# Screening in Contract Design:

# Evidence from the ACA Health Insurance Exchanges*

Michael Geruso† Timothy Layton‡ Daniel Prinz§

January 12, 2017

**Abstract**

We study insurers' use of prescription drug formularies to screen consumers in the ACA Health Insurance Exchanges. We begin by showing that consumer demand for certain drugs constitutes a strong and exploitable signal of individual-level profitability. Then, using a difference-in-differences strategy comparing Exchange formularies (where these selection incentives exist) to employer formularies (where they do not), we show that Exchange contracts are distorted in a manner consistent with screening. The findings, which demonstrate empirically how and why contracts offered in equilibrium can fail to optimally trade-off risk protection and moral hazard, are important for the continued reform of US healthcare.

*We thank Marika Cabral, Colleen Carey, Jeff Clemens, Leemore Dafny, Kurt Lavetti, Tom McGuire, Mark Shepard, Amanda Starc and seminar and meeting participants at Harvard Medical School, the Caribbean Health Economics Symposium, the United States Congress, and UT Austin for helpful feedback. We gratefully acknowledge financial support from center grants 5 R24 HD042849 and 5 T32 HD007081 awarded to the Population Research Center at the University of Texas at Austin by the Eunice Kennedy Shriver National Institute of Child Health and Human Development (Geruso), a center grant for work on "Selection Incentives in Health Plan Design" awarded by Pfizer (Geruso, Layton, Prinz), financial support from the National Institute of Mental Health (R01-MH094290, T32-019733) (Layton), and additional financial support from the Laura and John Arnold Foundation (Layton). No party had the right to review this paper prior to its circulation.

†University of Texas at Austin and NBER. Email: mike.geruso@gmail.com

‡Harvard Medical School. Email: layton@hcp.med.harvard.edu

§Harvard University. Email: dprinz@g.harvard.edu

# 1 Introduction

The Patient Protection and Affordable Care Act (ACA) of 2010 significantly altered the structure of the individual and small group health insurance markets in the United States. In establishing the new health insurance "Exchanges," the ACA created a system that largely resembles managed competition in Medicare Parts C and D and in health insurance markets throughout the OECD. Two hallmark features of these markets are (1) no consumer can be denied coverage and (2) plans cannot price discriminate based on an individual's health status. This ban against discrimination on pre-existing conditions continues to garner wide bipartisan public support, reflected in consumer polling in the years since the ACA's passage.[1] Indeed, the non-discrimination provisions are so overwhelmingly popular that congressional plans to repeal the ACA often explicitly highlight an intention to maintain protections for consumers with pre-existing conditions.[2]

Enforcing a policy of non-discrimination against the chronically ill can generate improvements in both equity and efficiency (Handel, Hendel and Whinston, 2015). But such reforms may also generate a relationship between non-contractible consumer characteristics and the underlying cost to the insurer of providing coverage. In such settings, two classes of distortions may arise. The first is a price distortion caused by adverse selection of consumers on price, as originally studied by Akerlof (1970).[3] The second—the focus of this paper—is a distortion of insurance contract features like risk protection and multidimensional quality. This type of distortion was first studied by Rothschild and Stiglitz (1976) and more recently applied to the context of modern health insurance by Glazer and McGuire (2000), Frank, Glazer and McGuire (2000), Azevedo and Gottlieb (2016), and Veiga and Weyl (2016). Under this type of distortion, insurers recognize that non-price features of the contract can act as screening mechanisms, inducing consumers to self-sort by profitability. The screening motivation drives a wedge between the contracts offered by insurers in equilibrium and the socially-optimal contract that efficiently trades off risk protection and moral hazard.[4]

[1] For example, a 2012 Reuters poll indicated that 82 percent of Americans favored banning insurance companies from denying coverage to people with pre-existing conditions. A 2014 Kaiser Family Foundation poll indicated that 70 percent of all respondents and 69 percent of Republicans favored protection against preexisting-condition exclusions.

[2] House Speaker Ryan's 2016 "Better Way" proposal emphasizes protecting patients with pre-existing conditions from coverage denials and coverage exclusions. In 2015, the Patient CARE Act proposed by Republican Senators Burr and Hatch, and Republican Representative Upton would have repealed the ACA's individual mandate but prohibited insurance companies from denying coverage or charging higher premiums to people with pre-existing conditions.

[3] For recent empirical applications, see Einav, Finkelstein and Cullen (2010), Handel, Hendel and Whinston (2015), and Hackmann, Kolstad and Kowalski (2015).

[4] Azevedo and Gottlieb (2016) show that this second class of distortion can actually be thought of as a version of the first class, where certain contracts would face complete death spirals if offered, resulting in their non-existence in equilibrium.

Although the theoretical importance of both types of distortions is well-established, empirical evidence has largely focused on price distortions. Distortions in other contract features are more difficult to identify because the dimensionality of the problem can be intractably large. A modern health insurance contract involves thousands of parameters, such as the copay for an in-network specialist visit or the formulary tier assignment of a particular immunosuppressant drug. In addition, screening behaviors do not necessarily push all dimensions of coverage or quality in a particular direction: For some dimensions, coverage may be inefficiently low while for others it may be inefficiently high. This complexity makes it difficult to identify whether an individual product characteristic or a summary measure of coverage generosity is consistent with the kind of socially efficient design that would result from a well-functioning market. For these reasons, the empirical study of contract distortions has been limited, and this limitation has been widely noted in the literature—for example, by Einav, Finkelstein and Levin (2010), Einav and Finkelstein (2011), and Azevedo and Gottlieb (2016). Only a handful of empirical studies have provided econometric evidence that (non-price) contract features respond to selection incentives. This prior empirical work has focused on the Medicare Part D prescription drug insurance market (Carey, 2016; Lavetti and Simon, 2016) and on distortions of hospital networks in the pre-ACA Massachusetts Exchange (Shepard, 2016).[5]

In this paper, we add to the small body of empirical evidence on non-price contract distortions. We examine the design of prescription drug formularies in the context of the individual health insurance markets that were a focus of the ACA. Even setting aside the popular and policy interest in the functioning of these markets, the setting is ideal for investigating the general phenomenon of contract distortions. Pharmaceuticals for managing chronic illness are likely to be among the most price-transparent and predictable medical goods that healthcare consumers encounter. This implies that formulary benefit design—i.e., how plans arrange prescription medication coverage into various cost-sharing tiers—may be particularly salient to consumers, and therefore particularly effective as a screening mechanism. Indeed, a recent case study of HIV drug coverage by Jacobs and Sommers (2015) shows that Exchange plans in several states appear to place an entire class of a commonly-prescribed HIV medications on a high cost-sharing tier, possibly in an attempt to avoid attracting patients with HIV. Such phenomena could be rationalized as a profit-maximizing strategy by firms

[5] A related literature considers insurance coverage distortions due not to selection, but due to the potential for drug and medical spending to offset each other and the feature that some markets separate these kinds of coverage. See Chandra, Gruber and McKnight (2010), Lavetti and Simon (2016), and Starc and Town (2015).

only if these patients were predictably unprofitable in spite of the risk adjustment system intended to neutralize such incentives.

We begin by systematically examining whether prescription drug utilization suggests a plausible screening mechanism for patient profitability. Using a large sample of health claims, we use the Exchange regulator's risk adjustment and reinsurance algorithms to simulate enrollee-specific net revenue. As a first result, we compare these simulated Exchange revenues to the observed costs across patient groups defined by prescription drug consumption. Drugs are partitioned into 220 standard therapeutic classes. We show that although risk adjustment and reinsurance neutralize selection incentives for the majority of drug classes, some classes are associated with consumer types that exhibit significant unprofitability. For example, a consumer taking a drug in the Biological Response Modifiers class is among the most unprofitable in our data. Such a consumer on average will generate $61,000 in claims costs but only $47,000 in net revenue after accounting for the (large) risk adjustment and reinsurance transfer payments to the plan enrolling her. This suggests insurers could potentially screen out these unprofitable types from their plans by placing these drugs on a high cost-sharing specialty tier and by raising the shadow price of drug access in other ways, such as requiring prior authorization before the drug will be covered.

After generating measures of consumer profitability at the drug class level, we ask whether insurers' actual equilibrium contracts appear to reflect these class-specific selection incentives. To do so, we turn to a unique dataset of plan formularies that covers every plan offered in the state and federal Exchanges in 2015, as well as a large sample of employer plans in that year. For both settings, we observe how drugs are arranged across the formulary tiers. The unique, disjointed structure of US healthcare allows us to compare equilibrium Exchange plan formularies (in which selection-related incentives for coverage distortions exist) to plan formularies for self-insured employers (in which these incentives do not exist) operating side-by-side in the same geographic markets. The comparison allows us to difference out some welfare-relevant considerations important in contract design, such as variation in consumer demand elasticities across drug classes (Einav, Finkelstein and Polyakova, 2016), in order to isolate contract distortions aimed at screening. Many of the employer and Exchange plans even utilize the same pharmacy benefits managers—who design the formularies, contract with pharmacies, and negotiate prices—allowing us in some specifications to hold constant unobservable features like the contract designer's institutional knowledge.

Using a difference-in-differences strategy that operates across drug classes and across employer and Exchange plans, we show that Exchange insurers design formularies to be differentially unattractive to unprofitable individuals. These results are not driven by the overall lower coverage generosity of Exchange plans. Instead, the pattern is that within a plan, drug classes used by less profitable consumers appear higher on the formulary tier structure (implying higher out-of-pocket costs) among Exchange plans. The pattern is particularly stark for the tails of the distribution of selection incentives. We find that drug classes in the upper 5% of the selection incentive distribution are about 30 percentage points (70 percent) more likely to be placed on a specialty tier, to face utilization management, or simply to not be covered—relative to the same drugs in employer plans. The associated out-of-pocket financial exposure can be significant. As we show, specialty tier coverage is likely to be governed by coinsurance rates rather than copays, a potential difference of thousands of dollars in annual out of pocket spending per consumer.[6] On the other hand, we show that drug classes that are used by consumers types who are *over*-compensated by the payment system, placing them in the bottom 5% of the selection incentive distribution, are covered relatively generously by Exchange plans.

After presenting our main results, we perform several extensions of our analysis to show that the contract design patterns we document among Exchange plans do not simply reflect insurers passing-through underlying drug costs to the consumer, or of nudging consumers toward lower-cost substitutes within a therapeutic class of alternatives. We show that while insurers do appear to place higher cost drugs on more restrictive tiers, they are sophisticated enough to react not only to overall cost heterogeneity across consumers but also to the *net* incentive generated by revenue heterogeneity embedded in the Exchange payment system. The practical implication is that even *cheap drugs* that are associated with *expensive patients* face high cost sharing or are left off formularies altogether.

We view our paper as filling an important gap in the literature on adverse selection in insurance markets. While several papers, including Frank, Glazer and McGuire (2000), Ellis and McGuire (2007), and Geruso and McGuire (2016), construct measures characterizing selection incentives that vary by service type or setting, only a small recent literature has been able to empirically document insurer *responses* to such incentives. Shepard (2016) investigates network benefit design in response to

[6]For a prescription from a class like Biological Response Modifiers (which we find to be particularly unprofitable) out-of-pocket consumer costs can exceed $1,000 per month in a typical Exchange Silver plan. Such costs could push consumers up to the out-of-pocket annual maximum, which in 2016 was $6,850 for an individual plan and $13,700 for a family plan.

selection incentives. Carey (2016) and Lavetti and Simon (2016) examine the use of formulary design to induce favorable selection in the context of Medicare Part D. Our paper connects most closely to these. Our study is unique in providing evidence on the ACA Exchanges, and the method we introduce is unique in allowing a direct prediction of where the contract distortions should occur, without requiring an intermediate mapping of contract parameters to consumer types. Our result that Exchange plan formularies respond most strongly to incentives to *deter* the tail of unprofitable enrollees provides an interesting contrast to finding of Lavetti and Simon (2016) that Medicare Advantage-Part D plan formularies respond most strongly to incentives to *attract* profitable enrollees. This suggests important differences between the widely studied Medicare Part D market and the understudied individual markets. Our paper is also unique in this small literature in investigating insurers' sophistication in responding to selection incentives. We show that, holding constant patient total costs or patient drug spending, insurers are sophisticated enough to react to the residual profitability implied by the payment system, and that the most popular drugs are differentially targeted. The insights regarding insurer sophistication carry the implication—predicted by theory, but often ignored in policy discussions—that selection incentives, and not merely high upstream pharmaceutical prices, are partly responsible for the high out-of-pocket drug costs faced by US consumers.

Our findings are immediately relevant for the continued evolution of the individual health insurance market in the US, which holds significant policy and popular interest, though the debate has for the most part outpaced research progress. The results here indicate that while the current regulatory framework goes a long way toward weakening insurer incentives to avoid unhealthy enrollees, some selection incentives remain and lead to an equilibrium in which the offered contracts expose consumers to significant drug cost-sharing risk. This issue is important to American consumers: An October 2016 Kaiser Family Foundation poll asked consumers about the top healthcare priorities for the next President and Congress. 74% of respondents agreed that "making sure that high-cost drugs for chronic conditions, such as HIV, hepatitis, mental illness and cancer, are affordable to those who need them" was a top priority. It was the most agreed-to statement in a list that included items like network adequacy, price transparency, cost-sharing subsidies for people with moderate incomes, and repealing the tax penalty for remaining uninsured. Underlining the concern about drug costs and access, the second most agreed-to priority was "government action to lower prescription drug prices." While high out-of-pocket costs for prescription drugs are clearly on the minds of consumers

and policymakers, the screening and selection incentives that we show are partially responsible for this problem are almost never discussed, with most attention directed at manufacturers' prices.

More generally, our findings connect to a broader recent literature investigating the role of private firms in delivering publicly funded or subsidized health benefits (e.g., Curto et al., 2014; Cabral, Geruso and Mahoney, 2014; Duggan, Gruber and Vabson, 2015; Einav, Finkelstein and Polyakova, 2016). The US Medicare and Medicaid programs, the US individual and small group markets, and the national health insurance programs of many members of the OECD have all come to increasingly rely on private insurance carriers to design and manage publicly funded or subsidized health benefits. Private carriers in these settings typically face minimum coverage rules and are prohibited from overt discrimination in the form of differential price-setting or coverage denial. We show that insurers may nonetheless be able to effectively discriminate and induce selection via benefit design. This carries both a distributional implication (the payment system error determines which patients face high cost sharing) and an overall efficiency cost (contracts that optimally balance moral hazard and financial risk protection across categories of services cannot exist in equilibrium). Understanding how this type of backdoor—which has featured prominently in the theory of adverse selection—functions in practice is critical to the continued reform of health insurance markets around the world.

## 2 Background

### 2.1 Conceptual Framework

The theory behind service-level selection in insurance contracts has been carefully developed elsewhere, including in Rothschild and Stiglitz (1976), Frank, Glazer and McGuire (2000), Glazer and McGuire (2000), Ellis and McGuire (2007), Veiga and Weyl (2016), and Azevedo and Gottlieb (2016). Our goal in this section is not to generate new theoretical insights, but merely to adapt some results from the prior literature to guide our empirical analysis. This section provides intuition for how the socially efficient contract, which trades-off the benefits of risk protection against the costs of moral hazard, compares to equilibrium contracts that are likely to arise given the type of selection incentives we document below as being empirically relevant in the ACA Exchange setting.

We start by following much of the prior literature in assuming that insurers offer a single contract that consists of a price $p$ and a coinsurance rate $1 - x$, so that $x \in [0, 1]$ is the portion of spending

paid by the insurer. In our context, this can be thought of as an insurance contract providing partial coverage for spending on one drug.[7] Each individual faces a distribution of potential drug spending with mean $\mu$ and variance $\sigma^2$. We most closely follow Veiga and Weyl (2016) in specifying an individual's expected cost to the insurer as the product of two components: a fixed component $\mu$, and a component $k(x)$ that varies with coverage and incorporates both the *direct* effect of coverage on insurer costs (a smaller $x$ implies that the insurer pays a smaller portion of the cost of the drug) and the *indirect* moral hazard effect (a smaller $x$ induces less consumption of the drug). Formally, $c^j = \mu k(x^j)$ is the expected cost to insurer $j$. We assume that the components are independent so that $k(x)$ does not vary with $\mu$.

Define $v$ as the product of the coefficient of absolute risk aversion and the variance of the spending distribution, $\sigma^2$, so that $v$ is related to the expected utility cost of anticipated risk. Veiga and Weyl (2016) show that under the assumption of CARA utility, willingness-to-pay for coverage $x$ is given by

$$u = \mu h(x) + v\psi(x), \tag{1}$$

where $\mu h(x)$ is the benefit the individual gets from insurer spending equal to $\mu k(x)$, and $v\psi(x)$ is the benefit the individual gets from the level of risk protection offered by the contract.

In this environment, with a distribution of consumer types defined by $f(\mu, v)$, social welfare can be described with the following expression:

$$W = \int_{\mu} \int_{v} f(\mu, v)[\mu h(x) + v\psi(x) - \mu k(x)] dv \, d\mu. \tag{2}$$

The additional term between Equations (1) and (2) is $\mu k(x)$, which captures the cost of coverage, including that due to moral hazard. It is straightforward to show that in order to maximize social welfare, the social planner would set coverage generosity $x^*$ to solve the following equality:

$$\psi'(x^*) = \phi(k'(x^*) - h'(x^*)), \tag{3}$$

where $\phi = \frac{E[\mu]}{E[v]}$. This is the classic trade-off between the benefits of risk protection, $\psi'(x^*)$, and the social cost of moral hazard, $k'(x^*) - h'(x^*)$, as first pointed out by Zeckhauser (1970) and Feldstein

[7] Empirically, we consider contracts with many such cost sharing parameters for many drugs, but the one parameter framework is common in the literature and sufficient to highlight the core intuitions here.

(1973).

We next consider insurer $j$'s choice of $x$ in a competitive health insurance market. We specify insurer $j$'s profit function as

$$\pi^j = \int_\mu \int_v f(\mu, v) D(x^j; \mu, v)[r(x^j, \mu, v) - \mu k(x^j)] dv \, d\mu, \tag{4}$$

where $D(x^j; \mu, v)$ is demand—the probability of enrollment in a plan with coinsurance rate $1 - x^j$ for an individual of type $(\mu, v)$. The term $r(x^j, \mu, v)$ is the payment the plan gets for an individual of type $(\mu, v)$, including risk adjustment, reinsurance, or any other regulatory transfer or payment. As above, $\mu k(x^j)$ denotes the cost of providing insurance.

The insurer sets the portion of spending it covers, $x^j$, to maximize profits. To understand the insurer's problem, we differentiate $\pi^j$ with respect to $x^j$:

$$\frac{\partial \pi^j}{\partial x^j} = \int_\mu \int_v f(\mu, v) \left[ \frac{\partial D(x^j; \mu, v)}{\partial x^j} \left( r(x^j, \mu, v) - \mu k(x^j) \right) \; - \; \mu k'(x^j) D(x^j; \mu, v) \right] dv d\mu. \tag{5}$$

The derivative consists of two components inside the brackets. The first component captures changes in demand (i.e. enrollment) due to a change in the portion of spending covered by the plan, $x^j$. The second component captures the change in plan spending among the existing enrollee population.[8]

The demand effect can be further decomposed to reveal two distinct demand-related consequences of a change in $x^j$. If we define $\bar{r} = E[r(\widehat{x^j}, \mu, v)]$ and $\bar{c} = E[\mu k(\widehat{x^j})]$ as the average net revenue and the average cost (for some fixed value $\widehat{x^j}$ of $x^j$) across the entire population, then:

$$\frac{\partial D(x^j; \mu, v)}{\partial x^j} \left( r(x^j, \mu, v) - \mu k(x^j) \right) = \underbrace{\frac{\partial D(x^j; \mu, v)}{\partial x^j} [\bar{r} - \bar{c}]}_{\text{More enrollees}} + \underbrace{\frac{\partial D(x^j; \mu, v)}{\partial x^j} \left[ \left( r(x^j, \mu, v) - \mu k(x^j) \right) - (\bar{r} - \bar{c}) \right]}_{\text{Different enrollees}}. \tag{6}$$

The "more enrollees" term above represents the change in insurer profits due to a change in the number of individuals of average profitability enrolled in the plan. This arises because consumers' willingness-to-pay for the plan, as described by Equation (1), varies with plan generosity. Importantly, this component is related to the social planner's problem because valuation in excess of cost will increase as $x^j$ converges to the social optimum. The "different enrollees" component reveals that

[8] The change in spending among existing enrollees is due to both the direct effect of the increase in the portion of spending covered by the plan and the indirect effect of the increase in the individual's total spending caused by moral hazard.

the insurer has an additional consideration in setting $x$, beyond trading off risk protection and moral hazard: The plan will attract marginal enrollees who may be differentially profitable to the insurer depending on their specific payments and costs.

Note that if the "different enrollees" term is zero, then the insurer solving the first order condition in Equation (6) under a symmetric competitive equilibrium will decrease the coinsurance rate $(1 - x)$ until the additional profits from enrolling more individuals equals the additional costs due to providing better coverage. This parallels the social planner's problem of trading off the benefits of risk protection with the cost of moral hazard.[9] In fact, Einav, Finkelstein and Polyakova (2016) show via simulation that the social planner's problem and that of the profit-maximizing firm coincide when the "different enrollees" term is zero, with both trading off the social costs and benefits of more generous insurance.

The possibility of screening types by setting the coinsurance rate thus represents a margin that drives a wedge between the level at which a profit-maximizing insurer sets the coinsurance rate and the socially efficient level. Though we merely sketch the intuition here, this result is shown rigorously by Glazer and McGuire (2000), Frank, Glazer and McGuire (2000), and Veiga and Weyl (2016), who also show that the size of the wedge is proportional to the covariance among marginal consumers between willingness-to-pay for coverage and the consumer's cost to the insurer. Ellis and McGuire (2007) devise a practical empirical metric that reflects this covariance, which we follow below when we empirically operationalize the insurer's selection incentive.

Several takeaways here are important for our analysis below: First, the model indicates that insurers should respond to the residual incentive net of the payment system (including risk adjustment and reinsurance), not the gross cost of an individual.[10] Second, the overall profitability of an indi-

[9] To see this, let the demand function be described as $D(x^j; \mu, v) = G(u^j = \mu h(x^j) + v\psi(x^j))$. This implies that $\frac{\partial D(x^j;\mu,v)}{\partial x^j} = G'[\mu h'(x^j) + v\psi'(x^j)]$. It is now straightforward to see that the same expression for the social benefit that enters the social planner's problem $(\mu h'(x^j) + v\psi'(x^j))$ also enters the insurer's profit maximization problem. It is also straightforward to see in Equation (6) above that the same expression for the social cost that enters the social planner's problem $(\mu k'(x^j))$ also enters the insurer's profit maximization problem. While the expressions differ in other ways, there are clear similarities that lead the level of coverage chosen by a profit maximizing insurer to mimic the level chosen by the social planner.

[10] This is true in the setting where all individuals choose a plan and no individual chooses uninsurance, because in that setting all new enrollees a plan acquires when lowering its coinsurance rate come from other plans. If these enrollees come from the uninsurance pool, both the net cost and the gross cost matter. To see this consider the case where a plan enrolls a previously unenrolled individual who has a chronic condition that is included in the risk adjustment formula. The plan's costs clearly increase when it enrolls this individual. The effect on plan revenues is less clear: While the individual will increase the plan's average risk score, he/she will also increase the average risk score in the market. Because the plan's payment is based on the ratio of the average risk score of its enrollees to the average risk score in the market, this implies that even if the risk adjustment formula perfectly captures the additional costs of this individual, the plan will not be fully

vidual to the insurer matters for the distortionary incentive, not just the individual's spending on the particular service (in our case, drug) in a multi-service contract. This means that if an expensive group of consumers uses a cheap drug, an insurer will want to inefficiently distort coverage to be poor for that cheap drug. Third, the extent of the contract distortion should scale with the size of the selection incentive.[11] Fourth, moral hazard, if correlated with the selection incentive, would confound estimates of contract distortions, because as revealed by Equation (6) and as shown by Einav, Finkelstein and Polyakova (2016) moral hazard plays a role in the insurer's decision over how to set $x^j$ independent of the selection motive. These items motivate the details of how we implement our empirical tests below. The moral hazard insight, in particular, motivates an in-depth examination below of whether our measures of the selection incentive correlate with class-specific price elasticities of demand.

Finally, the model makes clear that the welfare loss here does not arise specifically because consumers with chronic diseases have to pay "too much" for their drugs. While that is an important (and as we show, potentially sizable) distributional consequence of poor coverage for certain service types, the welfare loss arises because in equilibrium consumers cannot be adequately insured against the negative shock of transitioning to the poorly-covered chronic disease state.[12] This lack of risk protection affects the utility of *all* consumers with a non-zero probability of acquiring a disease requiring drug treatment, not just consumers who already have such a disease.

### 2.2 Regulatory Framework

The ACA contains several provisions aimed at curbing the use of benefit design as a means of selecting enrollees in the Exchanges. These fall into two broad categories. The first includes coverage mandates that directly constrain insurer benefit design.[13] Under the authority of the ACA, the Department of Health and Human Services (HHS) mandates a variety of essential health benefits (EHB). With respect to formularies, EHB regulations require that Exchange plans cover at least one drug in each therapeutic category and class of the United States Pharmacopeia (USP).[14] However, there is no

compensated for these additional costs. This implies that both the net and gross costs matter in this case.

[11] In the presence of adjustment costs, which Clemens, Gottlieb and Molnár (2015) show to be important in the setting of healthcare contracts, one might expect non-linear responses.

[12] This transition risk parallels the premium reclassification risk discussed by Handel, Hendel and Whinston (2015).

[13] These are in addition to the prohibitions against coverage denial or the use of medical underwriting in setting plan premiums.

[14] In states in which the designated "benchmark" Exchange plan covered more than one drug, plans were were required to cover at least the number of drugs covered by the benchmark plan in each category and class. Andersen (2016) shows

requirement on how such drugs must be tiered within a formulary, which is the primary margin of benefit design we examine in this paper.

The second category of adverse selection-related provisions includes payment adjustments that change the insurer's financial incentives with respect to selection. Whereas coverage mandates may compel insurers to act against their financial interests (e.g., benefit $x$ must be covered, regardless of its effects on profits), the payment adjustments change the insurer's underlying profit function (e.g., covering $x$ is no longer unprofitable). The two important payment adjustments in the ACA Exchanges are risk adjustment and reinsurance.[15]

Risk adjustment, which has become a ubiquitous feature in regulated health insurance markets in the US and much of the OECD, works by implementing a schedule of subsidies or transfers across insurers that are based on the diagnosed chronic health conditions of a particular insurer's enrollees. When functioning properly, risk adjustment makes all potential enrollees appear equally profitable to the plan, removing the incentive for insurers to attempt to cream skim via contract design (van de Ven and Ellis, 2000; Breyer, Bundorf and Pauly, 2011). Regardless of whether states created their own Exchanges or participated in the Federally Facilitated Marketplace, risk adjustment was implemented using the same HHS-HCC risk adjustment system.[16] This model was based on the one used to adjust payments to private Medicare plans in Part C (Medicare Advantage) since 2004.

In addition to mandatory risk adjustment, plans were also required to participate in a mandatory reinsurance program that in 2015 paid out 50% of the individual claims that exceeded an attachment point of $45,000 and fell below a cap of $250,000. The reinsurance operated separately from, and in addition to, the risk adjustment payment. While both sets of payments are based on individual-level characteristics, they were paid at the insurer level. The reinsurance subsidies were funded by health plan fees while the risk adjustment transfers were budget neutral. Risk adjustment transfers to plans

these EHB rules to be a binding constraint.

[15]Temporary risk corridors which insured insurers' overall plan profits were also in place from 2014 to 2016. These operated at the level of the plan, rather than at the level of the enrollee. Their purpose was to protect insurers from risk related to uncertainty around the average health status across the entire market rather than a particular insurer's draw of enrollees within the market.

[16]49 states and Washington, DC used the HHS-HCC system, which consists of a set of 128 payment factors (18 age-by-sex cells, 91 indicators for chronic conditions, and 19 interaction terms capturing interactions between different sets of conditions) and associated payment weights reflecting the incremental cost associated with the factors. The risk adjustment payment weights (or risk adjustment coefficients) were determined by CMS. Massachusetts was the only exception. Massachusetts used a risk adjustment model based on the HHS-HCC system, but estimated its own set of risk adjustment coefficients using claims data from the Massachusetts All-Payer Claims Database and from a subset of MarketScan claims data that was limited to enrollees in New England States. These fairly minor differences are unlikely to affect the implications of the model for individual or group-level profitability.

with sicker than average enrollees were paid for by transfers from plans with healthier than average enrollees. Together, these two payment adjustments altered the underlying financial incentives associated with the composition of a plan's enrollees.[17]

### 2.3 Selection Incentives under the ACA

Risk adjustment and reinsurance systems are generally imperfect, leaving significant shares of enrollee spending "unexplained" by the the transfer payment. The key feature of a well-functioning risk adjustment system is that though it may only explain a small fraction of the variance of healthcare spending, it explains much of the *predictable* variation along which insurers would otherwise be able to induce selection. As we discuss above, and as originally pointed out by Frank, Glazer and McGuire (2000) and Ellis and McGuire (2007) in the healthcare setting, to the extent that risk adjustment and reinsurance leave in place payment "errors" that are correlated with the predictable use of particular services, insurers have an incentive to distort benefits to attract or deter enrollment by consumers seeking coverage for those services. Therefore, the relevant questions are whether the risk adjustment and reinsurance systems of the Exchanges generate payment errors that are correlated with the predictable use of particular health care services, and whether these correlations are significant enough to induce insurers to distort coverage.

There are several reasons to suspect that the Exchange regulatory framework left significant selection incentives as well as sufficient scope for insurers to use formulary design as a tool for avoiding unprofitable patients. First, as early as the first public comment period for the proposed HHS-HCC algorithm now used in the Exchanges, there was concern that the risk adjustment model was not well suited to compensate insurers for the drug costs of their enrollees. Critics noted that the CMS-HCC algorithm on which the HHS-HCC algorithm was based was originally developed to adjust payments for only the non-drug portion of Medicare Part C plans, suggesting a potential problem when applied to drug-inclusive total costs in the Exchanges. Second, since the inception of the Exchanges in 2014, patient advocacy organizations have claimed, and the popular press has reported, that patients with some chronic conditions have faced significant barriers to drug access in Exchange plan formularies.[18] Third, the Centers for Medicare and Medicaid Services (CMS) has suggested

[17] See Centers for Medicare and Medicaid Services (2015) for additional detail on risk adjustment and reinsurance in the first years of the Exchanges.

[18] In 2014 a group of about 350 consumer advocacy groups expressed in an open letter to HHS that consumers with chronic conditions still faced important barriers, in particular in the area of prescription drugs. (http://www.

that by 2018, it will amend the risk adjustment algorithm in the Marketplaces to better capture drug spending, suggesting that drug-related selection incentives are viewed as an important issue by the regulator.[19] Finally, in the context of formulary design in Medicare Part D, both Carey (2016) and Lavetti and Simon (2016) show that in another market with a state-of-the-art risk adjustment system, insurers adjust benefits packages in response to the residual selection incentives. Taken together, there is reason to believe that Exchange insurers may be systematically designing formularies to induce selection. However, in the context of the ACA Exchanges, the prior literature has provided no econometric evidence on the issue.

# 3 Data

## 3.1 Formularies

We use a database from Managed Markets Insight & Technology (MMIT) that contains detailed formulary information for employer sponsored insurance (ESI) plans and plans offered in the ACA Exchanges.[20] The coverage of Exchange plan formularies in these data is remarkably complete: Totalling the enrollment data across the 501 plans in our sample yields 10.2 million covered lives. As a point of comparison, the Department of Health and Human Services reported that 11.7 million consumers selected plans for 2015, with 10.2 million effectuating that enrollment by paying premiums before March 31, 2015. The definition of an Exchange "plan" in this context aggregates the various metal-level products offered by the same carrier in the same market and sharing a formulary. For example, a carrier's gold, silver, and bronze variants on the same benefits package would be counted in our analysis as a single plan, as long as these variants all utilized a common formulary.[21]

The employer plan formulary data represent a large sample, covering about 3,200 plans and 47 million enrollees in self-insured ESI plans in 2015. This amounts to about a third of the universe of ESI enrollees.[22] Our focus on self-insured employers implies that this group does not include plans

theaidsinstitute.org/sites/default/files/attachments/IAmStillEssentialBurwellltr_0.pdf)

[19] "[W]e intend to propose that, beginning for the 2018 benefit year, prescription drug utilization data be incorporated in risk adjustment, as a source of information about individuals' health status and the severity of their conditions." (June 8, 2016 CMS Press Release, https://www.cms.gov/Newsroom/MediaReleaseDatabase/Fact-sheets/2016-Fact-sheets-items/2016-06-08.html)

[20] MMIT collects information on US plan formularies through agreements with insurance carriers, pharmacy benefit managers, pharmaceutical manufacturers and others.

[21] What would differ across such options would be the particular cost sharing (copay and coinsurance) amounts associated with each service and formulary tier. The different levels of cost sharing achieve different actuarial value targets.

[22] External sources, such as the Kaiser Family Foundation, estimate that approximately 150 million consumers were

from the "small group" ACA Exchange markets. For both settings, the data are a snapshot of plans operating in October 2015.

For each drug in each plan, the formulary data indicate the tier in which the drug appears. Drugs are coded at the level of a First Data Bank (FDB) drug identifier code, which is a minor aggregation from the 11-digit National Drug Code (NDC) directory.[23] In addition to a raw tier variable captured in the data, MMIT harmonizes tiering across plans.[24] Additional restrictions and exclusions, such as prior authorization and step therapy are also noted. These data do not provide the dollar cost-sharing amounts associated with each tier, only the tier itself: generic, preferred brand, non-preferred brand, etc.). For our purposes this coding of the data is sufficient, as it naturally aligns with our research design, which examines the relative tiering of various drugs within plans, not level differences in cost-sharing across plans. We also observe the pharmacy benefit manager (PBM) associated with each plan, the geographic coverage of the plan, and the number of beneficiaries covered. The PBM identifier allows us to compare the formularies of employer and Exchange plans that use the same PBM and to therefore hold many unobservables constant.

Table 1 presents summary statistics for the formulary data. Column (1) presents statistics for self-insured employer plans and column (2) presents statistics for Exchange plans. We list tiers from top to bottom in decreasing order of generosity. Drugs in the specialty tier have cost sharing higher than drugs in the covered/non-preferred brand tier, drugs in the covered/non-preferred brand tier have cost sharing higher than drugs in the preferred brand tier, and so on.[25] In order to illustrate the relationship between out-of-pocket consumer spending and tier, we import data made available by the Center for Consumer Information and Insurance Oversight (CCIIO) at CMS. The CCIIO public use files list the cost sharing details for each Exchange insurance product in each state. Whereas the MMIT data describe the mapping from individual drugs to formulary tiers, the CCIIO data describe

enrolled in ESI plans in 2015.

[23] Below, a "drug" means an FDB identifier. On average, an FDB drug identifier corresponds to five 11-digit NDC codes, which specify a labeler, product code, and package code. A "class" means one of the 257 therapeutic classes defined by the RED BOOK, unless otherwise stated.

[24] Plans set up their own formularies with a variety of different tiering structures. MMIT takes these tiering structures and synthesizes them into a unified structure that is common across plans. The unified tiers are generated by specialists who review the basic tiers as well as the specific drugs included in each tier. Among other benefits, the harmonization eliminates the possibility that "tier 1" indicates the lowest level of cost sharing in one formulary but the highest in another.

[25] Ordering of tiers such as "not listed," "medical," and "not covered" is less clear given that the coverage for these tiers is not transparent. Our conversations with the data provider, MMIT, indicated that the ordering in Table 1 is the most likely ordering of tiers by generosity. "Not listed" means that the plan likely covers the drug but they choose not to advertise it, "medical" means that the plan covers the drug but under the medical benefit rather than the drug benefit (likely implying higher cost sharing than the specialty tier), and "not covered" means the plan explicitly states that it will not pay for these drugs.

the mapping between these tiers and dollars of out-of-pocket costs. The two databases are not linkable at the level of individual plans, but CCIIO summary statistics at the level of the tier are presented in columns (3) and (4) of Table 1. Column (3) lists the mean copay associated with each tier among Silver-level Exchange products, conditional on a cost-sharing structure that only includes copays.[26] Column (4) indicates the unconditional probability that the tier faces a coinsurance regime.

The copays increase moving down the table, consistent with our ordering. Comparing copays alone significantly understates the differences in cost sharing across tiers because the probability that the drug is covered by coinsurance, which could generate much higher out-of-pocket costs, is also increasing significantly moving down the table. For expensive drugs, such as those treating multiple sclerosis or rheumatoid arthritis, drug prices may be several thousand dollars per month (Lotvin et al., 2014). For such drugs, consumer coinsurance payments could exceed $1,000 per month if placed on the specialty tier, compared to copayments on the order of $100 per month if placed on the non-preferred brand tier.

About one third of drugs are not listed in a typical plan's formulary. This is an issue not of missing data but of the benefit schedule not specifying to the consumer how each drug in the pharmacological universe is covered. Also, although categories like generic preferred, preferred brand, and specialty have clear vertical rankings, the assignment of some drugs to prior authorization and step therapy represents a qualitatively different type of restrictiveness. These assignments impose non-monetary hurdles to drug access. Prior authorization (PA) requires consumers to obtain special dispensation from the insurer for the drug to be covered, and step therapy (ST) requires patients to first demonstrate that alternative therapies are ineffective before coverage for the drug will be considered. Simon, Tennyson and Hudman (2009) show that the prior authorization and step therapy designations significantly affect access and consumption. For that reason, we group all drugs with a PA/ST designation into a separate, mutually exclusive category.

Not all plans utilize all tiers. For example, some plans do not have a non-preferred brand tier, while others do not have a generic preferred tier. To accommodate this, and to simplify exposition and analysis, we group the tiers into two categories: restrictive and not restrictive. This definition, indicated in Table 1, breaks at the level of the specialty drug tier. The specialty tier is a natural

[26] A significant fraction of Exchange consumers receive a cost sharing subsidy. Such consumers enroll in plan variants that adjust down the overall cost sharing, which often includes reducing the cost sharing associated with the formulary tiers. Such subsidies nonetheless preserve the rank ordering of financial exposure to cost sharing across tiers.

breaking point suggested by plan design, as column (4) of the table shows that plans switch from relatively generous copay-based cost-sharing to relatively ungenerous coinsurance at this tier. The break also reflects consumer complaints and regulator concerns about the use of the specialty tier, in particular, to discriminate against certain chronically ill types. For example, New York has banned the use of specialty tiers by plans in the state. Nonetheless, in our analysis we examine robustness to the choice of which tier defines the cutoff for the restrictive classification.

It is clear from Table 1 that employer and Exchange formularies differ in how they distribute drugs across tiers, with Exchange plans relying more heavily on the restrictive tiers. We illustrate these differences in formulary structure in more detail in Figure 1. In Panel A, we plot plan level histograms of the fraction of each plan's formulary that is placed on the restrictive tier (specialty or higher). In Panel B, we repeat the histogram for the fraction of each plan's formulary that is placed in the PA/ST category or is specifically called out as "not covered" (distinct from not listed). In both panels, it is clear that Exchange plans make much more extensive use of the restrictive tiers.[27] While the differences in ESI and Exchange generosity are important to note, our empirical strategy discussed below controls for differences between ESI and Exchange plans in overall generosity. The results are identified by differences in relative generosity across drug classes *within* plans.

The conceptual motivation in Section 2 suggests that plans will attempt to select against a patient type, rather than narrowly targeting one drug (among several alternatives) used to treat that type. Indeed, narrowly targeting some drugs within a class of potential substitutes is perfectly consistent with efficiently steering patients to more cost-effective options. In contrast, broadly restricting access to an entire therapeutic class of drugs cannot typically be rationalized by steering. To operationalize this idea, we organize prescription medications into therapeutic classes. We follow the standard categorization of therapeutic classes in the RED BOOK, a comprehensive industry drug dictionary. RED BOOK partitions the universe of prescription drugs into 257 mutually exclusive classes. These classes, which are intended to capture sets of substitutes, are the level at which we define the insurer's selection incentive. We measure restrictiveness in each class $c$ as the fraction of drugs in $c$ that are tiered specialty or higher. This is the main outcome variable below, though in some analyses, we limit attention to just the lowest-cost drugs within a class, or just the most popular drugs within a class. In a robustness exercise, we also re-run the analysis using an alternative classification system

[27] Note that the relative tiering is related to, but different from, the implied cost sharing.

designed by the American Hospital Formulary Service.

### 3.2 Claims Costs Data

To quantify the selection incentives implied by the Exchange payment scheme, we use administrative claims data for *non*-Exchange plans from the Truven Health MarketScan Research Database for years 2012 and 2013.[28] The MarketScan data contain inpatient, outpatient, and prescription drug claims from non-Exchange commercial plans. We apply several sample restrictions to the MarketScan data. Because our method, described below in Section 4, requires calculating the intertemporal correlation of spending, we restrict to the most recent sample available for which we can create a panel of total costs and drug utilization: We include consumers who were enrolled for all 12 months in 2013 and for at least 9 months in 2012 and have prescription drug and mental health coverage. We drop patients who had any negative payments or any capitated payments in either the inpatient or the outpatient file. The resulting sample includes 11.7 million consumers generating 143 million drug claims.

For this sample of consumers, we directly observe all information needed to calculate the total of inpatient, outpatient, and prescription drug spending, $C_i$, at the individual level. Also at the individual level, we observe all the information needed to simulate Exchange plan revenues. Patient diagnoses revealed in the claims provide the information necessary to calculate the risk adjustment subsidy $R_i^{RA}$. Total utilization can be used to determine the additional reinsurance payment $R_i^{Re}$, if any, implied by the Exchange regulations. Together, $R_i^{RA}$ and $R_i^{Re}$ describe the total regulatory transfer that would have occurred if each consumer in the non-Exchange claims data had generated their claims history while enrolled in an Exchange plan. These simulated payments are calculated precisely using the publicly-accessible algorithms that are supplied by the regulator for use by the participating plans. See Appendix 8 for details. We denote the total revenue (risk adjustment plus reinsurance plus premiums) as $R_i$.[29] Given $R_i$ and $C_i$ for each individual, we construct various measures of the relative profitability of individual patients.

An important feature of using non-Exchange claims data to measure the Exchange selection incentives is that it allows us to generate out-of-sample predictions for the costliness of patient types

[28] Access provided through the NBER. MarketScan claims data are collected from a selection of large employers, health plans, government, and public organizations.

[29] Premiums are assumed to equal average claims costs, ignoring loading. As Geruso and Layton (2015) show, in a symmetric competitive equilibrium with properly functioning risk adjustment, premiums would equal the market-level average costs.

that are not susceptible to contamination by feedback from the Exchange formulary designs. In other words, we develop measures of costliness and drug utilization in a setting where the utilization is not impacted by the contract distortion we are interested in studying.[30]

# 4 Research Design

We begin in this section by constructing various metrics of the residual selection incentives left in place by the ACA payment system. We then discuss our strategy for identifying the effects of these incentives on contract design.

## 4.1 Selection Incentive Measures

With patient-specific costs, $C_i$, and revenues, $R_i$, it is straightforward to characterize how patient profitability covaries with use of drugs in particular classes. We define $S_{mc}$ as the measure of the selection incentive, which varies across the therapeutic class of drugs, $c$, and the market setting, $m$. A market setting in this notation is employer sponsored insurance (ESI) or an ACA Health Insurance Exchange (HIX).

We begin by calculating the average costs and revenues associated with each class, respectively $\overline{C_c}$ and $\overline{R_c}$. These means are calculated over the set of consumers having non-zero drug consumption in the class and can be constructed only for the subset of therapeutic classes for which we observe claims in the MarketScan data. This removes classes like *toothpastes and floss* and *sunscreen agents* which are typically not covered by health plans. It also removes classes like *mumps*, which are extremely rare. This leaves 220 of the 257 therapeutic classes. For this set, we generate three alternative measures of the class-specific incentive for Exchange plans to distort coverage:

$$S_{HIX,c} = \begin{cases} \overline{C}_c - \overline{R}_c & \text{Cost-revenue difference,} \\ \dfrac{\overline{C}_c}{\overline{R}_c} & \text{Cost-to-revenue ratio,} \\ EM_c & \text{Ellis-McGuire predictable profitability.} \end{cases} \tag{7}$$

[30] In contrast, using data from the Exchange setting where insurers do face this incentive could create spurious correlation between our measure of the selection incentive and the equilibrium response to that incentive via formulary design. To see this point, consider the extreme case where providing *any* coverage for drug A results in a large increase in enrollment among a group of extremely unprofitable individuals. In such a setting, it is likely that no plan will provide coverage for drug A, resulting in low spending on drug A in the data (due to downward sloping demand) and therefore a muted relationship between spending on drug A and profitability.

In all cases, higher positive values of $S_{HIX,c}$ are associated with stronger incentives to inefficiently restrict coverage for the class. The first two measures are self-explanatory. The third measure is based on Ellis and McGuire (2007), who developed a theory of health plan benefit distortions in the presence of selection incentives. Ellis and McGuire (2007) show that a profit-maximizing insurer's incentive to distort coverage is defined by the following index:

$$EM_c = \underbrace{\frac{\sum_{i \in I_c} (\widehat{C_{ic}} - \overline{C_{ic}})^2}{\overline{C_c}}}_{predictability} \times \underbrace{\rho_c}_{predictiveness} . \tag{8}$$

In the first term of (8), predicted spending $\widehat{C_{ic}}$ reflects consumers' ability to forecast drug needs in class $c$ based on past use of drugs in any class. We regress 2013 spending in therapeutic class $c$ on a vector that contains dummies for the quartiles of spending in each of the therapeutic classes in 2012. We then predict 2013 spending in therapeutic class $c$ using the coefficients from this regression. Up to a normalization in the denominator, the predictability term is equivalent to the R-squared of that regression. It captures the extent to which spending in a therapeutic class next period is predictable by a consumer looking backward to his or her past spending (across all drugs). The predictiveness term, $\rho_c$, is defined as the correlation of individual-level profitability ($R_i - C_i$) and spending in therapeutic class $c$ in the same period ($\overline{C}_c$).

Like the other two measures, the Ellis-McGuire (E-M) measure considers the correlation between use of a service (a drug in our context) and profitability. Unlike the other two measures, it also considers the predictability of use of a drug. The intuition is that plans are most likely to distort benefits and services that are both predictive of higher insurer costs, and predictable in the sense that the consumer can anticipate his/her future demand for the drug when selecting a plan. Applied to our setting, drugs that treat chronic conditions are more predictable and thus more vulnerable to contract distortions by insurers aiming to avoid these patients. In contrast, there is little benefit in distorting coverage for a drug class for which consumers cannot anticipate need. For example, a local anesthetic may be an under-compensated drug, but because this would most likely be administered following a traumatic accident that is not predictable, the insurer faces little incentive to inefficiently distort coverage of this drug. A second important difference between the E-M measure and the others is that it effectively weights individuals by their spending on drugs in class $c$, giving more weight to the profitability of individuals with higher utilization of drugs in the class, under the intuition that

these higher-utilization consumers are likely to exhibit stronger demand for plans that offer more generous class-specific coverage. The other two measures effectively weight all individuals taking drugs in the class equally.

All three $S_{HIX,c}$ measures are based on the unconditional effect on plan profits of increasing coverage for a drug in class $c$—not on the partial effects that control for consumers' utilization of drugs in other classes. This is consistent with the model of Frank, Glazer and McGuire (2000) and of Ellis and McGuire (2007) and with the implementation of Lavetti and Simon (2016). The unconditional relationship correctly characterizes the incentives of interest here because it aligns most closely with the thought experiment of using formulary design as a screening mechanism to avoid enrollment by a patient type.[31] Relatedly, our approach captures all drug spending and all medical spending that is predicted by patients' demand for class $c$. This is motivated by the Section 2 discussion that insurers maximize over total profits, not the component of profits narrowly associated with drug costs.[32] Nonetheless, we investigate below the extent to which insurers appear to be unsophisticated in the sense of over-responding to class-specific costs, rather than the bottom line impact on (our proxies for) profits.

Insurers may approximate profit-maximizing behavior in ways that align with any of the three measures defined in (7). We report a main set of results with respect to each variant of $S_{mc}$ separately. Although the measures are correlated, they do contain some independent information. To give a sense of the information overlap, in Appendix Figure A1, we graph rank-rank scatterplots of the measures against each other. The rank correlation of the level and ratio variables is high. Both of these differ non-negligibly from the Ellis-McGuire measure.[33] For parsimony, in some specifications below we report only the two measures with the least overlap: the ratio and Ellis-McGuire measures.

[31] In contrast, the partial effects of drug utilization on spending would more closely align with the thought experiment of reducing costs associated with just one drug, holding enrollment and other drug utilization fixed. For additional intuition, consider two drug classes for which consumer utilization is highly correlated and where one of the two classes has a stronger relationship with profitability. In such a setting, an insurer has an incentive to restrict access to *both* of these drugs because coverage for both drugs affects demand for its plans among these unprofitable groups. The unconditional effects capture these dual incentives, while the conditional effects may not.

[32] For example, a consumer with HIV or MS knows at the time of enrollment that she will demand certain drugs in the coming plan year, and insurers may know that these patient types are expensive in terms of non-drug utilization (even net of risk adjustment).

[33] The axes range from rank 1 to rank 220, with rank 1 implying the strongest incentive to avoid enrollees using drugs in the class. The plots include one point for each of the 220 classes and show how the ordering of profitable and unprofitable classes compares across the measures. Panel A shows a high rank correlation between the level and ratio measures. Panels B and C show that the information content of the Ellis-McGuire measure differs, especially at ranks outside of the top few. Unlike the other two metrics, E-M explicitly accounts for what types of spending are predictable by consumers, and therefore potentially effective as tools for selection.

### 4.2 Regressions and Identification

Estimating the causal impact of screening incentives $S_{mc}$ on benefit generosity $Y_{mc}$ requires holding fixed any characteristics of drugs that could be correlated with $S_{mc}$ and are relevant for contract design for other reasons.[34] For example, consumer price elasticities of demand for the drug classes will impact benefit design because these play an important role in the formulary design problem of both the profit-maximizing insurer and the social planner. If drugs that are more elastically demanded also happen to be under-compensated in the risk adjustment payment scheme, then a profit-maximizing (as well as an efficient) response to moral hazard could be mislabeled as an inefficient selection-driven distortion.

To isolate the impact of selection incentives from other determinants of formulary generosity, we compare formulary design in the Exchange to formulary design in Employer plans. Exchange plans and employer-sponsored plans plausibly face similar considerations with respect to balancing coverage with consumer moral hazard, steering consumers to cost-effective options, and other considerations that could lead to an efficient design. However, the selection incentives differ significantly. Exchange plans can influence their enrollee composition by altering their formularies, but in self-insured employer-sponsored plans, the insurer (the employer) cannot avoid the costly enrollees (i.e. employees) in its firm.

We estimate difference-in-differences regressions of the following form:

$$Y_{mcj} = \beta[S_{HIX,c} \times HIX_j] + \gamma_c + \alpha_j + \epsilon_{mcj}. \tag{9}$$

$HIX_j$ is an indicator equal to one if plan $j$ is an Exchange plan and zero otherwise.[35] $\gamma_c$ are drug class fixed effects, and $\alpha_j$ are plan fixed effects. The parameter of interest in this equation is $\beta$, the correlation between the selection incentive and formulary generosity in Exchange plans after differencing out formulary generosity for the class among ESI plans. In most tables we present OLS estimates of (9), though we additionally present semi-parametric versions in several figures. To facilitate interpretation of $\beta$, in all regressions we standardize $S_{HIX,c}$ by subtracting the mean of the measure and dividing by its standard deviation. This places results for the various operationalizations of the

[34] In practice, generosity varies within class at the level of the drug. However, we are interested in the use of formularies to select classes of patients rather than steering across drugs within a class.

[35] Inclusion of the $HIX_j$ is redundant because $S_{mc}$ is zero for ESI plans. The notation is intended to emphasize that we allow the selection incentive to impact design in HIX plans only.

selection incentive on a comparable (z-score) scale. The estimation sample includes the universe of Exchange plans in 2015 and the large sample of employer plans described in Table 1. Observations are at the plan $\times$ state $\times$ class level. Data are weighted by covered lives within the plan, so that the estimates are representative of the Exchanges nationally for 2015. Standard errors are clustered at the level of the 220 drug classes.

Identification does not require that Exchange and employer plans are equally generous in practice or even that they should be equally generous in terms of socially optimal design. Plan fixed effects in Equation (9) address any differences in overall generosity between Exchange and employer plans, so that $\beta$ is identified by the differential slope $\frac{\partial Y_{mcj}}{\partial S_{mc}}$ within Exchange plans relative to within ESI plans. Our strategy also does not require that ESI plans achieve the social planner's optimum, just that the ACA-Exchange payment formula error does not generate any selection incentive in the self-insured ESI markets, where the ACA payment formula does not apply. Formally, we assume that $S_{mc}$ equals zero in the ESI setting.

The identifying assumption is plausible for several reasons. First, there is essentially no scope for selection by employer-insurers because the self-insured employer is the residual claimant on health care spending for *all* of the plans in an employee's choice set.[36] Second, there is no reason *a priori* to believe the characteristics like drug-specific demand elasticities vary between Exchange consumers and ESI enrollees in a way that would be correlated with the selection incentives generated by the ACA risk adjustment and reinsurance scheme. Although Exchange consumers may be more price-sensitive overall due to lower incomes and this would have implications for the overall level of optimal cost sharing, only differences across ESI and Exchange enrollees in demand elasticities that happened to be correlated with the over-/underpayment of risk adjustment and reinsurance would violate our identifying assumption. A related potential confounder is that even if demand elasticities do not differ importantly between the ESI and HIX market settings, HIX plans may for some reason be more responsive to those elasticities. We investigate this possibility directly, by examining the relationship between $S_{mc}$ and independent estimates of drug class-specific price elasticities of demand from Einav, Finkelstein and Polyakova (2016).

[36] It is possible in principle that employers attempt to induce exit from insurance coverage by employees with expensive conditions, or to offload these employees to a spouse's employer plan. We know of no study documenting such behavior, however.

# 5 Results

## 5.1 Evidence of Payment Error

We begin by showing that Exchange risk adjustment and reinsurance systematically overpay for some patient types and underpay for others. It is important to note that risk adjustment doesn't strive to explain all of the idiosyncratic variance in healthcare spending. Payment "errors" in the sense of payments that deviate from costs are problematic in this context only if they are exploitable for selection, such as being tied to demand for a particular medical good.[37]

We illustrate this idea in Figure 2, where we plot the the mean of total spending among consumers utilizing drugs in a class ($\overline{C_c}$) versus the mean of total simulated revenue among those consumers ($\overline{R_c}$). A dashed line at 45 degrees separates the space into over- and underpayments. Each scatterpoint corresponds to one of the 220 drug classes. Marker sizes reflect the relative number of consumers using drugs in the class. Patients associated with classes above the dashed line are profitable to avoid, because for these patients costs exceed Exchange reinsurance and risk adjustment revenue. In Figure 2 the majority of classes are clustered tightly around the 45-degree line, indicating that the payment system succeeds in neutralizing formulary selection incentives for the majority of potential enrollees. However, there are a small number of significant outliers, far off the diagonal. A few are labelled for illustration. The existence of such outliers establishes that risk adjustment payment "errors" are correlated with drug use, a key necessary condition for insurers to use formularies as screening devices.

How might these errors arise? One possibility, discussed by Carey (2016) in the context of Medicare Part D, is that the technology for treating a particular disease may have evolved after the risk adjustment system was calibrated, changing the association between the diagnoses that enter risk adjustment and patient costliness. Another (non-exclusive) possibility is that, even absent technological change, drug utilization comprises an informative signal of patient severity and cost after conditioning on diagnoses. The phenomenon of selecting patients by severity/costliness conditional on their risk-adjusted reimbursement has been shown to be empirically relevant in the context of physician

[37] Payment errors that are correlated with consumer "type" (geography, demographics, etc.) are also potentially problematic, but for subtly different reasons. The correlation between type and profitability generates incentives to avoid the type, but unless the type differentially uses a particular set of services, the tool of service-level selection or selection via benefit design is not feasible. Instead, these groups may be vulnerable to other forms of selection, such as via selective advertising, where the welfare consequences of selection are less clear. Investigation of these types of selection actions is beyond the scope of this paper.

and hospital coverage in Medicare Part C by Brown et al. (2014). And the specific concern that drug utilization may reveal exploitable severity information has been expressed by the Exchange regulator in discussing potential reforms to the payment system.[38]

Figure 3 provides an alternative view of the selection incentives. Here we plot histograms of the level, ratio, and Ellis-McGuire measures of $S_{mc}$ (without the z-score transformations) for the 220 classes. This class-level variation interacted with an Exchange indicator constitutes our identifying variation. All three panels show that risk adjustment appears to be working reasonably well in the Exchanges, with the majority of drug classes being essentially neutral with respect to selection incentives. In Panel A, the level difference measure is concentrated at zero, in Panel B the spending/revenue ratio is concentrated at one, and in Panel C the Ellis-McGuire measure is concentrated at zero (neutral). However, all three panels also confirm that important outliers exist.

Table 2 presents additional details on costs and revenues for the drug classes associated with the ten most profitable and ten least profitable groups. For this table, we restrict the list to classes that comprise at least 0.01% of drug claims. Column (3) lists the most popular drug in the indicated class, by count of users in our claims data. Column (4) displays the average of total healthcare spending associated with the class, $\overline{C_c}$. Column (5) displays the average simulated revenue, $\overline{R_c}$. A single consumer whose claims span several drug classes will contribute to multiple rows of the table (and to multiple points in Figure 2).

Figure 2 and Table 2 reveal several interesting patterns. Biological Response Modifiers are revealed to be a particularly unprofitable class. A consumer taking a Biological Response Modifier will on average generate $61,000 in claims costs but only $47,000 in net revenue after accounting for risk adjustment and reinsurance transfers. Table 2 shows that the most commonly filled prescription in the Biological Response Modifiers class in our claims data is Copaxone, which is used to treat and prevent relapse of multiple sclerosis (MS). This appears to corroborate external accounts: In November 2015, the National Multiple Sclerosis Society filed a comment with HHS's Office for Civil Rights explaining that "common health insurance practices that can discriminate against people with MS are formularies that place all covered therapies in specialty tiers." In this sense, even without leaning on our difference-in-differences regression framework, and despite relying on predictions made

[38] HHS writes in 45 CFR 153 (December 2016): "Drug utilization patterns can also provide information on the severity of the illness. The hierarchical condition categories (HCCs) already capture information about illness severity from diagnoses, but drugs can potentially measure the severity of illness within a given HCC. A patient may receive first, second, or third lines of treatment involving different medications that indicate increasing levels of severity."

completely out of the Exchange sample (these claims data come from ESI enrollees), the summary statistics here can rationalize the accounts in popular reporting and anecdotes from patient advocacy groups.

Other unprofitable classes in the "top ten" include Opiate Antagonists, which are used to treat opiate addiction, and two classes that treat infertility in women, a condition for which the risk adjustment algorithm does not provide compensation. One of these infertility-related classes, Gonadotropin-Releasing Hormone Antagonists, is called out in Figure 2. As far as we know, the strong selection incentives related to these drugs have not been previously noted. On the other hand, several of the *most profitable* classes in Table 2 treat cardiac conditions. Cardiac conditions were given close attention in Medicare's CMS-HCC risk adjustment algorithm on which the Exchange algorithm was based.[39]

### 5.2 Main Results

We start by illustrating our main results semi-parametrically. Figure 4 shows average generosity in Exchange and ESI plans for each ventile of the distribution of the selection incentive measures. To create the figure, we regress formulary restrictiveness on drug class fixed effects and plan fixed effects and then take averages of the residuals within each ventile of the selection incentive measure. This yields a semi-parametric analog of Equation (9). The left panels use the ratio measure of the selection incentive, and the right panels use the Ellis-McGuire measure. In the top panels, the horizontal axes are scaled to the count of the ventile (1 to 20). In the bottom panels, the horizontal axes reflect the mean value of the selection incentive within the ventile. Each ventile bin contains about 11 drug classes, and each class contains many individual drugs. The dashed lines in each panel correspond to OLS regressions over the scatters, separately for Exchange and ESI plans.

Figure 4 shows that across much of the middle of the distribution of selection incentives, employer and Exchange formulary restrictiveness are similar, though Exchange plans exhibit substantially more noise given the relatively small size of the universe of Exchange plans ($n = 501$). Formulary restrictiveness diverges significantly between employers and Exchanges at the highest ventiles

[39] Interestingly, the Antiviral therapeutic class that includes some HIV medications like nucleoside reverse-transcriptase inhibitors (NRTIs) is not associated with strong selection incentives by our measures. This need not conflict with the findings of Jacobs and Sommers (2015), who document apparent screening behavior around NRTIs in a case study of the formulary designations of these medications in several states. This is because the patient constituency of the RED BOOK-defined Antiviral class is large and diverse, containing many types of drugs beyond NRTIs. Our 220 drug classes are likely too aggregated to detect avoidance incentives associated with HIV-specific drugs that make up a minority of the Antiviral class.

(in the rightmost bins), with the Exchange plans providing much less generous coverage for the least profitable drugs. To put the scatterpoints in context, the 20th ventile, which is a clear outlier along both the horizontal and vertical axes, would include all "top ten" most unprofitable therapeutic classes from Table 2. For the drug classes where risk adjustment is predicted to systematically *overpay* relative to costs (in the leftmost bins), Exchange formularies on average provide relatively more generous coverage. However, it is clear that the largest distortionary incentives and the largest responses to those incentives occur in the direction of unprofitable patient types, which is the focus of most of our attention below.

Table 3 presents regression results corresponding to Equation (9). We report the difference-in-differences coefficient estimates for the interaction between the Exchange dummy, $HIX_j$ and the selection measure, $S_{mc}$. All regressions include plan and drug class fixed effects. The selection incentive variable, $S_{mc}$, varies across columns, as indicated in the column headers. In Panel A the dependent variable is the fraction of drugs within the class placed on the specialty tier or higher. This corresponds to the restrictive tier cutoff indicated in Table 1, and the measure used in Figure 4. In Panel B the dependent variable is the fraction of drugs within the class that require prior authorization or step therapy (PA/ST) or that are explicitly called out on the formulary as "not covered." These specifications explore the non-price hurdles that may be differentially used by Exchange plans. Given the possibility of non-linear effects suggested in Figure 4, we present both linear specifications and specifications that allow the relationship to be non-linear at the top ventile.[40]

Table 3 shows that the interaction between the Exchange indicator and the selection incentive measure(s) always yields a positive and statistically significant coefficient. The signs on the coefficients indicate that Exchange plans tend to provide less generous coverage (placement on a more restrictive tier) for drug classes where stronger selection incentives are generated by the payment system. Coefficients across the linear specifications in Panel A are similar, regardless of which of the three measures is used. The interpretation of the coefficient (0.045) is that a one standard deviation increase in the strength of the selection incentive increases the class-specific drugs assigned to a restrictive tier by about 4.5 percentage points in Exchange plans relative to employer plans. This is a substantial increase relative to a baseline restrictive tier use of 43% in employer plans and 59% in Exchange plans.

[40]These additionally include the regressor $HIX_j \times V^{20}$, where $V^{20}$ is the ventile 20 indicator.

For the difference and Ellis-McGuire measures, the non-linear specifications generate a better fit. The results in column 6 indicate that even controlling for a linear relationship between $S_{mc}$ and restrictiveness, drugs in the top ventile of the selection incentive measures face an additional 69 percent ($= \frac{.296}{.43}$) probability of being placed on a restrictive tier. Column 12 indicates that these same eleven drug classes face almost triple the probability of either being dropped from coverage entirely or of requiring step therapy or the insurer's prior authorization. For completeness, in Appendix Table A1, we report on a wider variety of non-linear specifications, coming closer to the semi-parametric plots in Figure 4.[41]

Another way to put these patterns in context is to note that an Exchange consumer choosing a drug in the top 10 percent of unprofitable drugs (by the ratio measure) would face a restrictive tier 76% of the time, while an ESI consumer would face a restrictive tier 45% of the time. For a drug in the bottom 10 percent, an Exchange consumer would face a restrictive tier 53% of the time, while an ESI consumer would face a restrictive tier 43% of the time.[42] These differences are economically sizable. Recall that the difference between a non-restrictive tier and a restrictive tier is generally associated with the change from copay-based to coinsurance-based cost-sharing (or to no coverage at all). Drugs in unprofitable classes like Biological Response Modifiers can be priced in excess of $4,000 per month (Lotvin et al., 2014). Thus, the out-of-pocket costs associated with even a 20% coinsurance rate would be an order of magnitude larger than copay-based cost sharing and would routinely push such patients to their annual out of pocket maximum.[43]

In summary, across the various parameterizations of the selection incentive and regression specifications, we find that Exchange plans are designing their drug formularies to offer differentially worse coverage for classes used by the most unprofitable individuals, consistent with the hypothesis that Exchange plan formularies are designed as screening devices.

### 5.3 How Sophisticated Do Insurers Appear?

It could be the case that insurers are naively responding to the *gross* profitability of potential enrollees in terms of claims costs, and not actually taking into account the fairly complex risk adjustment pay-

[41] These results show that for the Ellis-McGuire measure, the relationship is driven by the classes with the strongest incentives in both directions: positive coefficients for the top 15% of unprofitable drugs, along with negative coefficients for the 5% of drugs that are most profitable.

[42] These statistics are based on simple means within the sample, they are not derived from regression coefficients.

[43] In 2015 the out-of-pocket annual maximum could not exceed $6,600 for an individual Exchange plan and $13,200 for a family plan, though plans were free to set lower limits.

ments that determine *net* profitability. If patient costs were correlated with our selection incentive measures, then the findings above would still be consistent with insurers attempting to screen enrollees, but the interpretation regarding insurer sophistication would be very different and would lead to different policy responses.

The possibility that insurers intending to screen consumers via formulary design may have done so imperfectly is made more likely by the fact that in 2015 insurers still had little experience with both the population of Exchange enrollees and the risk adjustment system. This might have made it somewhat difficult to predict net profitability by group *ex ante*. In other words, formulary design to deter enrollment by unprofitable individuals is an equilibrium result, and the market we observe may not yet be in equilibrium.

We examine insurer sophistication in Table 4. Columns (1) and (2) repeat results from Table 3 for reference. Columns (3) and (4) additionally control for a naive selection incentive, where risk adjustment and reinsurance are not taken into account. Including this naive selection measure allows that insurers may perceive the drug-specific association with costs, but not the drug-specific association with revenues. Another related possibility is that Exchange insurers are responding not to net or gross profitability but simply to the costs of drugs within the specific class, ignoring the broader signal of overall (non-drug) costs. Columns (5) and (6) include controls for the cost of the drugs underlying the class. Each of the additional controls in Table 4 is interacted with the $HIX$ indicator to allow that Exchange plans may differentially respond to these measures relative to employer plans.

Table 4 suggests that HIX insurers do in fact differentially respond to various naive versions of the selection incentive, as the coefficients on these regressors reveal significance in the expected directions. However, the coefficient of interest on $HIX_j \times S_{mc}$ is robust to the inclusion of these controls. This implies that insurers respond to net profitability in addition to and independently from the gross measures of cost. In column (7) of the table, we simultaneously include the Ratio and E-M measures interacted with $HIX_j$. Both coefficients attenuate relative to specifications that include these regressors separately, but both remain highly significant. This implies that to the extent that the two measures capture different information sets regarding the insurer's selection incentives (see Figure A1), insurers respond to both information sets. In sum, it appears that relative to employer plans, Exchange plans limit coverage of drugs in response to both naive and sophisticated selection incentives.

Another way in which insurers may reveal sophistication is to specifically target drugs that will be most salient in dissuading unprofitable consumers from joining at the time of plan enrollment, while retaining within-class substitutes to encourage proper (and potentially cost-saving) disease management among the population of patients who nonetheless enroll. In Panel A of Table 5 we investigate the possibility that popular drugs within each class are more likely to be differentially relegated to restrictive tiers in Exchange plans when under-compensated by the payment system. In that table, we recalculate the dependent variable—the mean of the restrictive tier indicator within the class—over just the most popular drugs in each class. To do so, we rank each drug within each class according to the frequency of its consumption in the Marketscan data. We then calculate the restrictive tiering variable for only those drugs lying above a cutoff percentile, where the percentiles are weighted by consumption.[44] Columns (1) through (4) of Table 5 present results for the 75th and 90th percentiles of popularity. At both thresholds, coefficients are larger when focusing on the most popular drugs, compared to coefficients applying to the entire class from Table 3. When narrowly focusing on the 90th percentile of popular drugs within each class, the coefficient sizes approach twice the size of the main results. Thus, Exchange plans seem to limit coverage for popular drugs used by unprofitable enrollees more than they do for less popular drugs, though it is unclear whether this reflects insurers responding to salience biases of enrollees, or reflects insurers themselves displaying those same biases in formulary design.

## 6 Efficient Discrimination?

In this section, we provide further evidence that the tiering patterns we document in the Exchanges are not consistent with alternative non-selection-related explanations. Specifically, we show that our results cannot be rationalized by differential responses of Exchange plans to the availability of cost effective substitutes within therapeutic classes or to consumer price sensitivity across classes. We also show that the results are not driven by different pharmacy benefits managers across employer and Exchange markets. We focus here on differential responses of Exchange and employer plans because the inclusion of drug class fixed effects in our main analysis already controls for any similar response to these considerations across the two markets.

[44] For example, to compute the 75th percentile of popularity for a class in which one drug comprises 30% of the consumption share and seven other drugs each comprise 10% shares, the dependent variable would be computed only for the single drug with the 30% share.

### 6.1 Substitution to Cheaper Drugs and Generics

A profit maximizing insurer is incentivized to design its formulary to steer consumers to cheaper substitutes among alternatives with similar efficacy. This behavior is also likely to be socially efficient. Therefore, one potential explanation for our findings is that Exchange plans simply have a stronger interest than ESI plans in operating at the efficient frontier and therefore do a better job of steering patients to lower-cost alternatives within a class. There are two reasons that this is unlikely to explain our results. First, such an explanation would be difficult to motivate under a model in which employers providing ESI are profit maximizing. Such firms would have strong incentives to design an efficient health plan, allowing them to compensate workers with higher wages (Bhattacharya and Bundorf, 2009).[45] Second, there is no *a priori* reason why, even if steering incentives were stronger in the Exchanges overall, HIX plans' interest in steering would be *differentially* strong across classes in a way that is correlated with the error in the HHS risk adjustment and reinsurance scheme. Nonetheless, we can provide some direct evidence that efforts by Exchange insurers to incentivize efficient substitution are not driving our results.

To begin, we note that many of the drugs in classes with the strongest selection incentives have no generic equivalents. For example, the entire class of Biologic Response Modifiers contains not a single generic. In Appendix Table A4, we show that our results hold if we limit attention to classes without generics (28 classes), with less than 10% generics (49 classes), or with less than 25% generics (84 classes). Thus, our results cannot be driven by HIX plans using stronger nudges towards generics, as the results hold in the absence of a generic alternative.

We also show in Appendix Table A3 that our qualitative patterns hold if we look just within the generic drugs of a class or just within the branded drugs of a class. Using the same specification as in the main results (Table 3) but including only generic drugs in the measure of formulary restrictiveness, we show in Panel B of Table A3 that the selection incentive significantly predicts *restricted access to generics*. The way tiers are harmonized across the diverse formularies of our data does not mechanically allow for generics to be allocated to the specialty tier, so this result comes from HIX plans using non-price hurdles to restrict access to generics. This is consistent with supplemental summary statistics we present in Table A2, which show that HIX plans are ten times more likely than employer

[45] The alternative would be to offer an inefficient plan that generated the same utility at a higher cost, leading to lower wages, or to offer a plan at the same cost that generated lower utility.

plans to require prior authorization or step therapy for a generic, and are about twice as likely to not cover a generic on their formulary. For completeness, we also show that additionally controlling for the fraction of generic drugs within each class interacted with HIX does not alter results (Table A5).

Encouraging substitution toward lower cost alternatives may occur along other margins than brand versus generic. To investigate this possibility, in columns (5) through (8) of Table 5, we repeat the main analysis but restrict attention to just the cheapest (generic and branded) drugs within each class, as observed in the Marketscan data. This specification focuses on only the low-cost potential substitutes in each class. Table 5 shows that the results hold up to examining the tiering of only the least expensive 25% or 10% of drugs in each class. Coefficients are similar to the main results, indicating that even relatively cheap drugs that are associated with expensive patients are placed on high cost sharing tiers. Taken together, Tables 5, A2, A3, A4, and A5 support our claim that the contract designs we document do not merely reflect HIX plans pushing consumers to lower cost alternatives.

### 6.2 Tiering and Demand Elasticity

As we highlight in Section 2, moral hazard, reflected in demand elasticities, is a key consideration in the design of a socially-efficient contract (Glazer, Huskamp and McGuire, 2012). Einav, Finkelstein and Polyakova (2016) show that it is also a key consideration in a profit-maximizing insurer's formulary design. The class fixed effects in our regressions are intended to control for any class characteristics that are similar across ESI and Exchange settings, including own and cross-price elasticities. However, if ESI plans were differentially responsive to the same consumer price responsiveness, and if class-specific price elasticities happened to be correlated with class-specific payment errors generated by HHS's risk adjustment and reinsurance algorithms, then the tiering we identify in Exchange plans could be a result of profit maximizing insurers responding to the incentive to efficiently limit moral hazard rather than due to selection-related incentives. In this section, we provide some direct evidence against this possibility by incorporating external measures of consumer price elasticities.

We incorporate the class-specific demand elasticities estimated by Einav, Finkelstein and Polyakova (2016), who identify price sensitivity of prescription drug utilization by exploiting Medicare Part D's "donut hole" at which drug cost-sharing changes abruptly.[46] To map the EFP estimates into our anal-

[46]Econometrically identified estimates exist for own-price elasticities only. Nonetheless, as Glazer, Huskamp and McGuire (2012) show, cross-price responsiveness may matter as well.

ysis, we begin by re-organizing our data to match their therapeutic class grouping, developed by the American Hospital Formulary Service (AHFS).[47] Besides allowing us to import the EFP demand elasticities, this exercise demonstrates the robustness of our results to an alternative classification system.

Figure A2 plots the analog of Figure 2, using the 294 AHFS drug classes in place of the 220 REDBOOK classes used in the main analysis. As above, marker sizes reflect the relative number of consumers using drugs in each class, and the dashed line separates the space into profitable and unprofitable types. In the figure, a subset of the classes are indicated with blue markers. These are the 99 classes for which EFP generate demand elasticity estimates that we can match to our data.[48] For the whole sample of classes and for this demand elasticity subset in blue, there are significant outliers above the dashed line, mirroring Figure 2.

In Table A6, we replicate the main results using the AHFS classification. We generate our selection incentive measures exactly as above. In column (1) we include the full schedule of AHFS drug classes. In column (2) we restrict to only those classes for which we can directly control for a demand elasticity. In column (3) we add controls for the EFP estimate of class-specific elasticity interacted with the indicator for an Exchange plan. (The elasticity main effects are naturally absorbed by the class fixed effects.) We repeat this ordering of specifications for each of the three selection incentive measures and for both of the dependent variables from Table 3. The findings of Table A6 mirror those of Table 3 in that unprofitable classes are differentially assigned to restrictive tiers in Exchange plans. Most importantly, the addition of demand elasticity controls have essentially no effect on the coefficient estimates of interest. For completeness, Appendix Figure A3 plots the semi-parametric versions of the regressions.[49]

To better understand these results, we examine the correlation between the demand elasticity estimates and the selection incentive measures. Figure 5 graphs scatterplots of elasticity versus selection incentive by class. The three panels correspond to the three measures of $S_{mc}$. There is no significant correlation between the selection incentive generated by the payment system error and the demand elasticity. Taken together, Table A6 and Figures A2, 5, and A3 provide strong evidence

[47] For more information on differences across the classifications see Appendix 9.

[48] Einav, Finkelstein and Polyakova (2016) generate demand elasticities for 108 AHFS classes. We can match these classes and generate our selection incentive and tiering variables for 99 of these.

[49] The specifications using the Ellis-McGuire measures do not produce significant effects under the linear specification shown. Like the main results, however, there are significant non-linear effects for the E-M measure, concentrated among the most unprofitable classes.

that Exchange plans are not merely differentially responding to socially efficient profit-maximizing considerations regarding class-specific consumer moral hazard in a way that ESI plans are failing to do.

Finally, we explore sensitivity to excluding fertility-related classes. Table 2 showed that two of the ten classes associated with the least profitable patients were associated with infertility, a class for which one might expect especially high price sensitivity. To demonstrate that these particular classes are not driving the results, we re-estimate our main regressions excluding all fertility treatment classes. Results are reported in Appendix Table A7, and are almost numerically identical to our main results.

### 6.3 Contracting and Institutional Knowledge

Another possible explanation for our results is that the prices paid by insurers to drug manufacturers differ between Exchange plans and employer plans due to differences in plans' contracting with manufacturers.[50] To address this possibility to the extent possible in our data, we exploit the fact that essentially all insurers outsource price negotiations with drug manufacturers to a fairly small set of pharmacy benefit managers (PBMs). PBMs design the formularies, contract with pharmacies, and negotiate prices. In our data, we observe the PBM used by each plan, allowing us to construct a full set of PBM fixed effects. Let $\mathbf{1}(PBM_p)$ be an indicator equal to 1 if plan $j$ uses PBM $p$ and zero otherwise. We estimate a set of specifications where we interact the selection incentive with the PBM fixed effects ($\mathbf{1}(PBM_p) \times S_{mc}$) when estimating our coefficient of interest ($\beta$):[51]

$$Y_{mcj} = \beta[S_{HIX,c} \times HIX_j] + \sum_p \nu_p[\mathbf{1}(PBM_p) \times S_{HIX,c}] + \gamma_c + \alpha_j + \epsilon_{mcj} \tag{10}$$

Under this specification, $\beta$ is identified off of differences between Exchange plans and employer plans that use the same pharmacy benefits manager.

Table A8 displays these results, again separately for each selection measure. We present two versions. In columns (1) through (4), we estimate Equation (10) such that the $PBM_p$ variable is defined nationally. This implicitly compares, for example, Aetna's Exchange plans in New Jersey to Aetna's employer plans in New Jersey and elsewhere. In columns (5) through (8), we define $PBM_p$ at the state

[50] If upstream prices differ, then both profit-maximizing and (second-best) optimal consumer prices reflected in tiers, may also differ, following the intuition of Section 2.

[51] PBM main effects are absorbed by plan fixed effects.

level, so that the control is defined as $[\mathbf{1}(PBM_p) \times \text{state}_s \times S_{HIX,c}]$. Intuitively, in these specifications we are comparing reactions to the selection incentive in, for example, employer plans in Texas that contract with OptumRx (a PBM associated with United) to Exchange plans in Texas that contract with OptumRx. In all cases the results in Table A8 are robust to these controls, lending further support to our identifying assumption. These regressions address not only the bargaining power confounder, but provide additional evidence that the effect is not driven by responses to (or biased subjective beliefs about) consumer moral hazard that differ across insurance carriers.

# 7 Discussion

Our empirical findings have several important implications for the continued evolution of individual health insurance markets in the US. First, our results bear directly on a top concern among American consumers—high out-of-pocket prices for prescription drugs. High consumer prices are almost always assumed to be caused by upstream manufacturer prices. Our results confirm a clear, but often ignored, theoretical prediction: It is unprofitable patients rather than expensive patients (or patients that use expensive drugs) that will be the targets for high out-of-pocket costs by insurers. In a market setting with risk adjustment, claims costs and profitability can be decoupled. If the payment system were to generously compensate insurers who enrolled patients consuming expensive drugs, then in equilibrium, such patients could bear low out-of-pocket costs regardless of upstream drug prices. Upstream prices would be important in determining premiums, but their effect on the financial risk associated with transitioning to a health state that requires a particularly expensive treatment would be less extreme.

Second, we show that even in the face of coverage mandates such as Essential Health Benefits rules, insurers are able to effectively discriminate. Plans have many tools at their disposal to limit coverage, even while nominally complying with minimum coverage requirements. Our findings suggest that simply "strengthening" the list of mandated benefits will not solve the problems documented here. For example, if regulators restricted insurers' flexibility in setting cost-sharing—a popular proposal in the context of high patient drug costs—then plans could respond by relying more heavily on non-price barriers to access, such as step-therapy and prior authorization. If regulators then restricted the use of tools like prior authorization, plans could generate hurdles that were effectively invisible to the regulator, such as requiring consumers to use in-house mail-order pharmacies for particular

drugs and then making it very to difficult to work with those pharmacies.[52] Aside from plans' ability to partially avoid the spirit of such regulations, excessive minimum coverage requirements can have the negative welfare consequence of limiting the insurer's ability to optimally set coverage that trades off risk protection and moral hazard, which is a welfare-relevant tension in socially optimal insurance design (Pauly, 1968, 1974; Zeckhauser, 1970).

Third, the results here connect to the ongoing debate over reforms to individual markets in the US. Republican plans to repeal and replace the ACA often highlight a clear intention to protect patients with pre-existing chronic conditions, but these plans also favor reducing the regulatory interventions aimed at addressing selection and favor weakening minimum coverage requirements. Our findings suggest that weakening minimum coverage need not necessarily expose chronically-ill consumers to significant financial risk, but only so long as strong selection-related regulations, including risk adjustment and possibly reinsurance, are in place. The patterns of the selection-related contract distortions we document cast serious doubt on the notion that a less highly regulated payment system could effectively protect consumers with pre-existing conditions from discrimination.

Any comprehensive solution to implicit discrimination on health status must address the underlying perverse financial incentives. One possibility to address the distortions documented here is to refine the risk adjustment system to directly incorporate limited drug utilization information, including possibly interactions between drug utilization and medical diagnoses. HHS has suggested that by 2018, it will amend the risk adjustment algorithm used in the individual market in this way (Centers for Medicare and Medicaid Services, 2016). Incorporating drug utilization into the risk adjustment system would be novel in the US, and given experience with diagnosis-based risk adjustment, any such reform should be conducted with careful attention to game-ability by insurers, which has been shown to be an empirically relevant phenomenon (Geruso and Layton, 2015).[53] Nonetheless, because drug utilization actually requires real-world action by a patient at a pharmacy, rather than merely a paperwork edit by a physician's billing staff or by an insurer, it is possible that for some drug-class $\times$ diagnosis interactions, gaming may be more difficult than it is under current diagnosis-only risk adjustment.

[52] In a Congressional briefing on November 17, 2016, chronically ill patients reported on these types of hurdles. Patients also report that plans routinely deny claims for certain drugs even *after* granting prior authorization, causing significant financial uncertainty for patients while the claims process is resolved.

[53] Geruso and Layton (2015) show that patients' reported diagnoses are endogenous to the plan in which they are enrolled in the context of Medicare Advantage.

Issues surrounding formulary design-for-selection are likely to continue to be important in any future form of the individual markets, in Medicare Part D, and elsewhere, as long as there is scope for profitable selection. This is because drug costs for managing chronic illness are particularly transparent and predictable to consumers relative to other healthcare goods, making poor coverage for a drug particularly attractive as means of screening consumers. Finally, we emphasize that the cost of these contract distortions is not solely borne by patients with the targeted chronic conditions. As the long theoretical literature on service-level selection has noted, the incidence of this distortion falls on all consumers because contracts that optimally balance moral hazard and financial risk protection across categories of services cannot exist in equilibrium, and consumers may be left exposed to the financial risk of transitioning to a costly health state (Handel, Hendel and Whinston, 2015).

## References


**Akerlof, George A.** 1970. "The Market for "Lemons": Quality Uncertainty and the Market Mechanism." *Quarterly Journal of Economics*, 84(3): 488–500.

**Andersen, Martin.** 2016. "Constraints on Formulary Design Under the Affordable Care Act." University of North Carolina at Greensboro.

**Azevedo, Eduardo, and Daniel Gottlieb.** 2016. "Perfect Competition in Markets with Adverse Selection." *Econometrica*, Forthcoming.

**Bhattacharya, Jay, and M. Kate Bundorf.** 2009. "The Incidence of the Healthcare Costs of Obesity." *Journal of Health Economics*, 28(3): 649–658.

**Breyer, Friedrich, M. Kate Bundorf, and Mark V. Pauly.** 2011. "Health Care Spending Risk, Health Insurance, and Payment to Health Plans." In *Handbook of Health Economics*. Vol. 2, , ed. Mark V. Pauly, Thomas G. McGuire and Pedro P. Barros, 691–762. Elsevier.

**Brown, Jason, Mark Duggan, Ilyana Kuziemko, and William Woolston.** 2014. "How does risk selection respond to risk adjustment? Evidence from the Medicare Advantage program." *American Economic Review*, 104(10): 3335–3364.

**Cabral, Marika, Michael Geruso, and Neale Mahoney.** 2014. "Does Privatized Health Insurance Benefit Patients or Producers? Evidence from Medicare Advantage." National Bureau of Economic Research Working Paper 20470.

**Carey, Colleen.** 2016. "Technological Change and Risk Adjustment: Benefit Design Incentives in Medicare Part D." *American Economic Journal: Economic Policy*, Forthcoming.

**Centers for Medicare and Medicaid Services.** 2015. "Summary Report on Transitional Reinsurance Payments and Permanent Risk Adjustment Transfers for the 2014 Benefit Year." Department of Health and Human Services.

**Centers for Medicare and Medicaid Services.** 2016. "March 31, 2016, HHS-Operated Risk Adjustment Methodology Meeting Discussion Paper." Department of Health and Human Services.

**Chandra, Amitabh, Jonathan Gruber, and Robin McKnight.** 2010. "Patient Cost-sharing and Hospitalization Offsets in the Elderly." *American Economic Review*, 100(1): 193–213.

**Clemens, Jeffrey, Joshua D. Gottlieb, and Tímea Laura Molnár.** 2015. "The Anatomy of Physician Payments: Contracting Subject to Complexity." National Bureau of Economic Research Working Paper 21642.

**Curto, Vilsa, Liran Einav, Jonathan Levin, and Jay Bhattacharya.** 2014. "Can Health Insurance Competition Work? Evidence from Medicare Advantage." National Bureau of Economic Research Working Paper 20818.

**Duggan, Mark, Jonathan Gruber, and Boris Vabson.** 2015. "The Efficiency Consequences of Health Care Privatization: Evidence from Medicare Advantage Exits." National Bureau of Economic Research Working Paper 21650.

**Einav, Liran, Amy Finkelstein, and Jonathan Levin.** 2010. "Beyond Testing: Empirical Models of Insurance Markets." *Annual Review of Economics*, 2(1): 311–336.

**Einav, Liran, Amy Finkelstein, and Maria Polyakova.** 2016. "Private Provision of Social Insurance: Drug-specific Price Elasticities and Cost Sharing in Medicare Part D." National Bureau of Economic Research Working Paper 22277.

**Einav, Liran, Amy Finkelstein, and Mark R. Cullen.** 2010. "Estimating Welfare in Insurance Markets Using Variation in Prices." *Quarterly Journal of Economics*, 125(3): 877–921.

**Einav, Liran, and Amy Finkelstein.** 2011. "Selection in Insurance Markets: Theory and Empirics in Pictures." *Journal of Economic Perspectives*, 25(1): 115–138.

**Ellis, Randall P., and Thomas G. McGuire.** 2007. "Predictability and Predictiveness in Health Care Spending." *Journal of Health Economics*, 26(1): 25–48.

**Feldstein, Martin.** 1973. "The Welfare Loss of Excess Health Insurance." *Journal of Political Economy*, 81(2): 251–280.

**Frank, Richard G., Jacob Glazer, and Thomas G. McGuire.** 2000. "Measuring Adverse Selection in Managed Health Care." *Journal of Health Economics*, 19(6): 829–854.

**Geruso, Michael, and Thomas G. McGuire.** 2016. "Tradeoffs in the Design of Health Plan Payment Systems: Fit, Power and Balance." *Journal of Health Economics*, 47(1): 1–19.

**Geruso, Michael, and Timothy Layton.** 2015. "Upcoding: Evidence from Medicare on Squishy Risk Adjustment." National Bureau of Economic Research Working Paper 21222.

**Glazer, Jacob, and Thomas G. McGuire.** 2000. "Optimal Risk Adjustment in Markets with Adverse Selection: An Application to Managed Care." *American Economic Review*, 90(4): 1055–1071.

**Glazer, Jacob, Haiden A Huskamp, and Thomas G McGuire.** 2012. "A Prescription for Drug Formulary Evaluation: An Application of Price Indexes." *Forum for Health Economics & Policy*, 15(2): 1558–95.

**Hackmann, Martin B., Jonathan T. Kolstad, and Amanda E. Kowalski.** 2015. "Adverse Selection and an Individual Mandate: When Theory Meets Practice." *American Economic Review*, 105(3): 1030–1066.

**Handel, Benjamin R., Igal Hendel, and Michael D. Whinston.** 2015. "Equilibria in Health Exchanges: Adverse Selection vs. Reclassification Risk." *Econometrica*, 83(4): 1261–1313.

**Jacobs, Douglas B., and Benjamin D. Sommers.** 2015. "Using Drugs to Discriminate? Adverse Selection in the Insurance Marketplace." *New England Journal of Medicine*, 372(5): 399–402.

**Kautter, John, Gregory C. Pope, Melvin Ingber, Sara Freeman, Lindsey Patterson, Michael Cohen, and Patricia Keenan.** 2014. "The HHS-HCC Risk Adjustment Model for Individual and Small Group Markets under the Affordable Care Act." *Medicare & Medicaid Research Review*, 4(3): E1–E46.

**Lavetti, Kurt, and Kosali Simon.** 2016. "Strategic Formulary Design in Medicare Part D Plans." National Bureau of Economic Research Working Paper 22338.

**Lotvin, Alan M., William H. Shrank, Surya C. Singh, Benjamin P. Falit, and Troyen A. Brennan.** 2014. "Specialty Medications: Traditional and Novel Tools Can Address Rising Spending on These Costly Drugs." *Health Affairs*, 33(10): 1736–1744.

**Pauly, Mark V.** 1968. "The Economics of Moral Hazard: Comment." *American Economic Review*, 58(3): 531–537.

**Pauly, Mark V.** 1974. "Overinsurance and Public Provision of Insurance: The Roles of Moral Hazard and Adverse Selection." *Quarterly Journal of Economics*, 88(1): 44–62.

**Rothschild, Michael, and Joseph Stiglitz.** 1976. "Equilibrium in Competitive Insurance Markets: An Essay on the Economics of Imperfect Information." *Quarterly Journal of Economics*, 90(4): 629–649.

**Shepard, Mark.** 2016. "Hospital Network Competition and Adverse Selection: Evidence from the Massachusetts Health Insurance Exchange." National Bureau of Economic Research Working Paper 22600.

**Simon, Kosali, Sharon Tennyson, and Julie Hudman.** 2009. "Do State Cost Control Policies Reduce Medicaid Prescription Drug Spending?" *Risk Management and Insurance Review*, 12(1): 39–66.

**Starc, Amanda, and Robert J. Town.** 2015. "Internalizing Behavioral Externalities: Benefit Integration in Health Insurance." National Bureau of Economic Research Working Paper 21783.

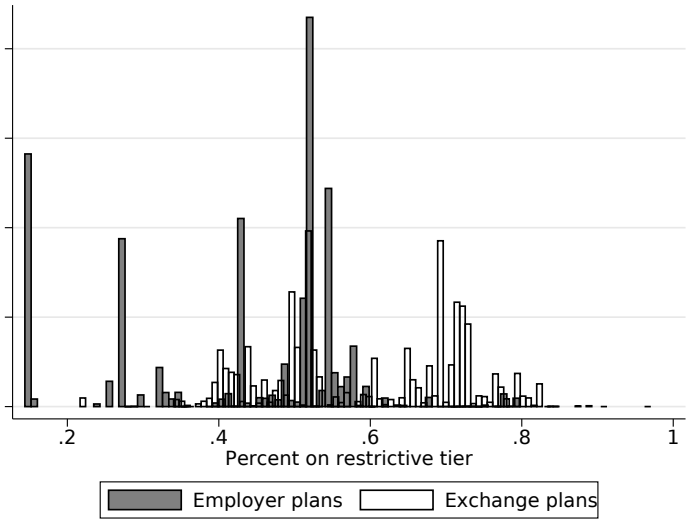
**van de Ven, Wynand P. M. M., and Randall P. Ellis.** 2000. "Risk Adjustment in Competitive Health Plan Markets." In *Handbook of Health Economics*. Vol. 1A, , ed. Anthony J. Culyer and Joseph P. Newhouse, 755–845. Elsevier.

**Veiga, André, and E. Glen Weyl.** 2016. "Product Design in Selection Markets." *Quarterly Journal of Economics*, 131(2): 1007–1056.

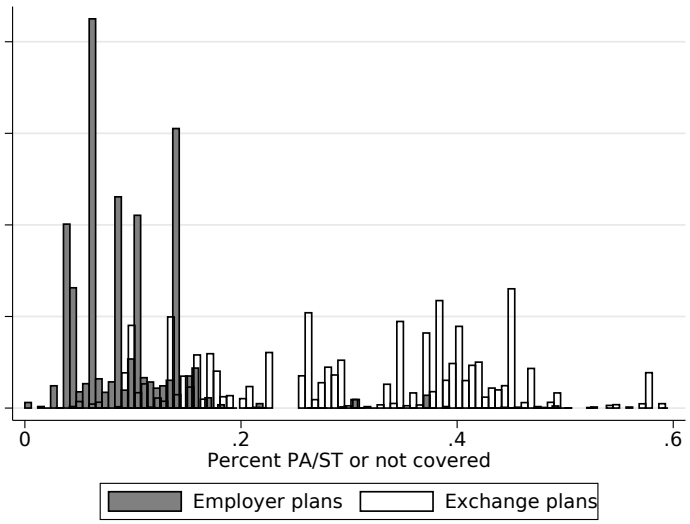
**Zeckhauser, Richard.** 1970. "Medical Insurance: A Case Study of the Tradeoff Between Risk Spreading and Appropriate Incentives." *Journal of Economic Theory*, 2(1): 10–26.

**Figure 1:** Formulary Data: Tiering in Employer and Exchange Plans

(A) Assignment to Restrictive Tier; Specialty or Higher

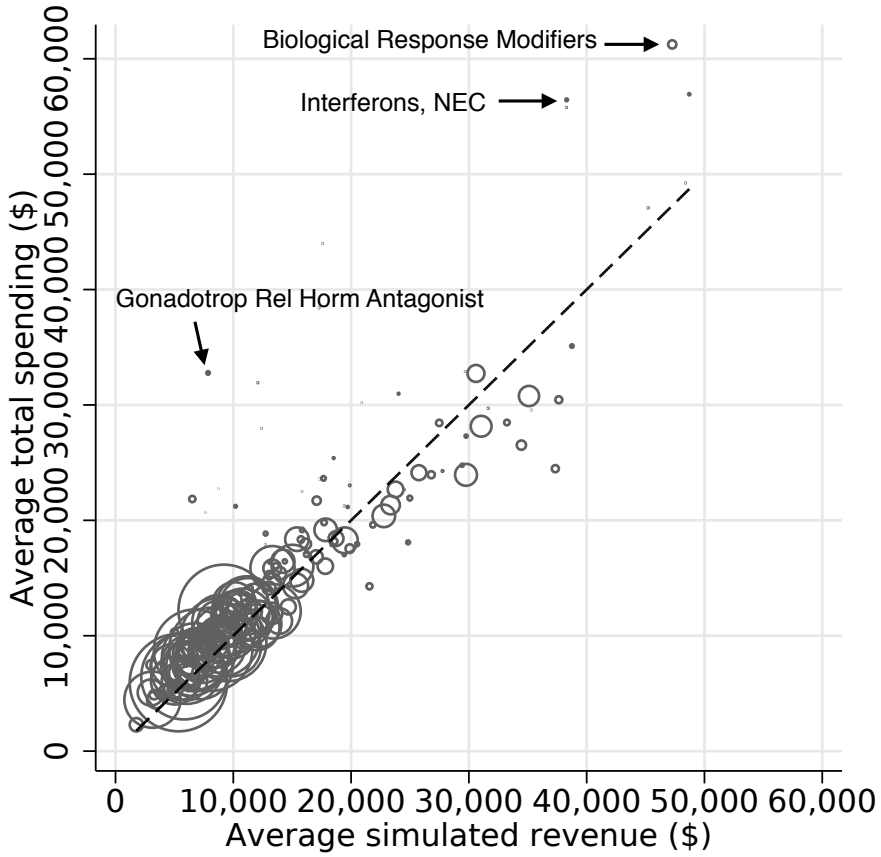



(B) Non-price Hurdles to Access




**Note:** Histograms indicate the fraction of drugs contained in restrictive tiers in employer and Exchange plans. Observations are plans. In Panel A, restrictive tiers are defined as the specialty tier or higher. See Table 1 for a complete ranked listing of the tiers. Panel B repeats the histogram for the fraction of drugs requiring prior authorization or step therapy (PA/ST) or explicitly listed in the formulary as not covered.

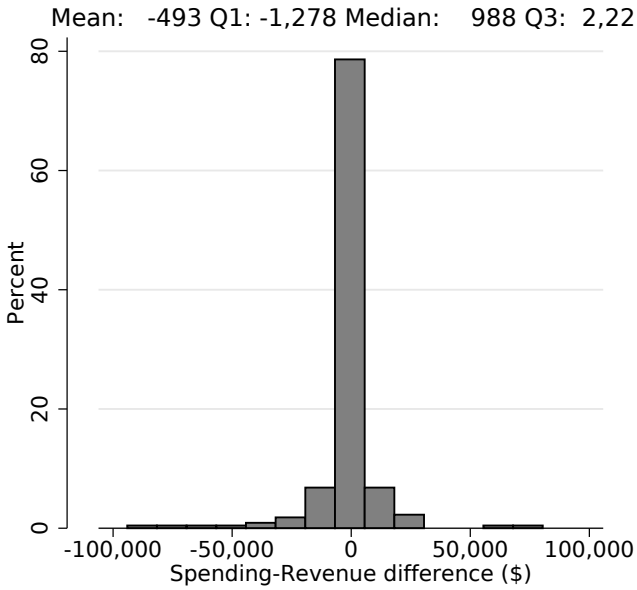
**Figure 2:** Actionable Selection Incentives Remain Net of Risk Adjustment



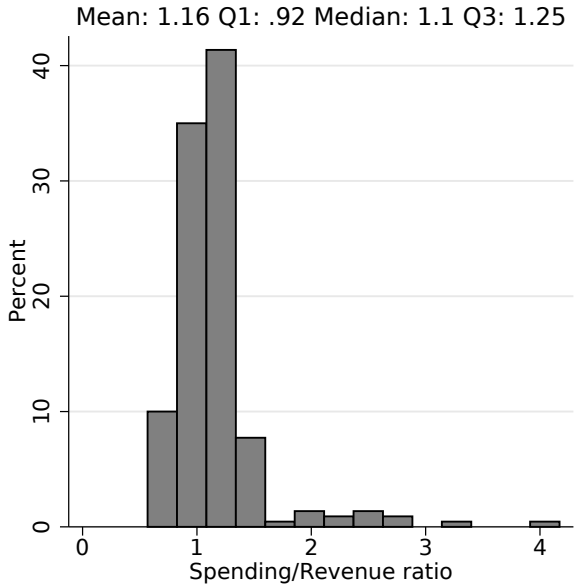

**Note:** Figure plots the relationship between healthcare spending and simulated revenue for each therapeutic class of drugs. Means are for total spending or revenue, calculated over the set of consumers who generate at least one drug claim in the class. Simulated revenue is calculated according to the HHS risk adjustment and reinsurance algorithms as described in the text. Each circle plots the spending and revenue means for a therapeutic class with marker sizes proportional to the number of consumers generating claims in the class. The dashed line at 45 degrees indicates the break even point.

**Figure 3:** Distributions of Selection Incentives Across Drug Classes

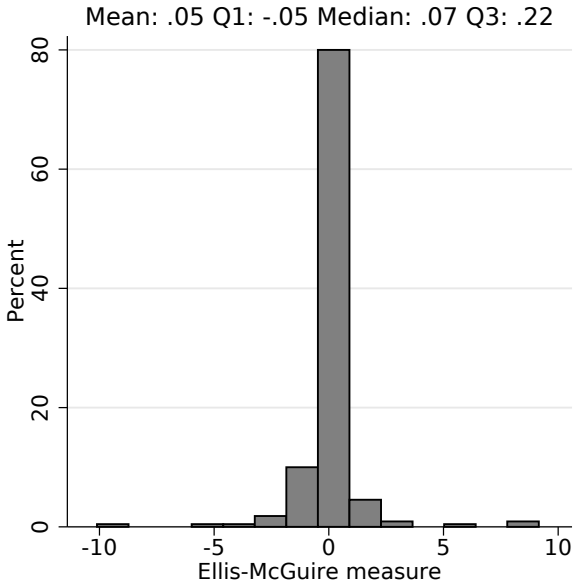
(A) Level Difference




(B) Spending/Revenue Ratio

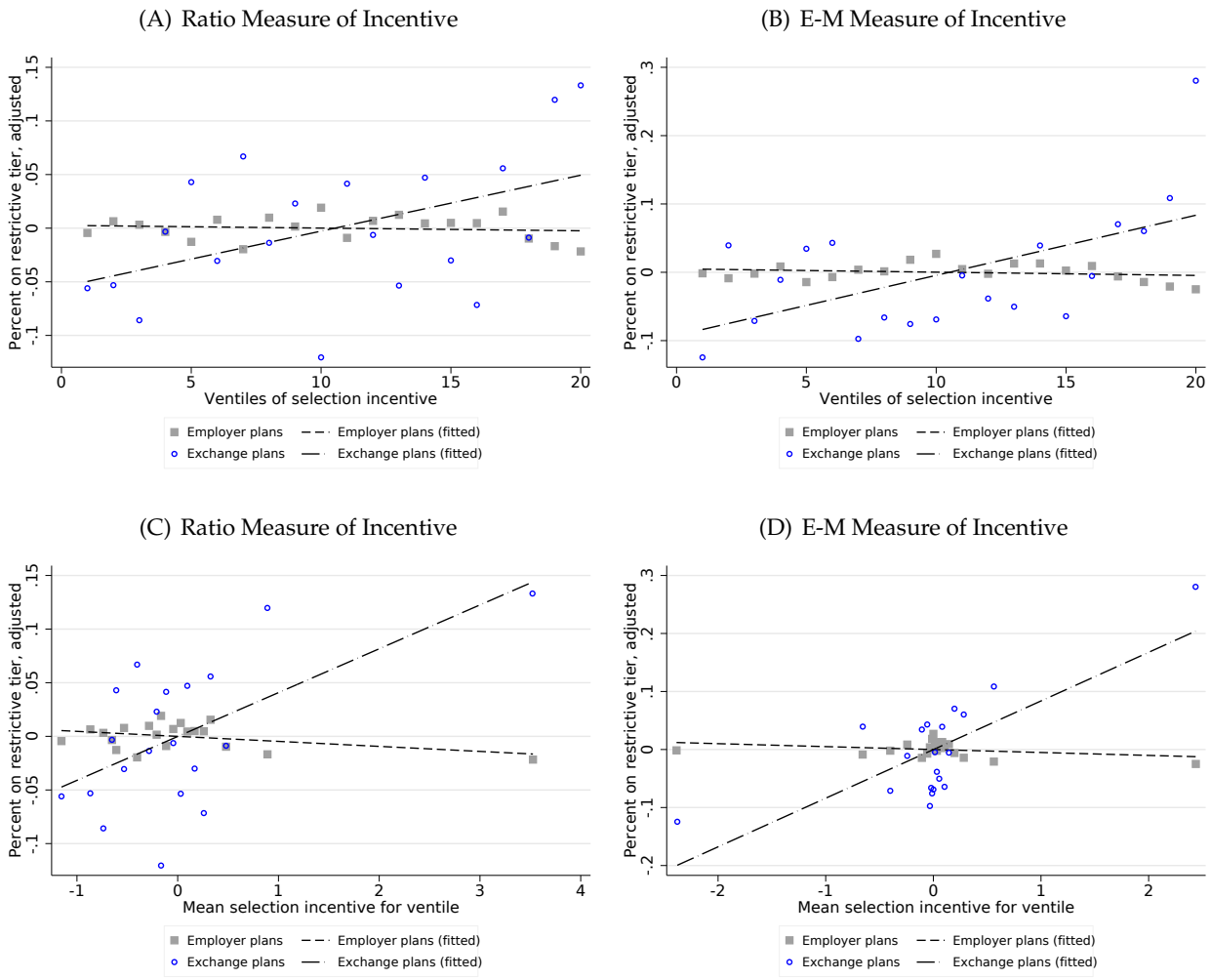



(C) Ellis-McGuire Measure



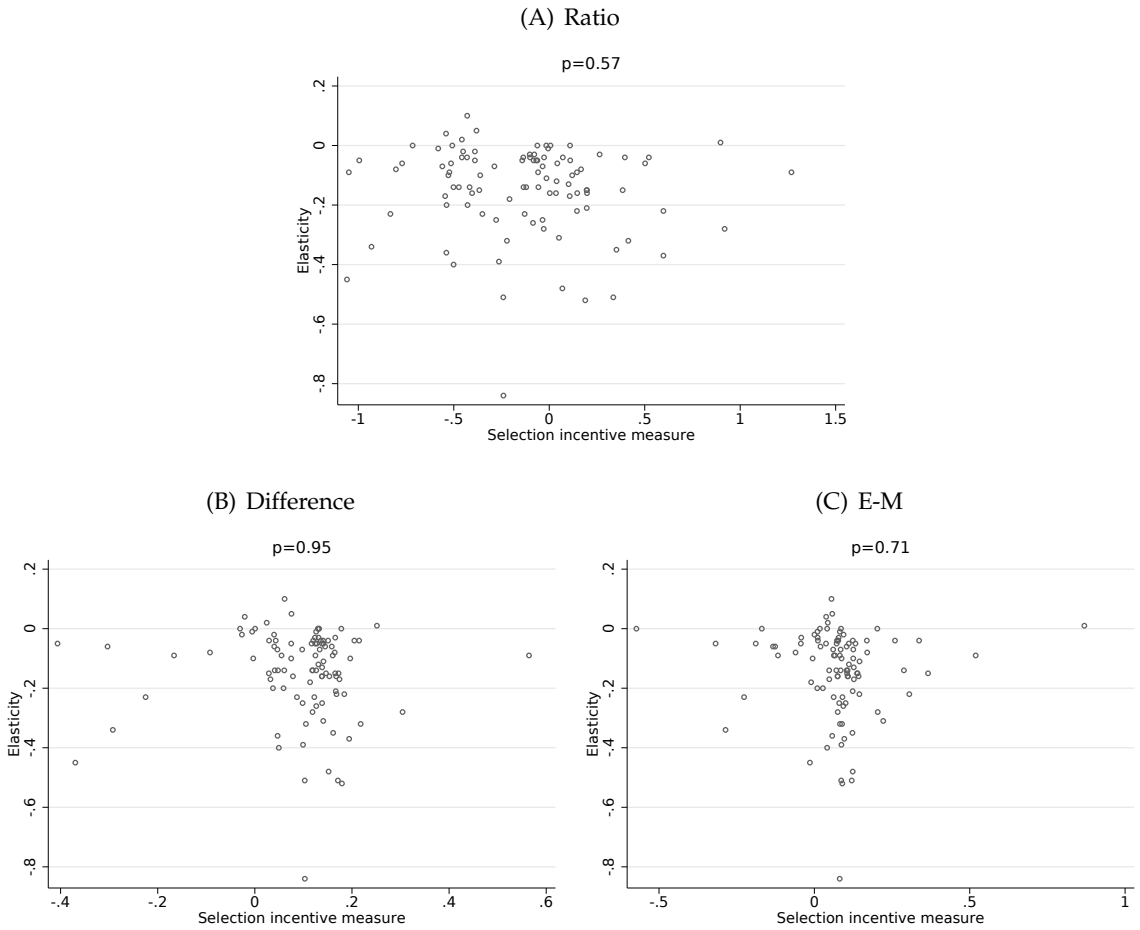

**Note:** Figure displays histograms of the selection incentives described by Equation (7). Panel (A) shows the distribution of the level difference measure. Panel (B) shows the distribution of the spending/revenue ratio, in which a value of 1 is neutral. Panel (C) shows the Ellis-McGuire selection incentive, in which a value of 0 is neutral. Although most classes have neutral or small associated incentives, important outliers exist.

**Figure 4:** Main Result: Selection Incentive and Restrictive Tiering in Two Markets




**Note:** Figure plots semi-parametric versions of the difference-in-differences regression described in Equation (9). Each point corresponds to a group of drugs within a ventile of the indicated selection incentive measure. To generate the position along the vertical axis, we find the residual from a regression of class-by-plan generosity on drug class fixed effects and plan fixed effects as in Equation (9). The left panels use the spending/revenue ratio selection incentive measure. The right panels use the Ellis-McGuire measure. The horizontal axes in the top panels are scaled by the ventile number. The horizontal axes in the bottom panels are scaled by the mean selection incentive value within the ventile. In each panel, the OLS regression line is plotted separately for Exchange and employer plans.

**Figure 5:** Class Selection Incentives Uncorrelated with Drug Class Demand Elasticities




**Note:** Figure plots scatters of the three selection incentive measures and estimates of class-specific demand elasticities from Einav, Finkelstein and Polyakova (2016). $p$-values correspond to the coefficient in a linear regression of the elasticities on the selection incentive measures.

**Table 1:** Summary Statistics: Formulary Tiering in Employer and Exchange Plans

| | Formulary Data | | CCIIO Cost-Sharing Data | |
|---|---|---|---|---|
| | Employer Plans | Exchange Plans | Mean Silver Copay, if no Coinsurance | Fraction Subject to Coinsurance |
| | (1) | (2) | (3) | (4) |
| Number of plans | 3194 | 501 | | |
| Covered lives per plan | 14,723 | 20,343 | | |
| Non-Retrictive Tiers Total: | **0.57** | **0.41** | | |
| Generic preferred | 0.21 | 0.17 | $10 | 11% |
| Generic | 0.00 | 0.05 | | |
| Preferred brand | 0.09 | 0.05 | $41 | 18% |
| Covered/ Non-preferred brand | 0.28 | 0.14 | $73 | 30% |
| Restrictive Tiers Total: | **0.43** | **0.59** | | |
| Specialty | 0.00 | 0.01 | $117 | 66% |
| Not listed | 0.33 | 0.27 | | |
| Medical | 0.00 | 0.01 | | |
| Prior Authorization/Step (PA/ST) | 0.01 | 0.10 | | |
| Not covered | 0.08 | 0.20 | | |
| Therapeutic Classes | 220 | 220 | | |

**Note:** Table lists formulary statistics separately for self-insured employer and Exchange plans in columns 1 and 2, respectively. The Exchange plans in column 2 cover the universe of Exchange formularies in 2015. The employer plans cover about one third of all consumers enrolled in an employer plan in 2015. Tiers are listed from top to bottom in order of increasing restrictiveness, though the Prior Authorization/Step Therapy (PA/ST) tier is horizontally differentiated by imposing non-price hurdles to access. "Not listed" means that the drug was not listed in the formulary, leaving some room for ambiguity with respect to coverage. "Not covered" means that the formulary affirmatively noted that the drug would not be covered by the plan. Tiers are harmonized across plans by the database creator, MMIT. Columns 3 and 4 are derived from a separate data source: the CCIIO public use files that describe Exchange plan attributes. In column 3, we calculate the mean copay associated with each tier in a sample limited to silver plans that charge only a copay (no coinsurance) in the relevant tier. Column 4 reports the fraction of plans that charge coinsurance at each tier.

**Table 2:** Actionable Selection Incentive: Drug Classes with the Largest Spending - Revenue Gaps

| Selection Rank | Class | Most Used Drug in Class | Per Capita Enrollee Spending | Per Capita Enrollee Revenue | Net Loss: Cost - Revenue | Ratio: Cost/ Revenue | Ellis-McGuire Measure |
|---|---|---|---|---|---|---|---|
| (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| **Largest Incentives to Avoid** | | | | | | | |
| 1 | Gonadotropins, NEC | Ovidrel | $21,848 | $6,522 | $15,326 | 3.3 | 0.3 |
| 2 | Biological Response Modifiers | Copaxone | $61,245 | $47,268 | $13,977 | 1.3 | 1.3 |
| 3 | Opiate Antagonists, NEC | naltrexone | $23,639 | $17,662 | $5,977 | 1.3 | 0.3 |
| 4 | Ovulation Stimulants, NEC | clomiphene citrate | $10,306 | $5,003 | $5,304 | 2.1 | 0.2 |
| 5 | Pituitary Hormones, NEC | desmopressin | $21,711 | $17,078 | $4,633 | 1.3 | 1.0 |
| 6 | Vitamin A and Derivatives, NEC | Claravis | $7,472 | $3,044 | $4,428 | 2.5 | 0.2 |
| 7 | Bioflavonoids and Comb, NEC | Metanx (algal oil) | $19,170 | $15,840 | $3,329 | 1.2 | 0.2 |
| 8 | Oxytocics, NEC | methylergonovine | $11,183 | $8,112 | $3,071 | 1.4 | 0.5 |
| 9 | Analg/Antipyr, Opiate Agonists | hydrocodone-acetaminophen | $12,214 | $9,212 | $3,001 | 1.3 | 0.8 |
| 10 | CNS Agents, Misc. | Lyrica | $18,369 | $15,405 | $2,965 | 1.2 | 1.3 |
| **Largest Incentives to Attract** | | | | | | | |
| 211 | Antineoplastic Agents, NEC | methotrexate sodium | $28,157 | $31,042 | -$2,885 | 0.9 | -0.4 |
| 212 | Multivit Prep, Multivit Plain | Folbic | $21,928 | $24,986 | -$3,058 | 0.9 | 0.0 |
| 213 | Coag/Anticoag, Anticoagulants | warfarin | $30,775 | $35,103 | -$4,328 | 0.9 | -0.5 |
| 214 | Cholelitholytic Agents, NEC | ursodiol | $28,481 | $33,232 | -$4,751 | 0.9 | -0.7 |
| 215 | Diuretics, Loop Diuretics | furosemide | $23,946 | $29,759 | -$5,813 | 0.8 | -0.7 |
| 216 | Ammonia Detoxicants, NEC | lactulose | $30,452 | $37,633 | -$7,181 | 0.8 | -0.6 |
| 217 | Anticonv, Hydantoin Derivative | phenytoin sodium extended | $14,284 | $21,559 | -$7,275 | 0.7 | -0.5 |
| 218 | Cardiac, Antiarrhythmic Agents | amiodarone | $26,519 | $34,461 | -$7,942 | 0.8 | -0.5 |
| 219 | Digestants and Comb, NEC | Creon | $44,621 | $56,971 | -$12,350 | 0.8 | -0.7 |
| 220 | Cardiac, Cardiac Glycosides | Digox | $24,480 | $37,338 | -$12,857 | 0.7 | -1.0 |

**Note:** Table lists costs and revenues associated with the drug classes that map to the most and least profitable consumers. Column 2 lists the drug class name. Column 3 lists the most popular drug in the indicated class, by count of users in our MarketScan claims data. Column 4 displays the average total healthcare spending associated with consumers who utilize any drug in the class, $\overline{C_c}$. Column 5 displays the average simulated revenue associated with consumers who utilize any drug in class, $\overline{R_c}$. A single consumer whose claims span several drug classes will contribute to multiple rows of the table. Columns 6 through 8 display for the listed classes the three selection incentive measures used in the analysis.

**Table 3:** Main Result: Selection Incentive Predicts Restrictive Design in Exchanges Relative to ESI

| | ***Panel A*** | | | | | |
|---|---|---|---|---|---|---|
| Dependent Variable: | Fraction of Class Tiered Specialty or Higher | | | | | |
| Selection Incentive Variable: | Ratio (Cost/Revenue) | | Difference (Cost - Revenue) | | Ellis-McGuire Measure | |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Exchange X Selection incentive | 0.046*** | 0.045** | 0.044** | 0.012 | 0.046*** | 0.010 |
| | (0.014) | (0.022) | (0.017) | (0.014) | (0.018) | (0.015) |
| Exchange X Selection incentive ventile 20 | | 0.006 | | 0.300*** | | 0.296*** |
| | | (0.105) | | (0.076) | | (0.089) |
| Therapeutic class FEs | X | X | X | X | X | X |
| Plan FEs | X | X | X | X | X | X |
| Therapeutic classes | 220 | 220 | 220 | 220 | 220 | 220 |
| Observations (plan X state X class) | 858,440 | 858,440 | 858,440 | 858,440 | 858,440 | 858,440 |

| | ***Panel B*** | | | | | |
|---|---|---|---|---|---|---|
| Dependent Variable: | Fraction of Class Tiered Prior Auth./Step Therapy/Not Covered | | | | | |
| Selection Incentive Variable: | Ratio (Cost/Revenue) | | Difference (Cost - Revenue) | | Ellis-McGuire Measure | |
| | (7) | (8) | (9) | (10) | (11) | (12) |
| Exchange X Selection incentive | 0.018* | 0.031** | 0.020* | 0.008 | 0.018* | -0.002 |
| | (0.011) | (0.016) | (0.011) | (0.011) | (0.010) | (0.014) |
| Exchange X Selection incentive ventile 20 | | -0.074 | | 0.108 | | 0.159** |
| | | (0.092) | | (0.083) | | (0.078) |
| Therapeutic class FEs | X | X | X | X | X | X |
| Plan FEs | X | X | X | X | X | X |
| Therapeutic classes | 220 | 220 | 220 | 220 | 220 | 220 |
| Observations (plan X state X class) | 858,440 | 858,440 | 858,440 | 858,440 | 858,440 | 858,440 |

**Note:** Table reports results from a series of regressions of formulary restrictiveness on the class-specific selection incentive. The coefficient of interest is on the interaction between an indicator for Exchange plans and the selection incentive variable, with the latter computed in the three ways described in Equation (7). The selection incentive used in each regression is indicated at the column header. In columns 1 through 6, the dependent variable is the fraction of drugs within the class placed on the specialty tier or higher. In columns 7 through 12, the dependent variable is the fraction of drugs within the class that require prior authorization or step therapy (PA/ST) or are explicitly listed in the formulary as "not covered." See Table 1 for a complete ranked listing of the tiers. All regressions include fixed effects for each of the therapeutic classes of drugs and fixed effects for each plan in the data. Observations are at the plan $\times$ state $\times$ class level. Standard errors are clustered at the the level of the therapeutic class (220 clusters). * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

**Table 4:** How Sophisticated is the Insurer Response to Selection Incentives?

| Dependent Variable: | Fraction of Class Tiered Specialty or Higher | | | | | | |
|---|---|---|---|---|---|---|---|
| Selection Incentive Variable: | Ratio | Ellis-McGuire | Ratio | Ellis-McGuire | Ratio | Ellis-McGuire | Ratio and E-M Simultaenoulsy |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
| Exchange X Selection incentive | 0.046*** | 0.046*** | 0.051*** | 0.041*** | 0.043*** | 0.025 | |
| | (0.014) | (0.018) | (0.015) | (0.013) | (0.013) | (0.019) | |
| Exchange X Average spending associated with class | | | 0.042*** | 0.041*** | | | |
| | | | (0.011) | (0.009) | | | |
| Exchange X Average in-class, drug-only spending | | | | | 0.047*** | 0.036** | |
| | | | | | (0.013) | (0.018) | |
| Exchange X Ratio measure | | | | | | | 0.038*** |
| | | | | | | | (0.014) |
| Exchange X Ellis McGuire measure | | | | | | | 0.039*** |
| | | | | | | | (0.017) |
| Therapeutic class FEs | X | X | X | X | X | X | X |
| Plan FEs | X | X | X | X | X | X | X |
| Therapeutic classes | 220 | 220 | 220 | 220 | 220 | 220 | 220 |
| Observations (plan X state X class) | 858,440 | 858,440 | 858,440 | 858,440 | 858,440 | 858,440 | 858,440 |

**Note:** Table reports results from a series of regressions of formulary restrictiveness on an interaction between an indicator for Exchange plans and the selection incentive. All regressions include fixed effects for each of the therapeutic classes of drugs and fixed effects for each plan in the data. Observations are at the plan $\times$ state $\times$ class level. Standard errors are clustered at the the level of the therapeutic class (220 clusters). See Table 3 for additional details. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

**Table 5:** Salience and Substitution: Popular Drugs and Cheap Drugs

| | ***Panel A*** | | | |
|---|---|---|---|---|
| Within-Class Subsample: | Most Popular Drugs in Class | | | |
| | 75th Percentile of Popularity or Higher | | 90th Percentile of Popularity or Higher | |
| Selection Incentive Variable: | Ratio (Cost/ Revenue) | Ellis-McGuire Measure | Ratio (Cost/ Revenue) | Ellis-McGuire Measure |
| | (1) | (2) | (3) | (4) |
| Exchange X Selection incentive | .061*** | .081*** | .074*** | .098*** |
| | (.022) | (.022) | (.025) | (.022) |
| Therapeutic class FEs | X | X | X | X |
| Plan FEs | X | X | X | X |
| Therapeutic classes | 188 | 188 | 156 | 156 |
| Observations (plan X state X class) | 733,576 | 733,576 | 608,712 | 608,712 |
| | ***Panel B*** | | | |
| Within-Class Subsample: | Least Expensive Drugs in Class | | | |
| | 25th Percentile of Cost or Lower | | 10th Percentile of Cost or Lower | |
| Selection Incentive Variable: | Ratio (Cost /Revenue) | Ellis-McGuire Measure | Ratio (Cost /Revenue) | Ellis-McGuire Measure |
| | (5) | (6) | (7) | (8) |
| Exchange X Selection incentive | 0.058*** | 0.051** | 0.061*** | 0.048** |
| | (0.015) | (0.020) | (0.015) | (0.020) |
| Therapeutic class FEs | X | X | X | X |
| Plan FEs | X | X | X | X |
| Therapeutic classes | 220 | 220 | 220 | 220 |
| Observations (plan X state X class) | 858,440 | 858,440 | 858,440 | 858,440 |

**Note:** Table reports results from a series of regressions of formulary restrictiveness on an interaction between an indicator for Exchange plans and the selection incentive. The dependent variable is the fraction of drugs in the plan $\times$ state $\times$ class tiered specialty or higher, as in Panel A of Table 3. In Panel A here, we limit the sample to the most popular drugs in each class when calculating the dependent variable. In columns 1 and 2, we limit the sample to the 75th percentile of popularity or higher *within each class* (and limit to classes with at least 4 drugs). In columns 3 and 4, we limit the sample to the 90th percentile of popularity or higher within each class (and limit to classes with at least 10 drugs). In Panel B we limit the sample to the least expensive drugs in each class when calculating the dependent variable. In columns 5 and 6, we limit the sample to the 25th percentile of drug prices and below in each class, and in columns 7 and 8, we limit the sample to the 10th percentile of drug prices and below in each class. When finding the least expensive drugs, we rank all drug claims in a class by cost, and make the sample cut at the appropriate point (25th percentile or 10th percentile) of the distribution of claim costs, including all drugs with any claims below the cutoff. All regressions include fixed effects for each of the therapeutic classes of drugs and fixed effects for each plan in the data. Observations are at the plan $\times$ state $\times$ class level. Standard errors are clustered at the the level of the therapeutic class (220 clusters). See Table 3 for additional details. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Online Appendix for:

# Screening in Contract Design: Evidence from the ACA Health Insurance Exchanges

## A Simulated payments

This section provides more detail on the simulated payments used to compute selection incentives and the HHS-HCC risk adjustment model. We define costs as the sum of all health care spending (inpatient, outpatient, and prescription drug) for person $i$ in a given year. We observe this in the Marketscan data. Revenues are not observed in the data and must be simulated. We simulate revenues according to Exchange plan payment formulas specified by the Department of Health and Human Services (HHS). Exchange plan revenues for plan $j$ consist of three components: premiums, $p_i j$, risk adjustment transfers, $R_i^{RA}$, and reinsurance payments $R_i^{Re}$. For risk adjustment transfers, we start by specifying a risk score, $r_i$, for each individual using the risk adjustment formula used in the Exchanges (Kautter et al., 2014). This formula assigns risk scores according to diagnoses in claims data. We use an individual's diagnoses from 2012 to assign his/her risk score. We then specify risk adjustment transfers according to the Exchange risk adjustment transfer formula:[54]

$$R_i^{RA} = \left(\frac{r_i}{\bar{r}} - 1\right)\bar{p},$$

where $\bar{r} = \frac{1}{n}\sum_{i=1}^{n} r_i$ and $\bar{p} = \frac{1}{n}\sum_{i=1}^{n} p_i j$ are the average risk score and average premium across all individuals in the market, respectively. Similarly, we define reinsurance payments as

$$R_i^{Re} = .8 \times \left(C_i - 60,000\right)$$

for claim costs above \$60,000.[55] We assume that reinsurance is funded by an actuarially fair per capita reinsurance premium, $\bar{re}$.[56] In words, the reinsurance payment is 80% of the individual cost above the \$60,000 attachment point minus the actuarially fair reinsurance premium equal to the average reinsurance payment. For premiums, we assume that competition forces all plans to charge a premium equal to the average cost in the market. We also assume a symmetric equilibrium so that all plans have the same premium and average cost:[57]

$$p_i j = \bar{C} = \frac{1}{n}\sum_{i=1}^{n} C_i,$$

[54] Note that risk adjustment transfers occur at the plan level, but in fact they are a sum of individual-level transfers. Here we specify the component of the plan's transfer attached to individual $i$.

[55] A policy with a cutoff of \$60,000 and a coinsurance rate of 0.8 was the originally announced reinsurance policy for the Exchanges. This was later adjusted *ex post* to a cutoff of \$45,000 and a coinsurance rate of 0.5. We use the originally announced policy, as insurers likely designed their formularies according to the announced policy rather than the one implemented *ex post*. In practice, there is little difference between the two policies for insurer incentives.

[56] In practice, the Exchange reinsurance program is also funded by a similar premium, but it is assessed across almost all covered lived in the US, not just across individuals in the Exchanges.

[57] Note that this assumption is not as strong as it may seem. If premiums are equal to a value different from average cost, this affects the profitability of all individuals equally, leaving relative profitability across individuals unchanged. The stronger assumption here is that individuals are all in plans that have the same premiums. However, our goal in this paper is not to assess differential incentives for different types of plans, as our data are insufficient for this type of analysis. Instead, we seek to assess the average incentive and the average insurer response to that incentive.

for all $i$ and $j$. Given these three components, we can then generate simulated revenues at the individual level as the sum of the three components which we then use to compute our selection incentive measures.

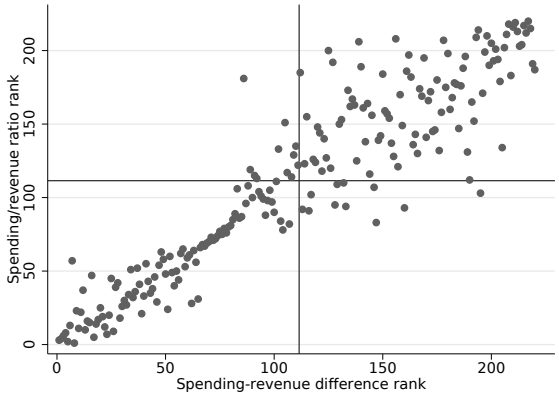
# B Therapeutic classifications

In most of the analyses presented in this paper we rely on the REDBOOK therapeutic classification that is also used in the Marketscan data. There are 257 classes in the REDBOOK classification, of which we analyze the 220 classes for which we are able to construct our selection incentive measures (because they are associated with claims in the Marketscan data) and that also appear in our formulary data. We also use another therapeutic classification system, the American Hospital Formulary Service (AHFS) 8-digit classification. There are 332 classes in the AHFS of which we analyze the 294 classes for which we are able to construct our selection incentive measures (because they are associated with claims in the Marketscan data) and that also appear in our formulary data. We also conduct analyses restricted to the 99 classes that we are able to match to the 108 “common" classes for which Einav, Finkelstein and Polyakova (2016) provide price elasticity measures.

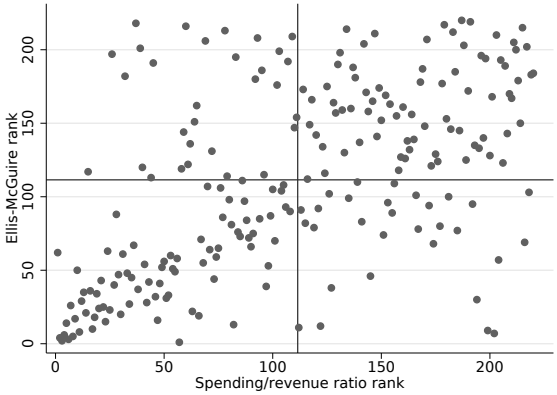
# C Additional Figures and Tables

**Figure A1:** Rank-Rank Correlations of the Three Selection Incentive Measures

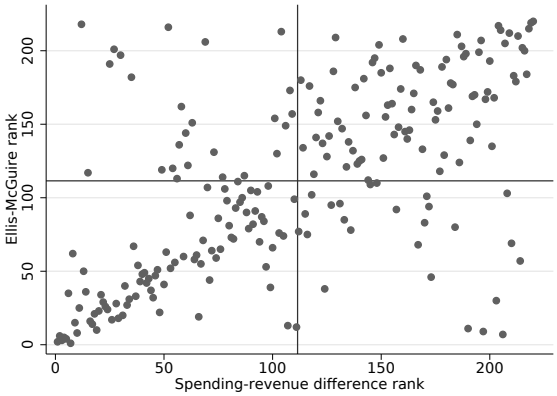
(A) Ratio and Levels




(B) E-M and Ratio

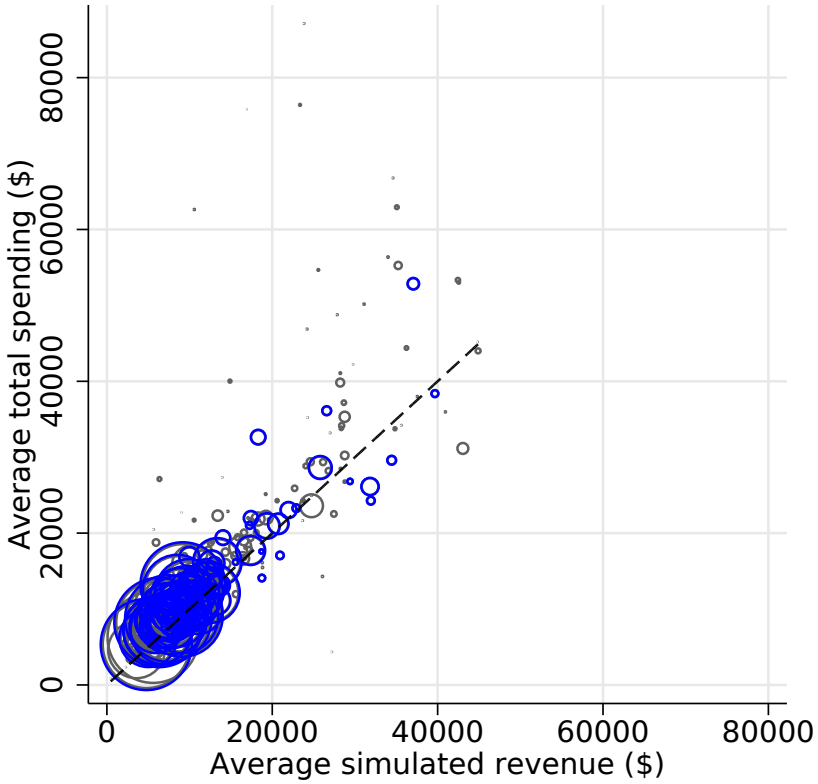



(C) E-M and Levels




**Note:** Figure plots rank-rank scatters of the three selection incentive measures discussed in Section 4.1. The axes range from rank 1 to rank 220, with rank 1 implying the strongest incentive to avoid enrollees. For each of the 220 classes, the scatterplot shows how the ordering of profitable and unprofitable classes compares across the measures. Panel A shows the rank correlation between the level and ratio measures. Panel B shows the rank correlation between the Ellis-McGuire and ratio measures. Panel C shows the rank correlation between the Ellis-McGuire and level measures.

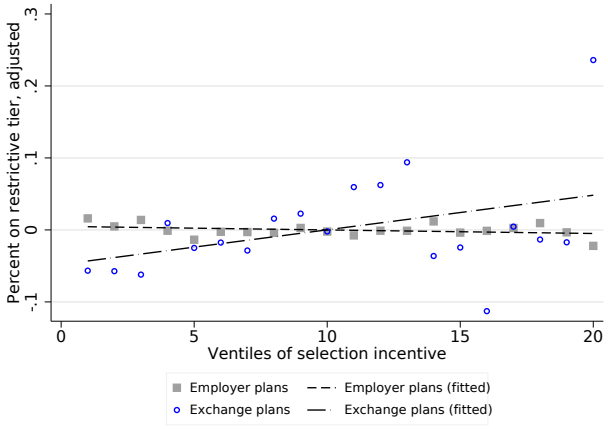
**Figure A2:** Selection Incentives, AHFS Classification



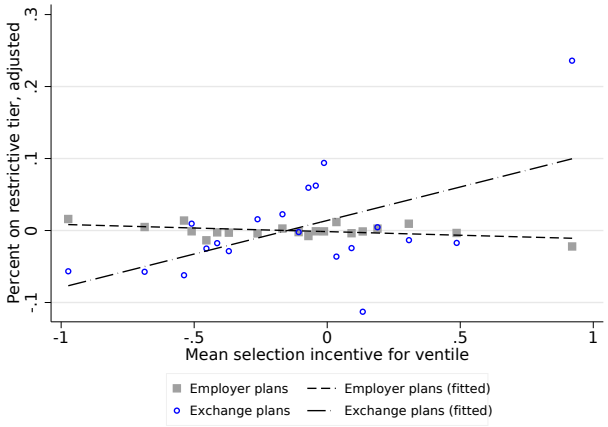

**Note:** Figure plots the relationship between healthcare spending and simulated revenue for each therapeutic class of drugs, as in Figure 2. Here, drugs are re-organized from REDBOOK classes into classes based on the AHFS classification. Blue circles indicate the classes for which Einav, Finkelstein and Polyakova (2016) estimate a demand elasticity that we can import to our analysis. See Figure 2 for additional notes.

**Figure A3:** Selection Incentive and Restrictive Tiering, AHFS Classification

(A) $S_c = \text{Ratio}_c$




(B) $S_c = \text{Ratio}_c$




**Note:** Figure plots semi-parametric versions of the difference-in-differences regression described in Equation (9). Figure repeats Figure 4, using the AHFS therapeutic classification of drugs in place of the RED BOOK classification. The horizontal axes in the top panels are scaled by the ventile number. The horizontal axes in the bottom panels are scaled by the mean selection incentive value within the ventile. In each panel, the OLS regression line is plotted separately for Marketplace and employer plans. See the Figure 4 notes for additional details.

## **Table A1:** Main Results with Alternative Functional Forms

| Dependent Variable: | Fraction of Class Tiered Specialty or Higher | | | | | Fraction of Class Tiered Prior Auth./Step Therapy/Not Covered | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Selection Incentive Variable: | Ratio (Cost/Revenue) | | | | | Ratio (Cost/Revenue) | | | | |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) |
| Exchange X Selection incentive | 0.046*** | 0.045** | 0.025 | 0.025 | | 0.018* | 0.031** | 0.027* | 0.036** | |
| | (0.014) | (0.022) | (0.022) | (0.023) | | (0.011) | (0.016) | (0.015) | (0.016) | |
| Exchange X Selection incentive ventile 20 | | 0.006 | 0.087 | 0.088 | 0.180** | | -0.074 | -0.054 | -0.092 | 0.042 |
| | | (0.105) | (0.107) | (0.111) | (0.070) | | (0.092) | (0.092) | (0.094) | (0.062) |
| Exchange X Selection incentive ventile 19 | | | 0.126 | 0.127 | 0.154* | | | 0.031 | 0.017 | 0.057 |
| | | | (0.085) | (0.086) | (0.080) | | | (0.074) | (0.074) | (0.070) |
| Exchange X Selection incentive ventile 18 | | | | 0.003 | 0.019 | | | | -0.071 | -0.045 |
| | | | | (0.057) | (0.054) | | | | (0.048) | (0.046) |
| Exchange X Selection incentive ventile 1 | | | | | -0.039 | | | | | -0.025 |
| | | | | | (0.056) | | | | | (0.035) |
| Selection Incentive Variable: | Difference (Cost - Revenue) | | | | | Difference (Cost - Revenue) | | | | |
| | (11) | (12) | (13) | (14) | (15) | (16) | (17) | (18) | (19) | (20) |
| Exchange X Selection incentive | 0.044** | 0.012 | 0.005 | 0.004 | | 0.020* | 0.008 | 0.008 | 0.009 | |
| | (0.017) | (0.014) | (0.013) | (0.013) | | (0.011) | (0.011) | (0.011) | (0.011) | |
| Exchange X Selection incentive ventile 20 | | 0.300*** | 0.325*** | 0.330*** | 0.337*** | | 0.108 | 0.109 | 0.104 | 0.123 |
| | | (0.076) | (0.076) | (0.076) | (0.066) | | (0.083) | (0.083) | (0.084) | (0.075) |
| Exchange X Selection incentive ventile 19 | | | 0.153* | 0.157* | 0.158** | | | 0.006 | 0.003 | 0.009 |
| | | | (0.080) | (0.080) | (0.079) | | | (0.062) | (0.062) | (0.061) |
| Exchange X Selection incentive ventile 18 | | | | 0.044 | 0.045 | | | | -0.034 | -0.031 |
| | | | | (0.035) | (0.035) | | | | (0.043) | (0.043) |
| Exchange X Selection incentive ventile 1 | | | | | -0.022 | | | | | -0.030 |
| | | | | | (0.055) | | | | | (0.041) |
| Selection Incentive Variable: | Ellis-McGuire Measure | | | | | Ellis-McGuire Measure | | | | |
| | (21) | (22) | (23) | (24) | (25) | (26) | (27) | (28) | (29) | (30) |
| Exchange X Selection incentive | 0.046*** | 0.010 | 0.002 | -0.001 | | 0.018* | -0.002 | -0.004 | -0.003 | |
| | (0.018) | (0.015) | (0.014) | (0.013) | | (0.010) | (0.014) | (0.015) | (0.015) | |
| Exchange X Selection incentive ventile 20 | | 0.296*** | 0.324*** | 0.340*** | 0.330*** | | 0.159** | 0.166** | 0.164** | 0.151** |
| | | (0.089) | (0.087) | (0.087) | (0.069) | | (0.078) | (0.079) | (0.079) | (0.067) |
| Exchange X Selection incentive ventile 19 | | | 0.154*** | 0.162*** | 0.155*** | | | 0.041 | 0.040 | 0.033 |
| | | | (0.054) | (0.054) | (0.053) | | | (0.050) | (0.050) | (0.048) |
| Exchange X Selection incentive ventile 18 | | | | 0.106* | 0.099* | | | | -0.012 | -0.018 |
| | | | | (0.056) | (0.055) | | | | (0.052) | (0.051) |
| Exchange X Selection incentive ventile 1 | | | | | -0.101* | | | | | -0.070* |
| | | | | | (0.055) | | | | | (0.036) |
| Therapeutic class FEs | X | X | X | X | X | X | X | X | X | X |
| Plan FEs | X | X | X | X | X | X | X | X | X | X |
| Therapeutic classes | 220 | 220 | 220 | 220 | 220 | 220 | 220 | 220 | 220 | 220 |
| Observations (plan X state X class) | 858,440 | 858,440 | 858,440 | 858,440 | 858,440 | 858,440 | 858,440 | 858,440 | 858,440 | 858,440 |

**Note:** Table reports results from a series of regressions of formulary restrictiveness on an interaction between an indicator for Exchange plans and the selection incentive. We repeat the results in Table 3 under a variety of alternative functional forms. All regressions include fixed effects for each of the therapeutic classes of drugs and fixed effects for each plan in the data. Observations are at the plan $\times$ state $\times$ class level. Standard errors are clustered at the the level of the therapeutic class (220 clusters). See Table 3 for additional details. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

**Table A2:** Additional Summary Statistics: Generic and Branded Tiering Separately

| | Branded Drugs Only | | Generic Drugs Only | |
|---|---|---|---|---|
| | Employer Plans | Exchange Plans | Employer Plans | Exchange Plans |
| | (1) | (2) | (3) | (4) |
| Non-Retrictive Tiers Total: | **0.56** | **0.30** | **0.60** | **0.61** |
| Generic preferred | 0.00 | 0.00 | 0.60 | 0.48 |
| Generic | 0.00 | 0.00 | 0.00 | 0.13 |
| Preferred brand | 0.12 | 0.08 | 0.00 | 0.00 |
| Covered/ Non-preferred brand | 0.44 | 0.22 | 0.00 | 0.00 |
| Restrictive Tiers Total: | **0.44** | **0.70** | **0.40** | **0.39** |
| Specialty | 0.00 | 0.01 | 0.00 | 0.00 |
| Not listed | 0.33 | 0.28 | 0.34 | 0.24 |
| Medical | 0.00 | 0.01 | 0.00 | 0.00 |
| Prior Authorization/Step (PA/ST) | 0.01 | 0.15 | 0.00 | 0.03 |
| Not covered | 0.10 | 0.25 | 0.06 | 0.11 |
| Therapeutic Classes | 218 | 218 | 192 | 192 |

**Note:** Table lists formulary statistics separately for self-insured employer and Exchange plans. Tiers are listed from top to bottom in order of increasing restrictiveness, though the Prior Authorization/Step Therapy (PA/ST) tier is horizontally differentiated by imposing non-price hurdles to access. Tiers are harmonized across plans by the database creator, MMIT. See notes to Table 1 for additional detail.

**Table A3:** Main Results Restricted to Generic-Only and Branded-Only Within Class

| | ***Panel A*** | | |
|---|---|---|---|
| Within-Class Subsample: | Branded Drugs Only | | |
| Selection Incentive Variable: | Ratio (Cost/ Revenue) | Difference (Cost - Revenue) | Ellis-McGuire Measure |
| | (1) | (2) | (3) |
| Exchange X Selection incentive | 0.033* | 0.041*** | 0.042*** |
| | (0.018) | (0.013) | (0.014) |
| Therapeutic class FEs | X | X | X |
| Plan FEs | X | X | X |
| Therapeutic classes | 218 | 218 | 218 |
| Observations (plan X state X class) | 850,636 | 850,636 | 850,636 |
| | ***Panel B*** | | |
| Within-Class Subsample: | Generic Drugs Only | | |
| Selection Incentive Variable: | Ratio (Cost /Revenue) | Difference (Cost - Revenue) | Ellis-McGuire Measure |
| | (4) | (5) | (6) |
| Exchange X Selection incentive | 0.040*** | 0.029* | 0.024 |
| | (0.013) | (0.015) | (0.019) |
| Therapeutic class FEs | X | X | X |
| Plan FEs | X | X | X |
| Therapeutic classes | 192 | 192 | 192 |
| Observations (plan X state X class) | 749,184 | 749,184 | 749,184 |

**Note:** Table reports results from a series of regressions of formulary restrictiveness on an interaction between an indicator for Exchange plans and the selection incentive. We repeat the results in Table 3, but alter the dependent variable. In Panel A, the dependent variable (fraction of drugs in class tiered specialty or higher) is calculated over branded products only. In Panel B, the dependent variable (fraction of drugs in class tiered specialty or higher) is calculated over generic products only. All regressions include fixed effects for each of the therapeutic classes of drugs and fixed effects for each plan in the data. Observations are at the plan $\times$ state $\times$ class level. Standard errors are clustered at the the level of the therapeutic class (220 clusters). See Table 3 for additional details. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

**Table A4:** Robustness: Stratifying by Fraction Generic in Class

| | ***Panel A*** | | |
|---|---|---|---|
| Subsample: | Classes with No Generics | | |
| Selection Incentive Variable: | Ratio (Cost/ Revenue) | Difference (Cost - Revenue) | Ellis-McGuire Measure |
| | (1) | (2) | (3) |
| Exchange X Selection incentive | .087** | .045* | .037** |
| | (.036) | (.024) | (.016) |
| Therapeutic class FEs | X | X | X |
| Plan FEs | X | X | X |
| Therapeutic classes | 28 | 28 | 28 |
| Observations (plan X state X class) | 109,256 | 109,256 | 109,256 |
| | ***Panel B*** | | |
| Subsample: | Classes with less than 10% Generics | | |
| Selection Incentive Variable: | Ratio (Cost /Revenue) | Difference (Cost - Revenue) | Ellis-McGuire Measure |
| | (4) | (5) | (6) |
| Exchange X Selection incentive | .083*** | .046* | .037** |
| | (.022) | (.024) | (.014) |
| Therapeutic class FEs | X | X | X |
| Plan FEs | X | X | X |
| Therapeutic classes | 49 | 49 | 49 |
| Observations (plan X state X class) | 191,198 | 191,198 | 191,198 |
| | ***Panel C*** | | |
| Subsample: | Classes with less than 25% Generics | | |
| Selection Incentive Variable: | Ratio (Cost /Revenue) | Difference (Cost - Revenue) | Ellis-McGuire Measure |
| | (4) | (5) | (6) |
| Exchange X Selection incentive | .065** | .047* | .048*** |
| | (.026) | (.027) | (.016) |
| Therapeutic class FEs | X | X | X |
| Plan FEs | X | X | X |
| Therapeutic classes | 84 | 84 | 84 |
| Observations (plan X state X class) | 327,768 | 327,768 | 327,768 |

**Note:** Table reports results from a series of regressions of formulary restrictiveness on an interaction between an indicator for Exchange plans and the selection incentive. We repeat the results in Table 3, but alter the sample of drug classes included in the regression. Panel A is restricted to classes containing no generics. Panel B is restricted to classes containing less than 10% generics. Panel C is restricted to classes containing less than 25% generics. All regressions include fixed effects for each of the therapeutic classes of drugs and fixed effects for each plan in the data. Observations are at the plan $\times$ state $\times$ class level. Standard errors are clustered at the the level of the therapeutic class (220 clusters). See Table 3 for additional details. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

**Table A5:** Robustness: Controlling for Exchange × Fraction Generic in Class

| Restrictive Tier Definition: | Specialty or Higher | | |
|---|---|---|---|
| Selection Incentive Variable: | Ratio | Diff. | E-M |
| | (1) | (2) | (3) |
| Exchange X selection incentive | .041*** | .035*** | .034** |
| | (.012) | (.014) | (.016) |
| Exchange X class fraction generic | -.26*** | -.25*** | -.24*** |
| | (.060) | (.064) | (.065) |
| Therapeutic class FEs | X | X | X |
| Plan FEs | X | X | X |
| Therapeutic classes | 220 | 220 | 220 |
| Observations (plan X state X class) | 858,440 | 858,440 | 858,440 |

**Note:** Table reports results from a series of regressions of formulary restrictiveness on an interaction between an indicator for Exchange plans and the selection incentive. We repeat the results in Table 3, controlling for the interaction of the Exchange indicator and the fraction of drugs in the class that are generic. All regressions include fixed effects for each of the therapeutic classes of drugs and fixed effects for each plan in the data. Observations are at the plan × state × class level. Standard errors are clustered at the the level of the therapeutic class (220 clusters). See Table 3 for additional details.* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

**Table A6:** Robustness: ESI-Exchange Differences Do Not Track Consumer Demand Elasticities

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | ***Panel A*** | | | | | | | | |
| Dependent Variable: | Fraction of Class Tiered Specialty or Higher | | | | | | | | |
| Selection Incentive Variable: | Ratio (Cost/Revenue) | | | Difference Measure | | | E-M Measure | | |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) |
| Exchange X Selection incentive | 0.037** | 0.098** | 0.097** | -0.004 | 0.349** | 0.348** | -0.006 | 0.228 | 0.226 |
| | (0.016) | (0.045) | (0.045) | (0.023) | (0.168) | (0.165) | (0.021) | (0.140) | (0.139) |
| Exchange X Elasticity | | | -0.053 | | | -0.066 | | | -0.059 |
| | | | (0.089) | | | (0.095) | | | (0.090) |
| Therapeutic class FEs | X | X | X | X | X | X | X | X | X |
| Plan FEs | X | X | X | X | X | X | X | X | X |
| Therapeutic classes | 294 | 99 | 99 | 294 | 99 | 99 | 294 | 99 | 99 |
| Observations (plan X state X class) | 1,147,188 | 386,298 | 386,298 | 1,147,188 | 386,298 | 386,298 | 1,147,188 | 386,298 | 386,298 |
| | ***Panel B*** | | | | | | | | |
| Dependent Variable: | Fraction of Class Tiered Prior Auth./Step Therapy/Not Covered | | | | | | | | |
| Selection Incentive Variable: | Ratio (Cost/Revenue) | | | Difference Measure | | | E-M Measure | | |
| | (10) | (11) | (12) | (13) | (14) | (15) | (16) | (17) | (18) |
| Exchange X Selection incentive | 0.006 | 0.065** | 0.065** | 0.006 | 0.248*** | 0.248*** | 0.006 | 0.105 | 0.105 |
| | (0.012) | (0.029) | (0.029) | (0.013) | (0.094) | (0.093) | (0.013) | (0.087) | (0.087) |
| Exchange X Elasticity | | | 0.001 | | | -0.008 | | | -0.005 |
| | | | (0.043) | | | (0.045) | | | (0.042) |
| Therapeutic class FEs | X | X | X | X | X | X | X | X | X |
| Plan FEs | X | X | X | X | X | X | X | X | X |
| Therapeutic classes | 294 | 99 | 99 | 294 | 99 | 99 | 294 | 99 | 99 |
| Observations (plan X state X class) | 1,147,188 | 386,298 | 386,298 | 1,147,188 | 386,298 | 386,298 | 1,147,188 | 386,298 | 386,298 |

**Note:** Table reports results from a series of regressions of formulary restrictiveness on an interaction between an indicator for Exchange plans and the selection incentive. To create this table, we use an alternative mapping of drugs to therapeutic classes generated by the American Hospital Formulary Service. This allows us to match classes to those for which Einav, Finkelstein and Polyakova (2016) estimate demand elasticities. In the third column of each set of three specifications, we additionally control for an interaction between these imported demand elasticities and the Exchange plan indicator. See text for full detail. All regressions include fixed effects for each of the therapeutic classes of drugs and fixed effects for each plan in the data. Observations are at the plan $\times$ state $\times$ class level. Standard errors are clustered at the the level of the therapeutic class (220 clusters). See Table 3 for additional details. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

**Table A7:** Robustness: Removing Fertility Treatment Classes from Analysis

| Restrictive Tier Definition: | Specialty or Higher | | |
|---|---|---|---|
| Selection Incentive Variable: | Ratio | Diff. | E-M |
| | (1) | (2) | (3) |
| Exchange X selection incentive | .046** | .041** | .046** |
| | (.020) | (.017) | (.018) |
| Therapeutic class FEs | X | X | X |
| Plan FEs | X | X | X |
| Therapeutic classes | 217 | 217 | 217 |
| Observations (plan X state X class) | 846,734 | 846,734 | 846,734 |

**Note:** Table reports results from a series of regressions of formulary restrictiveness on an interaction between an indicator for Exchange plans and the selection incentive. We repeat the results in Table 3, but remove the three therapeutic classes associated with fertility treatments. All regressions include fixed effects for each of the therapeutic classes of drugs and fixed effects for each plan in the data. Observations are at the plan $\times$ state $\times$ class level. Standard errors are clustered at the the level of the therapeutic class (220 clusters). See Table 3 for additional details.* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

**Table A8:** Robustness: Patterns Persist within Pharmacy Benefits Managers

| Dependent Variable: | Fraction of Class Tiered Specialty or Higher | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Selection Incentive Variable: | Ratio (Cost/Revenue) | | Ellis-McGuire Measure | | Ratio (Cost/Revenue) | | Ellis-McGuire Measure | |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| Exchange X Selection incentive | 0.041*** | 0.041* | 0.039** | 0.001 | 0.046*** | 0.047** | 0.042** | 0.003 |
| | (0.013) | (0.022) | (0.015) | (0.014) | (0.014) | (0.022) | (0.017) | (0.015) |
| Exchange X Selection incentive ventile 20 | | 0.003 | | 0.307*** | | -0.005 | | 0.316*** |
| | | (0.106) | | (0.091) | | (0.110) | | (0.093) |
| Therapeutic class FEs | X | X | X | X | X | X | X | X |
| Plan FEs | X | X | X | X | X | X | X | X |
| PBM FE X selection incentive | X | X | X | X | | | | |
| PBM FE X state X selection incentive | | | | | X | X | X | X |
| Therapeutic classes | 220 | 220 | 220 | 220 | 220 | 220 | 220 | 220 |
| Observations (plan X state X class) | 838,034 | 838,034 | 838,034 | 838,034 | 749,280 | 749,280 | 749,280 | 749,280 |

**Note:** Table reports results from a series of regressions of formulary restrictiveness on an interaction between an indicator for Exchange plans and the selection incentive. We repeat the results in Table 3, but add fixed effects for Pharmacy Benefits Managers (PBMs). All regressions include fixed effects for each of the therapeutic classes of drugs and fixed effects for each plan in the data. Observations are at the plan $\times$ state $\times$ class level. Standard errors are clustered at the the level of the therapeutic class (220 clusters). See Table 3 for additional details.* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$