We impose the following assumption on physicians' information sets:

$$\varepsilon_i \subseteq \mathcal{J}_i \tag{3}$$

where $\varepsilon_i = \{\varepsilon_{ij}\}_{j=1}^J$ and, for any two random vectors $\mathcal{A}_i$ and $\mathcal{B}_i$, we use $\mathcal{A}_i \subseteq \mathcal{B}_i$ to denote that the distribution of $\mathcal{A}_i$ conditional on $\mathcal{B}_i$ is degenerate. As indicated in equation (3), $\mathcal{J}_i$ includes the vector of idiosyncratic preferences ε_i, but other variables may also enter the information set. To clarify our notation, we distinguish $\mathcal{J}_i$ from $\mathcal{W}_i$, where the latter is the subset of $\mathcal{J}_i$ that includes all its elements except ε_i.

We impose two sets of assumptions on the distribution of ε_i. First, we assume that

$$F_\varepsilon(\varepsilon_i|\mathcal{W}_i) = F_\varepsilon(\varepsilon_i) = \exp(-\sum_{j=1}^{J} \exp(-\varepsilon_{ij})), \tag{4}$$

where $F_\varepsilon(\cdot)$ is the cumulative distribution function of ε_i. Equation (4) imposes ε_i is independent of all other elements of the physician's information set, as included in $\mathcal{W}_i$. It also imposes that ε_{ij} is *iid* across all drugs and follows a type I extreme value distribution with

location parameter equal to zero and scale parameter equal to one.[5] Second, we assume that

$$\mathbb{E}[u_i|\mathcal{J}_i] = \mathbb{E}[u_i|\mathcal{W}_i], \tag{5}$$

with $u_i = \{u_{ij}\}_{j=1}^J$. Equation (5) imposes that, once we condition on all other elements of the physician's information set, the vector ε_i does not provide any additional information that helps the physician forecast the utility terms in u_i. Equation (5) holds if $x_{ij} \subseteq \mathcal{W}_i$ for all j and, thus, the physician has full information on u_i. However, it also holds if the physician has imperfect information on x_{ij} for some drugs, as long as all information relevant to the physician's forecast of x_{ij} for those drugs is in $\mathcal{W}_i$.

Equations (1) to (5) imply that the probability the physician chooses drug j given $\mathcal{W}_i$ is

$$\mathcal{P}(d_{ij} = 1|\mathcal{W}_i) = \frac{\exp(\mathbb{E}[u(x_{ij}, \mu)|\mathcal{W}_i])}{\sum_{j'=1}^J \exp(\mathbb{E}[u(x_{ij'}, \mu)|\mathcal{W}_i])} \qquad \text{for } j = 1, \ldots, J. \tag{6}$$

Given the assumption of rational expectations, if we specify the content of the information set $\mathcal{W}_i$, then we can compute $\mathbb{E}[u(x_{ij}, \mu)|\mathcal{W}_i]$ for all j and our model becomes a multinomial logit model. Without any restriction on physicians' expectations, observing the choice probability $\mathcal{P}(d_{ij} = 1|\mathcal{W}_i)$ for all $j = 1, \ldots, J$ would generally not allow us to distinguish how the preference parameters, μ, or information set, $\mathcal{W}_i$, affect the treatment choices. As a middle ground, we assume that physicians' expectations are rational, and show in Section 4 that this assumption allows us to derive moment inequalities that are informative about both the preference parameters and the content of physicians' information sets.

Finally, we assume the researcher collects a random sample of N visits. For each sampled visit, the researcher observes the physician's choice set, the drug prescribed $d_i = \{d_{ij}\}_{j=1}^J$, the characteristics in the vector $x_i = \{x_{ij}\}_{j=1}^J$, and a vector of covariates $z_i = \{z_{ij}\}_{j=1}^J$ that may be used to predict x_i. Crucially, we do not impose that z_i is included in $\mathcal{W}_i$ and, more generally, we do not assume that the researcher observes the complete set $\mathcal{W}_i$ for any visit.[6]

The goal of estimation is to recover the value of the parameter vector μ and to learn about the content of the information sets $\{\mathcal{W}_i\}_{i=1}^N$. We use θ to denote the unknown parameter vector whose true value is μ and Θ to denote the parameter space. To infer the content of $\{\mathcal{W}_i\}_{i=1}^N$, we consider testing the null hypothesis that z_i belongs to $\mathcal{W}_i$ for every physician i in a group of interest; i.e., testing $H_0\colon z_i \subseteq \mathcal{W}_i$ for a subset of visits.

[5] In Section 9, we consider alternative models where we allow ε_{ij} to be correlated across drugs.

[6] In our empirical setting, the set of possible treatments is small and common to all physicians; thus, it is reasonable to assume that every physician's consideration set coincides with the actual choice set available. See Abaluck and Adams-Prassl (2021) and Barseghyan et al. (2021a,b) for methods for discrete choice analysis when the researcher does not observe the agent's consideration set.

4 Moment Inequalities

In this section, we show how to partially identify μ. We use two types of moment inequalities, odds-based and bounding inequalities, which we describe in sections 4.1 and 4.2, respectively. In Section 4.3, we discuss a connection between the two types of inequalities.

4.1 Odds-based Inequalities

For any two drugs j and j' in the choice set, any value of z_i in its support $\mathcal{Z}$, and any value of θ in the parameter space Θ, we define the following odds-based moment inequality

$$\mathbb{m}^o_{jj'}(z_i, \theta) \geqslant 0 \tag{7a}$$

with

$$\mathbb{m}^o_{jj'}(z_i, \theta) \equiv \mathbb{E}[d_{ij} \exp(-(u(x_{ij}, \theta) - u(x_{ij'}, \theta))) - d_{ij'} | z_i]. \tag{7b}$$

Here, we denote as Θ^o_0 the set of values of θ that satisfy equation (7) for every value of z_i in its support and all pairs of drugs j and j' in the physician's choice set:

$$\Theta^o_0 \equiv \{\theta \in \Theta \colon \mathbb{m}^o_{jj'}(z, \theta) \geqslant 0 \text{ for all } z \in \mathcal{Z},\ j = 1, \ldots, J, \text{ and } j' = 1, \ldots, J\}. \tag{8}$$

Theorem 1 establishes a sufficient condition for the true parameter value μ to belong to Θ^o_0.

Theorem 1 *Let μ be defined by equation (6). If $z_i \subseteq \mathcal{W}_i$, then $\mu \in \Theta^o_0$.*

Theorem 1 states that, when evaluated at the true parameter value, the inequality in equation (7) holds if z_i belongs to the information set $\mathcal{W}_i$ for every visit i in the population. This inequality holds regardless of both the value of z_i on which we condition and the drugs j and j' we use to define the moment, provided that both drugs are in the choice set. We provide a sketch of the proof of Theorem 1 below. The formal proof appears in Appendix B.1.

The moment inequality in equation (7) is a generalization to multinomial settings of the odds-based inequality introduced in Dickstein and Morales (2018) for binary choice models. To understand why the inequality in equation (7) holds for $\theta = \mu$, a key equation is

$$\mathbb{E}[d_{ij} \exp(-\mathbb{E}[\Delta u_{ijj'} | \mathcal{W}_i]) - d_{ij'} | \mathcal{W}_i] = 0, \tag{9}$$

with $\Delta u_{jj'} = u_{ij} - u_{ij'}$. Equations (1b) and (6) imply equation (9) for any j and j' in the physician's choice set and any $\mathcal{W}_i$. As d_{ij} is measurable in $\mathcal{J}_i$, $\exp(x)$ is convex in x, and

$\Delta u_{ijj'}$ is a mean-preserving spread of $\mathbb{E}[\Delta u_{ijj'}|\mathcal{J}_i]$, Jensen's inequality implies

$$\mathbb{E}[d_{ij}\exp(-\mathbb{E}[\Delta u_{ijj'}|\mathcal{J}_i])|\mathcal{J}_i] \leqslant \mathbb{E}[d_{ij}\exp(-\Delta u_{ijj'})|\mathcal{J}_i]. \tag{10}$$

Using equation (5), we can rewrite this inequality as:

$$\mathbb{E}[d_{ij}\exp(-\mathbb{E}[\Delta u_{ijj'}|\mathcal{W}_i])|\mathcal{J}_i] \leqslant \mathbb{E}[d_{ij}\exp(-\Delta u_{ijj'})|\mathcal{J}_i].$$

Furthermore, as $\mathcal{W}_i \subseteq \mathcal{J}_i$, applying the Law of Iterated Expectations (LIE), we obtain

$$\mathbb{E}[d_{ij}\exp(-\mathbb{E}[\Delta u_{ijj'}|\mathcal{W}_i])|\mathcal{W}_i] \leqslant \mathbb{E}[d_{ij}\exp(-\Delta u_{ijj'})|\mathcal{W}_i]. \tag{11}$$

The equality in equation (9) and the inequality in equation (11) jointly imply

$$\mathbb{E}[d_{ij}\exp(-\Delta u_{ijj'}) - d_{ij'}|\mathcal{W}_i] \geqslant 0.$$

Finally, for any $z_i \subseteq \mathcal{W}_i$, applying the LIE again, we conclude that

$$\mathbb{E}[d_{ij}\exp(-\Delta u_{ijj'}) - d_{ij'}|z_i] \geqslant 0. \tag{12}$$

This inequality equals the inequality in equation (7) for $\theta = \mu$, proving Theorem 1. Equation (12) holds with equality if equation (10) holds with equality. This outcome occurs if physicians have perfect information, meaning that $\mathbb{E}[\Delta u_{ijj'}|\mathcal{J}_i] = \Delta u_{ijj'}$. When physicians make expectational errors, Jensen's inequality implies the left-hand side of equation (12) is strictly positive. As we illustrate in a simulation exercise in Section 6, expectational errors make the moment inequality in equation (7) weaker and, as a result, the set of parameter values Θ_0^o larger.

4.2 Bounding Inequalities

For any two drugs j and j' in the physician's choice set, any value of z_i in its support $\mathcal{Z}$, and any function $e_{jj'}\colon \mathcal{Z} \times \Theta \to \mathbb{R}$, we define the following bounding moment inequality

$$\mathbb{m}^b_{jj'}(z_i, \theta, e_{jj'}(\cdot)) \leqslant 0 \tag{13a}$$

with

$$\mathbb{m}^b_{jj'}(z_i, \theta, e_{jj'}(\cdot)) \equiv$$
$$\mathbb{E}[d_{ij}\exp(-e_{jj'}(z_i,\theta))(1 + e_{jj'}(z_i,\theta) - (u(x_{ij},\theta) - u(x_{ij'},\theta))) - d_{ij'}|z_i]. \tag{13b}$$

The moment $\mathbb{m}^b_{jj'}(\cdot)$ depends on $e_{jj'}(z_i, \theta)$, which is a deterministic function of the observed vector z_i and the unknown parameter vector θ; $e_{jj'}(\cdot)$ may differ for any two drugs j and j'. We denote by e the set of $e_{jj'}(\cdot)$ for all pairs of drugs, and we denote by $\Theta^b_0(e)$ the set of values of θ that satisfy the inequality in equation (13) for every value of z_i in its support and for every two drugs j and j' in the physician's choice set. Formally,

$$\Theta^b_0(e) \equiv \{\theta \in \Theta \colon \mathbb{m}^b_{jj'}(z, \theta, e_{jj'}(\cdot)) \leqslant 0 \text{ for all } z \in \mathcal{Z},\ j = 1, \ldots, J, \text{ and } j' = 1, \ldots, J\}. \quad (14)$$

The following theorem, from Porcher et al. (2025), establishes a sufficient condition for μ to belong to $\Theta^b_0(e)$.

Theorem 2 *Let μ be defined by equation (6). If $z_i \subseteq \mathcal{W}_i$, then $\mu \in \Theta^b_0(e)$ for any set e of functions $e_{jj'} \colon \mathcal{Z} \times \Theta \to \mathbb{R}$.*

Theorem 2 indicates that, when evaluated at the true parameter value, the inequality in equation (13) holds if, for all visits in the population of interest, z_i belongs to $\mathcal{W}_i$, and the drugs j and j' are in the choice set. Importantly, this inequality holds regardless of the function $e_{jj'}(z_i, \theta)$ used to build it. We provide here a sketch of the proof of Theorem 2. The formal proof, which reproduces results in Porcher et al. (2025), is in Appendix B.2.

To show why the inequality in equation (13) holds for $\theta = \mu$, we start again from equation (9). However, the derivation of the bounding inequality uses a different implication of the convexity of $\exp(x)$ in x. In particular, as $\exp(x)$ is convex in x, a first-order approximation to it around any point bounds it from below. Thus, for any function $e_{jj'} \colon \mathcal{Z} \times \Theta \to \mathbb{R}$, the following inequality can be derived from the equality in equation (9):

$$\mathbb{E}[d_{ij} \exp(-e_{jj'}(z_i, \mu))(1 + e_{jj'}(z_i, \mu) - \mathbb{E}[\Delta u_{ijj'} | \mathcal{W}_i]) - d_{ij'} | \mathcal{W}_i] \leqslant 0, \quad (15)$$

where $e_{jj'}(z_i, \mu)$ is the point around which we take the first-order approximation. As d_{ij} is measurable with respect to $\mathcal{J}_i$ and $\Delta u_{ijj'}$ is a mean-preserving spread of $\mathbb{E}[\Delta u_{ijj'} | \mathcal{J}_i]$, the following equality holds for any $z_i \subseteq \mathcal{W}_i$:

$$\begin{aligned} &\mathbb{E}[d_{ij} \exp(-e_{jj'}(z_i, \mu))(1 + e_{jj'}(z_i, \mu) - \mathbb{E}[\Delta u_{ijj'} | \mathcal{J}_i]) | \mathcal{J}_i] = \\ &\mathbb{E}[d_{ij} \exp(-e_{jj'}(z_i, \mu))(1 + e_{jj'}(z_i, \mu) - \Delta u_{ijj'}) | \mathcal{J}_i]. \end{aligned}$$

Using equation (5) and applying the LIE, we obtain

$$\begin{aligned} &\mathbb{E}[d_{ij} \exp(-e_{jj'}(z_i, \mu))(1 + e_{jj'}(z_i, \mu) - \mathbb{E}[\Delta u_{ijj'} | \mathcal{W}_i]) | \mathcal{W}_i] = \\ &\mathbb{E}[d_{ij} \exp(-e_{jj'}(z_i, \mu))(1 + e_{jj'}(z_i, \mu) - \Delta u_{ijj'}) | \mathcal{W}_i]. \end{aligned} \quad (16)$$

The inequality in equation (15) and the equality in equation (16) jointly imply

$$\mathbb{E}[d_{ij} \exp(-e_{jj'}(z_i, \mu))(1 + e_{jj'}(z_i, \mu) - \Delta u_{ijj'}) - d_{ij'}|\mathcal{W}_i] \leqslant 0.$$

Finally, as $z_i \subseteq \mathcal{W}_i$, applying the LIE again, we conclude that

$$\mathbb{E}[d_{ij} \exp(-e_{jj'}(z_i, \mu))(1 + e_{jj'}(z_i, \mu) - \Delta u_{ijj'}) - d_{ij'}|z_i] \leqslant 0. \tag{17}$$

Given the expression for u_{ij} in equation (1b), this inequality equals the inequality in equation (13) for $\theta = \mu$, proving Theorem 2.

4.3 Connecting the Odds-based and Bounding Inequalities

Using the bounding inequality in equation (13) requires choosing the approximation point $e_{jj'}(z_i, \theta)$ for any two drugs j and j' and any value of the vectors z_i and θ. This choice is consequential for the size of the set of parameter values that satisfy the bounding inequality in equation (13). Given two drugs j and j' and a value of z_i, this set is minimized when $e_{jj'}(z_i, \theta)$ maximizes the moment in equation (13b) at each value of θ. As shown in Porcher et al. (2025), reproduced in Appendix B.3.1, the function $e_{jj'}(z_i, \theta)$ that attains this goal is

$$e_{jj'}(z_i, \theta) = \mathbb{E}[u(x_{ij}, \theta) - u(x_{ij'}, \theta)|z_i, d_{ij} = 1]. \tag{18}$$

Combining the inequality in equation (13) with the approximation point in equation (18), we can rewrite the resulting inequality, when evaluated at the true parameter value, as

$$\mathbb{E}[d_{ij} \exp(-\mathbb{E}[\Delta u_{ijj'}|z_i, d_{ij}]) - d_{ij'}|z_i] \leqslant 0. \tag{19}$$

In Appendix B.4, we show an alternative derivation of this inequality using steps analogous to those we use to derive the odds-based inequality; that is, we use Jensen's inequality to obtain the key inequality in the derivation. We provide a sketch of the derivation here.

For any $z_i \subseteq \mathcal{W}_i$, equation (9) and the LIE implies the following equality:

$$\mathbb{E}[d_{ij} \exp(-\mathbb{E}[\Delta u_{ijj'}|\mathcal{W}_i]) - d_{ij'}|z_i] = 0. \tag{20}$$

As $\exp(x)$ is convex in x and $\mathbb{E}[\Delta u_{ijj'}|\mathcal{J}_i]$ is a mean-preserving spread of $\mathbb{E}[\Delta u_{ijj'}|z_i, d_{ij}]$ for any $z_i \subseteq \mathcal{W}_i$, Jensen's inequality applies and we have

$$\mathbb{E}[d_{ij} \exp(-\mathbb{E}[\Delta u_{ijj'}|z_i, d_{ij}])|z_i, d_{ij}] \leqslant \mathbb{E}[d_{ij} \exp(-\mathbb{E}[\Delta u_{ijj'}|\mathcal{J}_i])|z_i, d_{ij}]. \tag{21}$$

Using equation (5) and applying the LIE to eliminate d_{ij} from the conditioning set, we have

$$\mathbb{E}[d_{ij}\exp(-\mathbb{E}[\Delta u_{ijj'}|z_i,d_{ij}])|z_i] \leqslant \mathbb{E}[d_{ij}\exp(-\mathbb{E}[\Delta u_{ijj'}|\mathcal{W}_i])|z_i]. \tag{22}$$

Equations (20) and (22) imply the inequality in equation (19), which holds with equality if equation (22) holds with equality. This is the case if (z_i, d_{ij}) contains as much information about ex-post utilities under drugs j and j' as the true information set; that is, if $\mathbb{E}[\Delta u_{ijj'}|\mathcal{W}_i] = \mathbb{E}[\Delta u_{ijj'}|z_i, d_{ij}]$. When $\mathcal{W}_i$ contains more information than (z_i, d_{ij}), Jensen's inequality implies that the left-hand side of equation (19) is strictly smaller than zero. Thus, as we illustrate in a simulation in Section 6, components of the physician's information set that are not incorporated in z_i make the inequality in equation (19) weaker and the set of parameters consistent with the bounding inequalities, Θ_0^b, larger.

5 Empirical Implementation

In this section, we describe how we use the odds-based and bounding moment inequalities in our empirical application. In Section 5.1, we describe the additional model restrictions we impose. In Section 5.2, we describe how we compute our moment inequality estimator.

5.1 Adapting the Drug Choice Model for Estimation

In our empirical application, we restrict the function $u(x_{ij}, \mu)$ in equation (1b) to be:

$$u(x_{ij}, \mu) = \kappa_j + \alpha p_{ij}, \tag{23}$$

where p_{ij} is the patient's realized out-of-pocket cost for drug j at visit i, α captures the physician's sensitivity to patient costs, and κ_j denotes the quality of treatment j. Quality here accounts for the drug's efficacy and side-effect profile. Given equation (23), we can write the expected utility term entering equation (6) as

$$\mathbb{E}[u(x_{ij}, \mu)|\mathcal{W}_i] = \kappa_j + \alpha\mathbb{E}[p_{ij}|\mathcal{W}_i], \qquad \text{for } j = 1, \ldots, J. \tag{24}$$

Thus, in our application, the covariate vector x_{ij} includes the out-of-pocket costs p_{ij} and the information set $\mathcal{W}_i$ includes variables that the physician uses to predict patient costs at visit i. Here, the subscript i on p_{ij} implies that drug prices may vary across visits, consistent with the statistics in Table 1. Similarly, the subscript i on $\mathcal{W}_i$ implies that physicians' information about drug prices may vary across visits. Specifically, we allow patients to

face different prices depending on their plan, and physicians to have different information about those prices based on the patient's plan.[7] Given the restriction in equation (23) and a normalization $\kappa_1 = 0$, the parameter vector to estimate is $\mu = (\alpha, \kappa_2, \ldots, \kappa_J)$. We denote as $\theta \equiv (\theta_\alpha, \theta_{\kappa_2}, \ldots, \theta_{\kappa_J})$ the unknown parameter whose true value is μ.

We note here an additional property of the inequalities described in Section 4. As shown in Appendix B.5, if the restriction in equation (23) holds, the odds-based moment inequality in equation (7) and the bounding moment inequality in equation (13) are valid even if the observed price difference between drugs is measured with error, as long as the error is mean zero conditional on the physician's information set. That is, denoting the true and the measured price differences between drugs j and j' for visit i as $\Delta p_{ijj'}$ and $\Delta\hat{p}_{ijj'}$, respectively, we derive moment inequalities that use information on $\Delta\hat{p}_{ijj'}$ for every visit i and drugs j and j' and that are valid if $\mathbb{E}[\Delta\hat{p}_{ijj'} - \Delta p_{ijj'}|\mathcal{J}_i] = 0$. Thus, these inequalities are valid even if the researcher does not observe the physician's price expectations or the actual prices.

5.2 Implementing the Moment Inequality Estimator

We combine odds-based and bounding moment inequalities for estimation. The conditional moment inequalities in equations (7) and (13) are defined for every ordered pair of drugs (j, j') and every value of z_i in its support. Because z_i is a continuous scalar variable in our setting, we transform these conditional moment inequalities into a finite set of unconditional moment inequalities to compute confidence sets for μ. Specifically, for each ordered pair of drugs (j, j'), and each instrument function $g^{(k)}_{jj'}\colon \mathcal{Z} \to [0, \infty)$ for $k = 1, \ldots, K$, we estimate the model parameters using the sample analogue of the odds-based moment inequality

$$\mathbb{E}[(d_{ij} \exp(-(\theta_{\kappa_j} - \theta_{\kappa_{j'}} + \theta_\alpha \Delta\hat{p}_{ijj'})) - d_{ij'}) g^{(k)}_{jj'}(z_i)] \geqslant 0, \tag{25}$$

and of the bounding moment inequality

$$\mathbb{E}[(d_{ij'} - d_{ij} \exp(-e_{jj'}(z_i, \theta))(1 + e_{jj'}(z_i, \theta) - (\theta_{\kappa_j} - \theta_{\kappa_{j'}} + \theta_\alpha \Delta\hat{p}_{ijj'}))) g^{(k)}_{jj'}(z_i)] \geqslant 0, \tag{26}$$

with $e_{jj'}(\cdot)$ equal to the expression in equation (18). Thus, given a choice set of size J and K instrument functions, we use $2J(J-1)K$ inequalities to compute a confidence set for J parameters: the drug fixed effects $(\kappa_2, \ldots, \kappa_J)$ and the price coefficient α.

To choose the instrument functions, we follow Andrews and Shi (2013), who indicate that, by using functions that partition the support of z_i, one can transform conditional

[7] Although physicians make choices in our set-up, patients may influence those choices. In particular, the drug qualities, κ, the price sensitivity, α, and the information set, $\mathcal{W}_i$, may reflect patient input.

moment inequalities into unconditional moment inequalities without loss of identification power. However, when implementing this approach, a tradeoff arises. On the one hand, a larger number of instrument functions—and thus a finer partition of the support—leads to a smaller identified set. On the other hand, increasing the number of instruments raises the computational burden and can reduce the precision of finite-sample inference, since each sample moment is computed using fewer observations. We describe the instrument functions we use in Appendix B.6, and present simulations that illustrate the finite sample properties of our inference procedure in Appendix C.4.1. Unless otherwise noted, we compute all confidence sets using the inference procedure for unconditional moment inequalities in Cox and Shi (2023); we describe how we implement this procedure in Appendix B.7.1.

To choose the instrument vector z_i, we focus on variables that are correlated with the observed prices. In particular, we use instruments defined by price averages; e.g., the average of current or lagged drug prices across all patients sharing the same insurance carrier or the same plan type. Given a choice of z_i and instrument functions, we build the sample analogue of the inequalities in equations (25) and (26) and use them both to infer the content of agents' information sets and to compute confidence sets for all model parameters. We infer that z_i is not contained in agents' information sets whenever we find an empty confidence set.[8]

6 Simulation

Before estimating our model on actual physician choices, we perform a simulation exercise. We design the simulation with the goal of comparing the properties of the maximum likelihood estimator (MLE) with our moment inequality estimator. We examine settings in which the researcher only partially observes the agent's information set or in which agents form expectations with error. Through this simulation, we first show that the MLE is inconsistent unless the researcher's assumed information set coincides *exactly* with the agent's information set. Conversely, consistent with theorems 1 and 2, the odds-based and bounding inequalities are satisfied at the true parameter value as long as the researcher correctly identifies a *subset* of the agent's information set. Second, we show that both the odds-based and bounding inequalities are useful for identifying parameters; i.e. neither type is redundant. Third, we discuss the size of the confidence set we find using the inequalities, and in

[8] An empty confidence set may also arise when a model assumption is violated, for example, if the rational expectations assumption does not hold. However, deviations from rationality do not yield empty confidence sets in the special case where, for every drug j, expectations of p_{ij} are biased by a constant γ_j. In this case, the expected utility in equation (24) is analogous, but with $\tilde{\kappa}_j = \kappa_j + \gamma_j$ replacing κ_j. As a result, the identified set is non-empty, but the estimated drug fixed effects now capture both drug qualities and the systematic bias in the drugs' expected prices.

particular when this set is likely to include many parameter values in addition to the true value. Finally, we show our inequalities can be used to test hypotheses about the variables agents know and use when forming expectations.

6.1 Simulation Set-up

We simulate data for $i = 1, \ldots, N$ observations, with $N = 4,000,000$, using the model described in sections 3 and 5.1 . By using one large sample, we illustrate the identification properties of our inequalities. In Appendix C.4.1, we present results using many samples of smaller size to explore the properties of our confidence sets.

Agents choose between three choices, $j = 1, 2, 3$, and we set the choice-specific fixed effects to $\kappa_1 = \kappa_2 = 0$ and $\kappa_3 = 1$, and the price coefficient to $\alpha = 1$. Unlike in the model in Section 3, we must specify the distribution of the price vector $p_i = \{p_{ij}\}_{j=1}^J$, and the content of the information set $\mathcal{W}_i$, as these elements determine the choice of the simulated agents.

For the price distribution, we impose that $p_{ij} = x_{1ij} + x_{2ij} + x_{3ij}$ for all i and j, with x_{kij} independent of $x_{k'i'j'}$ for $k \neq k'$, $i \neq i'$, or $j \neq j'$ and where x_{kij} is distributed uniformly with support of length $2\sigma_k$ for a given value of σ_k. In our simulations, we fix $\sigma_2 = 4$ and present results for different values of σ_1 and σ_3. For the agent's information set, we impose that $\mathcal{W}_i = (x_{1i}, x_{2i})$ for all i and, as a result of the assumed price distribution, we have $\mathbb{E}[p_{ij}|\mathcal{W}_i] = x_{1ij} + x_{2ij}$ for all i and j. The variable x_{3ij} thus represents the agent's expectational error.[9] Finally, we assume the researcher only observes the vector (d_i, p_i, x_{2i}) for each observation i. Thus, x_{1i} is a vector on which the agent conditions her decision but which the researcher does not observe.

We compute confidence sets using the inequalities in equations (25) and (26) for all possible drug pairs. Given a pair (j, j'), unless otherwise noted, we use as instrument functions $g_{jj'}^{(1)}(z_i) = \mathbb{1}\{\Delta x_{2ijj'} \geqslant 0\}$ and $g_{jj'}^{(2)}(z_i) = \mathbb{1}\{\Delta x_{2ijj'} < 0\}$ with $\Delta x_{2ijj'} = x_{2ij} - x_{2ij'}$. We also report maximum likelihood estimates and, unless we indicate otherwise, we compute these under the assumption that $\mathbb{E}[p_{ij}|\mathcal{W}_i] = x_{2ij}$. Thus, while the moment inequality estimator correctly assumes that x_{2i} belongs to agent i's information set, the MLE relies on x_{2i} being the only variable in that set, which is true only if $x_{1ij} = 0$; that is, if $\sigma_1 = 0$.

When computing both the moment inequality and the maximum likelihood estimates, we normalize the parameter θ_{κ_1} by fixing it to its true value; i.e., we set $\theta_{\kappa_1} = 0$. Appendix C.1 provides more details.

[9] The support of x_{kij} is $[\mu_{kj} - \sigma_k, \mu_{kj} + \sigma_k]$ for $k = 1, 2, 3$. We fix $\mu_{22} = -0.5$ and $\mu_{23} = -1$, and set $\mu_{kj} = 0$ for all other k and j. As a result, $\mathbb{E}[p_{ij}|\mathcal{W}_i] = x_{1ij} + x_{2ij} + \mathbb{E}[x_{3ij}|x_{1ij}, x_{2ij}]$, with $\mathbb{E}[x_{3ij}|x_{1i}, x_{2i}] = 0$ since x_{ki} is independent of $x_{k'i}$ for $k \neq k'$ and μ_{3j}, the mean of x_{3j}, equals zero.

6.2 Simulation Results

We report the main simulation results in Table 2. In case 1, we explore the scenario in which the researcher observes all variables on which the agent bases her decision (i.e., $\sigma_1 = 0$ and, thus, $x_{1i} = 0$ for all i) and agents make no expectational error (i.e., $\sigma_3 = 0$ and, thus, $x_{3i} = 0$ for all i). In this case, the maximum likelihood estimate coincides with the true parameter value, and the confidence sets defined either by the odds-based or the bounding moment inequalities include only the true parameter value.[10]

In case 2, we consider a scenario in which, as in case 1, the researcher observes the agent's information set (i.e., $\sigma_1 = 0$), but agents now make expectational errors (i.e, $\sigma_3 > 0$). The results show that neither the MLE nor the confidence set defined by the bounding moment inequalities is affected by the presence of expectational errors. Conversely, the confidence set defined by the odds-based moment inequalities is no longer a singleton, including the true value but also other values of the parameter vector.

In case 3, we consider a scenario in which agents make no expectational errors (i.e, $\sigma_3 = 0$) but the researcher only observes part of the agent's information set (i.e., $\sigma_1 > 0$). In this case, the MLE is biased towards zero. In contrast, the confidence sets defined by the odds-based and bounding inequalities contain the true value, but they both also include values other than the true value. The confidence sets defined by the odds-based and bounding moment inequalities are, however, distinct. The confidence set defined by the bounding inequality includes several points around the true value, while the set defined by the odds-based inequality only adds points that are so distinct from the true value that they fall outside the grid of points used to compute the results reported in Table 2.[11]

The differing behavior of the odds-based and bounding inequalities arises because of the different ways in which they handle the researcher's missing information on agents' true expectations. As discussed in Section 4.1, we build the odds-based inequality by replacing the unobserved expectation, $\mathbb{E}[\Delta u_{ijj'}|\mathcal{W}_i]$, entering the moment in equation (9) with the ex-post utility difference, $\Delta u_{ijj'}$. As the resulting moment is convex in the agent's expectational error, the bounds we recover when using odds-based inequalities alone increase in the dispersion in the expectational error. Hence, when $\sigma_3 > 0$, the confidence set defined by the odds-based inequalities alone will include values other than the true value.

As discussed in Section 4.3, we obtain the bounding inequality by replacing the unobserved expectation entering the moment in equation (9) with the proxy $\mathbb{E}[\Delta u_{ijj'}|z_i, d_{ij}]$.

[10] Table 2 contains projections of the 95% confidence sets for $(\kappa_2, \kappa_3, \alpha)$ computed following Cox and Shi (2023). In Appendix C.3, we present confidence sets computed following Andrews and Soares (2010).

[11] The odds-based inequalities sometimes yield a confidence set that includes values outside of the grid used to compute the results in Table 2. We indicate this outcome with an asterisk next to the label "Odds-based."

Table 2: Simulation Results - MLE and Confidence Intervals

Case	σ_1	σ_3	z_i	Estimator	MLE & Confidence Sets α	κ_2	κ_3
1	0	0	x_{2i}	MLE	1	0	1
				Odds-based	[1 , 1]	[0 , 0]	[1 , 1]
				Bounding	[1 , 1]	[0 , 0]	[1 , 1]
				Both	[1 , 1]	[0 , 0]	[1 , 1]
2	0	1	x_{2i}	MLE	1	0	1
				Odds-based*	[0.92 , 1.50]	[-0.33 , 0.33]	[0.67 , 1.33]
				Bounding	[1 , 1]	[0 , 0]	[1 , 1]
				Both	[1 , 1]	[0 , 0]	[1 , 1]
3	1	0	x_{2i}	MLE	0.91	0	0.91
				Odds-based*	[1 , 1]	[0 , 0]	[1 , 1]
				Bounding	[0.80 , 1.10]	[-0.30 , 0.30]	[0.70 , 1.30]
				Both	[1 , 1]	[0 , 0]	[1 , 1]
4	1	1	x_{2i}	MLE	0.92	0	0.91
				Odds-based*	[0.92 , 1.50]	[-0.48 , 0.50]	[0.65 , 1.50]
				Bounding	[0.80 , 1.10]	[-0.30 , 0.30]	[0.70 , 1.30]
				Both	[0.92 , 1.10]	[-0.33 , 0.30]	[0.70 , 1.30]
5	0	1	p_i	MLE	0.87	-0.03	0.87
				Odds-based	$\varnothing$	$\varnothing$	$\varnothing$
				Bounding	[0.87 , 0.87]	[-0.05 , -0.03]	[0.85 , 0.88]
				Both	$\varnothing$	$\varnothing$	$\varnothing$

Note: *Odds-based*, *Bounding*, and *Both* contain projections of 95% confidence sets computed as in Cox and Shi (2023). *Odds-based* indicates the confidence set is computed using the inequalities in equation (25); *Bounding* indicates the confidence set is computed using the inequalities in equation (26); *Both* indicates it is computed using both types of inequalities. Confidence sets are computed using a 3-dimensional grid whose sides are $[0.5, 1.5]$ (for α), $[-0.5, 0.5]$ (for κ_2) and $[0.5, 1.5]$ (for κ_3). We mark a row with an asterisk when the confidence set includes points outside the grid.

Because the agent's expectational error does not affect the agent's true expectation nor the proxy above, the bounding inequality is invariant to the distribution of expectational errors. Hence, in our simulation, the confidence set defined by the bounding inequality does not vary with σ_3. In contrast, the agent's true expectation and the researcher's proxy will differ when the researcher only partially observes the agent's information set. As the resulting bounding inequality is convex in this difference, the bounds we recover when using the bounding inequality alone increase in the dispersion in x_{i1} across observations. Hence, when $\sigma_1 > 0$, the confidence set defined by the bounding inequality includes values other than the true value.

To illustrate what we view as the most empirically relevant setting, in case 4 we generate a sample in which researchers do not observe all the variables agents use to form expectations (i.e., $\sigma_1 > 0$) *and* agents have imperfect information about the payoff variables they must forecast (i.e., $\sigma_3 > 0$). In this scenario, the MLE is biased downward, inheriting the bias present in case 3. The confidence sets defined by odds-based inequalities, bounding in-

equalities, or by combining both sets of inequalities, include the true value of the parameter. However, they all also include additional points. Case 4 illustrates the value of combining odds-based and bounding moments. On some dimensions of the parameter space, the odds-based moments generate a tighter boundary, and on other dimensions, the bounding moments do. Using the combination of the two types of inequalities leads to a strictly smaller confidence set than using either type of inequality alone.

Finally, in case 5, we consider a setting in which the researcher assumes the agent has strictly more information than the agent actually possesses. Specifically, the researcher assumes the agent has perfect information on all payoff-relevant variables and thus builds the moment inequalities for each drug pair (j, j') using the instrument functions $g^{(1)}_{jj'}(z_i) = \mathbb{1}\{\Delta p_{ijj'} \geqslant 0\}$ and $g^{(2)}_{jj'}(z_i) = \mathbb{1}\{\Delta p_{ijj'} < 0\}$. We also report maximum likelihood estimates under the assumption that $\mathbb{E}[p_{ij}|x_{2i}] = p_{ij}$. Table 2 shows that the confidence sets defined by the odds-based inequalities alone, or by both types of inequalities jointly, are empty.[12] Given theorems 1 and 2, we can therefore reject that agents have perfect information on prices. The MLE of α is asymptotically biased downwards and, given that we set the mean price for product $j = 1$ to be the lowest in our simulation, the downward bias in the price coefficient translates into a downward bias in the estimate of κ_j for $j = \{2, 3\}$.

7 Estimation Results

We now return to the diabetes setting and present estimates of the parameters of the prescription drug choice model described in sections 3 and 5.1. In Section 7.1, we present maximum likelihood estimates. To compute these estimates, we fully specify the content of each physician's information set. Here, we assume all physicians' information sets consist of a common scalar predictor of drug prices; we present results under different specifications of this common predictor. In Section 7.2, we use the moment inequality estimator described in Section 5.2 to compute bounds on the model parameters. To compute these bounds, we again specify a scalar price predictor that we assume belongs to the information sets of all physicians in the sample. However, we need not restrict the set of other variables that physicians use to predict price. In our empirical analysis, we present results for different specifications of the shared predictor. Finally, in Section 7.3, we look for heterogeneity in information sets, using moment inequalities to test whether certain price predictors are known only to a subset of physicians.

[12] As we show in Appendix C.3, the confidence set defined by the bounding inequalities is empty when we compute it following Andrews and Soares (2010). In unreported results, we find the set is also empty when we compute it following Cox and Shi (2023), but using four instrument functions instead of two.

7.1 Maximum Likelihood Estimates

Computing the MLE of physicians' preference parameters requires two steps. In the first step, we compute measures of the physician's price expectations for each medical visit and drug in the choice set. We do so by regressing the realized out-of-pocket costs on an information set we specify. In the second step, we compute maximum likelihood estimates of the model parameters conditioning on the expected out-of-pocket costs computed in the first step.[13] For our main analysis, we use data on treatment choices within the class of DPP-4 inhibitors.[14] As this class includes only three drugs during our sample period, we can write the MLE as

$$\underset{(\theta_\alpha,\theta_{\kappa_2},\theta_{\kappa_3})}{\operatorname{argmax}} \left\{ \sum_{i=1}^{N} \sum_{j=1}^{3} \mathbb{1}\{d_{ij}=1\} \ln \left(\frac{\exp(\theta_{\kappa_j} + \theta_\alpha \widehat{\mathbb{E}}[p_{ij}|z_i])}{\sum_{j'=1}^{3} \exp(\theta_{\kappa_{j'}} + \theta_\alpha \widehat{\mathbb{E}}[p_{ij'}|z_i])} \right) \right\}, \qquad \text{with } \theta_{\kappa_1} = 0. \quad (27)$$

Here, z_i is a scalar we assume coincides with physician i's information set, $\mathcal{W}_i$, and $\widehat{\mathbb{E}}[p_{ij}|z_i]$ is the predicted out-of-pocket costs for visit i and drug j obtained by projecting p_{ij} on z_i.

In Table 3, we report maximum likelihood estimates for different specifications of the price predictor z_i. Under the assumption of perfect information, we find an estimate of the price coefficient, α, equal to -0.47. If we instead assume providers form expectations using contemporaneous average prices at the drug-carrier level or drug-plan type level, we find estimates equal to -1.18 and -1.39, respectively. If we assume providers form expectations on out-of-pocket costs using only contemporaneous drug level price averages, the estimate of α equals -2.13. For each of these price predictors, if we assume physicians' information sets equal the same predictor but lagged by a year, the point estimate becomes smaller, reaching a minimum value of -2.96. In contrast, the estimates of the drug-specific fixed effects, κ_2 and κ_3, are more robust to the specification of the physician's information set.[15]

As the last column in Table 3 shows, different specifications of the physician's information set also imply different average elasticities of treatment choices with respect to expected out-of-pocket costs. Consistent with the heterogeneity in the estimates of α, we find larger

[13] We compute standard errors by bootstrapping. In each of 200 bootstrap samples, we compute the first-stage regression and the second-stage maximum likelihood estimates. We then compute standard errors for the maximum likelihood estimates as the standard deviation across the 200 sample estimates.

[14] In Appendix D.3, we study physicians' treatment choices among drugs in two classes: DPP-4 and SGLT2 inhibitors. In particular, we estimate a nested logit model, with each drug class forming a nest.

[15] Here, the distinct estimates of α do not translate into distinct estimates of κ_2 and κ_3 because the three drugs in the choice set have roughly similar average out-of-pocket costs, as we report in Table 1. In particular, regardless of the specified price predictor, we estimate the quality of Januvia (identified with the index $j = 2$) to be larger than that of Janumet ($j = 1$) and Tradjenta ($j = 3$). Although clinical trials and head-to-head studies generally find that the efficacy and side effects of DPP-4 drugs are more similar to each other relative to drugs in other classes (American Diabetes Association, 2017), they are not identical. For example, Tradjenta appears to be preferable for patients with kidney problems, but some reports find evidence of joint pain with Tradjenta (Kamatani et al., 2013; Nigro and Goldman, 2022).

Table 3: Maximum Likelihood Estimates

Information Set	κ_2	κ_3	α	Price Elast. (Janumet)
Current Prices	1.40	-0.26	-0.47	-0.58
	(0.03)	(0.04)	(0.03)	(0.04)
Average Current Prices	1.43	-0.30	-0.97	-1.19
By Drug-Plan Type-Carrier	(0.03)	(0.04)	(0.04)	(0.06)
Average Current Prices	1.43	-0.32	-1.18	-1.44
By Drug-Carrier	(0.03)	(0.04)	(0.05)	(0.06)
Average Current Prices	1.44	-0.18	-1.39	-1.69
By Drug-Plan Type	(0.03)	(0.05)	(0.12)	(0.15)
Average Current Prices	1.49	-0.13	-2.13	-2.58
By Drug	(0.04)	(0.05)	(0.17)	(0.21)
Lagged Prices	1.40	-0.24	-0.58	-0.71
	(0.03)	(0.04)	(0.05)	(0.06)
Average Lagged Prices	1.44	-0.23	-1.01	-1.23
By Drug-Plan Type-Carrier	(0.03)	(0.04)	(0.05)	(0.07)
Average Lagged Prices	1.46	-0.22	-1.25	-1.51
By Drug-Carrier	(0.03)	(0.04)	(0.06)	(0.07)
Average Lagged Prices	1.48	-0.02	-1.77	-2.11
By Drug-Plan Type	(0.03)	(0.05)	(0.13)	(0.16)
Average Lagged Prices	1.51	0.05	-2.96	-3.57
By Drug	(0.03)	(0.05)	(0.25)	(0.30)

Note: Columns labeled α, κ_2 and κ_3 present maximum likelihood estimates of the corresponding parameter computed following equation (27). The column labeled *Information Set* indicates the vector of observed covariates z_i used to build the log-likelihood function in equation (27). To illustrate the elasticity implied by the price coefficients, we report in the column labeled *Price Elast. (Janumet)* the in-sample average elasticity for Janumet, which corresponds to drug $j = 1$ in our choice set. We report standard errors using a bootstrap procedure that accounts for the two-step nature of our estimation.

elasticity estimates when the researcher assumes the physician forms expectations using coarser information. For example, for Janumet, the average elasticity to expected prices ranges from -0.58, when we assume physicians have perfect information on prices, to -3.57, when we assume physicians form expectations only using last year's average price by drug.

That the estimates of the price coefficient decrease in absolute value when we assume physicians form expectations using more detailed price predictors is consistent with our simulation results. Comparing cases 3 and 5 in Table 2, we find a larger attenuation bias in the estimated price coefficient when the researcher specifies an overly rich information set—including variables the agent did not actually use—than when the researcher omits variables the agent did use. Furthermore, comparing cases 5 and 5(b) in Table 2 and Appendix Table C.1, respectively, we find that the attenuation bias from specifying too large an information set is greater the more relevant those extra variables are for predicting prices.

Beyond distinct elasticity estimates, alternative assumptions on physician information sets also imply substantially different predictions under counterfactual market environments. To illustrate this result, we consider an intervention in which three major carriers, represent-

Table 4: Effect of a Reduction in Out-of-Pocket Costs on Janumet's Market Share

Information Set	(1) Counterfactual Share	(2) Counterfactual Share (Perfect Info. Est.)	(3) Counterfactual Share (Perfect Info. Prices)
Current Prices	21.66	21.66	21.66
Average Current Prices By Drug-Plan Type-Carrier	26.86	21.60	26.80
Average Current Prices By Drug-Carrier	29.37	21.58	29.16
Average Current Prices By Drug-Plan Type	30.64	21.33	30.63
Average Current Prices By Drug	38.25	21.28	36.87
Lagged Prices	21.74	20.76	22.86
Average Lagged Prices By Drug-Plan Type-Carrier	26.27	21.26	26.87
Average Lagged Prices By Drug-Carrier	28.36	21.17	29.13
Average Lagged Prices By Drug-Plan Type	31.83	20.83	33.27
Average Lagged Prices By Drug	44.00	20.83	42.30

Note: We compute the counterfactual shares reported in column 1 using maximum likelihood estimates and predicted prices consistent with the information set indicated in the corresponding row in the column labeled *Information Set*. We compute the counterfactual shares reported in column 2 using predicted prices consistent with the information set indicated in the corresponding row, but we combine these prices with the maximum likelihood estimates computed under the assumption of perfect information; i.e., those estimates reported in the first row of Table 3. The counterfactual shares reported in column 3 use the maximum likelihood estimates computed under the information set indicated in the corresponding row, but we use realized prices as predicted prices.

ing 55% of visits, negotiate a 50% reduction in Janumet's out-of-pocket costs for all patients enrolled in a plan offered by these carriers. Column 1 of Table 4 shows the model-predicted counterfactual share that Janumet captures after this price reduction. Depending on the assumed information set, this counterfactual share varies between 22% and 44%. Importantly, regardless of the assumed information set, all estimated models match the initial market share of 17.8%. Thus, the predicted increase in Janumet's prescription share ranges from slightly over four percentage points (under perfect information) to more than 26 percentage points (with the last year's drug-specific average prices as the information set).[16]

That the elasticities and counterfactual predictions from the maximum likelihood analyses differ across models highlights the importance of correctly specifying agents' information sets. This finding aligns with previous work on prescription drug choice. Carrera et al.

[16]Two factors explain the differences in counterfactual predictions. First, different informational assumptions yield different parameter estimates, as shown in Table 3. Second, different informational assumptions yield distinct changes in Janumet's expected price. Columns 2 and 3 in Table 4 illustrate the relative importance of the two factors: most of the variation in column 1 arises from differences in parameter estimates.

(2018), for example, show that the estimated price elasticity of an anti-cholesterol drug increases (in absolute value) in the period after its patent expires. An explanation for this finding is that the patent expiration generated an easily observed price change, improving physicians' information and thereby reducing the attenuation bias in the elasticity estimate that arises when physicians' expectations differ from actual prices. In Appendix Table D.2, we similarly report greater sensitivity to price in our setting after large price declines; physicians likely notice these price changes and incorporate them into their price expectations.

A natural way to identify the correct information set is to use model selection tests. Here, we implement the procedure in Vuong (1989) to compare the models whose estimates we report in Table 3. That is, for any two models, we compute a statistic that depends on the difference in the log-likelihood function evaluated at the corresponding maximum likelihood estimate. We then compare this statistic with the appropriate quantile from the standard normal distribution. In Appendix D.1, we provide more details and report the results. We find that the preferred model assumes that physicians form price expectations using contemporaneous average prices at the drug-carrier level. However, this testing approach can only compare the options we specify; if we fail to include the model that uses the true information set among our test options, the conclusion of such a test may be misleading.

7.2 Moment Inequality Estimates

We now use the moment inequality estimator described in Section 5.2 to compute confidence sets for the preference parameters in our drug choice model and, simultaneously, to test whether certain price predictors belong to the information set of every physician in the sample. We do so for each of the ten price predictors we previously used to compute maximum likelihood estimates, as listed in the first column in Table 3. In practice, we use all observations in the sample to estimate ten versions of our moment inequality model, where each version differs only in the price predictor we use as the instrument z_i. When estimating each of these models, we impose that all physicians share the same values of the preference parameters; i.e., of the price coefficient α and the drug fixed effects $(\kappa_1, \kappa_2, \kappa_3)$.[17]

We find that the 95% confidence set implied by the moment inequalities is empty for seven of the ten price predictors we consider. For each of the seven predictors for which the confidence set is empty, we reject the null hypothesis that all physicians rely on the

[17] In Appendix D.2, we explore the assumed independence between the vector of idiosyncratic preferences, ε_i, and the physician's expected prices. This assumption will be violated if, for example, patients anticipate their need for drug j and choose an insurance plan they know offers generous coverage for that drug. We re-estimate our model on a subsample that includes only observations for patients who are less likely to have switched plans due to drug coverage generosity. We find the projected confidence sets are similar regardless of whether we compute them using the full estimation sample or the restricted sample.

Table 5: Estimation Results - Moment Inequalities

Price Predictor	κ_2	κ_3	α	Elasticity
Average Current Prices By Drug	[1.35, 1.75]	[-0.70, 0.25]	[-4.35, -1.40]	[-5.27, -1.69]
Average Lagged Prices By Drug-Plan Type	[1.45, 1.65]	[-0.50, -0.20]	[-3.85, -1.70]	[-4.56, -2.03]
Average Lagged Prices By Drug	[1.35, 1.75]	[-0.75, 0.20]	[-4.15, -1.40]	[-5.05, -1.67]

Note: Columns labeled α, κ_2 and κ_3 present projected 95% confidence sets computed using the moment inequalities described in Section 5.2 and the inference procedure in Cox and Shi (2023). The column labeled *Price Predictor* indicates the vector of observed covariates z_i that we use as instruments in our moment inequalities. The column labeled *Elasticity* reports the 95% confidence interval for the average elasticity of Janumet's prescription share with respect to its expected price. The number of observations used in the estimation is 8,540. Confidence sets are computed using a grid for $(\kappa_2, \kappa_3, \alpha)$ with 31, 41, and 140 points, respectively, for a total of 177,940 points.

corresponding variable to forecast out-of-pocket costs.[18,19] Specifically, we reject the null hypotheses that, in all visits, physicians have information on either contemporaneous prices or the previous year's prices. We also reject that physicians know the most detailed price averages. For example, we reject that physicians know the contemporaneous or lagged average copayment by drug, plan type, and carrier when forecasting patient prices. Finally, we also reject that physicians know contemporaneous or lagged average copayments by drug, carrier, and year. This finding contrasts with the results of the Vuong (1989) tests described in Section 7.1, which identified contemporaneous average copayments at the drug-carrier level as the correct price predictor. In sum, our moment inequality estimates indicate that physicians use relatively coarse information on out-of-pocket costs when forming expectations about a given patient's price.

In Table 5, we report the projected confidence sets for the three predictors for which these confidence sets are non-empty. While we find that these sets contain similar parameter values, the sets are smallest when we use average lagged prices by drug and plan type as our instrument. This result is consistent with this aggregate price being a stronger predictor of current patient prices than either average lagged or average current prices by drug.[20]

Comparing the maximum likelihood estimates in Table 3 to the moment inequality confidence sets, we focus on the confidence set in the second row of Table 5, which uses the

[18]Since we test multiple hypotheses, we compute family-wise p-values following Holm (1979); see Appendix B.7.2. In our setting, the family-wise p-values for the tests with empty 95% confidence sets are below 5%.

[19]For all price predictors, we use the same set of instrument functions, as detailed in Appendix B.6.

[20]In Appendix Table D.3, we report the R^2 of linear regressions of current prices on the different price predictors we consider; the R^2 is indeed larger for the average lagged price by drug and plan type than for the current or lagged average price by drug. However, all three R^2 values are small. Thus, it may be that none of the ten predictors we study belongs to the physician's information set in *every* sampled visit, and that the non-empty confidence sets in Table 5 arise because the corresponding three predictors are only weakly correlated with observed prices. Indeed, the simulation results in Table 2 document that confidence sets grow larger the weaker the correlation between the predicted covariate, p_i, and the instrument, z_i.

strongest price predictor as an instrument. We observe that none of the maximum likelihood estimates falls within the projected confidence set along all three dimensions. In most cases, the maximum likelihood estimate of α is greater than -1.70, the upper bound of the projected confidence set. Thus, our moment inequality estimates generally suggest that physicians are more sensitive to expected prices than one would conclude based on our maximum likelihood estimates. When the maximum likelihood estimate of α falls inside the projected confidence set, the estimate of κ_3 is above -0.2, the upper bound of the confidence set projected on this parameter.

When comparing the confidence set for the elasticity of Janumet's market share to its expected price, reported in the last column of Table 5, to the corresponding maximum likelihood estimates, reported in the last column of Table 3, we find that the moment inequality confidence set includes elasticity values that are larger in absolute value than the corresponding maximum likelihood estimates. More specifically, the 95% confidence set for this elasticity ranges from -4.56 to -2.03, while the maximum likelihood estimates generally lie between 0.6 and -2.6. The one exception is when we assume physicians only know average lagged prices by drug; under this price predictor, the maximum likelihood estimate is -3.57.

7.3 Testing for Heterogeneity in Information sets

Here, we relax the assumption imposed in Section 7.2 and examine whether different subsets of physicians have distinct information sets. We do so by re-computing the moment inequality estimator described in Section 5.2 for different groups of physicians, testing whether specific price predictors belong to the information set of each group. In practice, we partition physicians into two subsamples according to five different criteria: medical specialty; graduation year; sex; ranking of the medical school attended; and the concentration of insurance plan types, measured using the Herfindahl index (HHI), in the patient's home region. For each subsample of physicians, we test whether all included physicians know each of the price predictors listed in the first column of Table 3. As we estimate our moment inequality model separately for different groups of physicians, we also allow the preference parameters to be heterogeneous across these groups. Hence, our findings on the heterogeneity of information sets between physician groups do not rely on the assumption that preference parameters are common among all physicians.

We report the results of the hypothesis tests in Table 6. Throughout, we follow Holm (1979) to adjust the p-values for the multiple hypotheses we test for each subset of physicians. We reject that a subset of physicians incorporates a particular price predictor into their price expectations when the adjusted p-value is below 5%.

Table 6: Testing for Heterogeneity in Information: Rejection at 5%

Set of physicians	Average current prices by			Average lagged prices by			N. Obs.
	drug plan type carrier	drug carrier	drug plan type	drug plan type carrier	drug carrier	drug plan type	
Endocrinologist	No	No	No	No	No	No	1,336
Primary Care Physician	Yes	Yes	Yes	Yes	Yes	No	7,204
Graduated Before 1996	Yes	Yes	Yes	Yes	Yes	No	3,503
Graduated After 1996	Yes	Yes	No	Yes	Yes	No	5,037
Female	Yes	Yes	Yes	No	No	No	4,889
Male	Yes	Yes	Yes	Yes	Yes	No	3,651
Lower-tier Medical School	Yes	Yes	Yes	Yes	Yes	No	7,307
Top-tier Medical School	Yes	No	No	No	No	No	1,233
High Plan Type HHI	Yes	Yes	Yes	Yes	Yes	No	4,296
Low Plan Type HHI	Yes	Yes	Yes	Yes	Yes	Yes	4,244

Note: We test the null hypotheses that the price predictor indicated in columns 2 to 7 belongs to the information set of every physician in the set indicated in column 1. The column labeled *N. Obs.* includes the number of claims in each set for the years 2012 to 2016. For each set of physicians, we adjust p-values using the procedure described in Appendix B.7.2. *No* indicates that we cannot reject at the 5% significance level the null hypothesis that the corresponding price predictor belongs to the information set of all physicians in the relevant set; *Yes* indicates that we reject this hypothesis.

Our results point to physicians' specialty training as the main source of heterogeneity in information.[21] In particular, when we compare endocrinologists to primary care physicians, we find clear differences in the set of predictors that each specialty incorporates into their price expectations. For endocrinologists, we cannot reject that they have access to information on current and lagged average prices at the drug-plan type-carrier level. Given this finding, it is not surprising that we also fail to reject that they have access to information on the current and lagged values of more aggregated average prices, including prices at the drug-plan type level and at the drug-carrier level. In contrast, our results suggest that primary care physicians have access to less detailed price information. In particular, we reject that they have access to five of the six predictors considered in Table 6. The only price predictor that we fail to reject in our tests for primary care physicians is the average lagged prices by drug and plan type.

We also find significant heterogeneity in information by physicians' sex and the rank of the medical schools they attended: female physicians possess better price information than male physicians, and physicians who attended a top-tier medical school appear to have better price information than those who attended lower-tier schools. In contrast, we find that the heterogeneity in information driven by the physician's graduation year and the concentration

[21] In unreported results, we show that, for each group of physicians we consider in Table 6, we reject the null hypothesis that *all* physicians in the group use contemporaneous or lagged individual prices to form their price forecast.

across plan types in the physician's market is more limited.[22]

While our testing procedure does not identify the mechanisms behind the heterogeneity in information, prior work offers potential explanations, including the differential adoption of information technology (Arrow et al., 2020) and learning-by-doing (Crawford and Shum, 2005; Chintagunta et al., 2009). Using a natural experiment, Doyle et al. (2010) suggest that training can affect the quality and costs of physician decisions; better training may similarly affect the physician's familiarity with patient costs.

In addition to heterogeneity in information, we also find significant heterogeneity in preference parameters between physician groups. Appendix Table D.5 reports projected confidence sets for the parameters κ_2, κ_3 and α for each physician subsample. For each group, we use as an instrument the most detailed price predictor for which we obtain a non-empty 95% confidence set for that group.[23] As with our testing of physician information sets, the sharpest differences in preference parameters arise when comparing across physician specialties. We find that primary care physicians are more sensitive to price: while the projected 95% confidence set for α is $[-0.60, -0.40]$ for endocrinologists, it is equal to $[-5.90, -1.40]$ for primary care physicians. The preferences for the quality terms κ_2 and κ_3 show more similarity and overlap between the two specialties and, more generally, across all subsamples of physicians we examine.

8 Effects of an Information Intervention

The estimates described in sections 7.2 and 7.3 suggest that physicians face information frictions when forecasting patient prices, and that these frictions differ between physician types. Here, we use our model to evaluate a counterfactual scenario in which policymakers or insurers provide physicians with perfect information on out-of-pocket costs for each patient they treat. In this scenario, physicians learn patient-specific prices for each drug at the point of prescribing, possibly through pop-up messages in electronic medical records (Desai et al., 2022). We implement our information intervention separately for endocrinologists and primary care physicians. As discussed in Section 7.3, these two groups of physicians differ significantly both in their baseline information and in their preference parameters. For each physician group, we then measure the model-implied counterfactual change in

[22]For example, physicians working in markets with a high concentration of plan types (high-HHI group) appear to have only slightly richer price information; the high-HHI group incorporates lagged prices by drug and plan type into their price expectations, while the low-HHI group does not.

[23]Because we adjust for multiple hypothesis testing, there are cases in which we fail to reject the null hypothesis that a price predictor belongs to the information set of the physicians in a subsample while the corresponding 95% confidence set is empty.

three outcomes: (a) the share of visits at which the physician prescribes the cheapest drug available for a patient; (b) the per-patient average out-of-pocket costs; and (c) the per-visit consumer surplus, measured in dollars.[24]

For the first two outcomes, we measure the effect of the information intervention by comparing the model-predicted counterfactual value against the baseline values observed in the data. Since we assume physicians have perfect information in the counterfactual scenario and we observe the outcome values in the current scenario, we can compute the counterfactual change in outcomes (a) and (b) without the need to fully specify the physician's information set—that is, the precise set of variables each physician uses to forecast price. This approach preserves consistency with our moment inequality estimation, where we only need to specify a subset of physicians' information sets.[25]

For outcome (c), consumer surplus, we do not observe the baseline value in the data. Instead, as we show in Appendix D.4, we can write the model-predicted expected change in consumer surplus from our intervention as:

$$\Delta S(\mu) = \frac{1}{\alpha}\mathbb{E}[\log(\sum_{j=1}^{J}\exp(\kappa_j + \alpha p_{ij})) - \log(\sum_{j=1}^{J}\exp(\kappa_j + \alpha \mathbb{E}[p_{ij}|\mathcal{W}_i]))]. \tag{28}$$

We cannot compute this expression, as it depends on the distribution of the information set $\mathcal{W}_i$ in the population of interest. However, as shown in Appendix D.4, we can compute an *upper* bound on $\Delta S(\mu)$ by identifying an observed vector z_i such that $z_i \subseteq \mathcal{W}_i$ for each observation i. We then replace the term $\mathbb{E}[p_{ij}|\mathcal{W}_i]$ in equation (28) with the expectation $\mathbb{E}[p_{ij}|z_i]$. Denoting the upper bound on the expected consumer surplus change in equation (28) as $\Delta\bar{S}(\mu)$, we can compute a confidence interval for this upper bound as

$$[\min_{\theta\in\hat{\Theta}} \Delta\bar{S}(\theta), \max_{\theta\in\hat{\Theta}} \Delta\bar{S}(\theta)], \tag{29}$$

where $\hat{\Theta}$ is a confidence set for μ. In practice, we specify the vector z_i using the most informative price predictor that we fail to reject as belonging to the physician's information set; for the confidence set $\hat{\Theta}$ in equation (29), we use the moment inequality confidence set that we find when using the same vector z_i as an instrument.[26]

[24] Consumer surplus in this setting is equal to the physician's utility, which incorporates patient interests.

[25] We lose this flexibility if we instead study a less-than-full-information counterfactual. In this case, the researcher would have to fully specify how the baseline information set potentially differs by physician type. However, in this case, we could use the tests described in Section 7.3 to guide the specification: for each physician subsample, we could exclude from the information set the price predictors rejected by the tests.

[26] How close the upper bound $\Delta\bar{S}(\mu)$ is to the consumer surplus change $\Delta S(\mu)$ depends on how well z_i approximates the information set $\mathcal{W}_i$. An indication that our choice of z_i is reasonable is that the predicted prescription shares under z_i are similar to the observed ones. We report this comparison in Appendix D.5.

Table 7: Information Intervention - By Physician Specialty

Price Predictor	Sample	Change in... Share Cheapest	Mean OOP Costs	Mean Consumer Surplus
Average Current Prices By Drug-Carrier-Plan Type	Endocrinologist	[1.4, 3.8]	[-2.0, -0.9]	[0.008, 0.018]
Average Lagged Prices By By Drug-Plan Type	Primary Care Physician	[6.8, 32.5]	[-8.7, -3.1]	[0.073, 0.224]

Note: The change in the *Share Cheapest* indicates a confidence interval for the percentage point change in the share of visits during which the physician prescribes the cheapest drug in the choice set relative to the observed baseline. We compute *Mean OOP Costs* as a sum across individuals and drugs of the model-implied prescription share multiplied by the corresponding price; it is measured in dollars per month. The change in *Mean Consumer Surplus* equals the change in the expected utility of office visits, averaged across all sample visits and re-normalized to be expressed in dollars per patient per month. We compute consumer surplus using equations (28) and (29).

In Table 7, we report the model-predicted effect of providing endocrinologists and primary care physicians with full information on patient costs. The differences in our predicted outcomes across both groups are due to differences in: (a) preferences, as reflected in Appendix Table D.5; (b) baseline information, as reflected in the different price predictor we use to compute physicians' expectations of drug prices in the surplus formula; and (c) the distribution of actual drug prices that their patients face. When computing the change in consumer surplus, we set the physician's information in the baseline scenario equal to the most informative price predictor that we fail to reject using the tests reported in Table 6.

As shown in Table 7, primary care physicians exhibit much larger changes in all three outcome measures. With complete information on patient prices, the share of their patients taking the cheapest drug increases by between 6.8 to 32.5 percentage points, and those patients save an average of \$3.1 to \$8.7 per month in out-of-pocket costs.[27] In contrast, patients of endocrinologists see savings of only \$0.9 to \$2.0 in their average monthly costs after the intervention, with the share of patients on the cheapest drug increasing by only 1.4 to 3.8 percentage points.

For consumer surplus, we again find that the intervention translates into larger gains for patients of primary care physicians. They experience gains between 7 and 22 cents, while endocrinology patients experience gains between 1 and 2 cents. The surplus gains are smaller than the mean out-of-pocket cost savings for both groups. This gap reflects quality differences across drugs; in response to new price information, physicians sometimes shift patients from higher-quality expensive drugs to lower-quality cheap drugs, which reduces the effect of price information on surplus.

[27]From a baseline average out-of-pocket cost of \$43.9 per month, these savings translate to a 7.1%–19.8% reduction in out-of-pocket costs. This magnitude is similar to the savings found in a randomized control experiment in Desai et al. (2022); they estimate that exposing physicians to real-time information led to an 11.2% (95% CI, 6.4% to 15.7%) reduction in out-of-pocket costs for patients of the treated physicians.

Overall, our results illustrate the value of accounting for heterogeneity in information sets and preferences when determining which groups to target when designing an intervention. In our example, targeting primary care physicians results in greater out-of-pocket costs savings and surplus gains than targeting endocrinologists. This difference arises because primary care physicians have less information than endocrinologists in the baseline scenario and because they are more sensitive to price. Targeting an intervention rather than implementing it across all providers is worthwhile when implementation involves nontrivial costs.

We emphasize that our estimates of physicians' preferences and information may partly reflect patient circumstances. For example, the finding that endocrinologists are less sensitive to price may reflect that their patients tend to have more severe conditions and, therefore, care more about drug quality relative to price. However, regardless of the source of these differences, our policy implication remains the same: improving price information will have greater effects when targeted toward physician types with more price sensitivity or worse information.

Finally, we note that our counterfactual predictions hold out-of-pocket costs fixed. That is, we do not account for potential supply-side responses to the information intervention. Price competition may intensify as physicians become more informed about patient prices, leading to larger gains in consumer surplus, as shown in Bronnenberg et al. (2015). Alternatively, price information can leak from physicians to insurers, increasing price transparency and potentially allowing tacit price collusion, as in Albæk et al. (1997), ultimately leading to higher overall drug prices and lower consumer surplus.[28]

9 Moment Inequalities for Alternative Choice Models

In Section 4, we describe moment inequalities that are valid under the assumptions of the multinomial logit model introduced in Section 3. In that framework, we assume that the physician's idiosyncratic preference for drug j at visit i, ε_{ij}, is independent and identically distributed across drugs and follows a type I extreme value distribution; see equation (4). Here, we show that the general approach we use to derive moment inequalities for multinomial logit models can be applied more broadly. In particular, we derive moment inequalities that bound the parameters of both nested logit models and multinomial probit models. These models allow idiosyncratic preferences to be correlated across choices.

[28]Our counterfactual predictions also hold the total number of prescriptions for DPP-4 inhibitors fixed. However, an intervention that discloses information on the prices of treatments in this class may lead to substitution across dual-therapy diabetes classes, and thus to changes in the aggregate prescription share of the DPP-4 inhibitor class. As a result, the information intervention we consider may lead to larger gains in consumer surplus than we predict.

While the specific moments entering our inequalities for the nested logit and multinomial probit models differ from those for the multinomial logit model, the derivation procedure follows a common approach. Similarly to the multinomial logit case, we begin by deriving a moment equality (for the nested logit model) or inequality (for the multinomial probit model) where the physician's unobserved price expectation enters through a function that is globally convex. We then construct two sets of inequalities that do not depend on the physician's expectations. To obtain odds-based moment inequalities, we replace the unobserved expectations with observed prices; to derive bounding moment inequalities, we take a first-order approximation to the convex function and subsequently replace the physician's expected prices with the observed ones.

In Appendix E.1, we describe in detail the moment inequalities we derive for the nested logit model. We construct three types of inequalities. The first two types correspond to the odds-based and bounding moment inequalities introduced in equations (7) and (13), respectively, but now we restrict their application to the comparison of choice pairs in the same nest. Through these inequalities, we partially identify the ratio of the price coefficient and the choice-specific fixed effects to the parameter that determines the correlation between the idiosyncratic preferences for choices in the same nest. We then bound this extra parameter using a third type of inequality, which compares the inclusive values of any two nests. In Appendix E.1.4, we present simulation results illustrating the properties of our moment inequalities. These inequalities point-identify all the parameters of the nested logit model when (a) the agent has no expectational error, and (b) the researcher fully observes the agent's information set. In all other cases, our moment inequalities only partially identify the parameters of the nested logit model.

In Appendix E.2, we describe in detail the moment inequalities we derive for the multinomial probit model, where idiosyncratic preferences are potentially correlated across choices. In particular, we use two types of inequalities that are similar to the odds-based and bounding moment inequalities introduced in equations (7) and (13), respectively, for the multinomial logit model. However, a crucial difference is that the inequalities for the multinomial probit model compare the expected utility of two choices without conditioning on one of those two choices being selected. For this reason, the moment inequalities for the multinomial probit model are weaker than those for the multinomial logit model. The simulation results presented in Appendix E.2.4 illustrate an important implication of this difference: even if the agent has no expectational error and the researcher fully observes the agent's information set, our moment inequalities only partially identify the parameters of the multinomial probit model. Thus, the use of the moment inequalities described in Appendix E.2 for the estimation of a multinomial probit model may result in a loss of identification power.

10 Conclusion

We study the estimation of discrete choice models in which the decision-maker forms expectations about choice characteristics, such as price, but the researcher does not observe these expectations. Our estimation procedure combines the bounding inequalities in Porcher et al. (2025) with a novel type of moment inequality that generalizes the odds-based inequality in Dickstein and Morales (2018) to settings with more than two choices. We require researchers to specify only a subset of the information agents use to form their expectations. This approach contrasts with traditional estimation approaches, in which researchers must specify the exact information set agents use to forecast choice characteristics.

We apply our estimation procedure to analyze the choice of diabetes treatment. We conclude that physicians do not have perfect information on the treatment costs their patients face. Instead, our findings suggest that most physicians use relatively broad averages of out-of-pocket costs—for example, last year's average price at the drug-plan type level—when forming expectations about a patient's true out-of-pocket costs. In addition, our moment inequality estimates indicate that physicians are more sensitive to expected patient costs than previous full-information models suggest.

Our estimated model predicts that an information intervention has the potential to guide prescribing choices toward more cost-effective options, but that the effect of the intervention is heterogeneous across physicians. In particular, we estimate that the effect of providing physicians with perfect information on patient- and drug-specific out-of-pocket costs is smaller for endocrinologists than for primary care physicians. We can explain this difference by our findings that endocrinologists already have more precise price information before the intervention and are less sensitive to patient prices in their prescribing behavior.

Electronic "pop-ups" in the provider's medical chart, as studied in Desai et al. (2022), could thus help steer prescribing toward cheaper drugs. However, given the pecuniary and non-pencuniary costs of implementing these interventions—in terms of health system dollars and provider time—our evidence suggests a value of targeted interventions or academic detailing (Soumerai and Avorn, 1990). Here, sharing price information with less specialized physicians, who are less likely to know the true out-of-pocket costs of the patient, would have a larger effect on the costs patients realize.

Finally, the analysis in our empirical application relies on the assumption that physicians' idiosyncratic preferences are independent across drugs and follow a type I extreme value distribution. However, we show that one may similarly derive moment inequalities that partially identify the parameters of nested logit and multinomial probit models even if the researcher only partially observes agents' information sets. In all settings, our moment

inequality estimation approach relies on the assumption that a decision maker's expectations are rational, which is admittedly restrictive and potentially consequential. We leave for future research the study of the potential use of moment inequalities to estimate discrete choice models when researchers only partially observe agents' information sets and agents' expectations deviate from rationality.

References

Abaluck, Jason and Abi Adams-Prassl, "What do Consumers Consider Before They Choose? Identification from Asymmetric Demand Responses?," *The Quarterly Journal of Economics*, 2021, *136* (3), 1611–1663.

Albæk, Svend, Peter Møllgaard, and Per B. Overgaard, "Government-Assisted Oligopoly Coordination? A Concrete Case," *The Journal of Industrial Economics*, 1997, *45* (4), 429–443.

American Diabetes Association, "Pharmacologic Approaches to Glycemic Treatment," *Diabetes Care*, 2017, *40* (Supplement 1), S64–S74.

Andrews, Donald W. K. and Gustavo Soares, "Inference for Parameters Defined by Moment Inequalities Using Generalized Moment Selection," *Econometrica*, 2010, *78* (1), 119–157.

_ **and Xiaoxia Shi**, "Inference Based on Conditional Moment Inequalities," *Econometrica*, 2013, *81* (2), 609–666.

Arrow, Kenneth J., L. Kamran Bilir, and Alan Sorensen, "The Impact of Information Technology on the Diffusion of New Pharmaceuticals," *American Economic Journal: Applied Economics*, 2020, *12* (3), 1–39.

Baicker, Katherine and Dana Goldman, "Patient Cost-Sharing and Healthcare Spending Growth," *Journal of Economic Perspectives*, 2011, *25* (2), 47–68.

Barseghyan, Levon, Francesca Molinari, and Matthew Thirkettle, "Discrete Choice under Risk with Limited Consideration," *American Economic Review*, 2021, *111* (6), 1972—-2006.

_ **, Maura Coughlin, Francesca Molinari, and Joshua C. Teitelbaum**, "Heterogeneous Choice Sets and Preferences," *Econometrica*, 2021, *89* (5), 2015—-2048.

Bombardini, Matilde, Bingjing Li, and Francesco Trebbi, "Did US Politicians Expect the China Shock?," *American Economic Review*, 2023, *113* (1), 174–209.

Bronnenberg, Bart J., Jean-Pierre Dubé, Matthew Gentzkow, and Jesse M. Shapiro, "Do Pharmacists Buy Bayer? Informed Shoppers and the Brand Premium," *The Quarterly Journal of Economics*, 2015, *130* (4), 1669–1726.

Brown, Zach Y and Jihye Jeon, "Endogenous Information and Simplifying Insurance Choice," *Econometrica*, 2024, *92* (3), 881–911.

Canay, Ivan A, Gastón Illanes, and Amilcar Velez, "A User's Guide for Inference in Models Defined by Moment Inequalities," *Journal of Econometrics*, 2023, p. 105558.

Carrera, Mariana, Dana P. Goldman, Geoffrey Joyce, and Neeraj Sood, "Do Physicians Respond to the Costs and Cost-Sensitivity of their Patients?," *American Economic Journal: Economic Policy*, 2018, *10* (1), 113–52.

Centers for Disease Control and Prevention, "National Diabetes Statistics Report," https://www.cdc.gov/diabetes/data/statistics-report/index.html 2022. Accessed: 2023-08-21.

Ching, Andrew T., "Consumer Learning and Heterogeneity: Dynamics of Demand for Prescription Drugs after Patent Expiration," *International Journal of Industrial Organization*, 2010, *28* (6), 619–638.

_ **and Masakazu Ishihara**, "Measuring the Informative and Persuasive Roles of Detailing on Prescribing Decisions," *Management Science*, 2012, *58* (7), 1374–1387.

Chintagunta, Pradeep K., Renna Jiang, and Ginger Z. Jin, "Information, Learning, and Drug Diffusion: the Case of Cox-2 Inhibitors," *Quantitative Marketing and Economics*, 2009, *7* (4), 399–443.

CMS, "The Doctors and Clinicians National Downloadable File," https://data.cms.gov/provider-data/dataset/mj5m-pzi6#data-table 2011-2016.

_ , "National Plan and Provider Enumeration System," https://download.cms.gov/nppes/NPI_Files.html 2011-2016.

Congressional Budget Office, "Prescription Drugs: Spending, Use, and Prices," https://www.cbo.gov/system/files/2022-01/57050-Rx-Spending.pdf 2022. Accessed: 2023-08-21.

Cox, Gregory and Xiaoxia Shi, "Simple Adaptive Size-Exact Testing for Full-Vector and Subvector Inference in Moment Inequality Models," *The Review of Economic Studies*, 2023, *90* (1), 201–228.

Crawford, Gregory S. and Matthew Shum, "Uncertainty and Learning in Pharmaceutical Demand," *Econometrica*, 2005, *73* (4), 1137–1173.

Currie, Janet, Anran Li, and Molly Schnell, "The Effects of Competition on Physician Prescribing," *mimeo*, 2025.

Desai, Sunita M., Alan Z. Chen, Jiejie Wang, Wei-Yi Chung, Jay Stadelman, Chris Mahoney, Adam Szerencsy, Lisa Anzisi, Ateev Mehrotra, and Leora I. Horwitz, "Effects of Real-time Prescription Benefit Recommendations on Patient Out-of-Pocket Costs: A Cluster Randomized Clinical Trial," *JAMA Internal Medicine*, 2022, *182* (11), 1129–1137.

Dickstein, Michael J., "Physician vs. Patient Incentives in Prescription Drug Choice," 2018. mimeo.

_ **and Eduardo Morales**, "What do Exporters Know?," *The Quarterly Journal of Economics*, 07 2018, *133* (4), 1753–1801.

Doyle, Joseph J., Steven M. Ewer, and Todd H. Wagner, "Returns to physician human capital: Evidence from patients randomized to physician teams," *Journal of Health Economics*, 2010, *29* (6), 866–882.

Eizenberg, Alon, "Upstream Innovation and Product Variety in the U.S. Home PC Market," *Review of Economic Studies*, 2014, *81*, 1003–1045.

Goldman, Dana P., Geoffrey F. Joyce, and Yuhui Zheng, "Prescription Drug Cost Sharing: Associations With Medication and Medical Utilization and Spending and Health," *JAMA*, 07 2007, *298* (1), 61–69.

Grennan, Matthew, Kyle R. Myers, Ashley Swanson, and Aaron Chatterji, "No Free Lunch? Welfare Analysis of Firms Selling Through Expert Intermediaries," *The Review of Economic Studies*, 2024, *92* (4), 2537–2577.

Handel, Ben and Jonathan Kolstad, "Health Insurance for "Humans": Information Frictions, Plan Choice, and Consumer Welfare," *American Economic Review*, 2015, *105* (8), 2449–2500.

_ **and Joshua Schwartzstein**, "Frictions or Mental Gaps: What's Behind the Information We (Don't) Use and When Do We Care?," *Journal of Economic Perspectives*, 2018, *32* (1), 155–178.

_ **, Jonathan Kolstad, and Johannes Spinnewijn**, "Information Frictions and the Welfare Consequences of Adverse Selection," *Review of Economics and Statistics*, 2019, *101* (2), 326–340.

Ho, Katherine, "Insurer-Provider Networks in the Medical Care Market," *American Economic Review*, 2009, *99* (1), 393–430.

_ **and Ariel Pakes**, "Hospital Choices, Hospital Prices, and Financial Incentives to Physicians," *American Economic Review*, 2014, *104* (12), 3841–3884.

Holm, Sture, "A Simple Sequentially Rejective Multiple Test Procedure," *Scandinavian Journal of Statistics*, 1979, *6* (2), 65–70.

Holmes, Thomas J., "The Diffusion of Wal-Mart and Economies of Density," *Econometrica*, 2011, *79* (1), 253–302.

Houde, Jean-François, Peter Newberry, and Katja Seim, "Nexus Tax Laws and Economies of Density in E-Commerce: A Study of Amazon's Fulfillment Center Network," *Econometrica*, 2023, *91* (1), 147–190.

Iizuka, Toshiaki, "Physician Agency and Adoption of Generic Pharmaceuticals," *American Economic Review*, 2012, *102* (6), 2826–2858.

Illanes, Gastón, "Switching Costs in Pension Plan Choice," *mimeo*, 2017.

Ito, Koichiro, "Do Consumers Respond to Marginal or Average Price? Evidence from Nonlinear Electricity Pricing," *American Economic Review*, 2014, *104* (2), 537–63.

Kaiser Family Foundation, "Employer Health Benefits: 2022 Annual Survey," 2022.

Kamatani, Naoto, Taiya Katoh, Yoshikuni Sawai, Hitoshi Kanayama, Naoyuki Katada, and Mitsuyasu Itoh, "Comparison between the clinical efficacy of linagliptin and sitagliptin," *Journal of Diabetes and Endocrinology*, 2013, *4* (4), 51–54.

Kline, Brendan and Elie Tamer, "Recent Developments in Partial Identification," *Annual Review of Economics*, 2023, *15*, 125–150.

_ , **Ariel Pakes, and Elie Tamer**, "Moment Inequalities and Partial Identification in Industrial Organization," in Kate Ho, Ali Hortacsu, and Alessandro Lizerri, eds., *Handbook of Industrial Organization*, Vol. 4, Elsevier, 2021, pp. 345–431.

Maini, Luca and Fabio Pammolli, "Reference Pricing as a Deterrent to Entry: Evidence from the European Pharmaceutical Market," *American Economic Journal: Microeconomics*, 2023, *15* (2), 345–383.

Manski, Charles F., "Nonparametric Estimation of Expectations in the Analysis of Discrete Choice Under Uncertainty," in William Barnett, James Powell, and George Tauchen, eds., *Nonparametric and Semiparametric Methods in Econometrics and Statistics*, Cambridge: Cambridge University Press, 1991.

Morales, Eduardo, Gloria Sheu, and Andrés Zahler, "Extended Gravity," *The Review of Economic Studies*, 2019, *86* (6), 2668–2712.

Nigro, Stefanie C. and Jennifer D. Goldman, "Linagliptin-Induced Arthralgia," *Clinical Diabetes*, 2022, *40* (1), 109–112.

Oregon Health Authority, "All Payer All Claims Data," https://www.oregon.gov/oha/HPA/ANALYTICS/Pages/All-Payer-All-Claims.aspx 2011-2016.

Pakes, Ariel, "Alternative Models for Moment Inequalities," *Econometrica*, 2010, *78* (6), 1783–1822.

_ , **Jack Porter, Katherine Ho, and Joy Ishii**, "Moment Inequalities and Their Application," *Econometrica*, 2015, *83* (1), 315–334.

Porcher, Charly, Eduardo Morales, and Thomas Fujiwara, "Measuring Information Frictions in Migration Decisions: A Revealed-Preference Approach," *mimeo*, 2025.

Schnell, Molly, "Physician Behavior in the Presence of a Secondary Market: The Case of Prescription Opioids," *mimeo*, 2025.

Shrank, William H., Henry N. Young, Susan L. Ettner, Peter Glassman, Steven M. Asch, and Richard L. Kravitz, "Do the Incentives in 3-tier Pharmaceutical Benefit Plans Operate as Intended? Results from a Physician Leadership Survey," *American Journal of Managed Care*, 2005, *11* (1), 16–22.

Soumerai, Stephen B. and Jerry Avorn, "Principles of Educational Outreach ('Academic Detailing') to Improve Clinical Decision Making," *JAMA*, 1990, *263* (4), 549–556.

Vuong, Quang H., "Likelihood Ratio Tests for Model Selection and Non-Nested Hypotheses," *Econometrica*, 1989, *57* (2), 307–333.

Wollmann, Thomas G., "Trucks Without Bailouts: Equilibrium Product Characteristics for Commercial Vehicles," *American Economic Review*, 2018, *108* (6), 1364–1406.