NBER WORKING PAPER SERIES

PATIENT COSTS AND PHYSICIANS' INFORMATION

Michael J. Dickstein
Jihye Jeon
Eduardo Morales

Working Paper 32014
http://www.nber.org/papers/w32014

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
January 2024

We thank Stephane 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 for his very useful discussion. We also thank the staff at the Oregon Health Authority for assistance using the APAC dataset. All errors are our own. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Patient Costs and Physicians' Information
Michael J. Dickstein, Jihye Jeon, and Eduardo Morales
NBER Working Paper No. 32014
January 2024
JEL No. I11,I13,L0,L15

ABSTRACT

Health insurance plans in the U.S. increasingly use price mechanisms to steer demand for prescription drugs. The effectiveness of these incentives, however, depends both on physicians' price sensitivity and their knowledge of patient prices. We develop a moment inequality model that allows researchers to identify agents' preferences without fully specifying their information. Applying this model to diabetes care, we find that physicians lack detailed price information and are more price-elastic than full-information models imply. We predict that providing physicians detailed information on prices at the point of prescribing can save patients 12-23% of their out-of-pocket costs for diabetes treatment.

Michael J. Dickstein
New York University
Stern School of Business
Kaufman Management Center,7-78
44 West 4th Street
New York, NY 10012
and NBER
michael.dickstein@nyu.edu

Eduardo Morales
Department of Economics
Princeton University
291 Julis Romo Rabinowitz Building
Princeton, NJ 08544
and NBER
ecmorale@princeton.edu

Jihye Jeon
Boston University
Department of Economics
270 Bay State Road
Boston, MA 02215
jjeon@bu.edu

An online appendix is available at http://www.nber.org/data-appendix/w32014

1 Introduction

Real per-capita annual spending on prescription drugs increased from $140 to over $1,000 in the U.S. between 1980 and 2018, more than doubling the overall growth in health care spending during the same period (CBO, 2022). In response, as a means 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. For example, in 2000, 22% of employer insurance plans had a single copayment level for all drugs, and only 27% had three or more tiers (KFF, 2005). In contrast, by 2022, 85% of employer plans had at least three tiers (KFF, 2022).

The success of price incentives in shifting demand for prescription drugs depends both 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, the out-of-pocket cost for an expensive branded drug increases and usage remains high relative to lower cost options, is the lack of switching because the branded medication has higher effectiveness, because physicians and their patients are price-inelastic, or because physicians are unaware of the relative differences in out-of-pocket costs?

We develop a model that allows us to estimate both physicians' sensitivities to out-of-pocket costs and the value they place on a drug's efficacy. Importantly, we do so allowing physicians to vary unobservably in the information they use to form expectations about drug- and patient-specific out-of-pocket costs. In addition, we use the model to test whether physicians with different training or with different prescribing experience possess more or less information about patient prices. Combining estimates of physicians' preferences with inferences about their information, we evaluate the effects of policies to inform physicians about out-of-pocket costs at the point of prescribing.

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 because of both the size of the affected population and the rapid growth in treatment costs. In the U.S., 37 million people lived with diabetes in 2019 (CDC, 2022), and the direct medical costs of diabetes totaled $237 billion in 2017, rising 26% above 2012 inflation-adjusted levels (ADA, 2017).

Using Oregon's All Payer All Claims database for the years 2012-2016, 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 demographic information. Using these data, we start by documenting two facts about out-of-pocket costs. First, there is significant dispersion in monthly 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 $40 and $50, depending on the drug. The interquartile range of out-of-pocket costs is large relative to the mean, ranging between $20 and $30 for a drug in a given year, when computed across the distribution of insurers and plans. If we compute the same statistic across drugs and plans for a given insurer, the interquartile range of out-of-pocket costs is also substantial, lying between $16 and $43 for the top four insurers.

As a second key fact, we observe 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 patients' lowest-cost option only 38% of the time, roughly the same as if they had chosen randomly. If instead physicians selected the cheapest alternative at each visit, patients would save roughly $10 per month on average. Furthermore, for some plans, patients would see much larger savings, on the order of $50 to $100 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. While we assume the physician's expectations are rational, we allow the information set to vary flexibly across physicians and office 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, generalize the approach in Dickstein and Morales (2018) to settings with more than two choices. We build the second set, labeled "bounding" moments, as in Fujiwara et al. (2023). We show formally that 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 conduct a simulation to illustrate the properties of both our moment inequalities and alternative full-information estimation approaches. We have four key conclusions. First, following Manski (1991, 2004), we show maximum likelihood estimates of preference parameters are inconsistent when the researcher incorrectly specifies the agent's information set. The bias grows in the degree to which agents' true expectations differ from the expectations implied by the researcher's assumed information set. Second, consistent with our formal analysis, our inequalities yield an identified set that contains the true parameter value whenever our instruments—i.e. the

variables we assume agents know—form a *subset* of agents' true information sets. While now containing the true parameter, the size of the identified set nonetheless grows in the degree to which agents' true expectations differ from the expectations implied by the researcher's instruments. Third, we show that if our instruments coincide with the agent's complete information set, the identified set defined by our inequalities includes only the true parameter. 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 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 our 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 our estimated model to predict the effect of an intervention that provides patient-specific price information to physicians at the point 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 -0.53 when we assume providers know the patient's out-of-pocket costs for each drug. When we instead assume providers form expectations with less information—specifically, with information only on average prices by drug and insurance plan type—we find an analogous own-price-elasticity equal to -1.77. If instead we assume physicians know only drug-specific averages of last year's prices, the same elasticity equals -3.77.

Importantly, the distinct 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, who collectively account for around half of the patients in our sample, decide to cut a drug's out-of-pocket costs by 50%. Depending on which information set the researcher assumes, the empirical model predicts the now cheaper product's prescription share will increase anywhere from 3.5 percentage points to 23.5 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, estimate relatively inelastic demand.

This sensitivity to informational assumptions, in both parameter estimates and counterfactual predictions, motivates our move to a moment inequality framework. With our inequalities, we test assumptions about physicians' information sets before evaluating counterfactual policies. In our setting, we reject the null hypothesis that physicians have perfect information on out-of-pocket costs. Instead, our data and model suggest physicians in-

corporate only coarse price information in their choices; specifically, we fail to reject that physicians form expectations on out-of-pocket costs using either contemporaneous or lagged average prices for each drug, or lagged average prices for each drug by insurance plan type. By concluding that physicians form expectations about patient costs using only these aggregate price measures, we depart from the more 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).

Using our moment inequality model, we evaluate the effect of an intervention to change physicians' information. Here, we quantify the effect of moving from a setting in which physicians form expectations using broad price averages—specifically, those averages that our moment inequality estimates suggest compose a subset of physicians' information sets—to a counterfactual setting in which physicians know actual patient prices. When physicians possess more detailed information, we predict a reduction in out-of-pocket costs of roughly \$5 to \$10 per month, from a baseline of \$46 per month. However, the 12 to 23% decline in costs leads to a smaller surplus gain of between \$0.10 and \$0.24 per patient per month. Here, the gap between the cost savings and surplus gains is due to differential drug quality. With better information on patient prices, physicians switch patients from high quality expensive drugs toward lower quality cheaper drugs.

Finally, we show the cost savings and surplus gains per patient from our informational intervention would be smaller if we provided information only to endocrinology specialists. Our estimates show that endocrinologists begin with better information on prices and are less elastic with respect to expected prices; as a consequence, providing these specialists better price information leads to an average reduction in monthly out-of-pocket costs of only \$2 to \$5. Thus, if providing physicians with patient-specific prices at the point of prescribing is costly, our analysis suggests a value of targeting this provision toward general practitioners.[1]

Our paper relates to several research areas. First, we contribute to a literature that explores heterogeneity in physicians' drug treatment decisions. This literature studies the influence of several factors, including financial incentives (Iizuka, 2012; Dickstein, 2018), advertising or detailing (Ching and Ishihara, 2012; Grennan et al., 2021), as well as the interplay with secondary markets (Schnell, 2022) and competitive forces (Currie et al., 2023). Our results emphasize the role that physician information plays in prescribing behavior: we show that estimates of price responsiveness are biased if researchers mis-specify physicians' information sets, and we present evidence that indicates physicians' price information is

[1]Our finding that physicians are heterogeneous in their information and price sensitivity may reflect differential training or a different population of patients: if endocrinologists treat more severely ill patients, for example, their choices may reflect a preference for efficacy over price. Nonetheless, under either source of heterogeneity, the per-patient savings from interventions that target primary care physicians are greater.

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 Carrera et al. (2018) and Desai et al. (2022), who show that providers' prescribing behavior reacts to informational shocks on out-of-pocket costs.

Second, our research relates 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, 2023), 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 literature in that we do not micro-found physicians' information sets, but rather apply moment inequalities to infer their content.

Finally, we contribute to a literature that uses moment inequalities to estimate agents' preference parameters. 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). Previous 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), Fujiwara et al. (2023), and Houde et al. (2023). Our contribution is to generalize the odds-based moment inequalities introduced in Dickstein and Morales (2018), and subsequently applied in Bombardini et al. (2023), to discrete choice settings with more than two, and possibly many, choices. In our model, we allow for individual- and choice-specific unobserved preference heterogeneity, while also permitting agents to have unobserved expectations over a product characteristic, like price. In this way, our approach can handle many discrete choice settings in which agents face uncertain product attributes.

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 and 4, we present a model of prescription drug choice and introduce the inequalities we use for estimation. In Section 5, we present simulation results comparing the properties of our moment inequality estimator to those of maximum likelihood estimators. We present our estimates and counterfactual results in sections 6 and 7, and test for heterogeneity in Section 8. Section 9 concludes.

2 Empirical Setting and Data

Our analysis focuses on care for type 2 diabetes patients. In Section 2.1, we describe the treatments typically prescribed for diabetes patients. 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

The treatment of type 2 diabetes often begins with a diagnosis based on abnormal test results for either fasting plasma glucose or hemoglobin A1c. The treatment's goal is to achieve a particular A1c level (ADA, 2017). Treatment usually starts with metformin, a generic drug. If the patient fails to achieve the A1c goal after three months, the clinician may start the patient on dual combination therapy. At this stage, physicians must choose among several alternative drug classes, including DPP-4 inhibitors, GLP-1 agonists, and SGLT2 inhibitors. We focus on this set of drug classes in our analysis. The choice across these classes depends on medical factors.[2] We treat the choice of class as given exogenously by the patient's health status, and focus on the choice of treatment within class as a function of price and efficacy.

2.2 Data

We use data from Oregon's All-Payers All-Claims (APAC) database for the years 2011-2016. Our sample includes both medical and prescription drug claims for patients with private insurance through the individual insurance market and through group insurance.[3] 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. In addition, we complement the information on a physician's background in the APAC data by validating the physician's characteristics in two public registries that contain information on the provider's specialty, gender, and medical school graduation year.[4]

Sample creation. To form our sample, we include only claims where the identity of providers, patients, and treatments verify certain restrictions. First, we restrict our sample to claims in which the physician's specialty is one that typically provides primary care for diabetes patients, including family medicine, internal medicine, pediatrics, obstetrics and gynecology, and endocrinology. We also include non-physician providers, such as nurse practitioners, who can prescribe treatments for diabetes patients.[5] Second, we focus only on patients who both receive a diagnosis of type 2 diabetes and who are prescribed a drug in one of three treatment classes: DPP-4 inhibitors, GLP-1 agonists, and SGLT2 inhibitors.

[2]The classes differ in their efficacy, risk of hypoglycemia, likelihood for weight gain, and other side effects.

[3]In March 2016, the US Supreme Court, in Gobeille v. Liberty Mutual Insurance Company, created an exemption that allows self-insured plans to opt out of reporting their claims to a state's all-payers database. Following that decision, we lose claims for a portion of self-insured plans.

[4]Specifically, we use the National Plan and Provider Enumeration System registry and the Doctors and Clinicians National Downloadable File.

[5]Throughout the paper, we use "physician" as shorthand for both medical doctors and other providers who prescribe treatments in our sample.

Finally, we additionally restrict our sample in several minor ways, such as excluding drug observations that reflect refills rather than active choices by a physician. In Appendix A.1, we detail the steps we follow to build our sample.

The resulting sample includes 184,783 claims prescribed by 4,990 providers for a set of 35,721 patients. Close to 70% of providers are primary care physicians with specialties in internal medicine or family medicine; 2% of physicians are endocrinologists with expertise in metabolic diseases like diabetes; and the remaining 28% are non-physician providers or providers with other specialties including obstetrics and gynecology. Although endocrinologists make up a small fraction of providers in our sample, they account for a larger share of claims: on average, endocrinologists record 51 claims per quarter in our sample, relative to 14 per quarter for primary care physicians, and only 9 per quarter for non-physician providers or physician providers with other specialties.

The average patient age is 55 years old. Among the insurance plan type options, 46% of patients enroll in a preferred provider organization (PPO) plan and 33% choose either a health maintenance organization (HMO) or point-of-service (POS) plan. The remaining 21% have self-insured plans. Among the set of insurers, the largest carrier enrolls 25% of the patients, and the top four carriers by patient volume jointly account for 70% of patients.

Drug choice set. In the three drug classes we consider, the set of choices available to physicians varies from three to four drug treatments. In all three classes, the most popular treatment accounts for approximately 70% of non-refill drug claims, with the remaining 30% distributed roughly equally across all other available treatments.

We use copayments as our measure of the out-of-pocket costs patients face when filling a prescription. Unlike coverage for inpatient and outpatient services, it is rare for the plans in our sample to use either deductibles or coinsurance for prescription drug spending.[6] When specifying the copayment levels for each drug and plan, however, we face a missing data problem. We do not observe the full drug formulary at the plan level; using our claims data, we can only infer copayments using observations from patients who filled a prescription. To generate the full cost list for a plan, we employ a random forest model that uses our observed data to impute missing drug prices for all plans and years. We provide more detail on this imputation in Appendix A.2.

In Table 1, we report summary statistics of the distribution of copayments. The statistics show significant heterogeneity in out-of-pocket costs. Looking within drug, and focusing on the drug Janumet in the DPP-4 inhibitor class as an example, we see in panel A that the mean monthly price is close to $42 with a standard deviation of $28 per month. The implied coefficient of variation is thus 0.67; across the drugs in the classes we study, the coefficient

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

Table 1: Distribution of Monthly Out-Of-Pocket Costs

Panel A: By Drug - Variation Across Plans

Drug Class	Drug	Mean	St. Dev.	IQR
DPP-4 Inhibitors	Janumet	41.76	28.18	20.46
	Januvia	44.29	29.53	26.71
	Tradjenta	46.52	22.53	23.08
GLP-1 Agonists	Bydureon	50.42	39.62	38.30
	Byetta	52.11	40.50	37.25
	Trulicity	59.18	37.16	37.17
	Victoza	61.68	44.62	45.32
SGLT2	Farxiga	64.43	51.97	65.78
	Invokana	57.63	40.11	61.49
	Jardiance	56.06	34.44	42.67

Panel B: By Carrier - Variation Across Plans and Drugs

Drug Class	Carrier	Mean	St. Dev.	IQR
DPP-4 Inhibitors	A	42.16	15.69	15.89
	B	36.77	22.70	23.51
	C	79.95	31.72	43.42
	D	32.23	13.51	17.69
GLP-1 Agonists	A	51.59	24.58	17.01
	B	35.53	21.78	23.46
	C	111.05	39.28	38.70
	D	37.32	19.78	21.84
SGLT2	A	49.58	22.19	21.42
	B	44.19	35.06	43.03
	C	104.54	39.36	52.55
	D	32.07	18.02	22.80

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 out-of-pocket 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 all plans within a drug; while panel B reflects variation across plans and drugs within a carrier.

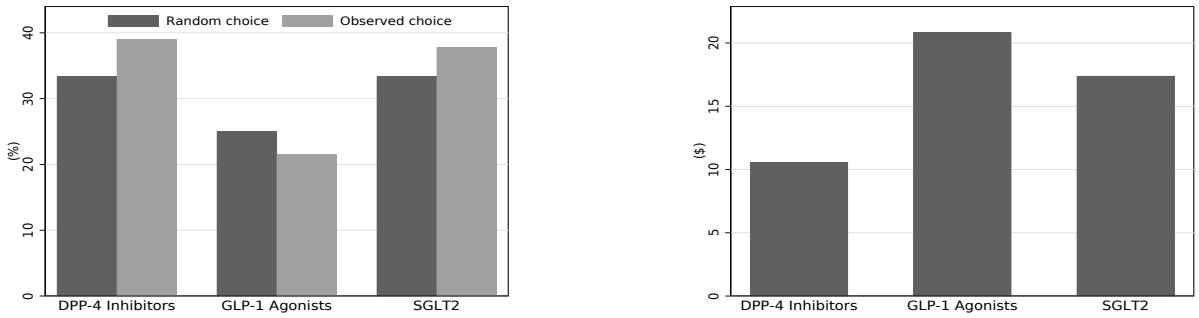
of variation is similar, ranging from 0.48 to 0.81. In panel B, we report statistics of the price variation across drugs and plans offered by each of the top four carriers. The dispersion in prices within a carrier is slightly smaller than the price dispersion within drug, with coefficients of variation that vary between 0.35 and 0.79, depending on the carrier and class.

With significant heterogeneity in copayments, both across insurance plans for a given drug as well as 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 the left panel of Figure 1, we show physicians indeed prescribe the cheapest drug at roughly the same rate as if they had chosen the prescribed drug randomly. In the right panel, we show that, relative to a world in which physicians always choose the cheapest drug within a class, patients face additional out-of-pockets costs per month of between $10 (for DPP-4 Inhibitors) and $20 (for GLP-1 Antagonists) on average under the observed distribution of choices.

The choice pattern in Figure 1 could arise due to physician preferences for efficacy over price, or could reflect physicians' lack of information about a given patient's out-of-pocket costs for each drug in the choice set. The heterogeneity in prices reported in Table 1 may also reflect some noise due to our need to predict copayments for drug-plan pairs not observed in the data. To better understand physicians' information and their preferences, we move next to present a model that accounts for the possibility that physicians possess imperfect information about patient prices as well as for possible quality differences across drugs in the choice set. In addition, in Section 4, we outline an estimation approach that can account for classical measurement error in our observed price measure.

Figure 1: Out-of-pocket Costs and Physicians' Choices



(a) Prob. Choosing Cheapest Drug

(b) Monthly Cost Difference

Notes: In panel (a), by drug class, we report the observed probability of choosing the drug with the lowest out-of-pocket cost. We compare this observed probability to a hypothetical setting in which the physician chooses each drug with equal probability. In panel (b), we report the mean difference in monthly copayments between the observed drug chosen for a patient and the cheapest drug available to that patient in the class.

3 Model of Prescription Choice

We model a physician's choice of prescription drug within a class 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

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

$$u_{ij} = \kappa_j + \alpha p_{ij}, \tag{1b}$$

where p_{ij} denotes the out-of-pocket cost of treatment j for the patient at visit i, α captures the physician's sensitivity to patient costs, and κ_j and ε_{ij} capture the common and idiosyncratic components of the quality of treatment j, respectively.[7] Defining a binary 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 denotes the cardinality of the set of drugs that the physician could have prescribed at visit i; $\mathcal{J}_i$ denotes the physician's information set at visit i; and $\mathbb{E}[\cdot|\mathcal{J}_i]$ is a conditional expectation operator reflecting the physician's beliefs. We assume physicians' expectations are rational and, 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 office visits of interest. We impose the following assumptions on physicians' information sets:

$$\mathcal{J}_i = (\mathcal{W}_i, \varepsilon_i), \tag{3a}$$

$$(\alpha, \kappa) \subseteq \mathcal{W}_i, \tag{3b}$$

where $\kappa = \{\kappa_j\}_{j=1}^J$, $\varepsilon_i = \{\varepsilon_{ij}\}_{j=1}^J$ 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 (3a), the information set $\mathcal{J}_i$ thus includes the vector of idiosyncratic shocks, ε_i, and all variables in $\mathcal{W}_i$. Equation (3b) imposes that $\mathcal{W}_i$ includes the price sensitivity α, and the drug quality terms κ, but other variables may also enter that set.

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 $F_\varepsilon(\cdot)$ denotes the cumulative distribution function of ε_i. Equation (4) imposes that ε_i is independent of all other elements of the physician's information set, as included in $\mathcal{W}_i$. The equation also imposes that, for any visit i, ε_{ij} is independent and identically distributed across all $j = 1, \ldots, J$, and follows a type I extreme value distribution with location parameter equal to zero and scale parameter equal to one. Second, we assume that

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

[7] We use quality to denote the drug's efficacy, side effect profile, or the ease of prescribing the drug.

where $p_i = \{p_{ij}\}_{j=1}^J$. Equation (5) imposes that, once we condition on all other elements of the physician's information set, the vector of idiosyncratic shocks ε_i does not provide any additional information that helps the physician forecast the patient's out-of-pocket costs.[8]

Equations (1), (3), and (5) imply

$$\mathbb{E}[\mathcal{U}_{ij}|\mathcal{J}_i] = \mathbb{E}[u_{ij}|\mathcal{W}_i] + \varepsilon_{ij} = \kappa_j + \alpha\mathbb{E}[p_{ij}|\mathcal{W}_i] + \varepsilon_{ij}, \qquad \text{for } j = 1, \ldots, J, \tag{6}$$

where the first equality is implied by equations (1a), (3a), and (5), and the second equality is implied by equations (1b) and (3b). Equations (2) and (4) further imply that we can write the probability that drug j is prescribed given $\mathcal{W}_i$ as

$$\mathcal{P}(d_{ij} = 1|\mathcal{W}_i) = \frac{\exp(\kappa_j + \alpha\mathbb{E}[p_{ij}|\mathcal{W}_i])}{\sum_{j'=1}^J \exp(\kappa_{j'} + \alpha\mathbb{E}[p_{ij'}|\mathcal{W}_i])} \qquad \text{for } j = 1, \ldots, J. \tag{7}$$

If we assume physicians have perfect information on prices and, thus, $\mathbb{E}[p_i|\mathcal{W}_i] = p_i$, our model becomes a multinomial logit model with choice-specific fixed effects. If we place no restrictions on how physicians form their price predictions, observing the conditional probabilities, $\mathcal{P}(d_{ij} = 1|\mathcal{W}_i)$ for all $j = 1, \ldots, J$, generally does not allow a researcher to distinguish how a physician's preference parameters (κ, α) or information set, $\mathcal{W}_i$, affect her treatment choices. As a middle ground, we show in Section 4 that the assumption that physicians' expectations are rational is enough to learn both about the content of physicians' information set as well as about the value of their preference parameters.

To clarify how we will later apply our model to study diabetes drug choices, we comment on two additional features of the model. First, although formally physicians choose the prescription drug in our set-up, their patients' information and preferences may influence their choices. Thus, the drug qualities, $\{\kappa_j\}_{j=1}^J$, the price sensitivity, α, and the physician's information on prices, $\mathcal{W}_i$, may partly reflect patient input. Second, we emphasize again that the subindex i on p_{ij} implies that drug j's out-of-pocket costs may vary across visits, consistent with the statistics in Table 1. Similarly, the subindex i on $\mathcal{W}_i$ implies that physicians' information about drug prices may vary across visits. Thus, in our setting, we allow patients to face different drug prices depending on their plan, and we allow physicians to have different information about the out-of-pocket costs of patients on different plans.[9]

[8] If $p_i \subseteq \mathcal{W}_i$ and, thus, the physician at visit i has perfect information on p_i, equation (5) naturally holds. However, equation (5) is also compatible with the physician having imperfect information on p_i; it simply requires that all information relevant to the physician's forecast of p_i is included in $\mathcal{W}_i$.

[9] Conversely, equations (1) and (4) impose restrictions on the distribution of drug qualities. Quality is the sum of the component κ, common across all visits, and the idiosyncratic component, ε_i. The common quality component matches the empirical setting we study: when we condition on specific classes of diabetes treatments, the patients treated will have similar health statuses and similar treatment effects under drug j.

Finally, we assume the researcher collects a random sample of N visits. For each visit in the sample, we assume the researcher observes the drug prescribed, $d_i = \{d_{ij}\}_{j=1}^J$; the out-of-pocket costs for each drug, $p_i = \{p_{ij}\}_{j=1}^J$; and, a set of variables that may be used to predict these out-of-pocket costs, $z_i = \{z_{ij}\}_{j=1}^J$. Here, z_{ij} is a vector of covariates correlated with p_{ij} that may belong to $\mathcal{W}_i$; e.g., z_{ij} may include the average out-of-pocket cost of drug j across subsets of insurance plans. Crucially, we do not assume the researcher observes the complete set $\mathcal{W}_i$ for any visit.

Given a normalization $\kappa_1 = 0$, the goal of estimation is to recover the value of the parameters $\{\kappa_j\}_{j=2}^J$ and α, and to learn about the content of the information sets $\{\mathcal{W}_i\}_{i=1}^N$. To acquire knowledge about $\{\mathcal{W}_i\}_{i=1}^N$, the researcher can test the null hypothesis that z_i belongs to the information set $\mathcal{W}_i$ of every physician in a group of interest; i.e., the researcher tests $H_0 \colon z_i \subseteq \mathcal{W}_i$ for a subset of visits. To simplify the notation, we use $\theta \equiv (\theta_\alpha, \theta_{\kappa_2}, \ldots, \theta_{\kappa_J})$ to denote the unknown parameter vector, Θ to denote the parameter space, and $\theta^* \equiv (\alpha, \kappa_2, \ldots, \kappa_J)$ to denote the true parameter value, which is determined by equation (7).

4 Moment Inequalities

In this section, we show how to partially identify θ^*. 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 how we use these inequalities to compute a confidence set for θ^*.

4.1 Odds-based Inequalities

For any two drugs j and j', any value of z_i in its support, and any $\theta \in \Theta$, we define the following odds-based moment inequality

$$\mathbb{m}^o_{jj'}(z_i, \theta) \geqslant 0 \tag{8a}$$

with

$$\mathbb{m}^o_{jj'}(z_i, \theta) \equiv \mathbb{E}[d_{ij} \exp(-(\theta_{\kappa_j} - \theta_{\kappa_{j'}} + \theta_\alpha \Delta p_{ijj'})) - d_{ij'} | z_i], \tag{8b}$$

and $\Delta p_{ijj'} = p_{ij} - p_{ij'}$. We denote as Θ^o_0 the set of values of θ that jointly satisfy the inequality in equation (8) for every value of z_i in its support and all pairs of drugs j and j' in the physician's choice set. Formally, denoting as $\mathcal{Z}$ the support of z_i, we define

$$\Theta^o_0 \equiv \{\theta \in \Theta \colon \mathbb{m}^o_{jj'}(z, \theta) \geqslant 0 \text{ for all } z \in \mathcal{Z},\ j = 1, \ldots, J, \text{ and } j' = 1, \ldots, J\}. \tag{9}$$

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

Theorem 1 *Let $\theta^* \equiv (\alpha, \kappa_2, \ldots, \kappa_J)$ be defined by equation (7) and the normalization $\kappa_1 = 0$. If $z_i \subseteq \mathcal{W}_i$, then $\theta^* \in \Theta_0^o$.*

Theorem 1 indicates that, when evaluated at the true parameter value, the inequality in equation (8) holds if z_i belongs to the information set $\mathcal{W}_i$ for every visit i in the population of interest. This inequality holds regardless of the value of z_i on which we condition, and holds for any two drugs j and j' we may use to define the moment, as long as these drugs could have been prescribed by all healthcare providers in the population of interest. We provide an intuitive explanation of Theorem 1 below. The formal proof appears in Appendix B.1.

The moment inequality in equation (8) is a generalization to multinomial settings of the odds-based inequality introduced in Dickstein and Morales (2018) for binary choice models.[10] To understand why the inequality in equation (8) holds for $\theta = \theta^*$, a key equation is

$$\frac{\mathbb{E}[d_{ij}|\mathcal{W}_i]}{\mathbb{E}[d_{ij'}|\mathcal{W}_i]} = \exp(\Delta\kappa_{jj'} + \alpha\mathbb{E}[\Delta p_{ijj'}|\mathcal{W}_i]), \tag{10}$$

with $\Delta\kappa_{jj'} = \kappa_j - \kappa_{j'}$. Equation (7) implies equation (10) for any drugs j and j' in the physician's choice set and any information set $\mathcal{W}_i$. Reordering the terms in equation (10), we obtain

$$\mathbb{E}[d_{ij}\exp(-(\Delta\kappa_{jj'} + \alpha\mathbb{E}[\Delta p_{ijj'}|\mathcal{W}_i]) - d_{ij'}|\mathcal{W}_i] = 0. \tag{11}$$

As the moment function in equation (11) is convex in the unobserved expectation $\mathbb{E}[\Delta p_{ijj'}|\mathcal{W}_i]$ and physicians' expectational errors are mean zero (as implied by the assumption of rational expectations), Jensen's inequality implies that

$$\mathbb{E}[d_{ij}\exp(-(\Delta\kappa_{jj'} + \alpha\Delta p_{ijj'}) - d_{ij'}|\mathcal{W}_i] \geqslant 0. \tag{12}$$

This inequality also holds if the variable $\Delta p_{ijj'}$ is affected by classical measurement error in prices, as classical measurement errors and expectational errors have the same properties in our setting. Finally, applying the Law of Iterated Expectations, we conclude that

$$\mathbb{E}[d_{ij}\exp(-(\Delta\kappa_{jj'} + \alpha\Delta p_{ijj'}) - d_{ij'}|z_i] \geqslant 0, \tag{13}$$

for any $z_i \subseteq \mathcal{W}_i$, proving Theorem 1 in this way.

[10] A restriction imposed in the multinomial model in Section 3 that is not imposed in the binary model in Dickstein and Morales (2018) is the requirement that ε_i follows a type I extreme value distribution. In the binary choice case, Dickstein and Morales (2018) show one can derive inequalities analogous that in equation (8) if the distribution of $\varepsilon_{ij} - \varepsilon_{ij'}$ is log-concave; thus, ε_{ij} can follow multiple distributions, including the normal distribution.

To gain intuition on why inequalities of the type in equation (8) provide non-trivial bounds, consider the following specific cases of the general inequality in equation (8):

$$\mathbb{E}[d_{i1}\exp(-(-\theta_{\kappa_2}+\theta_\alpha\Delta p_{i12}))-d_{i2}|z_i]\geqslant 0, \tag{14a}$$

$$\mathbb{E}[d_{i2}\exp(-(\theta_{\kappa_2}+\theta_\alpha\Delta p_{i21}))-d_{i1}|z_i]\geqslant 0, \tag{14b}$$

where we have imposed the normalization $\kappa_1 = 0$ and, as a reminder, θ_α and θ_{κ_2} are unknown parameters with true values α and κ_2, respectively. The function $\exp(x)$ goes to 0 as x goes to $-\infty$; thus, given a value of θ_α, equation (14a) provides a finite lower bound on θ_{κ_2} and, similarly, equation (14b) provides a finite upper bound on θ_{κ_2}. Theorem 1 guarantees that, if $\theta_\alpha = \alpha$, κ_2 belongs to the interval defined by these bounds.

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

$$\mathbb{m}^b_{jj'}(z_i,\theta,e_{jj'}(\cdot))\geqslant 0 \tag{15a}$$

with

$$\mathbb{m}^b_{jj'}(z_i,\theta,e_{jj'}(\cdot))\equiv$$
$$\mathbb{E}[d_{ij'}-d_{ij}\exp(-e_{jj'}(z_i,\theta))(1+e_{jj'}(z_i,\theta)-(\theta_{\kappa_j}-\theta_{\kappa_{j'}}+\theta_\alpha\Delta p_{ijj'}))|z_i]. \tag{15b}$$

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, and may vary by pair of drugs j and j'. Defining e as the set of functions that includes the function $e_{jj'}(\cdot)$ for all drug pairs (i.e., $e = \{e_{jj'}(\cdot)\}_{j=1,j'=1}^{J,J}$), we denote as $\Theta^b_0(e)$ the set of values of θ that jointly satisfy the inequality in equation (15) for every value of z_i in its support, and every pair of drugs j and j' in the physician's choice set. Formally,

$$\Theta^b_0(e)\equiv\{\theta\in\Theta\colon \mathbb{m}^b_{jj'}(z,\theta,e_{jj'}(\cdot))\geqslant 0 \text{ for all } z\in\mathcal{Z},\ j=1,\dots,J, \text{ and } j'=1,\dots,J\}. \tag{16}$$

Regardless of the set e used to build the moment inequalities in equation (15), the following theorem establishes a sufficient condition for the true parameter value θ^* to belong to $\Theta^b_0(e)$.

Theorem 2 *Let $\theta^*\equiv(\alpha,\kappa_2,\dots,\kappa_J)$ be defined by equation (7) and the normalization $\kappa_1 = 0$. If $z_i\subseteq\mathcal{W}_i$, then $\theta^*\in\Theta^b_0(e)$ 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 (15) holds if, for every visit i in the population of interest, z_i belongs to $\mathcal{W}_i$ and the two drugs j and j' could have been prescribed by the corresponding healthcare provider. Importantly, this inequality holds regardless of the set of functions e used to build the bounding moment inequalities, as long as they are deterministic functions of z_i and the unknown parameter vector θ. We provide an intuitive explanation of Theorem 2 below. The formal proof appears in Appendix B.2.

The bounding inequality in equation (15) was first introduced in Fujiwara et al. (2023) for the case in which $e_{jj'}(z_i, \theta) = \check{e}$ for a constant $\check{e} \in \mathbb{R}$. We show the inequality in equation (15) holds more generally for any $e_{jj'} \colon \mathcal{Z} \times \Theta \to \mathbb{R}$ and, by increasing the set of functions $e_{jj'}(\cdot)$ we consider, we obtain bounding inequalities that yield tighter bounds on θ^*. To understand why the inequality in equation (15) holds for $\theta = \theta^*$, we first multiply equation (11) by -1, obtaining the equality

$$\mathbb{E}[d_{ij'} - d_{ij} \exp(-(\Delta\kappa_{jj'} + \alpha\mathbb{E}[\Delta p_{ijj'}|\mathcal{W}_i])|\mathcal{W}_i] = 0. \tag{17}$$

As $-\exp(-x)$ is concave in x, a first-order approximation to it around any point bounds it from above. Thus, for any $e_{jj'} \colon \mathcal{Z} \times \Theta \to \mathbb{R}$, we derive from equation (17) the inequality:

$$\mathbb{E}[d_{ij'} - d_{ij} \exp(-e_{jj'}(z_i, \theta^*))(1 + e_{jj'}(z_i, \theta^*) - (\Delta\kappa_{jj'} + \alpha\mathbb{E}[\Delta p_{ijj'}|\mathcal{W}_i]))|\mathcal{W}_i] \geqslant 0, \tag{18}$$

where $e_{jj'}(z_i, \theta^*)$ is the point around which the first-order approximation is taken. If $z_i \subseteq \mathcal{W}_i$, properties of rational expectations imply the sign of the inequality in equation (18) is preserved when introducing the price difference, $\Delta p_{ijj'}$, in place of the unobserved expectation, $\mathbb{E}[\Delta p_{ijj'}|\mathcal{W}_i]$. In this way, we obtain the inequality:

$$\mathbb{E}[d_{ij'} - d_{ij} \exp(-e_{jj'}(z_i, \theta^*))(1 + e_{jj'}(z_i, \theta^*) - (\Delta\kappa_{jj'} + \alpha\Delta p_{ijj'}))|\mathcal{W}_i] \geqslant 0. \tag{19}$$

This inequality also holds if the variable $\Delta p_{ijj'}$ is affected by classical measurement error in prices. Finally, applying the Law of Iterated Expectations, we conclude that

$$\mathbb{E}[d_{ij'} - d_{ij} \exp(-e_{jj'}(z_i, \theta^*)(1 + e_{jj'}(z_i, \theta^*) - (\Delta\kappa_{jj'} + \alpha\Delta p_{ijj'}))|z_i] \geqslant 0. \tag{20}$$

for any $z_i \subseteq \mathcal{W}_i$, proving Theorem 2 in this way.

To gain intuition on why the bounding inequalities provide non-trivial bounds, consider the following specific cases of the general type of inequality introduced in equation (15):

$$\mathbb{E}[d_{i2} - d_{i1} \exp(-e_{i12}(z_i))(1 + e_{i12}(z_i) - (-\theta_{\kappa_2} + \theta_\alpha \Delta p_{i12}))|z_i] \geqslant 0, \tag{21a}$$

$$\mathbb{E}[d_{i1} - d_{i2}\exp(-e_{i21}(z_i))(1 + e_{i21}(z_i) - (\theta_{\kappa_2} + \theta_\alpha \Delta p_{i21}))|z_i] \geqslant 0, \tag{21b}$$

where, for simplicity, we use a function $e_{jj'}(\cdot)$ that is constant in θ. The moments in equations (21a) and (21b) are linearly decreasing and increasing, respectively, in θ_{κ_2}. Thus, given a value of θ_α, equations (21a) and (21b) identify finite upper and lower bounds on θ_{κ_2}, respectively. Theorem 2 guarantees that, if $\theta_\alpha = \alpha$, κ_2 belongs to the interval defined by these bounds.

For any two drugs j and j', using the inequality in equation (15) for estimation requires choosing a function $e_{jj'}(\cdot)$. This choice is consequential for the size of the identified set $\Theta_0^b(e)$ and, as shown in Appendix B.3.1, $\Theta_0^b(e)$ is minimized when, for every pair of drugs j and j',

$$e_{jj'}(z_i, \theta) = e^*_{jj'}(z_i, \theta) \quad \text{with} \quad e^*_{jj'}(z_i, \theta) = \theta_{\kappa_j} - \theta_{\kappa_{j'}} + \theta_\alpha \mathbb{E}[\Delta p_{ijj'}|z_i, d_{ij} = 1]. \tag{22}$$

Furthermore, if $e_{jj'}(\cdot) = e^*_{jj'}(\cdot)$ for every pair of drugs j and j', and the vector z_i is such that

$$\mathbb{E}[p_i|z_i] = \mathbb{E}[p_i|\mathcal{W}_i], \tag{23}$$

the inequalities in equation (15) point identify θ^*. That is, defining $e^* = \{e^*_{jj'}(\cdot)\}_{j=1,j'=1}^{J,J}$, it holds that $\Theta_0^b(e^*) = \theta^*$. We prove this result formally in Appendix B.3.2.

As we show in Section 5, given a vector z_i of observed predictors of the choice characteristic p_i, a maximum likelihood estimator (MLE) that uses $\mathbb{E}[\Delta p_{ijj'}|z_i]$ as a proxy for the unobserved expectation $\mathbb{E}[\Delta p_{ijj'}|\mathcal{W}_i]$ is a consistent estimator of θ^* if and only if equation (23) holds—that is, if and only if the researcher correctly specifies the agent's expectations. The MLE is inconsistent otherwise. The advantage of using the inequalities in equation (15) together with the approximation points in equation (22) is that (a) these inequalities yield an identified set that always contains the true parameter value (see Theorem 2) and, (b) when equation (23) holds for every j and j', the identified set shrinks to include only the true parameter value. Thus, in settings in which the MLE is a consistent estimator of θ^*, using the inequalities in equation (15) instead does not entail a loss of identification power. In settings in which the MLE is not a consistent estimator of θ^*, these inequalities still yield an identified set that includes θ^*.[11]

4.3 Using Inequalities for Estimation

We combine odds-based and bounding moment inequalities for estimation. The inequalities in equations (8) and (15) are defined for every ordered pair of drugs (j, j') and every value

[11] As we show in Section 5, when the MLE is inconsistent, $\Theta_0^b(e)$ may not include the *plim* of the MLE.

of z_i in its support. In our setting, z_i is continuous and, thus, exploiting the information contained in all these inequalities is computationally challenging.[12] Instead, we compute confidence sets for θ^* using a finite number of unconditional moment inequalities implied by the conditional ones in equations (8) and (15). Specifically, for each ordered pair of drugs (j, j'), and each instrument function $g_k \colon \mathcal{Z} \to [0, \infty)$ in a set $\mathcal{G}_K = \{g_k(\cdot)\}_{k=1}^K$, we use the odds-based moment inequality

$$\mathbb{E}[(d_{ij} \exp(-(\theta_{\kappa_j} - \theta_{\kappa_{j'}} + \theta_\alpha \Delta p_{ijj'})) - d_{ij'}) g_k(z_i)] \geqslant 0, \tag{24}$$

and the bounding moment inequality

$$\mathbb{E}[(d_{ij'} - d_{ij} \exp(-e^*_{jj'}(z_i, \theta))(1 + e^*_{jj'}(z_i, \theta) - (\theta_{\kappa_j} - \theta_{\kappa_{j'}} + \theta_\alpha \Delta p_{ijj'})))g_k(z_i)] \geqslant 0, \tag{25}$$

where $e^*_{jj'}(\cdot)$ is defined in equation (22). 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 a vector with J elements: the drug fixed effects $(\kappa_2, \ldots, \kappa_J)$ and the price coefficient α. In our application, z_i is a scalar and every instrument function $g_k(\cdot)$ is an indicator function; we describe in Appendix B.4 the instrument functions we use. In our baseline results, we compute confidence sets for θ^* following the inference procedure for unconditional moment inequalities in Cox and Shi (2023); we describe in Appendix B.5 our implementation of this procedure.

5 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 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 moment 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 moment inequalities are useful for identifying parameters; i.e. neither type of moment is redundant. Third,

[12] Andrews et al. (2022) and Cox and Shi (2023) contain computationally convenient procedures for subvector inference in conditional moment inequality settings in which the nuisance parameters enter linearly and the associated covariates depend only on the instruments. No parameter enters linearly in the moment in equation (8), and all parameters are multiplied by d_{ij} (which does not belong to the vector of instruments) in the moment in equation (15), Those procedures are not applicable in our context.

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

5.1 Simulation Set-up

We simulate data for $i = 1, \ldots, N$ observations, with $N = 4,000,000$, using the model in Section 3. We assume agents choose between three choices, $j = \{1, 2, 3\}$, and we set the choice-specific quality levels 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 need to specify the data generating process of the price vector p_i, and the content of the information set $\mathcal{W}_i$, as these determine the choices of the simulated observations.

For every choice j, we impose the following data generating process for price: $p_{ij} = x_{1ij} + x_{2ij} + x_{3ij}$, with x_{kij} independent of both ε_i and $x_{k'i'j'}$ for $k \neq k'$, $i \neq i'$, or $j \neq j'$, and distributed uniformly with a support that increases in a parameter σ_k.[13]

We impose that the agent's price expectations depend only on x_{1i} and x_{2i}, with $x_{ki} = \{x_{kij}\}_{j=1}^{3}$ for all i and $k = 1, 2, 3$. That is, $\mathcal{W}_i = (\kappa, \alpha, x_{1i}, x_{2i})$ and thus $\mathbb{E}[p_{ij}|\mathcal{W}_i] = x_{1ij} + x_{2ij}$ for $j = 1, 2, 3$ and all i. The agent thus does not have perfect foresight, and x_{3i} represents her expectational error.[14] We also assume that, for each observation, the researcher only observes (d_i, p_i, x_{2i}). Thus, x_{1i} captures variables on which the agent conditions her decision but which the researcher does not observe.

Unless otherwise noted, we compute confidence sets using the inequalities in equations (24) and (25) for all six possible drug pairs and the following two instrument functions:

$$g_1(x_{2i}) = \mathbb{1}\{\Delta x_{2ijj'} \geqslant 0\} \qquad \text{and} \qquad g_2(x_{2i}) = \mathbb{1}\{\Delta x_{2ijj'} < 0\}. \tag{26}$$

with $\Delta x_{2ijj'} = x_{2ij} - x_{2ij'}$. We also report MLEs computed as

$$\underset{(\theta_\alpha, \theta_{\kappa_2}, \theta_{\kappa_3})}{\operatorname{argmax}} \left\{ \sum_{i=1}^{N} \sum_{j=1}^{3} \mathbb{1}\{d_{ij} = 1\} \ln \left(\frac{\exp(\theta_{\kappa_j} + \theta_\alpha x_{2ij})}{\sum_{j'=1}^{3} \exp(\theta_{\kappa_{j'}} + \theta_\alpha x_{2ij'})} \right) \right\}, \qquad \text{with } \theta_{\kappa_1} = 0.$$

Thus, while the inequality estimates correctly assume that x_{2i} belongs to physician i's in-

[13] 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. Thus, choices decline in mean price in order from $j = 1$ to $j = 3$. We fix the length of the support of x_{2ij} to equal 8 (i.e., $\sigma_2 = 4$) and present results for different values of σ_1 and σ_3.

[14] In detail, $\mathbb{E}[p_{ij}|\mathcal{W}_i] = x_{1ij} + x_{2ij} + \mathbb{E}[x_{3ij}|x_{1ij}, x_{2ij}]$, with $\mathbb{E}[x_{3ij}|x_{1i}, x_{2i}] = 0$ because x_{ki} is independent of $x_{k'i}$ for $k \neq k'$ and μ_{3j}, the mean of x_{3j}, equals zero.

formation set, the MLEs impose that $\mathbb{E}[p_{ij}|\mathcal{W}_i] = x_{2ij}$, which is correct only when $\sigma_1 = 0$.

5.2 Simulation Results

We report the main simulation results in Table 2.[15] 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 MLE coincides with the true parameter vector, and the confidence sets defined by the odds-based and the bounding moment inequalities both include only one parameter value, the true one.

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 are 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 the 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., $\sigma_1 > 0$). When the true information set is (x_{1i}, x_{2i}) but the research assumes it includes only x_{2i}, the MLE is asymptotically biased downwards by a scaling factor that decreases in σ_1.[16] In the presence of unobserved elements of the agent's true information set, the confidence sets defined by the odds-based and by the bounding inequalities nonetheless contain the true value, θ^*. However, they also include other values of the parameter vector; the number of additional points included in the confidence sets increases in the importance of the unobserved variable, x_{1i}, in the agent's price expectations. For example, in case 3(a), when σ_1 is small and, thus, the unobserved element x_{1i} contributes little to agents' price expectations, the confidence set we obtain when combining the odds-based and the bounding moment inequalities includes only the true parameter value. In contrast, when σ_1 becomes larger in case 3(b), the resulting confidence set grows and includes points beyond the true parameter value.

In case 4, we consider a setting 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., $\sigma_3 > 0$). This case, which is likely the most empirically relevant setting, combines the complications present separately in cases 2 and

[15] The confidence sets reported in Table 2 are computed following Cox and Shi (2023). In Table C.3 in Appendix C.3.3, we present analogous confidence sets computed following Andrews and Soares (2010).

[16] That is, as σ_1 increases, the parameter estimates move towards zero unless the true parameter value is zero. If the true parameter value is zero, as it is for κ_2, its MLE remains consistent.

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(a)	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]
3(b)	2	0	x_{2i}	MLE	0.75	0	0.75
				Odds-based*	[1 ,1]∪[1.15 , 2.50]	[-1.50 , 1.50]	[-0.50 , 2.50]
				Bounding	[0.50 , 1.50]	[-1 , 1]	[0 , 1.95]
				Both	[1 ,1]∪[1.15 , 1.50]	[-0.15 , 0.15]	[1 , 1.40]
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	∅	∅	∅
				Bounding	[0.87 , 0.87]	[-0.05 , -0.03]	[0.85 , 0.88]
				Both	∅	∅	∅

Note: *MLE* denotes the maximum likelihood estimate. *Odds-based*, *Bounding*, and *Both* contain projections on each parameter of 95% confidence sets computed as in Cox and Shi (2023). *Odds-based* indicates the confidence set is computed using inequalities of the type in equation (24); *Bounding* indicates the confidence set is computed using inequalities of the type in equation (25); *Both* indicates it is computed using both types of inequalities. In cases 1 to 4, we use the instrument functions in equation (26). In case 5, we use the instrument functions $g_1(p_i) = \mathbb{1}\{\Delta p_{ijj'} \geqslant 0\}$ and $g_2(p_i) = \mathbb{1}\{\Delta p_{ijj'} < 0\}$. In all cases other than 3(b), 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). In case 3(b), we use a grid whose sides are $[-0.5, 2.5]$ (for α), $[-1.5, 1.5]$ (for κ_2) and $[-0.5, 2.5]$ (for κ_3). We mark with an asterisk when the confidence set includes points outside the grid.

3. We observe the MLE is asymptotically biased downwards, inheriting the bias present in case 3. The confidence sets defined by odds-based inequalities, bounding inequalities, or by both sets of inequalities together, include the true parameter value, but also other values.

Considering the results from cases 1 through 4 together, we emphasize two additional properties of our inequalities. First, we cannot conclude that the confidence sets defined by the odds-based moment inequalities only, or by the bounding inequalities only, are always weakly smaller than the confidence sets defined by both types of inequalities; in fact, in some settings (e.g., in cases 3(b) and 4), combining the inequalities leads to a strictly smaller set.

Thus, in our empirical setting, we use both types of inequalities together. Second, if $\sigma_1 > 0$ or $\sigma_3 > 0$, the confidence sets defined using only the odds-based inequalities include points outside of the grid, as indicated by the asterisks in Table 2. Conversely, the confidence sets defined using only the bounding moment inequalities are always contained within the boundaries of the grid. As we discuss in Appendix Sections C.1 and C.3, we find this outcome because, if $\sigma_1 > 0$ or $\sigma_3 > 0$, the odds-based moments are globally convex in the parameter θ_α and tend to ∞ as θ_α goes to ∞ or $-\infty$. Conversely, the bounding moments are always globally concave in θ_α and tend to $-\infty$ as θ_α goes to ∞ or $-\infty$.[17]

Finally, in case 5, we consider a setting in which the researcher wrongly assumes that the agent has more information than she possesses. Specifically, we consider the case in which the researcher assumes that the agent has perfect information on all payoff-relevant variables and, consequently, builds the moment inequalities using the following two instrument functions

$$g_1(p_i) = \mathbb{1}\{\Delta p_{ijj'} \geqslant 0\} \qquad \text{and} \qquad g_2(p_i) = \mathbb{1}\{\Delta p_{ijj'} < 0\}, \tag{27}$$

instead of those in equation (26). We compute the maximum likelihood estimates as

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

As shown in Table 2, the confidence sets defined by the odds-based moment inequalities alone, or by both types of inequalities jointly, are empty.[18] 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 1 to be the lowest in our simulation, the downward bias in the price coefficient translates into a downward bias in the choice-specific fixed effects of all other options in the choice set.

Table 2 presents projections of the 95% confidence set for $(\kappa_2, \kappa_3, \alpha)$ on each of three dimensions separately. In Appendix C.2, we show in figures projections of the confidence set on the two-dimensional space (κ_2, α) and on the two-dimensional space (κ_3, α). In Appendix C.3, we present simulation results for additional values of σ_1 and σ_3, and plot the odds-based and bounding moments as a function of θ_α while holding θ_{κ_2} and θ_{κ_3} at their true values.

[17] As a consequence, given real numbers a_1, a_2, a_3 and a_4, the confidence set defined by the odds-based inequalities may be of the form $(\infty, a_1] \cup [a_2, \infty)$ (as in case 2) or of the form $(\infty, a_1] \cup [a_2, a_3] \cup [a_4, \infty)$ (as in cases 3(a) and 3(b)). Conversely, the confidence set defined by the bounding inequalities is always of the form $[a_1, a_2]$. Thus, with a sufficiently large grid, this confidence set will always be included in the grid.

[18] As we show in Table C.3 in Appendix C.3.3, the confidence set defined by the bounding moment inequalities is also empty when it is computed following the procedure in Andrews and Soares (2010). In unreported results, we find it is also empty when computed following the procedure in Cox and Shi (2023) but using four instrument functions instead of the two instrument functions in equation (26).

6 Estimation Results

We now use the model in Section 3 to study a physician's choice of diabetes treatment. In Section 6.1, we present maximum likelihood estimates of the model parameters; computing these estimates requires assumptions on the exact content of the physician's information set. In Section 6.2, we relax these informational assumptions and present estimates that use the moment inequalities described in Section 4. All results we present in this section and subsequent sections cover the class of DPP-4 inhibitors. We also restrict attention to the prescription decisions of primary care physicians and endocrinologists.

6.1 Maximum Likelihood Estimates

We start by computing maximum likelihood estimates of physicians' preference parameters. Following Manski (1991), this estimation approach requires the researcher to first compute measures of the physician's price expectations for every medical visit and drug in the choice set. We compute these measures 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 using a multinomial logit specification. Our specification conditions on drug-specific fixed effects and the expected out-of-pocket costs computed in the first step. Formally, building on the model-implied choice probability in equation (7), we compute

$$\underset{(\theta_\alpha,\theta_{\kappa_2},\theta_{\kappa_3})}{\operatorname{argmax}} \left\{ \sum_{i=1}^{N}\sum_{j=1}^{3} \mathbb{1}\{d_{ij}=1\} \ln\left(\frac{\exp(\theta_{\kappa_j} + \theta_\alpha \widehat{\mathbb{E}}[p_{ij}|z_i])}{\sum_{j'=1}^{J}\exp(\theta_{\kappa_{j'}} + \theta_\alpha \widehat{\mathbb{E}}[p_{ij'}|z_i])} \right) \right\}, \qquad \text{with } \theta_{\kappa_1}=0. \quad (28)$$

Here, z_i is a vector of observed variables assumed to coincide 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 from the first stage linear regression of p_{ij} on z_i. The log-likelihood function in equation (28) has only two fixed effects to estimate because the class of DPP-4 inhibitors includes only three drugs during our sample period. We compute parameter estimates as in equation (28) for different assumptions on the physician's information sets; i.e., for different vectors $\{z_i\}_{i=1}^N$. We report the estimates in Table 3.

Under the assumption of perfect information, we find an estimate of the price coefficient, α, equal to -0.43. When we instead assume providers form expectations using contemporaneous average out-of-pocket costs at the drug-carrier level or drug-plan type level, we find a coefficient equal to -1.21 and -1.45, respectively. If we assume providers form expectations on each patient's out-of-pocket costs using only contemporaneous drug-year price averages, the estimate of the price coefficient equals -2.09. Finally, if we assume that physicians'

Table 3: Estimation Results - MLE

Information Set	α	κ_2	κ_3	Price Elast. (Janumet)
Perfect Information	−0.43	1.40	−0.26	−0.53
	(0.03)	(0.03)	(0.04)	(0.04)
Average Current Prices By Drug-Plan Type-Carrier-Year	−1.02	1.44	−0.30	−1.24
	(0.04)	(0.03)	(0.04)	(0.06)
Average Current Prices By Drug-Carrier-Year	−1.21	1.44	−0.31	−1.47
	(0.05)	(0.03)	(0.04)	(0.06)
Average Current Prices By Drug-Plan Type-Year	−1.45	1.46	−0.17	−1.77
	(0.11)	(0.03)	(0.04)	(0.13)
Average Current Prices By Drug-Year	−2.09	1.51	−0.10	−2.52
	(0.14)	(0.03)	(0.04)	(0.17)
Lagged Prices	−0.67	1.40	−0.23	−0.82
	(0.04)	(0.03)	(0.04)	(0.05)
Average Lagged Prices By Drug-Plan Type-Carrier-Year	−1.04	1.44	−0.21	−1.26
	(0.05)	(0.03)	(0.04)	(0.06)
Average Lagged Prices By Drug-Carrier-Year	−1.27	1.46	−0.20	−1.54
	(0.05)	(0.03)	(0.04)	(0.07)
Average Lagged Prices By Drug-Plan Type-Year	−1.69	1.47	−0.02	−2.03
	(0.12)	(0.03)	(0.04)	(0.15)
Average Lagged Prices By Drug-Year	−3.09	1.50	0.06	−3.77
	(0.22)	(0.03)	(0.04)	(0.27)

Note: Columns labeled α, κ_2 and κ_3 present maximum likelihood estimates of the corresponding parameter computed following equation (28). The column labeled *Information Set* indicates the vector of observed covariates z_i used to build the log-likelihood function in equation (28). 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.

information sets equal these same averages but lagged by a year, the point estimates move further away from zero, reaching a minimum value of −3.09. In our setting, the estimates of the choice-specific fixed effects κ_2 and κ_3 are generally more robust to the specification of the physician's information set.[19]

In sum, we find the maximum likelihood estimate of the price coefficient decreases (in absolute value) when we assume physicians form price expectations using more detailed information. This pattern is consistent with results in Dickstein and Morales (2018). In that setting, when the researcher specifies too large an information set, including variables the agent did not use when forming her expectations, the parameter on the mis-measured expectation is asymptotically biased toward zero. Our simulation results in Table 2 and in Appendix C.3 show similar bias patterns.

As the last column in Table 3 shows, the distinct estimates of κ_2, κ_3 and, especially,

[19]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.

α, 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 elasticity estimates when the researcher assumes the physician forms expectations using coarser information. For example, in the case of the Janumet, the in-sample average elasticity of Janumet's share with respect to its expected price grows in absolute value from -0.53, when we assume physicians have perfect information on prices, to -3.77, when we assume physicians form their price predictions using only last year's average price by drug. The average own-(expected) price elasticities for other drugs in the choice set exhibit similar heterogeneity across models, depending on how we specify the physician's information set.

Moreover, different assumptions on physician information sets also imply substantially different predictions under counterfactual market environments. To illustrate these differences, we consider an intervention in which three carriers negotiate a better acquisition price for Janumet. These three carriers represent 55% of all sample visits. We impose that, as a result of the negotiation, Janumet's out-of-pocket cost falls 50% for all patients enrolled in a plan offered by these carriers. In column 1 of Table 4, we show the counterfactual share that Janumet captures after this price reduction. Depending on the assumed information set, this counterfactual share varies between roughly 21% and 41%. Importantly, regardless of the assumed information set, all estimated models match the initial market share of 17.8%. Thus, the predicted change in Janumet's prescription share ranges from slightly over 3 percentage points under the model that assumes workers have perfect information on prices, to more than 23 percent points under the model that assumes physicians form expectations using only last year's drug-specific average prices.

There are two reasons why models that differ in the assumed information set will generate different predicted counterfactual shares. First, when we estimate our model with different informational assumptions, we find distinct estimates of the physician's preference parameters, as shown in Table 3. Second, different informational assumptions yield distinct changes in Janumet's *expected* price, because our counterfactual price changes filter through the physician's information set into her expectations. For example, if we assume physicians use carrier-specific average drug prices in their expectations, their expected price differs in the counterfactual across patients with different carriers. Columns 2 and 3 in Table 4 illustrate how these two factors influence the counterfactual predictions. From column 3 in particular, we see that the differing predictions by model follow mostly from differences in the estimated preference parameters.

That the parameter estimates, implied elasticities, and counterfactual predictions differ across models illustrates the importance of correctly specifying agents' information sets. One potential way to identify the physician's true information set is to use model selection

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)
Perfect Information	21.24	21.24	21.24
Average Current Prices By Drug-Plan Type-Carrier-Year	26.97	21.21	27.43
Average Current Prices By Drug-Carrier-Year	28.87	21.22	29.73
Average Current Prices By Drug-Plan Type-Year	30.47	20.95	31.29
Average Current Prices By Drug-Year	34.90	20.95	37.68
Lagged Prices	23.61	20.54	22.58
Average Lagged Prices By Drug-Plan Type-Carrier-Year	26.87	20.96	26.88
Average Lagged Prices By Drug-Carrier-Year	28.88	20.86	28.88
Average Lagged Prices By Drug-Plan Type-Year	31.87	20.56	31.62
Average Lagged Prices By Drug-Year	41.38	20.48	45.94

Note: All models reproduce the initial observed market share of Janumet, equal to 17.83%. We compute the counterfactual market share in column 1 using both the maximum likelihood estimates $(\hat{\theta}_\alpha, \hat{\theta}_{\kappa_1}, \hat{\theta}_{\kappa_2})$ and the predicted prices that correspond to the information set indicated in the row label. The counterfactual shares reported in column 2 use the predicted prices that correspond to the information set indicated in the row label, but combine them with the maximum likelihood estimates computed under the assumption of perfect information (i.e., those 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 row label, but use predicted prices that correspond to the assumption that physicians have perfect information on prices.

tests. We implement the testing procedure in Vuong (1989) to compare the models whose estimates we report in Table 3. As we show in Appendix D.1, this procedure selects the model that assumes physicians form price expectations using the contemporaneous average price at the drug-carrier-year level. We note, however, that 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 from such a test may be misleading. We provide an alternative approach to testing information sets in Section 6.2 below.

6.2 Moment Inequality Estimates

We now use the moment inequalities described in Section 4.3 to estimate the parameters of the drug choice model in Section 3. We consider the same collection of potential information

sets that we used with the maximum likelihood approach, as listed in Table 3. Importantly, while the maximum likelihood estimator is consistent only if the researcher's assumed information set coincides with the entire vector of information that the physician uses to form expectations, our moment inequality approach does not require the researcher to specify the entire vector. When using a vector of price predictors within our inequalities, those predictors need only compose a subset of the information the physician uses in her forecast. For each information set we consider, we compute 95% confidence sets implementing the inference procedure in Cox and Shi (2023).

We structure our discussion of the moment inequality estimates to provide two broad insights into diabetes care. First, we show that physicians use relatively aggregate information on out-of-pocket costs when forming expectations about a given patient's price. Second, we show that physicians exhibit greater elasticity with respect to expected out-of-pocket costs than one might conclude based on estimates from full-information models.

To support the notion that physicians use relatively coarse information to form expectations, we combine our inequalities with the same ten sets of potential instruments listed in Table 3. The resulting 95% confidence sets are empty for seven of them. We report in Table 5 the projected confidence sets for the three instruments for which these confidence sets are not empty. Because we test multiple hypotheses, we compute family-wise adjusted p-values following Holm (1979), as described in Appendix B.5.2. In our setting, the p-values for the tests that yield empty 95% confidence sets remain below 5% even after adjusting for family-wise testing. Thus, for all seven variables listed in Table 3 for which the corresponding confidence set is empty, we reject the null hypothesis that the physician uses the corresponding variables to forecast out-of-pocket costs.[20]

In practical terms, in this testing we reject the assumption that all physicians have perfect information on either contemporaneous prices or last year's prices. We also reject that providers, as a whole, know the most detailed averages of out-of-pocket costs. For example, we reject that physicians know the contemporaneous or lagged average copayment by drug, plan type, and carrier when they forecast patient prices. Finally, we also reject that physicians know contemporaneous or lagged average copayments by drug, carrier, and year. This last finding contrasts with the results of the Vuong (1989) tests described in

[20]For each information set we test, we use the same number of instruments and the same instrument functions, as detailed in Appendix B.4. Differences in the value of an instrument functions $g_k(z_i)$ across specifications thus reflect differences in the value of the instrument z_i for each observation i. E.g., when testing whether physicians know carrier-specific prices at the drug and year level, patients with carrier A vs. B will have different values of the instrument. When testing whether they know only drug- and year-specific average prices, patients have the same value of the instrument regardless of their carrier. In this way, the fact that we reject the more specific price averages here is a reflection of the value of the instrument and not the number of instruments or the choice over instrument functions.

Table 5: Estimation Results - Moment Inequalities

Information Set	α	κ_2	κ_3	Price Elast. (Janumet)
Average Current Prices By Drug-Year	$[-4.30, -1.20]$	$[1.40, 1.70]$	$[-0.65, 0.10]$	$[-5.10, -1.40]$
Average Lagged Prices By Drug-Plan Type-Year	$[-3.55, -1.40]$	$[1.40, 1.65]$	$[-0.65, -0.15]$	$[-3.92, -1.53]$
Average Lagged Prices By Drug-Year	$[-4.10, -1.10]$	$[1.40, 1.75]$	$[-0.50, 0.25]$	$[-4.60, -1.20]$

Note: Columns labeled α, κ_2 and κ_3 present projected 95% confidence sets computed using the moment inequalities described in Section 4.3 and the inference procedure in Cox and Shi (2023). The column labeled *Information Set* indicates the vector of observed covariates z_i that we use as instruments in our moment inequalities. The column labeled *Price Elast. (Janumet)* reports the 95% confidence interval for the in-sample average elasticity of Janumet's prescription share with respect to its expected price. Janumet is one of the three DPP-4 inhibitors in our sample.

Section 6.1, which singled out average out-of-pocket costs at the drug-carrier-year level as the preferred information set among the alternatives we tested.

The confidence sets described in Table 5 are similar across the instruments that lead to non-empty confidence sets. For example, when we assume physicians form expectations using an information set that includes last year's average price at the drug-plan type-year level, we obtain a confidence set for the price coefficient, α, between -3.55 and -1.40. When we estimate our model assuming physicians use last year's average price at the drug-year level to form expectations, the confidence set for the parameter α includes values from -4.10 to -1.10.

Our confidence intervals of α provide the second key insight on diabetes care: physicians may be more sensitive to expected out-of-pocket costs than the estimates from full information models suggest. Specifically, for the three information vectors that we fail to reject using our moment inequality approach, we can compare the 95% confidence set for α, reported in Table 5, to the corresponding point estimate in Table 3. When we impose the assumption that an information set forms only a subset of the physician's information rather than the complete set, we find elasticities of market shares with respect to expected prices that can be much higher than the level the corresponding maximum likelihood estimates imply. As an example, the maximum likelihood estimator of α equals -1.69 when we assume the physician forms price expectations using only last year's average prices at the drug-plan type-year level. If we instead assume this same information forms only a subset of the physician's information, the 95% confidence interval ranges from -3.55 to -1.40. In terms of elasticities, the maximum likelihood estimates imply an elasticity of Janumet's market share with respect to its expected price equal to -2.03, while the moment inequality confidence set implies an analogous elasticity between -3.92 and -1.53.

7 Policy Discussion: Informational Intervention

The estimates described in Section 6.2 suggest physicians face substantial information frictions when forecasting patient prices. In this section, we use our model to predict outcomes were policymakers or insurers to provide physicians with perfect information on out-of-pocket costs for each patient they treat. In this counterfactual setting, physicians learn the patient's specific prices for each drug at the point of prescribing, possibly through pop-up messages in their electronic medical record (Desai et al., 2022). As a result of this information intervention, we measure the model-implied change in each drug's market share; the total fraction of treatments chosen that are the cheapest available for a patient; the average per patient realized out-of-pocket costs; and, consumer surplus measured in dollars per patient.

We measure the impact of the informational intervention under each of the three informational vectors that we failed to reject using the testing procedure in Section 6.2; see Table 5. A limitation of our counterfactual analysis relates to the assumptions we now impose on these three information vectors. In estimation, we required only that physicians know *at least* the variables in the specified vector; e.g., last year's average prices by drug. Here, we instead assume that physicians' information set prior to the informational intervention includes *only* this information variable that we failed to reject. That is, for each of the three information sets we consider, we form pre-intervention price expectations by regressing realized prices on a constant and the covariates that we assume form the physician's information set.[21] We then combine these price expectations with the set of parameter values in the relevant confidence set we found with our moment inequality model. For example, we compute the lower limit of the confidence interval for drug j's initial prescription share as

$$\min_{\hat{\theta}\in\hat{\Theta}} \left\{ \sum_{i=1}^{N} \frac{\exp(\hat{\theta}_{\kappa_j} + \hat{\theta}_\alpha \widehat{\mathbb{E}}[p_{ij}|z_i])}{\sum_{j'=1}^{J} \exp(\hat{\theta}_{\kappa_{j'}} + \hat{\theta}_\alpha \widehat{\mathbb{E}}[p_{ij'}|z_i])} \right\}, \qquad \text{with } \hat{\theta}_{\kappa_1} = 0, \tag{29}$$

where $\hat{\theta} = (\hat{\theta}_\alpha, \hat{\theta}_{\kappa_2}, \hat{\theta}_{\kappa_3})$, $\hat{\Theta}$ is the 95% confidence set for $\theta^* = (\alpha, \kappa_2, \kappa_3)$ and, $\widehat{\mathbb{E}}[p_{ij}|z_i]$ is the predicted price for visit i and drug j computed via a regression of p_i on z_i. We compute the upper limit in a similar fashion, replacing the minimization over $\hat{\theta} \in \hat{\Theta}$ with a maximization.

For each drug and baseline information set we consider, Table 6 reports the initial prescription shares as well as the change in these shares that results from the information intervention. The initial prescription shares observed in the data equal 17.8%, 68.7% and 13.5% for Janumet, Januvia, and Tradjenta, respectively. Although the confidence intervals for these shares reported in Table 6 are relatively narrow, they all generally include the ob-

[21] We do not face this limitation when computing the counterfactual shares: because physicians have perfect information on prices in this counterfactual setting, their expected prices are uniquely determined.

Table 6: Effects of an Informational Intervention - Product Market Shares

Information Set	Drug	Initial Share	Change in Share
Average Current Prices By Drug-Year	Janumet	[16.1, 20.3]	[0.9, 5.1]
	Januvia	[65.8, 72.6]	[−10.0, −3.7]
	Tradjenta	[9.0, 15.6]	[2.8, 5.5]
Average Lagged Prices By Drug-Plan Type-Year	Janumet	[16.8, 19.9]	[1.0, 3.9]
	Januvia	[70.3, 74.8]	[−11.9, −5.4]
	Tradjenta	[7.1, 11.1]	[4.3, 8.0]
Average Lagged Prices By Drug-Year	Janumet	[15.2, 18.9]	[1.1, 5.9]
	Januvia	[68.6, 73.9]	[−13.8, −4.1]
	Tradjenta	[8.8, 13.3]	[3.0, 7.9]

Note: The column *Initial Share* contains 95% confidence intervals for each drug's model-predicted share in the sample under the information set specified in the row. The column *Change in Share* contains a 95% confidence interval for the percentage point change in each drug's model-predicted share when changing physicians' information set from the one specified in the row to perfect information. Janumet, Januvia, and Tradjenta are the three DPP-4 Inhibitor products available in our sample period.

served shares. The final column in Table 6 shows the change in prescription shares for each product as we provide physicians full information on patient prices. Across all specifications, we see the shares of Janumet and Tradjenta increase, while the share of Januvia decreases.

We see this change in share precisely because physicians can now form better expectations of patient prices. When physicians only have access to aggregate price information, as in our baseline, these rational physicians nonetheless form expectations that are correct on average. However, given that there exists important variation in patient prices around that average, better information allows physicians to update their expectations to reflect the entire distribution of patient prices. Thus, the shifts in shares for Janumet, Januvia, and Tradjenta reported in the final column in Table 6 partly reflect the relative frequency with which patients' actual prices are below the expected price that generates the baseline shares.

From a policy perspective, the changes in shares reported in Table 6 do not translate immediately into useful measures of patient outcomes. In Table 7, we instead compute more direct measures of the effectiveness of the intervention from the perspective of patients.

First, we show that the share of patients receiving the cheapest drug in *their* choice set increases significantly following the intervention. Across different specifications, we find the share of patients receiving the cheapest drug jumps roughly 11 to 30 percentage points when we provide price information to providers, relative to a baseline of 35%. This shift suggests that the informational frictions that physicians face at baseline are substantial, and that their price elasticity is sufficiently large to generate changes in prescribing behavior as more information becomes available.

Second, we show in Table 7 that, as physicians shift to prescribing cheaper drugs in

Table 7: Effects of an Informational Intervention - Out-of-Pocket Costs and Surplus

Information Set	Change in... Share Cheapest	Mean OOP Costs	Mean Consumer Surplus
Average Current Prices By Drug-Year	$[11.6, 31.0]$	$[-10.1, -5.5]$	$[0.093, 0.237]$
Average Lagged Prices By Drug-Plan Type-Year	$[13.1, 26.0]$	$[-9.7, -6.3]$	$[0.090, 0.191]$
Average Lagged Prices By Drug-Year	$[11.0, 30.6]$	$[-10.5, -5.4]$	$[0.079, 0.231]$

Note: In this table, we report, for three different outcome variables, the effect of changing physicians' information from the set indicated in each row to perfect information. The change in the *Share Cheapest* column reports a confidence interval for the percentage point change in the share of all sample visits in which the physician prescribes the cheapest drug. We compute the *Mean OOP Costs* as a sum across individuals and drugs of the model-implied prescription share multiplied by the corresponding true price, measured in dollars per month. The change in *Mean Consumer Surplus* corresponds to 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.

the patient's choice set, the patient's average out-of-pocket spending falls. We predict the average patient's monthly spending to fall between \$5.5 and \$10.5 from an initial average spending of \$46 per month, or between 12 and 23%. This prediction is similar to experimental evidence from Desai et al. (2022), who found an 11.2% reduction in out-of-pocket costs after an intervention in which physicians in the treatment group received real-time patient cost information during office visits.

Finally, the last column in Table 7 reports the change in consumer surplus under the intervention, expressed in dollars per patient per month. We predict a change in surplus of between \$0.09 and \$0.25 per patient per month. Comparing this change in surplus with the change in average out-of-pocket costs, we see that looking only at changes in cost overestimates the actual welfare gains. Here, the overestimates stem from differential product quality. As we show in Table 6, the information intervention shifts overall demand from Januvia toward Janumet and Tradjenta. However, as we report in Table 5, Januvia, which corresponds to the index $j = 2$ in our model, has the highest effective quality, as proxied by the choice-specific effects $\{\kappa_j\}_{j=1}^3$. Relative to a normalized quality for Janumet, $\kappa_1 = 0$, the 95% projected confidence set for Januvia's quality, κ_2, lies roughly between 1.4 and 1.7. Tradjenta's effect, κ_3, includes mostly negative values. Thus, by providing precise information on prices, physicians may change their prescribing behavior towards lower-quality drugs, which limits the gains in consumer surplus.

8 Testing for Heterogeneity in Information

We compute the moment inequality estimates in Section 6.2 under the assumption that a particular variable belongs to every physician's information set. The results in Table 5 are

thus compatible with the claim that every physician knows average contemporaneous and lagged prices by drug and year, and average lagged prices by drug, plan type, and year. In this section, we examine whether specific subsets of physicians have access to more precise information about prices. We also consider whether heterogeneity in information suggests a benefit from targeting an informational intervention.

We test for heterogeneous information sets as a function of four observable provider characteristics: medical specialty, graduation year, gender, and recent prescribing experience with drugs in our specific diabetes class. For each subgroup of physicians defined according to these observables, we test whether the physicians in the corresponding group know last year's average price at the drug-carrier-plan type-year level. In Section 6.2, we found that the 95% confidence set for the model parameters is empty when we assume that every physician knows and uses this information in forming expectations. Thus, our goal in this analysis is to look for evidence that some categories of physicians know this more specific price average, even when not all of them do. Throughout, we correct our p-values using the family-wise adjustment of Holm (1979) to handle multiple hypothesis testing.

When splitting physicians by gender, and when splitting them into groups of equal size based on their graduation year and prescribing experience, we find that, for all these groups, we reject the null that physicians in the group know last year's average price at the drug-carrier-plan type-year level. However, when splitting physicians by specialty, we fail to reject that endocrinologists know the more specific price averages. Conversely, primary care physicians appear to use only more aggregate price information in their treatment choice.

Given the evidence that endocrinologists form price expectations using more detailed information, we re-evaluate the effect of our informational intervention. In our re-assessment, we assume endocrinologists form expectations in the initial scenario, before the intervention, using the more precise information. We report the results in Table 8.

In panel A, we show that endocrinologists' drug choices are less sensitive to expected prices than the estimates we found when pooling all physicians together. Here, the 95% projected confidence set for an endocrinologist's coefficient on expected price, α, is $[-1.15, -0.50]$; the analogous confidence sets for all physicians, reported in Table 5, included more negative values of α, implying greater elasticity to expected prices.

In panel B, we evaluate the effect of providing perfect information on patient-specific prices to endocrinologists. We implement this evaluation in two settings. First, we assign endocrinologists a baseline information set that includes only lagged average drug prices by drug, plan type, and year. In this setting, we also use the parameter estimates we obtain when pooling all physicians together—i.e., the estimates in the second row of Table 5. We use this scenario to illustrate how the counterfactual predictions under our original assumptions

Table 8: Estimation Results & Informational Intervention - Endocrinologists

Panel A: Confidence Sets for Preference Parameters

Information Set	α	κ_2	κ_3
Average Lagged Prices By Drug-Carrier-Plan Type-Year	$[-1.15, -0.50]$	$[1.25, 1.60]$	$[-0.45, -0.05]$

Panel B: Outcomes of an Informational Intervention

Information Set	Change in... Share Cheapest	Mean OOP Costs	Mean Consumer Surplus
Average Lagged Prices By Drug-Plan Type-Year	$[12.8, 26.9]$	$[-9.6, -6.1]$	$[0.073, 0.168]$
Average Lagged Prices By Drug-Carrier-Plan Type-Year	$[4.3, 10.1]$	$[-4.7, -2.3]$	$[0.014, 0.056]$

Note: In panel A, columns labeled α, κ_2 and κ_3 present projected 95% confidence sets computed using the moment inequalities described in Section 4.3 and the inference procedure in Cox and Shi (2023). The column labeled *Information Set* indicates the vector of observed covariates z_i that we use as an instrument in our moment inequalities. In panel B, 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. 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* corresponds to 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.

apply for the subset of endocrinology visits. In this case, as we show in the first row of panel B, the effect of the intervention for the subset of endocrinologists is similar to the predictions discussed in Section 7 for the average physician.

We next predict the same outcomes in a second setting in which we assume endocrinologists have more precise price information at baseline. We also employ the endocrinology-specific parameter estimates in our predictions, which show relatively lower price sensitivity. As we report in the second row of panel B, endocrinologists now respond to the intervention with a smaller change in behavior. In this scenario, the share of office visits in which the physician chooses the cheapest drug increases by only 4 to 10 percentage points. As a consequence, the average patient's monthly out-of-pocket costs decrease between \$2.3 and \$4.7, and the average consumer surplus increases minimally, from 1 to 6 cents per patient per month.

This example illustrates that heterogeneity in both preference parameters and initial information sets can play an important role in determining the effect of an informational intervention. Here, patients of endocrinologists gained little in terms of out-of-pocket cost savings and consumer surplus relative to patients visiting the wider pool of providers. Our

findings suggest a value from estimating potentially heterogeneous information sets and preferences by subgroups: with these parameters, policymakers can better target interventions toward those visits where physicians are likely to alter their drug choice. This targeting seems particularly useful for interventions like providing price information at the point of prescribing, which may require non-trivial costs to implement.

We emphasize again that our preference measurement could encompass both patient and physician preferences. For example, our finding that endocrinologists are less sensitive to price might reflect their expertise, but might also reflect that their patient population suffers from more severe illness; these patients, and therefore their physicians, might care more about drug effectiveness relative to price. From a policy perspective, however, this distinction does not change the implications of our measurement. Informational interventions that focus on drug prices will nonetheless generate a smaller behavior change when they target endocrinologists, whether that change is driven by physician or by patient preferences.

9 Conclusion

We develop a new moment inequality estimation procedure that allows researchers to estimate preference parameters in discrete choice settings in which the decision-maker must form expectations about a product characteristics—here, price. Our procedure applies in settings with arbitrarily large choice sets. Importantly, our tool requires the researcher to specify only a subset of the information that agents use to form their expectations. This approach contrasts with traditional maximum likelihood approaches, where the researcher must specify the exact information set the agent uses to forecast product characteristics.

We apply our estimation procedure to study the choice of diabetes treatment. Pairing our model with medical claims data, we conclude that physicians do not have perfect information on the out-of-pocket costs that their patients face for each of the drugs in the choice set. Instead, we find 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 suggest physicians may be more sensitive to patient costs than prior full-information models would suggest.

Applying our estimates in a context in which competing insurers map drugs to multiple tiers with distinct out-of-pocket costs, we find an information intervention can steer prescribing patterns towards more cost-effective treatment options. In a counterfactual experiment in which we give all physicians in our sample perfect information on patient- and drug-specific out-of-pocket costs, we find average costs fall 12 to 23% for diabetes patients.

However, the effect of this intervention is smaller when such detailed price information is provided to endocrinologists. The relatively smaller predicted response reflects the finding that these specialists possess more precise information on prices initially, and that they are less sensitive to prices in their prescribing behavior.

Electronic "pop-ups" in the provider's medical chart, as trialed in a single healthcare system in Desai et al. (2022), could thus help steer prescribing towards cheaper drugs, particularly for those physicians least likely to know the patient's true out-of-pocket costs. However, given the pecuniary and non-pecuniary costs of these interventions—in terms of health system dollars and provider time and hassle costs—our evidence suggests a value of targeted academic detailing (Soumerai and Avorn, 1990). Here, sharing price information with less specialized physicians can have an important effect on the costs patients realize, and on overall healthcare spending.

References

American Diabetes Association, "Pharmacologic Approaches to Glycemic Treatment," *Diabetes Care*, 2017, *40* (Supplement 1), S64–S74. [1, 6]

Andrews, Isaiah, Jonathan Roth, and Ariel Pakes, "Inference for Linear Conditional Moment Inequalities," *The Review of Economic Studies*, 2022, *Forthcoming.* [17]

Baicker, Katherine and Dana Goldman, "Patient Cost-Sharing and Healthcare Spending Growth," *Journal of Economic Perspectives*, June 2011, *25* (2), 47–68. [4]

Bombardini, Matilde, Bingjing Li, and Francesco Trebbi, "Did US Politicians Expect the China Shock?," *American Economic Review*, January 2023, *113* (1), 174–209. [5]

Brown, Zach Y. and Jihye Jeon, "Endogenous Information and Simplifying Insurance Choice," *mimeo*, 2023. [5]

Canay, Ivan, Gaston Illanes, and Amilcar Velez, "A User's Guide for Inference in Models Defined by Moment Inequalities," *mimeo*, 2023. [5]

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. [5]

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. [1]

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. [5]

_ **and Masakazu Ishihara**, "Measuring the Informative and Persuasive Roles of Detailing on Prescribing Decisions," *Management Science*, 2012, *58* (7), 1374–1387. [4]

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. [5]

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. [1]

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. [11, 22, 24, 17, 19, 20, 21, 26, 27, 32]

Crawford, Gregory S. and Matthew Shum, "Uncertainty and Learning in Pharmaceutical Demand," *Econometrica*, 2005, *73* (4), 1137–1173. [5]

Currie, Janet, Anran Li, and Molly Schnell, "The Effects of Competition on Physician Prescribing," *mimeo*, 2023. [4]

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*, 11 2022, *182* (11), 1129–1137. [5, 28, 30, 34]

Dickstein, Michael J., "Physician vs. Patient Incentives in Prescription Drug Choice," 2018. mimeo. [4]

_ **and Eduardo Morales**, "What do Exporters Know?," *The Quarterly Journal of Economics*, 07 2018, *133* (4), 1753–1801. [13, 28, 2, 5, 13, 23]

Eizenberg, Alon, "Upstream Innovation and Product Variety in the U.S. Home PC Market," *Review of Economic Studies*, 2014, *81*, 1003–1045. [5]

Fujiwara, Thomas, Charly Porcher, and Eduardo Morales, "A Revealed-Preference Approach to Measuring Information Frictions in Migration," *mimeo*, 2023. [2, 5, 15]

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. [4]

Grennan, Matthew, Kyle Myers, Ashley Swanson, and Aaron Chatterji, "No Free Lunch? Welfare Analysis of Firms Selling Through Expert Intermediaries," *mimeo*, 2021. [4]

Handel, Ben and Jonathan Kolstad, "Health Insurance for "Humans": Information Frictions, Plan Choice, and Consumer Welfare," *American Economic Review*, 2015, *105* (8), 2449–2500. [5]

_ **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. [5]

_ **, Jonathan Kolstad, and Johannes Spinnewijn**, "Information Frictions and the Welfare Consequences of Adverse Selection," *Review of Economics and Statistics*, 2019, *101* (2), 326–340. [5]

Ho, Katherine, "Insurer-Provider Networks in the Medical Care Market," *American Economic Review*, 2009, *99* (1), 393–430. [5]

_ **and Ariel Pakes**, "Hospital Choices, Hospital Prices, and Financial Incentives to Physicians," *American Economic Review*, 2014, *104* (12), 3841–3884. [5]

Holm, Sture, "A Simple Sequentially Rejective Multiple Test Procedure," *Scandinavian Journal of Statistics*, 1979, *6* (2), 65–70. [13, 26, 31]

Holmes, Thomas J., "The Diffusion of Wal-Mart and Economies of Density," *Econometrica*, 2011, *79* (1), 253–302. [5]

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. [5]

Iizuka, Toshiaki, "Physician Agency and Adoption of Generic Pharmaceuticals," *American Economic Review*, 2012, *102* (6), 2826–2858. [4]

Illanes, Gastón, "Switching Costs in Pension Plan Choice," *mimeo*, 2017. [5]

Kaiser Family Foundation, "Employer Health Benefits: 2005 Annual Survey," 2005. [1]

_ , "Employer Health Benefits: 2022 Annual Survey," 2022. [1]

Kline, Brendan and Elie Tamer, "Recent Developments in Partial Identification," *Annual Review of Economics*, 2023, *forthcoming.* [5]

_ **, 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*, 2021. [5]

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. [5]

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. [2, 22]

_ , "Measuring Expectations," *Econometrica*, 2004, *5* (9), 1329–1376. [2]

Morales, Eduardo, Gloria Sheu, and Andrés Zahler, "Extended Gravity," *The Review of Economic Studies*, 2019, *86* (6), 2668–2712. [5]

Pakes, Ariel, "Alternative Models for Moment Inequalities," *Econometrica*, 2010, *78* (6), 1783–1822. [5]

_ **, Jack Porter, Katherine Ho, and Joy Ishii**, "Moment Inequalities and Their Application," *Econometrica*, 2015, *83* (1), 315–334. [5]

Schnell, Molly, "Physician Behavior in the Presence of a Secondary Market: The Case of Prescription Opioids," *mimeo*, 2022. [4]

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. [5]

Soumerai, Stephen B. and Jerry Avorn, "Principles of Educational Outreach ('Academic Detailing') to Improve Clinical Decision Making," *JAMA*, 1990, *263* (4), 549–556. [34]

Vuong, Quang H., "Likelihood Ratio Tests for Model Selection and Non-Nested Hypotheses," *Econometrica*, 1989, *57* (2), 307–333. [29, 30, 25, 26]

Wollmann, Thomas G., "Trucks Without Bailouts: Equilibrium Product Characteristics for Commercial Vehicles," *American Economic Review*, 2018, *108* (6), 1364–1406. [5]