

Screening in Contract Design: Evidence from the ACA Health Insurance Exchanges[†]

By MICHAEL GERUSO, TIMOTHY LAYTON, AND DANIEL PRINZ*

We study insurers' use of prescription drug formularies to screen consumers in the ACA Health Insurance exchanges. We begin by showing that exchange risk adjustment and reinsurance succeed in neutralizing selection incentives for most, but not all, consumer types. A minority of consumers, identifiable by demand for particular classes of prescription drugs, are predictably unprofitable. We then show that contract features relating to these drugs are distorted in a manner consistent with multidimensional screening. The empirical findings support a long theoretical literature examining how insurance contracts offered in equilibrium can fail to optimally trade off risk protection and moral hazard. (JEL D82, G22, H51, I13, I18)

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 Organisation for Economic Co-operation and Development (OECD). Two hallmark features of these markets are that no consumer can be denied coverage and that plans cannot price discriminate based on an individual's health status. This ban against price discrimination on preexisting conditions continues to play a central role in debates over the future of

*Geruso: Department of Economics, University of Texas at Austin, 2225 Speedway, Austin, TX 78712, and NBER (email: mike.geruso@gmail.com); Layton: Department of Health Care Policy, Harvard Medical School, 180 Longwood Avenue, Boston, MA 02115, and NBER (email: layton@hcp.med.harvard.edu); Prinz: PhD Program in Health Policy, Harvard University, 14 Story Street, Cambridge, MA 02138 (email: dprinz@g.harvard.edu). Kate Ho was coeditor for this article. We thank Naoki Aizawa, Marika Cabral, Colleen Carey, Jeff Clemens, Leemore Dafny, Kurt Lavetti, Tom McGuire, Rita Santos, Mark Shepard, Amanda Starc, Gábor Virág, and seminar and meeting participants at the American-European Health Economics Study Group, the University of Arizona, the Caribbean Health Economics Symposium, the Federal Trade Commission, Harvard Medical School, Hunter College, the Kellogg Healthcare Markets Conference, the NBER Insurance Working Group Meeting, the NBER Summer Institute Health Care Meeting, the Penn-LDI Health Insurance Exchange Conference, the Tulane Conference on Behavior and Innovation in Healthcare, the US Congress, the University of Texas at Austin, and the Hungarian Academy of Sciences Institute of Economics Summer Workshop for helpful feedback. We gratefully acknowledge financial support from center grants P2C HD042849 and 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). The content is solely the responsibility of the authors and does not necessarily represent the official views of any funder. No party had the right to review this paper prior to its circulation.

[†]Go to <https://doi.org/10.1257/pol.20170014> to visit the article page for additional materials and author disclosure statement(s) or to comment in the online discussion forum.

the individual markets. Proposals to repeal the ACA have often explicitly highlighted an intention to maintain protections for consumers with preexisting conditions.

Enforcing a policy of no price 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).¹ 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 Frank, Glazer, and McGuire (2000); Glazer and McGuire (2000); Veiga and Weyl (2016); and Azevedo and Gottlieb (2017).² 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 incentive 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. Although the theoretical importance of both types of distortions is well-established, empirical evidence has largely focused on price distortions.

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 reformed by the ACA. Pharmaceuticals for managing chronic illness are likely to be among the most price-transparent and predictable medical goods that health care 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 tool.³

We begin our study of this issue by systematically examining whether prescription drug utilization represents a plausible screening mechanism for patient profitability. Identifying patient profitability in this setting requires accounting for the regulatory transfers aimed at compensating plans for enrolling costly consumers. Risk-adjustment transfers and reinsurance were introduced into the individual markets by the ACA with this purpose. These payment mechanisms imply that an insurer's net revenues can vary substantially across enrollees who pay the same premiums. Drawing on a large sample of health claims, we use the exchange regulator's risk-adjustment and reinsurance algorithms to simulate enrollee-specific net revenue

¹ For recent empirical applications, see Einav, Finkelstein, and Cullen (2010); Hackmann, Kolstad, and Kowalski (2015); and Handel, Hendel, and Whinston (2015).

² Azevedo and Gottlieb (2017) shows that this second class of distortion can actually be thought of as a version of the first class, where the contract space is large and certain contracts in that space would face complete death spirals if offered, resulting in their nonexistence in equilibrium.

³ In a study of HIV medication access, Jacobs and Sommers (2015) shows that exchange plans in several states have placed an entire class of commonly prescribed HIV medications, including generic medications, on a high cost-sharing tier. Such benefit design choices are plausibly an attempt by plans to avoid attracting enrollees with HIV. In addition, numerous lawsuits have been filed against insurers by patient advocacy groups for similar formulary design choices.

from data on expenditures and diagnoses. We compare the simulated exchange revenues to the directly observed claims costs, yielding estimates of person-specific implied profits. To understand the potential for screening unprofitable patient types on the basis of drug coverage, we group patients by prescription drug consumption in various therapeutic classes.⁴

As a first result, we show that while there is significant variation in expected insurer *costs* for individuals taking drugs in different therapeutic classes, expected insurer *profits* are similar across the vast majority of patient types, in line with the regulatory goal. For example, consider a consumer who fills a prescription for a *vasodilating agent* to treat angina, a symptom of coronary artery disease. In our data, such a consumer has expected annual medical spending around \$24,000, which is far above premiums. But that consumer generates revenues of around \$26,000 after accounting for the regulatory transfer payments. These transfers represent a regulatory success in that they neutralize the incentive for an insurer to discourage enrollment of this patient type—for example, by restricting access to these drugs. In fact, we find the average relationship between total medical spending and profitability across drug-class types to be approximately zero.

Despite that for most consumer types selection incentives are neutralized, we also find that “payment errors” exist for a small number of drug classes. In some cases, costs exceed revenues and in others vice versa. For these outlier cases, the insurer screening incentives can be significant. A consumer taking a drug in the *biological response modifiers* class (such as Copaxone, which treats multiple sclerosis), is among the most unprofitable in our data. Such a consumer on average will generate \$61,000 in costs but only \$47,000 in net revenue after accounting for the large risk-adjustment and reinsurance transfer payments to the plan enrolling her.

The existence of these types of payment errors provides a natural experiment by which we can test how insurers respond to the signal of profitability embedded in a consumer’s demand for a particular drug class. To leverage this experiment, we use formulary data covering every plan offered in the state and federal exchanges in 2015. We generate measures of coverage generosity for each drug class for each exchange plan. We then examine formulary design for drug classes where the payment system is working well (i.e., average revenues match average costs) and for drug classes where the payment system is working poorly (i.e., average revenues do not match average costs). We compare these patterns to formulary designs in large, self-insured employer plans, which do not face the same screening incentives. This difference-in-difference design allows us to control for drug-class characteristics that are difficult to measure but fixed across the two market settings, such as cost-effectiveness.

We find that insurers respond to payment system errors by designing formularies to be differentially unattractive to unprofitable groups. These results are not driven by the overall lower coverage generosity of exchange plans. Instead, conditional on an exchange plan’s overall generosity, drug classes used by less profitable consumers

⁴Rather than focus on individual drug products, the universe of drugs is partitioned into groups of standard therapeutic classes. This allows us to differentiate between a plan attempting to screen out a patient and a plan attempting to steer a patient to lower cost or higher cost-effectiveness alternatives within a class of substitutes.

appear higher on the formulary tier structure (implying higher out-of-pocket costs) or are subject to more non-price barriers to access such as prior authorization. The pattern is particularly stark for the tails of the distribution of selection incentives. We find that drug classes in the upper 5 percent 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 other drugs in the same plan and 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, implying a potential difference of thousands of dollars in annual out-of-pocket spending per consumer.⁵ Although throughout the paper we describe the screening problem in terms of insurer behavior (in designing formularies), it is important to recognize that the patterns we uncover involve market forces that are beyond any insurance carrier's ability to control. These contract distortions are a reaction to a market failure that has not been sufficiently counteracted by a regulatory response.

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 nudging consumers toward lower-cost substitutes. Consistent with the screening hypothesis, we show that the relationship between payment errors and formulary restrictiveness is strongest for the most popular drugs within an unprofitable class, possibly because coverage for such drugs may be particularly salient in the consumer's plan enrollment decision.

Our paper contributes both specific insights into the functioning of the exchange risk-adjustment system and broader insights on the use of contract features as screening mechanisms. In the narrow context of the ACA exchanges, we show that exchange risk adjustment and reinsurance neutralize the selection incentives associated with most consumer types that are signaled by drug demand. We also document, for the first time, several important facts about the design of exchange formularies and how these compare to formularies in employer plans. In particular, we show that exchange plans are far more likely to use utilization management to constrain pharmaceutical access, possibly in part due to the ACA's cost-sharing subsidies, which constrain insurers' options in setting out-of-pocket cost-sharing.

More broadly, our work connects to a long literature on screening in selection markets and the notion of service-level selection. While several papers (Frank, Glazer, and McGuire 2000; Ellis and McGuire 2007; Geruso and McGuire 2016; Layton et al. 2017) construct measures characterizing selection incentives that vary by service type or setting, only a small recent literature (Shepard 2016; Carey 2017, 2018; Decarolis and Guglielmo 2017; Lavetti and Simon 2018) has been able to empirically document insurer *responses* to such incentives in terms of contract

⁵For 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.

design.⁶ Our work most closely aligns with that of Carey (2017) and Lavetti and Simon (2018), which both examine formulary design in Medicare Part D. Like those papers, we find that the generosity of drug coverage tracks the profitability of various consumer health types.⁷ In contrast to prior work, we find that exchange plan formularies appear to use nonmonetary aspects of plan design in the form of utilization management, suggesting important differences between the widely studied Medicare Part D market and the understudied individual markets.⁸

In addition to presenting the first econometric evidence of screening in the exchanges, our findings extend the existing literature by providing new insights regarding insurers' sophistication in responding to selection incentives. We show that insurers appear to look beyond drug-specific costs when setting cost-sharing schedules. Unlike in Medicare Part D stand-alone plans, which cover only drugs, drug expenditure is a minority share of total health care spending in the plans we study. Therefore, savvy insurers would restrict access to even *cheap drugs* that are associated with patients who are expensive net of risk adjustment. This is what we find, with plans restricting access to lower cost brand drugs and generics when demand for those drugs predicts patients who are unprofitable.⁹ These 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 in the individual health insurance market. It is unprofitable enrollees, rather than costly ones, who are likely to bear high out-of-pocket spending risk.

The remainder of the paper proceeds as follows. We begin in Section I by briefly reviewing the theory of service-level selection and by describing the regulatory environment. In Section II, we describe the data, and in Section III, we evaluate the performance of the ACA's risk-adjustment and reinsurance programs in neutralizing selection incentives. Section IV describes our research design, and Section V reports our findings of contract distortions in the exchanges. Section VI shows our results are not easily explained by alternative hypotheses regarding efficiently steering patients to more cost-effective alternatives. Section VII performs a simple counterfactual analysis of the effects of dropping reinsurance from the exchange

⁶Carey (2017) and Lavetti and Simon (2018) empirically investigate a high-dimensional service-level screening problem of the type described by Frank, Glazer, and McGuire (2000). Other work in the area of contract distortions has focused on a single screen. In the context of a pre-ACA Massachusetts Exchange, Shepard (2016), none investigates the inclusion of expensive "star" hospitals in plans' networks as a screening mechanism. Decarolis and Guglielmo (2017) studies how overall plan generosity induces differential enrollment in privatized Medicare (Parts C and D), collapsing plan generosity into a single dimension. Other related strands of research investigate insurers' use of advertising to achieve favorable selection (Aizawa and Kim 2018) and more direct forms of discrimination that do not necessarily operate via benefit design (Kuziemko, Meckel, and Rossin-Slater 2013).

⁷A related literature considers insurance coverage distortions in formularies 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) and Starc and Town (2015).

⁸Methodologically, the technique we introduce to identify unprofitable consumer types differs from the prior empirical literature in allowing a direct prediction of where the contract distortions should occur without requiring an intermediate mapping of contract parameters to variables included in the risk-adjustment algorithm. This is a subtle but important point because it allows us to identify patient types who face discrimination but whose chronic conditions are not included in the risk-adjustment formula. In our empirical context, this includes women seeking fertility treatments.

⁹For example, due to high inpatient and outpatient spending that isn't fully compensated by risk adjustment and reinsurance.

markets. Section VIII concludes with a discussion of policy implications and potential solutions.

I. Background

A. Conceptual Framework

The theory behind insurance contract distortions due to the screening incentives 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 (2017). Our goal in this section is not to generate new theoretical insights. Rather, we discuss how this body of theory applies to the setting we study: prescription drug formularies among exchange plans. We refer the reader to Geruso and Layton (2017) for a more comprehensive treatment of this literature.

Consider consumers of types $c \in C$, who vary in both expected health care spending and in demand for particular classes of medical services. For simplicity, assume a one-to-one mapping of consumer types to health care services, so that c can be thought of as service types. Insurers offer contracts that consist of service- or type-specific coinsurance rates, $1 - x_c$, with $x_c \in [0, 1]$ being the portion of spending paid by the insurer. It is straightforward to show that in a static one-period setting, the social planner would maximize social welfare by setting each coinsurance rate, $1 - x_c^*$, to balance the benefit of risk protection against the social cost of moral hazard (Zeckhauser 1970, Feldstein 1973). In this same static setting, if insurers $j \in J$ can set type-specific premiums p_c^j and restrict enrollment into a given contract to consumers of a particular type, then competition causes the type-specific profit-maximizing contracts, (p_c^j, x_c^j) , to be identical to the socially optimal contract (p_c^*, x_c^*) .^{10,11}

However, the consequences of competition change when, as in the ACA exchanges, all consumers in a market are combined in a single-risk pool and insurers cannot directly discriminate via setting type-specific premiums or via restricting particular contracts to particular types.¹² As we model in detail in online Appendix A, relative to the social planner's contract design problem, the profit-maximizing insurer now has an additional consideration in how it sets coinsurance rates, x_c^j : By varying coverage for service c , the plan will attract marginal enrollees who may be differentially profitable to the insurer depending on their type-specific costs relative to the uniform premium. Thus, the plan now has an interest not only in providing optimal risk protection for a fixed set of enrollees, it must also consider the set of enrollees its benefits package attracts.

¹⁰ Note that we assume that insurers have full information about consumer types, i.e., there is no asymmetric information in the model. Type-specific contracts thus need not specify coinsurance rates for services other than the service used by the consumer type (service c).

¹¹ When considering a dynamic or multi-period setting, transitions between health states (or types c) would lead to different optimal copayments (Handel, Hendel, and Whinston 2015).

¹² In this case, the insurer offers contracts specifying the full vector of type-specific coinsurance rates, \mathbf{x}_c .

The possibility of screening consumers by setting a schedule of coinsurance rates x_c that are differentially attractive across consumer types drives a wedge between the profit-maximizing coverage levels in the single-risk pool setting and the socially efficient level of coverage. Risk adjustment can affect the size of this wedge by shifting the relative profitability of different groups. With risk adjustment, it is not the comparison between the cost of the consumer type and the uniform premium that motivates the distortionary movement of the coinsurance rates away from the optimal rates. Rather, it is the comparison between the cost of the consumer type and the uniform premium *plus* any risk-adjustment transfer the insurer receives for the type. Given the right set of risk-adjustment transfers, insurers could theoretically be induced to offer the socially optimal contract. In the presence of risk-adjustment “payment errors,” wedges between the socially optimal coinsurance rates and the equilibrium rates will remain. Though we merely sketch the intuition here, this result is shown rigorously by Frank, Glazer, and McGuire (2000); 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 (net of risk adjustment) cost to the insurer. Ellis and McGuire (2007) devises 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, although the theoretical literature has primarily focused on settings in which the only revenue associated with enrollees is premiums, it is straightforward to observe that when additional revenues or transfer payments are present (such as risk adjustment and reinsurance, described below), insurers should respond to the residual incentive net of the payment system, not the gross cost of an individual. Second, in a multiservice contract, the overall profitability of an individual to the insurer matters for the distortionary incentive, not just the individual’s spending on the particular service—in our case, drugs. This means that if an unprofitable group of consumers desires access to 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.¹³ Fourth, moral hazard, if correlated with the selection incentive, would confound estimates of contract distortions because it plays a role in the insurer’s decision over how to set x_c independent of the screening motive. These items motivate the details of how we implement our empirical tests below.

B. Regulatory Environment

The ACA contains several provisions aimed at curbing the use of benefit design as a means of screening out enrollees. These fall into two broad categories. The first includes coverage mandates that directly constrain insurer benefit design.¹⁴

¹³This is because plans are balancing the screening motivation against the motivation to satisfy consumer preferences. In the presence of adjustment costs, which Clemens, Gottlieb, and Molnár (2017) shows to be important in the setting of health care contracts, one might expect nonlinear responses.

¹⁴These are in addition to the prohibitions against coverage denial or the use of medical underwriting in setting plan premiums.

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).¹⁵ However, there is no 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.¹⁶

Risk adjustment, which has become a ubiquitous feature in regulated health insurance markets in the United States 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 consumer types appear approximately 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.¹⁷ 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 percent 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 broadly assessed health plan fees, while the risk-adjustment transfers

¹⁵In states in which the designated "benchmark" exchange plan covered more than one drug, plans were required to cover at least the number of drugs covered by the benchmark plan in each category and class. Andersen (2017) shows these EHB rules to be a binding constraint.

¹⁶Temporary risk corridors which insured insurers' overall plan profits were also in place from 2014 to 2016, though not fully funded. 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.

¹⁷Forty-nine 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 the Centers for Medicare and Medicaid Services (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.

were budget neutral within the market.¹⁸ Risk-adjustment transfers to plans 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.¹⁹

Another feature of the exchange regulation that may be important to understanding the screening phenomenon we study is the cost-sharing reduction (CSR) program. CSRs affect out-of-pocket spending for low-income consumers by raising the effective actuarial value of silver (70 percent AV) plans to 73 percent, 87 percent or 94 percent, as a function of household income.²⁰ A little over half of exchange consumers were receiving CSRs during our study period, 2015. Because the higher actuarial values are achieved by setting lower deductibles, copays and coinsurance compared to the levels set for other consumers nominally enrolled in the same plan, CSRs may have affected insurer's ability to screen via cost-sharing. We discuss this possibility in detail in Section V.

C. 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 health care 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 health care 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 insurers, in fact, react to these signals by distorting coverage.

There are several reasons to suspect that the exchange regulatory framework left in place significant selection incentives as well as sufficient scope for insurers to use formulary design as a tool for avoiding unprofitable patients. First, 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.²¹ Second, CMS has suggested that by 2018, it will amend the risk-adjustment algorithm in the

¹⁸ Reinsurance, which was funded by fees imposed on all health insurance issuers and self-insured employer plans (and so not limited to exchange plans), was in place from 2014 to 2016.

¹⁹ See Centers for Medicare and Medicaid Services (2015) for additional detail on risk adjustment and reinsurance in the first years of the exchanges.

²⁰ See DeLeire et al. (2017) for more information and an analysis of how the CSR subsidy affects consumer-plan choice.

²¹ 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.theaidsinstitute.org/sites/default/files/attachments/IAmStillEssentialBurwellltr_0.pdf).

exchanges to better capture drug spending, suggesting that drug-related selection incentives are viewed as an important issue by the regulator.²² Finally, in the context of formulary design in Medicare Part D, both Carey (2017) and Lavetti and Simon (2018) 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.

II. Data

A. Formularies

We use a database from Managed Markets Insight and Technology (MMIT) that contains detailed formulary information for employer sponsored insurance (ESI) plans and plans offered in the ACA exchanges.²³ 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.²⁴

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.²⁵ Our focus on self-insured employers implies that this group does not include plans 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.²⁶ In addition to a raw tier variable captured in the data, MMIT harmonizes tiering across plans.²⁷ Additional restrictions and exclusions, such as

²²“We propose to use prescription drug utilization data to improve the predictive ability of our risk adjustment models beginning for the 2018 benefit year.” (June 8, 2016 CMS Press Release, <https://www.cms.gov/newsroom/fact-sheets/proposed-hhs-notice-benefit-and-payment-parameters-2018>).

²³MMIT collects information on US plan formularies through agreements with insurance carriers, pharmacy benefit managers, pharmaceutical manufacturers, and others.

²⁴What 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.

²⁵External sources, such as the Kaiser Family Foundation, estimate that approximately 150 million consumers were enrolled in ESI plans in 2015.

²⁶Below, 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 REDBOOK, unless otherwise stated.

²⁷Plans 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,

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 describes 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 higher cost-sharing than drugs in the covered/non-preferred brand tier, drugs in the covered/non-preferred brand tier have higher cost-sharing than drugs in the preferred brand tier, and so on.²⁸ 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 the mapping from these tiers to 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 through 8 of Table 1. We report the statistics separately for standard Silver plans, as well as for the 87 percent and 94 percent actuarial value variants on these plans that are available to CSR-eligible consumers. Columns 3 through 5 list the mean copay associated with each tier among Silver-level exchange products, conditional on a cost-sharing structure that includes only copays for the indicated tier. Columns 6 through 8 report the unconditional probability that the plan assigns a drug in the indicated tier to a coinsurance regime.

The copays increase moving down the table (regardless of the CSR variant), 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

the harmonization eliminates the possibility that “tier 1” indicates the lowest level of cost-sharing in one formulary but the highest in another.

²⁸ 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 the drug. Note that as long as all of the tiers we classify as “restrictive” are more restrictive than all of the tiers we classify as “nonrestrictive,” our analysis is valid. The precise ordering of tier restrictiveness within the restrictive and nonrestrictive categories is not important in our design.

TABLE 1—SUMMARY STATISTICS: FORMULARY TIERING IN EMPLOYER AND EXCHANGE PLANS

	Formulary data		CCIIO cost-sharing data – Silver plans					
	Employer plans (1)	Exchange plans (2)	Mean copay, if no coinsurance			Fraction of plans subjecting tier to coinsurance		
			No cost-sharing reduction (CSR) (3)	CSR 87 percent actuarial value (4)	CSR 94 percent actuarial value (5)	No cost-sharing reduction (CSR) (6)	CSR 87 percent actuarial value (7)	CSR 94 percent actuarial value (8)
Number of plans	3,194	501						
Covered lives per plan	14,580	18,229						
Nonrestrictive tiers total:	0.57	0.41						
Preferred generic	0.21	0.17	\$10	\$7	\$5	11%	7%	7%
Generic	0.00	0.05						
Preferred brand	0.09	0.05	\$41	\$29	\$24	18%	13%	13%
Covered/non-preferred brand	0.28	0.14	\$73	\$54	\$45	30%	30%	30%
Restrictive tiers total:	0.43	0.59						
Specialty	0.00	0.01	\$117	\$81	\$61	66%	61%	61%
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						

Notes: This 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 ambiguity with respect to coverage. “Not covered” means that the formulary affirmatively noted that the drug would not be covered by the plan. Columns 3 through 8 are derived from a separate data source: the CCIIO public-use files that describe plan attributes for the universe of exchange plans in 2015. Columns 3 through 5 list the mean copay associated with each tier among Silver-level exchange products, conditional on a cost-sharing structure that only includes copays for the indicated tier. Columns 6 through 8 report the unconditional probability that the plan assigns the indicated tier to a coinsurance regime. Statistics are calculated separately for standard Silver plans, 87 percent actuarial value CSR Silver plans, and 94 percent actuarial value CSR Silver plans. The CCIIO data do not distinguish between generic and preferred generic, so in columns 3 through 8, these are combined into a single row.

tier, compared to copayments on the order of \$100 per month if placed on the non-preferred brand tier.²⁹

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 nonmonetary 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) shows 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.

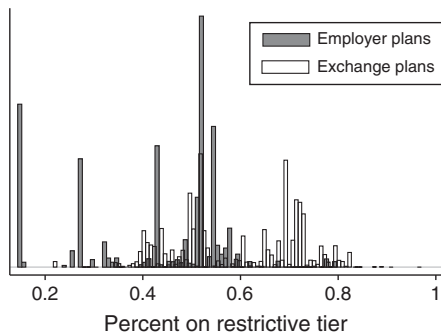
²⁹Carey (2017) shows evidence in Medicare Part D consistent with plans using the copay/coinsurance margin as a screening device.

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 breaking point suggested by plan design, as columns 6 through 8 of the table show that Silver plans switch from relatively generous copay-based cost-sharing to relatively ungenerous coinsurance at this tier. Online Appendix Table A4 reports additional summary statistics for Bronze, Gold, and Platinum plans. That table confirms that this large jump in the probability of facing coinsurance occurs at the specialty tier across all metal levels. Defining the restrictiveness cutoff at the specialty tier 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, NY 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. One of the most obvious differences is that exchange plans are about twice as likely to explicitly list drugs as “not covered” (distinct from not listed) and about ten times as likely to gate-keep drug access via prior authorization or step therapy. 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. In both panels, it is clear that exchange plans make much more extensive use of the restrictive tiers. It is also clear that there is more heterogeneity in restrictiveness of formularies across exchange plans, given the larger spread of these distributions in the figure. While the differences in ESI and exchange generosity are themselves of interest, 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 I 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 REDBOOK, a comprehensive industry drug dictionary. REDBOOK partitions the universe of prescription drugs into 257 mutually exclusive classes. In Section III, we restrict attention to the 220 classes of these for which we observe claims data. These classes, which are intended to capture sets of substitutes, are the level at which we define the insurer’s selection

Panel A. Assignment to restrictive tier; specialty or higher



Panel B. Non-price hurdles to access

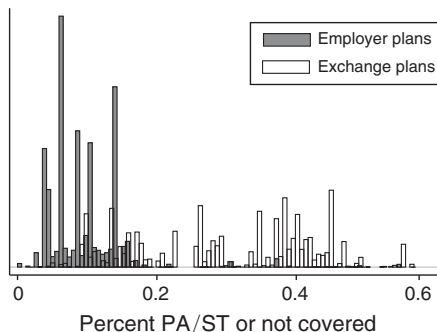


FIGURE 1. FORMULARY DATA: TIERING IN EMPLOYER AND EXCHANGE PLANS

Notes: 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.

incentive.³⁰ 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 rerun the analysis using an alternative classification system designed by the American Hospital Formulary Service.

B. 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 2013.³¹ 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 IV, requires calculating the inter-temporal 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.

³⁰ We use drug class as a proxy for substitutability despite that not all drugs within each of these 220 classes will be perfectly mutually substitutable.

³¹ Access is provided through NBER. MarketScan claims data are collected from a selection of large employers, health plans, government, and public organizations.

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 reported 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 online Appendix B for details. We denote the total revenue (risk adjustment plus reinsurance plus premiums) as R_i .³²

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 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.³³ On the other hand, our cost calculations are in sample in the sense that similar commercial claims data were used by HHS in its original calibration of the risk-adjustment coefficients.

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. Denote the average costs and revenues associated with some class c , respectively, as \bar{C}_c and \bar{R}_c . These means are calculated over the set of consumers having nonzero drug consumption in the class. In practice, we can create the measures of average revenue and average cost only for the 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.

III. Screening Incentives in the Exchanges

Figure 2 gives the first broad look at the extent to which exchange risk adjustment and reinsurance succeed in neutralizing the screening incentives associated with various drug classes. In panel A, we plot the mean of total simulated revenue (premiums plus risk adjustment plus reinsurance) among consumers flagged as consuming

³² Premiums are assumed to equal average claims costs, ignoring loading. As Geruso and Layton (forthcoming) shows, in a symmetric competitive equilibrium with properly functioning risk adjustment, premiums would equal the market-level average costs.

³³ 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.

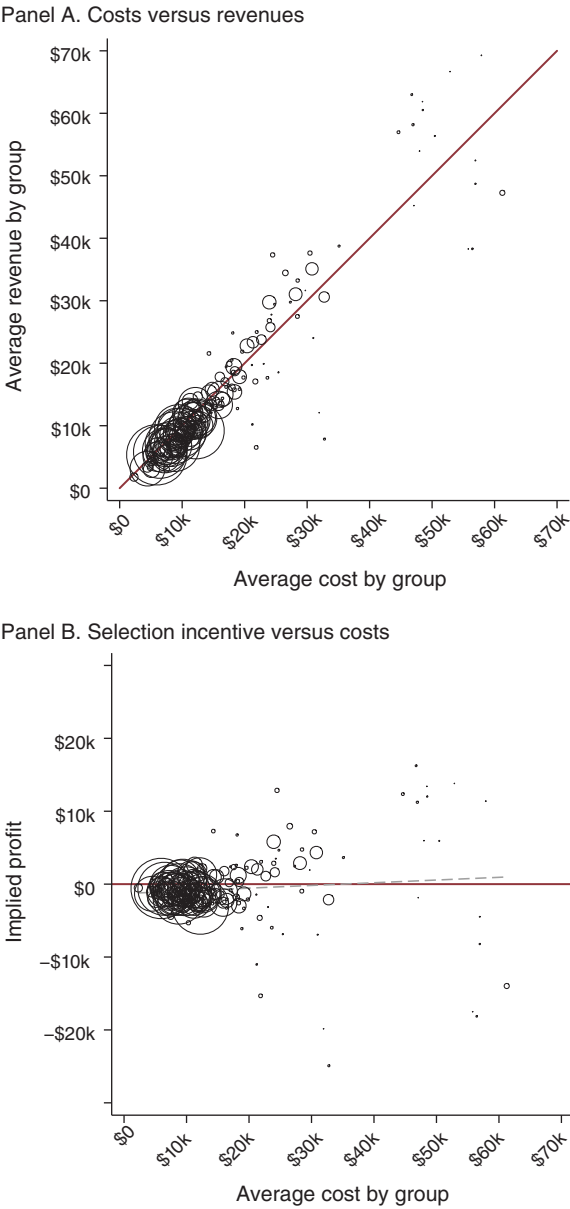


FIGURE 2. ACTIONABLE SELECTION INCENTIVES REMAIN NET OF RISK ADJUSTMENT

Notes: This figure shows the relationship between health care spending and simulated revenue for each therapeutic class of drugs for which spending and revenue are less than \$70,000. Means are for total spending, revenue, or profit, 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. In panel A, the line at 45 degrees indicates the break-even point. In panel B, a horizontal solid line at zero indicates breakeven, and a dashed line plots the linear regression coefficient, weighted by the number of consumers represented in each class.

a drug in class c versus the mean of total cost among those consumers. A line at 45 degrees separates the space into over- and underpayments. Each scatterpoint corresponds to one of the 220 drug classes in our sample.³⁴ Marker sizes reflect the relative number of consumers using drugs in the class. Enrollees associated with classes below the 45-degree line are unprofitable because for these consumers costs exceed total revenue in expectation.³⁵

Most classes in panel A of Figure 2 are tightly clustered around the 45-degree line, indicating that the payment system succeeds in neutralizing formulary screening incentives for the vast majority of potential enrollees. For example, consider a consumer using a vasodilating agent to treat angina, a symptom of coronary artery disease. The average patient that consumes a drug in the *vasodilating agents* class generates \$24,129 in annual costs. Uniform (non-underwritten) premiums, here calculated as equal to the average claims costs in our sample, would amount to only \$4,200. Such a patient would be significantly unprofitable absent some other revenue or transfer payment. In the exchanges, risk adjustment and reinsurance would generate expected transfer payments of \$17,878 and \$3,680, respectively, to a plan enrolling a consumer flagged as taking this type of drug. This generates a total of \$25,758 in revenues, eliminating the insurer's financial incentive to avoid the average consumer of this type.

Nonetheless, there are a small number of significant outliers, far off the diagonal. The existence of outliers in Figure 2 establishes that risk-adjustment payment "errors" are correlated with drug use, a key necessary condition for insurers to use formularies as screening devices. From a practical policy perspective, our results in Section V describe the extent to which consumers whose conditions place them in these outlier groups are exposed to higher out-of-pocket costs and other barriers to access. From a theoretical perspective, the existence of these outliers allows us to test theoretical predictions from the literature on service-level selection (Frank, Glazer, and McGuire 2000) and to observe insurer sophistication in responding to these complex financial incentives that include several cost components (drug utilization, inpatient, and outpatient care) and revenue streams (premiums, reinsurance, and risk adjustment). As we explain in Section IV, we do this by comparing formulary design for the drug classes falling far from the 45-degree line to formulary design for the classes on or near the 45-degree line.

How might these payment errors arise? One possibility, discussed by Carey (2017) 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 (nonexclusive) possibility is that, even absent technological change,

³⁴ Here, we restrict the axes to $< \$70,000$. In online Appendix Figure A1, we zoom out in panels A and B to capture the small number of outliers, and we zoom in panels C and D to provide a clearer view of the cluster of points closer to the origin.

³⁵ Payment errors that are correlated with other consumer "types" (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 screening 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 screening actions is beyond the scope of this paper.

drug utilization comprises an informative signal of patient severity and cost after conditioning on diagnoses. In general, there is no reason to expect that drug utilization—or any other variable not included in the risk-adjustment calibration—would be perfectly orthogonal to profitability.³⁶ This could be due to certain drug-treated conditions being left out of the model or due to incomplete diagnosis coding (the input to the ACA risk-adjustment formula) for some drug-treated conditions. Indeed, 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.³⁷

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 attention to classes for which at least 0.05 percent of sample consumers had a claim. Column 1 indicates the REDBOOK class, and column 2 lists the most popular drug in the class, by count of users in our claims data. Column 3 displays the average of total health care spending associated with the class, \overline{C}_c . Column 4 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 one of these drugs 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 of the screening phenomenon of interest: 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-difference regression framework, and despite relying on predictions made 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.³⁸

³⁶The 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 and hospital coverage in Medicare Part C by Brown et al. (2014) and others.

³⁷HHS 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."

³⁸In online Appendix C, we investigate these selection incentives at the level of individual drug products, rather than therapeutic classes. We show that there is within-class variation in profitability that is comparable in size to the across-class variation shown in Figure 2. However, we find that plan formulary designs much more closely track the variation in profitability associated with therapeutic classes than they track the variation associated with individual drug products. This suggests that insurers are attempting to avoid patient *types*—who may substitute between alternative drug therapies—rather than targeting individual drugs. Such behavior would be consistent with work by Jacobs and Sommers (2015) on the case of HIV drug coverage in several state exchanges. They show that

TABLE 2—ACTIONABLE SELECTION INCENTIVE: DRUG CLASSES WITH THE LARGEST SPENDING - REVENUE GAPS

Class (1)	Most used drug in class (2)	Conditions treated by most used drug (3)	Per capita enrollee spending (4)	Per capita enrollee revenue (5)	Difference: cost – revenue (6)	Ratio: cost/ revenue (7)	Ellis- McGuire measure (8)
<i>Largest incentives to avoid</i>							
Gonadotropins, NEC	Ovidrel	Infertility in women	\$21,848	\$6,522	\$15,326	3.3	0.3
Biological response modifiers	Copaxone	Relapsing multiple sclerosis	\$61,245	\$47,268	\$13,977	1.3	1.3
Opiate antagonists, NEC	Naltrexone	Substance abuse disorders	\$23,639	\$17,662	\$5,977	1.3	0.3
Ovulation stimulants, NEC	Clomiphene citrate	Infertility in women	\$10,306	\$5,003	\$5,304	2.1	0.2
Pituitary hormones, NEC	Desmopressin	Diabetes insip., hemophilia A	\$21,711	\$17,078	\$4,633	1.3	1.0
Vitamin A and derivatives, NEC	Claravis	Severe nodular acne	\$7,472	\$3,044	\$4,428	2.5	0.2
Analg/antipyr, opiate agonists	Hydrocodone- acetamin.	Moderate to severe pain	\$12,214	\$9,212	\$3,001	1.3	0.8
CNS agents, misc.	Lyrica	Nerve pain; fibromyalgia; seizure	\$18,369	\$15,405	\$2,965	1.2	1.3
Mydriatics EENT, NEC	Atropine	Poisonings; presurgical preparations	\$12,895	\$10,018	\$2,877	1.3	0.0
Androgens and comb, NEC	AndroGel	Low testosterone	\$12,023	\$9,335	\$2,688	1.3	0.3
<i>Largest incentives to attract</i>							
Antineoplastic agents, NEC	Methotrexate sodium	Various cancers; various autoimmune diseases	\$28,157	\$31,042	–\$2,885	0.9	–0.4
Multivit prep, multivit plain	Folbic	Vitamin deficiency	\$21,928	\$24,986	–\$3,058	0.9	0.0
Coag/anticoag, anticoagulants	Warfarin	Blood clots; stroke prevention	\$30,775	\$35,103	–\$4,328	0.9	–0.5
Cholelitholytic agents, NEC	Ursodiol	Primary biliary cirrhosis; gallstones	\$28,481	\$33,232	–\$4,751	0.9	–0.7
Diuretics, loop diuretics	Furosemide	Edema due to heart, liver, kidney disease; high blood pressure	\$23,946	\$29,759	–\$5,813	0.8	–0.7
Ammonia detoxicants, NEC	Lactulose	Complications of liver disease	\$30,452	\$37,633	–\$7,181	0.8	–0.6
Anticonv, hydantoin derivative	Phenytoin sodium ext.	Seizures; heart arrhythmias; neuropathic pain	\$14,284	\$21,559	–\$7,275	0.7	–0.5
Cardiac, antiarrhythmic agents	Amiodarone	Heart arrhythmias	\$26,519	\$34,461	–\$7,942	0.8	–0.5
Digestants and comb, NEC	Creon	Chronic pancreatitis; cystic fibrosis; pancreatic cancer	\$44,621	\$56,971	–\$12,350	0.8	–0.7
Cardiac, cardiac glycosides	Digox	Heart arrhythmias; heart failure	\$24,480	\$37,338	–\$12,857	0.7	–1.0

Notes: This table lists costs and revenues associated with the drug classes that map to the most and least profitable consumers. Column 1 lists the drug-class name. Column 2 lists the most popular drug in the indicated class by count of users in our MarketScan claims data. Column 3 indicates some conditions treated by the drug in column 2. Column 4 displays the average total health care spending associated with consumers who utilize any drug in the class, \bar{C}_c . Column 5 displays the average simulated revenue associated with consumers who utilize any drug in class, \bar{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.

Other unprofitable classes in the “top ten” include *opiate antagonists*, which are used to treat substance abuse disorders, and two classes that treat infertility in women, a condition that does not enter the risk-adjustment algorithm.³⁹ As far as

insurers restricted access to all HIV drugs, not merely the products within the category that predicted the most unprofitable patients.

³⁹Unlike related studies by Carey (2017, 2018) and Lavetti and Simon (2018), our method for identifying unprofitable consumer types, illustrated in the figure, does not rely on a mapping of drugs to diagnoses. This allows us to predict where unfavorable drug coverage should occur, even among conditions like infertility that are not included in the risk-adjustment formula. This method corresponds directly with the theoretical models of the service-level selection literature (Frank, Glazer, and McGuire 2000).

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 cardiovascular conditions. Cardiovascular conditions were given close attention in Medicare's CMS-HCC risk-adjustment algorithm on which the exchange algorithm was based.⁴⁰

Although some of the most unprofitable patient types in Table 2 are high cost, some of the most profitable patient types, who insurers have incentives to attract, are high cost as well. In fact, we find that there is no strong relationship between profitability and utilization. Panel B of Figure 2 plots the implied profits (vertical deviations from the 45-degree line in panel A) against average costs, again grouping consumers according to drug utilization in various therapeutic classes. The figure makes it clear that payment errors exist in both directions (over- and underpayment) and at all levels of patient severity. In aggregate, these errors net to about zero across groups with no strong trend along the horizontal axis.⁴¹ The plotted linear regression coefficient is only marginally significant ($p = 0.07$) and has a positive slope, indicating that patients in our sample with higher health care spending are on average *more* likely to be profitable under the 2015 exchange payment parameters. Overall, exchange risk adjustment and reinsurance break the tight link between profitability and patient costs.

Interestingly, most of the cost associated with patients whose drug use flags them as unprofitable does not arise from the drug expenditures of those patients. Figure 4 decomposes costs into inpatient, outpatient, and drug costs for each class and then groups classes into twenty ventiles ranked by the strength of the selection incentive (revenue minus costs), with the classes with the strongest selection incentives (i.e., most unprofitable) being represented by the bars on the far right of the figure. Drug costs are further split into medications within the class and outside of it. (A patient who takes a diabetes medication may also be taking medicine for a heart condition.) The figure shows that across all groups of classes, drugs make up a small fraction of total patient costs, usually less than 30 percent. Spending on drugs within the class that defines the patient type is even smaller, on average only 6 percent. Although both in- and out-of-class drug spending are higher for the most unprofitable classes, within-class spending never rises above 19 percent of total costs. Thus, demand for a particular therapeutic class of drugs is a signal correlated with profitability even though the drugs themselves are not the primary drivers of patient costs. We

⁴⁰Interestingly, 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 REDBOOK-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.

⁴¹It has to be the case that the mean error in the risk-adjustment algorithm is zero, as the algorithm arises from an OLS regression in which the dependent variable is costs, and regression coefficients are normalized so that the mean-cost enrollee yields no transfer payment. The mean error would be exactly zero if our analysis were run on exactly the sample of patients over which the algorithm was calibrated. However, it need not be the case that the unweighted mean *group* error be equal to zero, nor that the relationship between group-level costs and profits be equal to zero.

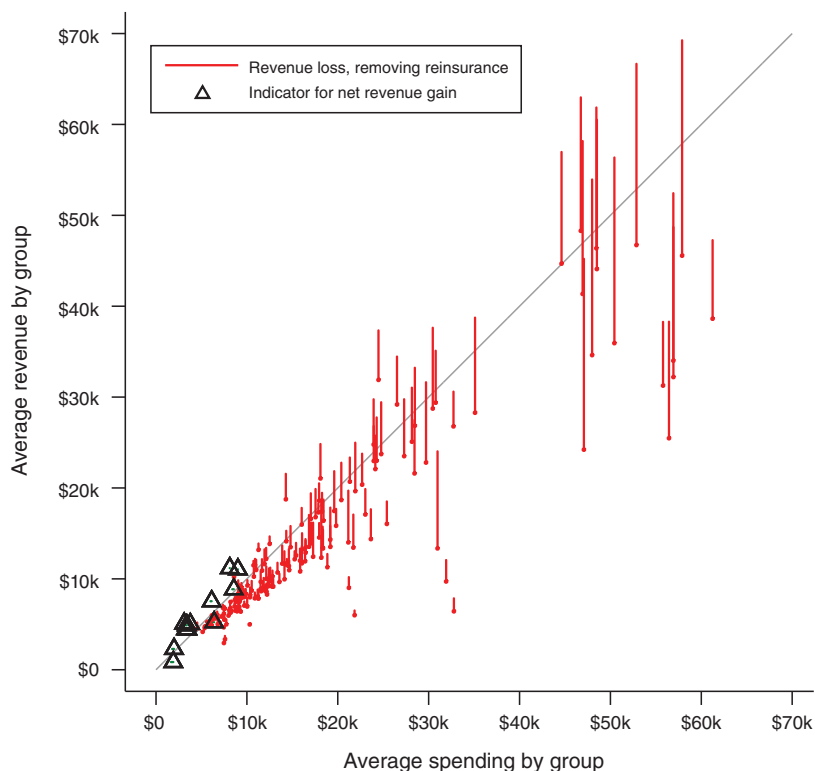


FIGURE 3. CHANGING INCENTIVES DUE TO REMOVAL OF REINSURANCE

Notes: This figure plots how the relationship between health care spending and simulated revenue changes with the removal of reinsurance. Vertical lines describe the loss in net revenue associated with each class that is implied by the removal of reinsurance. The bottom point of each line plots revenue without reinsurance (dropping both the reinsurance premiums and payouts). The top point of each line plots revenue with reinsurance included. For a small set of therapeutic classes, each containing a large number of lower utilization enrollees, removing reinsurance generates a net gain in revenue. These are indicated with triangle markers. Simulated revenue is calculated according to the HHS risk adjustment and reinsurance algorithms as described in the text. The line at 45 degrees indicates the break-even point. See Figure 2 for additional notes.

examine below the extent to which plans are savvy in restricting access to cheap drugs that predict unprofitable patients.⁴²

Finally, we note that reinsurance plays an important role in combatting selection incentives. Figure 3 shows how the profitability associated with the same classes would change if reinsurance were eliminated in a budget-neutral manner.⁴³ The figure follows the structure of Figure 2, but rather than plotting a single

⁴² We also plot in online Appendix Figure A2 a version of Figure 2 that removes in-class drug spending from the total costs associated with the class. Relative to Figure 2, the markers migrate left (toward higher profitability) as a component of cost is removed. Online Appendix Figure A2 shows that in-class spending is substantial enough to visibly alter the scatterplot for the highest cost classes. Nonetheless, substantial variation in profitability remains, indicating the in-class drug spending is not the sole driver of across-class differences in profitability.

⁴³ The ACA used non-budget neutral reinsurance to subsidize the exchange markets. Our simulations, in contrast, fund reinsurance via actuarially fair reinsurance premiums paid by plans on all enrollees. This allows us to isolate the effects of removing reinsurance from the effects of removing a market-wide subsidy.

marker for each drug class, vertical lines connect two points that correspond to simulated revenue for the class with and without reinsurance. Thus, the length of these lines shows the loss in net revenue associated with the loss of reinsurance. By the nature of the budget neutrality of our simulated elimination of reinsurance, many smaller classes with expensive patients lose a large amount of per-patient revenue by receiving less in reinsurance payouts, while a few larger classes with lower cost patients (along with the very large set of patients with no drug utilization) gain a small amount of per-patient revenue by paying less in reinsurance premiums. We depict the small set of therapeutic classes that become more profitable with the elimination of reinsurance (each containing a large number of lower utilization enrollees) with triangle markers.⁴⁴ Figure 3 makes clear that reinsurance is nontrivially contributing to mitigating the adverse screening incentives for the particularly high-cost groups. In Section VII, we discuss how formulary design might be expected to adapt following the removal of reinsurance from these markets in 2017.

IV. Research Design

We build on the findings of the last section, constructing alternative 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.

A. Selection-Incentive Measures

We generate three alternative measures of the class-specific incentive, S_c , for exchange plans to distort coverage:

$$(1) \quad S_c = \begin{cases} \overline{C}_c - \overline{R}_c & \text{cost-revenue difference,} \\ \frac{\overline{C}_c}{\overline{R}_c} & \text{cost-to-revenue ratio,} \\ EM_c & \text{Ellis-McGuire predictable profitability.} \end{cases}$$

In all cases, higher positive values of S_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), which developed a theory of health plan benefit distortions in the presence of adverse selection on service-level benefits. Ellis and McGuire (2007) shows that a profit-maximizing insurer's incentive to distort coverage is defined by the following index:

$$(2) \quad EM_c = \underbrace{\frac{\sum_{i \in I_c} (\hat{C}_{ic} - \overline{C}_c)^2}{\overline{C}_c}}_{\text{predictability}} \times \underbrace{\rho_c}_{\text{predictiveness}}.$$

⁴⁴The positive vertical movement of these points is small enough to be imperceptible in the figure.

The first term of (2) 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 spending in each of the therapeutic classes in 2012. We then predict 2013 spending in each therapeutic class (\hat{C}_{ic}) using the coefficients from this regression. Up to a normalization in the denominator, the predictability term is equivalent to the R^2 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 second term, ρ_c , captures what Ellis and McGuire (2007) refers to as "predictiveness," and it is defined as the correlation between spending in therapeutic class c and individual-level profitability ($R_i - C_i$) in the same period.

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, though it uses the correlation between profitability and a continuous measure of use (total spending on drugs in the class) rather than between profitability and an indicator for any positive spending on a drug in the class. 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 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.

All three S_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 (2018). 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.⁴⁵ Relatedly, our approach captures all drug spending and all medical spending that is predicted by patients' demand for class c , as a patient taking an expensive drug in one class may be likely to have higher consumption in other classes or in nondrug spending. In fact, Figure 4 shows that in-class spending on drugs tends to be the smallest component of spending, even for the consumers taking drugs for which the insurer faces the strongest selection incentives. Savvy insurers would take this into account, maximizing over total profits, not focusing narrowly on one component of costs. Nonetheless, we investigate below the

⁴⁵ 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.

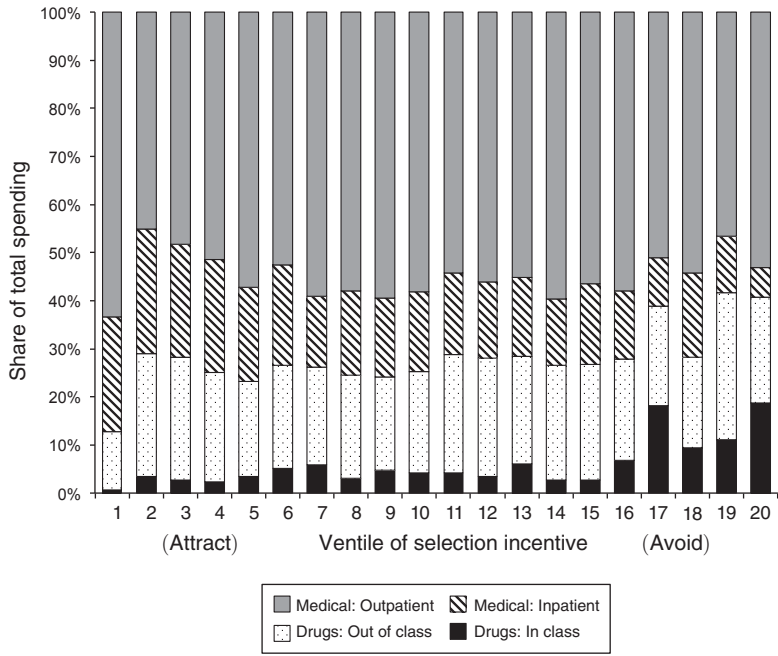


FIGURE 4. DETERMINANTS OF ENROLLEE COSTS BY SELECTION-INCENTIVE STRENGTH

Notes: This figure decomposes total enrollee costs into inpatient, outpatient, and drug costs. Drug costs are divided according to whether the drug is inside or outside of the defining therapeutic class. Each of the 220 therapeutic classes are ranked according to the strength of the selection incentive and then binned into 20 ventiles of the incentive measure. Classes are associated with increasingly unprofitable patients moving from left to right.

extent to which insurers appear to be unsophisticated in the sense of overresponding to class-specific costs, rather than the bottom line impact on (our proxies for) profits.

Online Appendix Figure A3 provides a summary of the three alternative measures of the selection incentive. There, we plot histograms of the level, ratio, and Ellis-McGuire measures of S_c for the 220 classes. This class-level variation constitutes our identifying variation. Consistent with Figure 2, 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, providing the necessary conditions for us to test how insurers design formularies in response to payment errors.

Insurers may approximate profit-maximizing behavior in ways that align with any of the three measures defined in (1). Although the measures are correlated, they do contain some independent information. To give a sense of the information overlap, in online Appendix Figure A4, 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.

B. Regression Model

To test how insurer formulary design responds to payment errors, we exploit two forms of variation in the selection incentive. First, we leverage variation in the selection incentive across drug classes within a plan. Figure 2 shows the extent of this variation. Second, we leverage variation in the selection incentive between the exchange and employer-sponsored insurance markets. Exchange plans and employer-sponsored plans plausibly face similar considerations in balancing risk protection against consumer moral hazard, in steering consumers to cost-effective options, and in responding to other phenomena relevant for efficient benefit design. However, selection incentives differ significantly in the two settings. Even if large, self-insured employers were able to significantly influence their enrollee pool (and there are reasons to believe the scope for such behavior is small), these employer plans do not face the exchange risk-adjustment and reinsurance payment scheme, and so they do not face the screening incentives, S_c , that comprise the identifying variation here.⁴⁶ Thus, we identify insurer formulary design responses to payment errors by comparing the difference in formulary design for drug classes with strong selection incentives to classes with weak selection incentives in exchange plans to the same difference for self-insured employer plans.

To implement this comparison across drug classes and market settings, we estimate difference-in-difference regressions of the following form:

$$(3) \quad Y_{cj} = \beta[S_c \times HIX_j] + \gamma_c + \alpha_j + \epsilon_{cj}.$$

The variable HIX_j is an indicator equal to one if plan j is an exchange plan and zero otherwise.⁴⁷ The γ_c terms are drug-class fixed effects, and α_j are plan fixed effects. The estimation sample includes the universe of exchange plans in 2015 and the large sample of employer plans described in Table 1, with employer plans primarily serving to identify the drug-class fixed effects. Observations are at the plan \times state \times class level. Thus, the dependent variable for formulary restrictiveness, Y_{cj} , describes the fraction of drugs within class c in plan j that were placed on any of the tiers we label as restrictive. The primary measure of formulary restrictiveness is the fraction of drugs in c that are tiered specialty or higher, which includes being left off of formularies altogether, though we examine results for other definitions as described below. See Table 1 for the complete rank ordering of tiers.

The parameter of interest in this equation is β , the correlation between the selection incentive and formulary generosity in exchange plans after differencing out this same correlation among ESI plans. In most tables, we present OLS estimates of (3), though we additionally present semi-parametric versions in several figures. To facilitate interpretation of β , in all regressions we standardize S_c by subtracting its mean and dividing by its standard deviation. This places results for the various

⁴⁶In self-insured employer-sponsored plans, the insurer (the employer) cannot so easily avoid its costly enrollees, as these enrollees are closely tied to the firm/plan via the employment relationship. Because of this, we effectively assume that self-insured employers face no selection incentives.

⁴⁷Inclusion of the HIX_j is redundant because S_c is zero for ESI plans. The notation is intended to emphasize that we allow the selection incentive to impact design in HIX plans only.

operationalizations of the selection incentive on a comparable (z -score) scale. 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.

C. Identification

The ESI population is a natural choice for a control group here. The MarketScan sample of ESI plans, in particular, is what CMS chose to use when calibrating risk adjustment for the exchanges, determining it to be the most comparable population for which claims data was available. Nonetheless, identifying exchange plans' responses to the screening incentive does not require that exchange and employer plans are equally generous in practice or that they should be equally generous in terms of socially optimal design. Plan fixed effects in equation (3) address any differences in overall generosity between exchange and employer plans, so that β is identified by the differential slope $\partial Y_{c,j}/\partial S_c$ within exchange plans relative to within ESI plans. Our strategy also does not require that ESI benefits arrangements achieve the social optimum. Instead, the key assumption is that the exchange payment formula error, which differs across drug classes, is not correlated with some other drug-class characteristic that is relevant for formulary design *and* that differs for a given drug class between the exchange and employer markets. Under the assumption that S_c is orthogonal to other employer-exchange differences influencing formulary design, what we identify is whether exchange insurers are responsive to the financial incentives embedded in the exchange payment errors. If one further assumed that employer plan formularies are (approximately) optimal, then one could further interpret the responsiveness of exchange plans to the screening incentives as deviations away from the optimal contract. Here, however, our intention is only to understand whether—and with what degree of apparent sophistication—insurers respond to these incentives, not to characterize the optimal contract.

To make the identifying assumption concrete, consider a case that would violate it. Assume that the selection incentives in the exchanges (S_c) were correlated with health care costs in employer plans, so that the consumer types that were under-compensated by the exchange payment scheme were the same consumer types who were disproportionately expensive to care for when enrolled in employer plans. (Empirically, this is not the case, as panel B of Figure 2 shows that there is no correlation between S_c and health care costs in employer plans, but the thought experiment highlights the important assumptions.) One might imagine that employers with wage rigidities would wish to discriminate against high medical-cost workers, as high-cost types would drive up total compensation. These firms might use insurance design to help facilitate this employment discrimination. If this were the case, our estimates would be biased. In this case, the bias would cause us to underestimate the impact of the exchange screening incentive, as would any phenomenon leading to correlation between S_c in the exchanges and reduced generosity for class c in the employer setting.

Another potentially relevant issue with respect to identification is whether exchange and employer-enrolled consumers differ in their price elasticities of

demand for prescription drugs. They may. But only *differences* across ESI and exchange enrollees in *class-specific* demand elasticities that happened to be correlated with the *class-specific* errors in the payment formula would violate our identifying assumption. This seems unlikely a priori, especially given evidence presented in Section VI that there is no correlation between our measures of the screening incentives and a set of estimates for class-specific demand elasticities from the Medicare Part D market.⁴⁸

V. Effects on Formulary Designs

A. Main Results

We start by illustrating the basic correlations underlying our main results semi-parametrically. Figure 5 shows how the generosity of coverage differs between exchange and ESI plans according to the strength of the screening incentive associated with each drug class. Each of the three screening incentive measures are presented. To create each figure, we regress formulary restrictiveness on drug-class fixed effects and plan fixed effects and compute the residuals. We then group therapeutic classes c into ventiles according to the strength of each selection-incentive measure S_c . The vertical axes in each panel measure the means of the residuals within each ventile, separately for employer and exchange plans.⁴⁹ In the plots on the left, the horizontal axes measure the mean of S_c within the bin, and in the plots on the right, the horizontal axes are scaled by the ventile bin number (1–20). Residuals corresponding to exchange plans exhibit substantially more noise given the relatively small size of the universe of exchange plans ($n = 501$).⁵⁰

Because risk adjustment succeeds in neutralizing selection incentives for most classes, many of the scatterpoints in panels A, C, and E of Figure 5 are clustered near neutral ($S_c = 0$) along the horizontal axes. Where the incentives diverge from neutrality, so does benefit design. Formulary restrictiveness is significantly different between employers and exchanges at the highest ventiles (in the rightmost bins), with the exchange plans providing much less generous coverage for the drugs that are signals of very unprofitable enrollees. 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, though these differences are smaller than the differences for the classes with the strongest incentives to screen out enrollees.

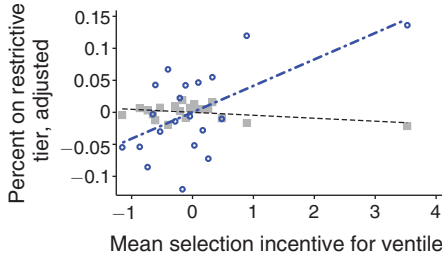
Table 3 presents regression results corresponding to equation (3). We report the difference-in-difference coefficient estimates for the interaction between the

⁴⁸ A related idea is that exchange plans may for some reason be more responsive to the *same* demand elasticities displayed by consumers in employer plans. We investigate this possibility directly, by examining the relationship between S_c and independent estimates of drug-class-specific price elasticities of demand from Einav, Finkelstein, and Polyakova (2018).

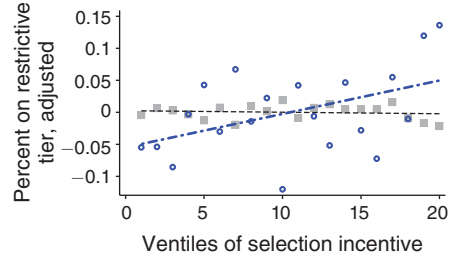
⁴⁹ Each ventile contains around 11 classes. The classes, in turn, contain many individual drugs, but analyses in the paper operate at the level of classes, not at the level of individual drugs.

⁵⁰ Recall from Section II that the definition of an exchange “plan” in this context aggregates various insurance products offered by the same carrier in the same market that share a common formulary.

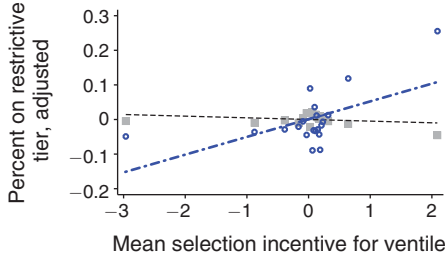
Panel A. Ratio measure of incentive



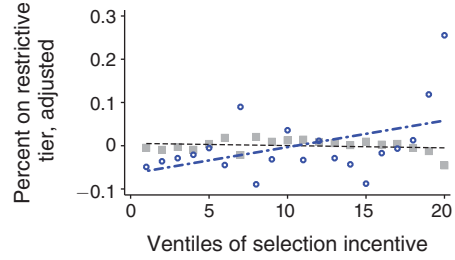
Panel B. Ratio measure of incentive



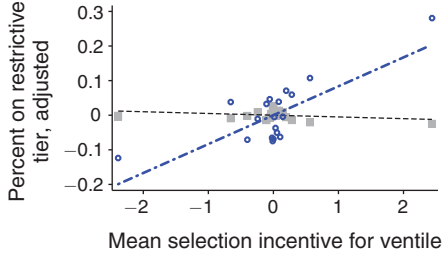
Panel C. Difference measure of incentive



Panel D. Difference measure of incentive



Panel E. E-M measure of incentive



Panel F. E-M measure of incentive

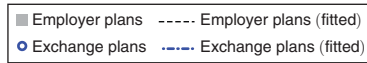
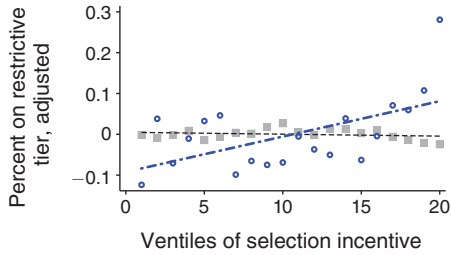


FIGURE 5. RESIDUAL PLOTS OF RESTRICTIVE TIERING VERSUS SELECTION INCENTIVES

Notes: This figure plots residuals from a regression of formulary restrictiveness on drug-class fixed effects and plan fixed effects: $Y_{cj} = \gamma_c + \alpha_j + \epsilon_{cj}$. To generate the plots, therapeutic classes c are grouped into ventiles according to the strength of the selection-incentive measure S_c . We then take the means of the residuals within each ventile, separately for employer and exchange plans. The vertical axes plot these means. The horizontal axes in the left column correspond to the mean of the selection incentive, normalized as a z -score. The horizontal axes in the right column correspond to the ventile number with ventile 20 including the ninety-fifth to one-hundredth percentiles of classes by incentives to avoid. In each panel, an OLS regression line is plotted separately for exchange and employer plans.

exchange dummy and the selection measure, $HIX_j \times S_c$. All regressions include plan and drug-class fixed effects. The selection-incentive variable, S_c , 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 5. In panel B, the dependent variable is the fraction of drugs within the class

TABLE 3—MAIN RESULT: SELECTION INCENTIVE PREDICTS RESTRICTIVE DESIGN IN EXCHANGES RELATIVE TO ESI

Dependent variable:	Fraction of class tiered specialty or higher					
	Ratio		Difference		Ellis-McGuire	
	(cost/revenue)		(cost – revenue)		measure	
Selection-incentive variable:	(1)	(2)	(3)	(4)	(5)	(6)
<i>Panel A</i>						
Exchange	0.046	0.045	0.044	0.012	0.046	0.010
× Selection incentive	(0.014)	(0.022)	(0.017)	(0.014)	(0.018)	(0.015)
Exchange		0.006		0.300		0.296
× Selection-incentive ventile 20		(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 × state × class)	858,440	858,440	858,440	858,440	858,440	858,440
Dependent variable:	Fraction of class tiered prior auth./step therapy/not covered					
	Ratio		Difference		Ellis-McGuire	
	(cost/revenue)		(cost – revenue)		measure	
Selection-incentive variable:	(7)	(8)	(9)	(10)	(11)	(12)
<i>Panel B</i>						
Exchange	0.018	0.031	0.020	0.008	0.018	−0.002
× Selection incentive	(0.011)	(0.016)	(0.011)	(0.011)	(0.010)	(0.014)
Exchange		−0.074		0.108		0.159
× Selection-incentive ventile 20		(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 × state × class)	858,440	858,440	858,440	858,440	858,440	858,440

Notes: This 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 (1). 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 × state × class level. Standard errors are clustered at the level of the therapeutic class (220 clusters).

that require prior authorization or step therapy (PA/ST) or that are explicitly called out on the formulary as “not covered.” Given the possibility of nonlinear effects suggested by the residuals in Figure 5, we present both linear specifications and specifications that allow the relationship to be nonlinear at the top ventile.⁵¹

Table 3 shows that exchange plans tend to provide less generous coverage (more frequent placement on the specialty tier or higher) for drug classes where stronger selection incentives are generated by the payment system. The interpretation of the coefficient in column 1 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 4.6 percentage points in exchange plans relative to employer plans. This is a substantial increase relative to a baseline restrictive tier use of 43 percent in

⁵¹ These additionally include the regressor $HIX_j \times V^{20}$, where V^{20} is an indicator for the class having a selection incentive in the top 5 percent of the distribution of S_c .

employer plans and 59 percent in exchange plans. Coefficients across the linear specifications in panel A are similar, regardless of which of the three incentive measures is used as S_c . For the difference and Ellis-McGuire measures, the nonlinear specifications generate a better fit. The results in column 6 indicate that even controlling for a linear relationship between S_c and restrictiveness, drugs in the top ventile of the selection-incentive measures face an additional 69 percent ($= 0.296/0.43$) probability of being placed on a specialty tier or higher.

One way to put these patterns in context is to consider the difference in formulary structure associated with moving from the ninetieth percentile of profitability among therapeutic classes to the tenth percentile. Employer plan formularies, which are not subject to the exchange payment scheme generating S_c , don't track this incentive: for employer plans the 90/10 selection-incentive difference is associated with going from 43 percent of drugs being placed on a restrictive tier to 45 percent. Among exchange plans, the same 90/10 difference is associated with an increase in restrictive tiering from 53 percent to 76 percent.⁵² To benchmark these effects, we note that the implied difference-in-differences, $(76 - 53) - (45 - 43) = 21$, is a bit larger than the 16 percentage point gross difference in mean plan generosity between employer and exchange plans shown in Table 1.

Although the vast majority of consumers will not fill prescriptions in the most unprofitable classes, these differences can be economically sizable for the consumers affected. Recall that the difference between a nonrestrictive 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 percent coinsurance rate would be an order of magnitude larger than copay-based cost-sharing and could routinely push such patients to their annual out-of-pocket maximum. Such out-of-pocket costs are relevant not only for consumers who are currently chronically ill; they also imply that currently healthy consumers cannot be adequately insured against the negative shock of transitioning to one of the poorly covered chronic disease states. Additionally, these results suggest that if there were no risk adjustment in this market, which would dramatically increase the screening incentives for many drug classes, many additional consumers would face much higher out-of-pocket costs for their drugs.

In panel B of Table 3, we investigate whether plans tailor more than just cost-sharing amounts in responding to S_c . One reason to do so might be the ACA's cost-sharing reduction subsidies (CSRs). About half of exchange enrollees during our time period were enrolled in a CSR variant of a Silver plan. Importantly, CSR-eligible consumers enroll in nominally the same plans as non-CSR consumers and face the same formulary structures. The CSR variants differ in that subsidies paid by the government reduce the consumer's out-of-pocket exposure and increase the actuarial value of the plan. (Table 1 shows that copays at each formulary tier are smaller in CSR plans relative to the corresponding "standard" Silver plan.) This feature limits

⁵² These statistics are based on simple means within the sample, they are not derived from regression coefficients. The ratio measure is used to rank the classes by profitability here.

insurers' ability to rely on consumer cost-sharing—either as a potentially efficient response to moral hazard or as a socially inefficient screening tool. Perhaps this is one reason why, as shown in Table 1, 10 percent of all drugs in exchange plans were subject to prior authorization or step therapy, compared to less than 1 percent among employer plans.

Panel B of Table 3 shows that this overall difference in utilization management varies by drug class and is strongly correlated with patient profitability. Drugs in unprofitable classes are more likely to not be covered or to face prior authorization or step therapy requirements. From column 7, a one standard deviation increase in the strength of the selection incentive increases the percent of drugs that are PA/ST or not covered by 1.8 percentage points on a base of 30 percent. Our findings for these outcomes contrast with results from the Medicare Part D setting, which find no positive relationship between unprofitability and exclusion from the formulary.⁵³ Our findings with respect to utilization management and formulary inclusion provide a warning and counterpoint against policy reactions to the discrimination/screening problem that would attempt to inhibit discriminatory behavior by constraining plans' ability to set cost-sharing. Plans have many (potentially difficult to observe) dimensions along which to increase the shadow price of accessing care.

For completeness, in online Appendix Table A1, we report on a wider variety of nonlinear specifications. In online Appendix Figure A5, we plot semi-parametric versions of the regressions.⁵⁴ Each of these alternative specifications yields the same pattern of results. In all cases, tiering effects are large and significant at the highest ventiles of the selection incentive, which contain the most unprofitable enrollees. The special insight from the nonlinear specifications is that effects do not appear to be symmetric, as the magnitude and statistical significance of effects at the lowest ventiles (containing the most profitable enrollees) are sensitive to specification.⁵⁵ In summary, 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.

B. Insurer Sophistication

It could be the case that insurers are naïvely responding to the *gross* costs of potential enrollees and not actually taking into account the fairly complex risk-adjustment payments that determine *net* profitability. However, panel B of Figure 2 suggests that insurers must be at least somewhat sophisticated in looking beyond

⁵³ Lavetti and Simon (2018) finds small, insignificant effects on utilization management and opposite-signed effects on the probability that a drug is excluded from the formulary. Carey (2017) finds no effect on the probability of exclusion from the formulary and does not examine impacts on the utilization management margins. Nonetheless, low-income subsidy enrollees in Part D are similarly shielded from copays.

⁵⁴ Online Appendix Figure A5 plots binned means of the 220 regression coefficients β_c from the regression $Y_{cj} = \sum_{c \in C} \beta_c \times [S_c \times HIX_j \times I_c] + \gamma_c + \alpha_j + \epsilon_{cj}$, where Y is the fraction of drugs assigned to a restrictive tier and I_c is an indicator for the class. The classes, c , are binned into ventiles of the strength of the screening incentive, S_c , and the means of β are plotted against the means of S_c for each ventile.

⁵⁵ For the most theoretically grounded measure of the selection incentive (Ellis-McGuire), the relationship is significant in both directions with positive coefficients for the top three ventiles (most unprofitable) and a negative coefficient on the bottom ventile (most profitable).

TABLE 4—HOW SOPHISTICATED IS THE INSURER RESPONSE TO SELECTION INCENTIVES?

Selection-incentive variable:	Ratio (1)	Diff. (2)	Ellis-McGuire (3)	Ratio (4)	Diff. (5)	Ellis-McGuire (6)
<i>Panel A. Implied profits and total-costs horserace</i>						
Exchange × Selection incentive	0.051 (0.015)	0.049 (0.016)	0.041 (0.013)	0.062 (0.017)	0.064 (0.018)	0.051 (0.016)
Exchange × Average total cost associated with class	0.042 (0.011)	0.042 (0.014)	0.041 (0.009)			
Exchange × [Indicators for 20 total-cost bins]				X	X	X
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 × state × class)	858,440	858,440	858,440	858,440	858,440	858,440
Selection-incentive variable:	Ratio (7)	Diff. (8)	Ellis-McGuire (9)	Ratio (10)	Diff. (11)	Ellis-McGuire (12)
<i>Panel B. Implied profits and drug-costs horserace</i>						
Exchange × Selection incentive	0.043 (0.013)	0.024 (0.017)	0.025 (0.019)	0.047 (0.011)	0.036 (0.017)	0.028 (0.012)
Exchange × Average drug-only cost associated with class	0.047 (0.013)	0.038 (0.016)	0.036 (0.018)			
Exchange × [Indicators for 20 drug-cost bins]				X	X	X
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 × state × class)	858,440	858,440	858,440	858,440	858,440	858,440
Selection-incentive variable:	Ratio (13)	Diff. (14)	Ellis-McGuire (15)	Ratio (16)	Diff. (17)	Ellis-McGuire (18)
<i>Panel C. Profits, drug costs, and total costs simultaneously</i>						
Exchange × Selection incentive	0.045 (0.014)	0.049 (0.021)	0.049 (0.024)	0.052 (0.012)	0.027 (0.019)	0.024 (0.011)
Exchange × Average total cost associated with class	0.007 (0.013)	0.042 (0.024)	0.039 (0.029)			
Exchange × Average drug-only cost associated with class	0.046 (0.018)	0.001 (0.029)	−0.003 (0.037)			
Exchange × [Indicators for 20 total cost bins]				X	X	X
Exchange × [Indicators for 20 drug-cost bins]				X	X	X
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 × state × class)	858,440	858,440	858,440	858,440	858,440	858,440

Notes: This 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 × state × class level. Standard errors are clustered at the level of the therapeutic class (220 clusters). See Table 3 for additional details.

costs, as there is no strong relationship between costs and profitability overall. In other words, if insurers were attempting to screen out high-cost enrollees without taking into account risk-adjustment and reinsurance payments, we would not observe the correlations we document in Figure 5 between restrictiveness and profitability. In this section, we dig deeper into issues of insurer sophistication.

Table 4 presents a series of “horserace” regressions testing whether formulary design more closely tracks the expected costs or expected profits of potential enrollees. In panel A, we add to the main regression specification controls for the average

total health care spending associated with the therapeutic class c , interacted with the exchange indicator. In panel B, we add controls for the average spending on drugs, again interacted with the exchange indicator. In panel C, we include controls for both average total health care spending associated with the class and average drug-only spending associated with the class. In the leftmost three columns within each panel, we control linearly for the interaction(s). One might think of the coefficients on these additional controls as measuring a naïve version of the selection incentive, in which risk adjustment and reinsurance are not taken into account by exchange plans, or in which the exchange insurer is narrowly focused on the specific costs of the drugs, rather than on the broader signal of profitability implied by medication use. In the rightmost three columns in each panel, we flexibly control for the interaction with indicators for 20 bins of costs.⁵⁶

Table 4 suggests that HIX insurers do in fact respond to gross total and drug-only costs in setting formularies, but simultaneously respond to the expected profits. In columns 1 through 3 and 7 through 9, coefficients on the controls for total and drug costs are statistically significant and in the direction of more restrictive tiering for more expensive patients.⁵⁷ This suggests that relative to employer plans, exchange plans develop formularies that are differentially less generous for the drugs used by the more expensive enrollees. Nonetheless, the coefficients of interest (on $\text{Exchange} \times \text{Selection incentive}$) are generally robust to the inclusion of these controls. In online Appendix Table A2, we present a wider array of specifications that flexibly control for costs in various ways, yielding very similar results. Insurers thus appear to respond to net profitability in addition to, and independently from, gross costs. The fact that there is any response to gross drug costs (relative to employer plans) is consistent both with unsophisticated screening attempts or with the idea that insurers are attempting to cut costs by offering poor coverage for whole classes of expensive drugs.

Although we have framed these results in terms of sophistication, another way to view the Table 4 and online Appendix Table A2 regressions is as a robustness exercise. By estimating effects within sets of drug classes associated with similar spending (essentially within narrow vertical slices of the scatterplot in Figure 2), we rule out the possibility that the results are driven by a correlation between the selection incentive and costs coupled with some differential response to costs in exchange plans relative to employer plans.⁵⁸

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. In panel A of Table 5, we investigate the possibility that popular drugs within each class are more likely to be relegated to restrictive tiers in exchange plans when under-compensated by the payment system. In this

⁵⁶ The specifications are otherwise identical to the linear specifications in panel A of Table 3.

⁵⁷ Simultaneously including drug and nondrug costs in panel C yields less precise coefficients on these two cost controls as they are correlated.

⁵⁸ For example, if exchange plans had an interest in placing a larger burden for expensive treatments on patients, and if costs were correlated with S_c , we could mistake this behavior as screening. These specifications rule out this possibility. Another type of confounder ruled out is that demand elasticities that are differentially increasing in costs across the employer and exchange patient populations, and exchange plans are merely responding to these demand elasticity differences.

TABLE 5—SALIENCE AND SUBSTITUTION: POPULAR DRUGS AND CHEAP DRUGS

Within-class subsample:						
Most popular drugs in class						
Selection-incentive variable:	Seventy-fifth percentile of popularity or higher			Ninetieth percentile of popularity or higher		
	Ratio (1)	Diff. (2)	Ellis-McGuire (3)	Ratio (4)	Diff. (5)	Ellis-McGuire (6)
<i>Panel A</i>						
Exchange × Selection incentive	0.061 (0.022)	0.051 (0.028)	0.081 (0.022)	0.074 (0.025)	0.060 (0.034)	0.098 (0.022)
Therapeutic class FEs	X	X	X	X	X	X
Plan FEs	X	X	X	X	X	X
Therapeutic classes	188	188	188	156	156	156
Observations (plan × state × class)	733,576	733,576	733,576	608,712	608,712	608,712
Least expensive drugs in class						
Selection-incentive variable:	Twenty-fifth percentile of cost or lower			Tenth percentile of cost or lower		
	Ratio (7)	Diff. (8)	Ellis-McGuire (9)	Ratio (10)	Diff. (11)	Ellis-McGuire (12)
<i>Panel B</i>						
Exchange × Selection incentive	0.058 (0.015)	0.049 (0.019)	0.051 (0.020)	0.061 (0.015)	0.047 (0.019)	0.048 (0.020)
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 × state × class)	858,440	858,440	858,440	858,440	858,440	858,440

Notes: This 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 × state × 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, 2, and 3, we limit the sample to the seventy-fifth percentile of popularity or higher within each class (and limit to classes with at least four drugs). In columns 4, 5, and 6, we limit the sample to the ninetieth percentile of popularity or higher within each class (and limit to classes with at least ten drugs). In panel B, we limit the sample to the least expensive drugs in each class when calculating the dependent variable. In columns 7, 8, and 9, we limit the sample to the twenty-fifth percentile of drug prices and below in each class, and in columns 10, 11, and 12, we limit the sample to the tenth 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 (twenty-fifth percentile or tenth 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 × state × class level. Standard errors are clustered at the level of the therapeutic class (220 clusters). See Table 3 for additional details.

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.⁵⁹ Columns 1 through 6 of Table 5 present results for the seventy-fifth and ninetieth percentiles of popularity. At both thresholds, coefficients are larger

⁵⁹For example, to compute the seventy-fifth percentile of popularity for a class in which one drug comprises 30 percent of the consumption share and seven other drugs each comprise 10 percent shares, the dependent variable would be computed only for the single drug with the 30 percent share.

when focusing on the most popular drugs, compared to coefficients applying to the entire class from Table 3. When narrowly focusing on the ninetieth 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 of these drugs for consumers, or reflects insurers themselves displaying a salience bias toward these drugs in formulary design.

All of this evidence is consistent with insurers designing their formularies to screen out unprofitable enrollees. We note that an important input to the theory connecting screening incentives to formulary design is plan demand response of unprofitable consumers to drug coverage. We do not provide direct evidence of this type of relationship between demand, unprofitability, and formulary design. Indeed, this would be difficult to do in the exchange setting as there is no available plan-level enrollment data linkable to patient health conditions and utilization.⁶⁰ Instead, we interpret our results as indirect evidence of the phenomenon described in Section I. We also interpret the salience results described in Table 5 as further evidence for this relationship, given that the effects are stronger for the more popular drugs in each class.

Additionally, the related literature has shown in other contexts a clear relationship between drug coverage and demand for a given insurance plan. In the context of Medicare Part D, Heiss et al. (2013) shows that plan demand in year $t + 1$ is responsive, albeit imperfectly, to the implied cost-sharing of that plan for the individuals' year t prescriptions. Han and Lavetti (2017) provides evidence that enrollment decisions on the Traditional Medicare/Medicare Advantage (MA) margin respond to formulary designs in MA-Part D plans. Of course, consumers may underweight the out-of-pocket costs implied by formularies relative to premiums (Abaluck and Gruber 2011) and may exhibit significant inertia (Ho, Hogan, and Scott Morton 2017; Carey 2018), but even weak consumer responsiveness may make it worthwhile for insurers to be attentive to screening considerations in benefit design.⁶¹

⁶⁰ Even if such data were available, and even if consumers' plan demand was responsive to formulary design, the proper empirical test is not obvious: in a partial equilibrium sense, plans with more generous benefits for unprofitable types should attract them. But it may be the case that in general equilibrium, no plan will wish to offer such benefits, as deviations from that skimpy-coverage strategy would be unprofitable. The question hinges on the nature of the equilibrium and, in particular, whether a separating equilibrium (by drug-class enrollee type) exists in practice. For example, Frank, Glazer, and McGuire (2000) considers a symmetric equilibrium in which all plans offer poor coverage for services demanded by unprofitable types. In that model, there is no net sorting of consumer types across plans. The only common prediction across the models in this literature (Rothschild and Stiglitz 1976; Frank, Glazer, and McGuire 2000; Glazer and McGuire 2000; Veiga and Weyl 2016; Azevedo and Gottlieb 2017) is a reduction in the average generosity of coverage for services that predict unprofitable patients. This is the empirical test Tables 3 and 4 implement.

⁶¹ Regardless of the actual consumer-plan demand elasticity with respect to formulary design, one could interpret the results here as revealing that plans believe consumers will be somewhat responsive to the cost-sharing. For example, consistent with such beliefs, Aetna launched exchange plans for 2016 that were specifically marketed toward diabetics. In these plans, Aetna set low cost-sharing for specialist visits linked to diabetes management in an apparent effort to attract diabetics (Andrews 2015).

VI. Discrimination or Efficient Design?

In this section, we provide evidence that our results cannot be rationalized by differential responses of exchange plans to the availability of cost-effective substitutes within therapeutic classes or to differences in consumer price sensitivity across classes. We also show that the results are not driven by different pharmacy benefit 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 settings.

A. Substitution to Cheaper Drugs and Generics

Insurers have a cost-saving interest in steering consumers to less expensive medications 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).⁶² 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.⁶³

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 online Appendix Table A6, we show that our results hold if we limit attention to classes without generics (28 classes), with less than 10 percent generics (49 classes), or with less than 25 percent generics (84 classes). Thus, our results cannot be driven by HIX plans using stronger nudges toward generics, as the results hold in the absence of a generic alternative.

We also show in online Appendix Table A5 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 when measuring the formulary restrictiveness, we show in panel B of online Appendix Table A5 that the selection incentive significantly predicts

⁶² 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.

⁶³ One possibility that we can not directly address in these data is that inter-temporal spending spillovers (Fang and Gavazza 2011) may be internalized differently in ESI versus HIX plans, due to different rates of enrollee turnover. This could cause drug coverage to differ across plan types. This would pose a problem only if such drug-class-specific dynamic incentives correlated with the error in the HHS risk-adjustment and reinsurance scheme.

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 online Appendix Table A3, which show that HIX plans are ten times more likely than employer 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 (online Appendix Table A7).

Encouraging substitution toward lower cost alternatives may occur along other margins than brand versus generic. To investigate this possibility, in columns 7 through 12 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 percent or 10 percent 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, Table 5 and online Appendix Tables A3, A5, A6, and A7 support the idea that the contract designs we document do not merely reflect HIX plans pushing consumers to lower cost alternatives.

B. Demand Elasticity

As we highlight in Section I, moral hazard, reflected in demand elasticities, is a key consideration in a profit-maximizing insurer's formulary design problem (Einav, Finkelstein, and Polyakova 2018). It is also an important component of the design of a socially efficient health insurance contract (Glazer, Huskamp, and McGuire 2012). 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, a problem for identification could arise if ESI plans were differentially responsive to the same consumer price elasticities, and if these class-specific price elasticities happened to be correlated with class-specific payment errors.

With this in mind, 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. We reestimate our main regressions excluding all fertility treatment classes in online Appendix Table A8. The coefficients of interest are almost numerically identical to our main results.

We also import external measures of class-specific demand elasticities estimated by Einav, Finkelstein, and Polyakova (2018), who identify price sensitivity of prescription drug utilization by exploiting Medicare Part D's "donut hole" at which drug cost-sharing changes abruptly. Online Appendix Figure A6 graphs scatterplots of these elasticities versus the strength of the selection incentive, revealing

that there is no significant correlation between S_c and consumer price sensitivity across classes. We describe this exercise in detail in online Appendix D, where we also show that our main results are robust to directly controlling for class-specific demand elasticities in our regressions.

C. 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.⁶⁴ 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 one 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 ($\mathbf{1}(PBM_p) \times S_c$) when estimating our coefficient of interest (β):⁶⁵

$$(4) \quad Y_{cj} = \beta[S_c \times HIX_j] + \sum_p \nu_p [\mathbf{1}(PBM_p) \times S_c] + \gamma_c + \alpha_j + \epsilon_{cj}.$$

Under this specification, β is identified off of differences between exchange plans and employer plans that use the same pharmacy benefit manager.

Online Appendix Table A9 displays these results, again separately for each selection measure. We present two versions. In columns 1 through 4, we estimate equation (4) 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 control for PBM-by-state fixed effects, so that the control is defined as $[\mathbf{1}(PBM_p) \times state_s \times S_c]$. 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 online Appendix Table A9 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 different responses to (or different biased subjective beliefs about) consumer moral hazard across different plan benefits architects.

⁶⁴ 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 I.

⁶⁵ PBM main effects are absorbed by plan fixed effects.

VII. Reinsurance Counterfactual

In 2017, the reinsurance program that had operated during the first three years of the exchanges' existence ended.⁶⁶ In Section III, we discuss how insurer screening incentives change with the removal of the reinsurance program, holding all else fixed. Intuitively, reinsurance absorbs some part of the underpayment errors that risk adjustment leaves behind, so its removal makes incentives worse for the expensive classes that were already associated with unprofitable consumer types. In this section, we simulate how insurers might be expected to alter their formularies in response to the removal of reinsurance. We proceed with an important caveat. Because insurers were possibly learning about their enrollee pools, costs, and profits over time in the first years of these new markets, because there was significant policy uncertainty regarding the regulatory and payment system rules for the exchanges in 2017, and because these markets were in apparent disequilibrium in 2017, with insurers exiting markets in several states, we present these counterfactuals not as predictions for 2017, but as an alternative approach to providing intuition for the magnitudes of the screening phenomenon.

Table 6 shows how predicted restrictiveness differs with and without reinsurance. To compute these measures, we use the empirical models estimated in the odd columns of Table 3. The “with reinsurance” columns in Table 6 simply generate predicted values from the regressions, using the same measures of the screening incentive S_c used to estimate the model. For the “without reinsurance” columns, we recalculated the screening incentives to not include reinsurance payments or fees (see Figure 3) and used these alternative values of S_c along with the original coefficient estimates to generate predictions.⁶⁷ In panel A, we take the mean across all drug classes to produce the overall average predicted percent of drugs on a restrictive tier. In panel B, we describe the distribution of changes in formulary generosity under the counterfactual. Percentiles in panel B are at the level of the drug class, with classes ordered according to the change in the fraction of drugs in the class predicted to be on the restrictive tier when moving from a payment system with reinsurance to one without it. Positive values in panel B indicate more drugs on a restrictive tier under the no-reinsurance policy regime.

The results reported in panel A of Table 6 suggest that the counterfactual removal of reinsurance in the 2015 markets would lead to an average increase in the percent of drugs on a restrictive tier of around 2.5 percentage points and an average increase in the percent of drugs either excluded from the formulary entirely or with prior authorization or step therapy requirements of a little over 1 percentage point. This is true despite it being the case that for some therapeutic classes, the removal of (budget neutral) reinsurance reduces the overpayment, and brings the class closer to neutrality (the 45-degree line in Figure 3). To put these effects of reinsurance in context, recall from Table 1 that for exchange plans around 59 percent of drugs appear

⁶⁶ As described in detail in Section IB, the ACA's mandatory reinsurance program paid plans for outlying individual enrollees with high realized spending.

⁶⁷ Consistent with the actual functioning of the exchanges, we do not alter the risk-adjustment algorithm when removing reinsurance. Parameters for these payment mechanisms were set independently, despite that optimization requires setting them jointly (Layton and McGuire 2017).

TABLE 6—COUNTERFACTUAL: PREDICTED FORMULARY RESTRICTIVENESS WITH AND WITHOUT REINSURANCE						
	Fraction of drugs on restrictive tiers					
	Ratio		Difference		Ellis-McGuire	
	With reins	Without reins	With reins	Without reins	With reins	Without reins
<i>Panel A. Mean effects</i>						
Percent on restrictive tier	0.56	0.59	0.46	0.48	0.46	0.49
Percent PA/ST/NC	0.17	0.18	0.13	0.14	0.13	0.14
	Fraction of drugs on restrictive tiers (without reinsurance minus with reinsurance)					
	Ratio		Difference		Ellis-McGuire	
	With reins	Without reins	With reins	Without reins	With reins	Without reins
<i>Panel B. Distributional effects</i>						
First percentile	−0.01		0.00		0.00	
Fifth percentile	0.00		0.00		0.00	
Twenty-fifth percentile	0.01		0.00		0.00	
Median	0.01		0.00		0.00	
Seventy-fifth percentile	0.02		0.01		0.01	
Ninety-fifth percentile	0.06		0.10		0.10	
Ninety-ninth percentile	0.15		0.45		0.70	

Notes: This table reports results from a counterfactual exercise in which we simulate the removal of reinsurance. To perform the simulation, we first estimate the relationship between formulary restrictiveness and the screening measures using the model presented in Table 3. We then use the coefficients from the model to predict formulary restrictiveness for each drug class with reinsurance. We then recalculate the screening incentive S_c without reinsurance and use the same coefficient estimates with the new S_c s to predict formulary restrictiveness without reinsurance. We present results separately for formulary restrictiveness measured by (i) the fraction of drugs tiered specialty or higher and (ii) the fraction with a non-price barrier such as prior authorization, step therapy, or the drug not being covered on the formulary (PA/ST/NC). The table reports means with and without reinsurance for each of the three screening measures, as well as statistics describing the distribution in the predicted change in restrictiveness with and without reinsurance.

on a restrictive tier while for employer plans around 43 percent of drugs appear on a restrictive tier, a difference of 16 percentage points, implying that reinsurance accounts for a sizable portion of the difference between ESI and exchange formulary restrictiveness. The results in panel B of Table 6 suggest that the small overall mean difference masks significant heterogeneity. Classes at the ninety-fifth percentile of the distribution of changes experience increases in restrictiveness between 6 and 10 percentage points and classes at the ninety-ninth percentile experience increases in restrictiveness between 15 and 70 percentage points. These changes represent significant shifts in access to the associated drugs for exchange plan enrollees, suggesting that reinsurance plays an important role in the ACA plan payment system’s ability to maintain access to drugs for some groups of enrollees, given the current risk-adjustment system.

VIII. Discussion

We find that the ACA’s risk-adjustment and reinsurance transfer payments are successful in neutralizing selection incentives for most, but not all, consumer types. Where payment errors exist, plans offer poor coverage for the medications demanded by the corresponding patients. The distortions we document currently affect a relatively small number of consumers because only a small set of consumer types

are unprofitable, net of the ACA's payment adjustments.⁶⁸ Our findings show that insurers are responsive to complex screening incentives where they exist, implying that it is only a well-functioning risk-adjustment and reinsurance system that keeps market forces from eroding access to care for all costly patients. Further, the cost of these contract distortions is not solely borne by patients currently burdened by the targeted chronic conditions. While that is an important and potentially sizable distributional consequence, a welfare loss arises because in equilibrium consumers cannot be adequately insured against the negative shock of transitioning to the poorly covered chronic disease states.⁶⁹ This lack of risk protection can affect the utility of all consumers.

These 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) who are likely to bear high out-of-pocket costs. 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.⁷⁰

Our results also bear on the use of essential health benefits (EHB) rules as a means of guaranteeing health care access for certain services or patient types.⁷¹ The ACA's EHBs mandate coverage of items and services in ten benefit categories, including prescription drugs. However, we find that even while complying with these coverage mandates, plans design formularies that are significantly unfavorable to consumer types that are unprofitable.⁷² Our findings thus suggest that simply “strengthening” the list of mandated benefits, such as by mimicking Medicare Part D's protected class regulations, will not solve the problems documented here.⁷³ Plans have many tools at their disposal to limit coverage because their products are highly multidimensional. 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 making it difficult to work with those

⁶⁸ Some previous literature has also found the overall consequences of adverse selection to be small (Cardon and Hendel 2001).

⁶⁹ This transition risk parallels the premium reclassification risk discussed by Handel, Hendel, and Whinston (2015).

⁷⁰ Upstream prices would nonetheless 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.

⁷¹ Setting aside questions of efficacy in combatting selection, excessive minimum coverage requirements can have the negative welfare consequence of limiting the insurer's ability to 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).

⁷² See Andersen (2017) for further description and for evidence that the number of drugs covered is affected by the EHBs.

⁷³ Medicare Part D requires plans to cover a minimum of two drugs per USP class and designates six protected classes for which plans are required to cover all drugs.

pharmacies. As a corollary, our findings suggest that weakening or eliminating minimum coverage requirements, which has been proposed in various Republican plans to repeal and replace the ACA, need not expose chronically ill consumers to significant financial risk. But this is true only if strong selection-related regulations, including risk adjustment and possibly reinsurance, are in place.⁷⁴

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 (Cosby 2016). Incorporating drug utilization into the risk-adjustment system would be novel in the United States at the national level. Given experience with diagnosis-based risk adjustment, any such reform should be conducted with careful attention to impacts on administrative burden, impacts on cost-control incentives, and the potential for gaming by insurers, which has been shown to be an empirically relevant phenomenon (Geruso and Layton forthcoming).⁷⁵ 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.

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. Our results demonstrate that in such settings, ensuring non-discrimination requires more than prohibiting overt discrimination and mandating minimum essential health benefits. It requires getting the plan payment incentives right. Therefore, despite the inherent imperfections in using risk scores to govern transfer payments across plans with different features and cost structures (Einav et al. 2016), despite risk adjustment's distortionary incentives to code patients intensely (Geruso and Layton 2015), and despite risk adjustment's weakened incentives to restrain overall health care spending relative to pure capitation (Geruso and McGuire 2016), risk adjustment and reinsurance play a crucial role in mitigating screening and selection problems endemic to competitive insurance markets.

REFERENCES

- Abaluck, Jason, and Jonathan Gruber.** 2011. "Choice Inconsistencies among the Elderly: Evidence from Plan Choice in the Medicare Part D Program." *American Economic Review* 101 (4): 1180–1210.
- Aizawa, Naoki, and You Suk Kim.** 2018. "Advertising and Risk Selection in Health Insurance Markets." *American Economic Review* 108 (3): 828–67.
- Akerlof, George A.** 1970. "The Market for 'Lemons': Quality Uncertainty and the Market Mechanism." *Quarterly Journal of Economics* 84 (3): 488–500.

⁷⁴ See Layton and McGuire (2017) for a discussion of how risk adjustment and reinsurance in the exchanges may be modified within the current regulatory framework to generate payments that better match person-level costs for outlier enrollees.

⁷⁵ Geruso and Layton (2015) shows that patients' reported diagnoses are endogenous to the plan in which they are enrolled in the context of Medicare Advantage. Drug utilization may be similarly endogenous.

- Andersen, Martin.** 2017. "Constraints on Formulary Design under the Affordable Care Act." *Health Economics* 26 (12): e160–78.
- Andrews, Michelle.** 2015. "New Health Plans Offer Discounts for Diabetes Care." *Kaiser Health News*, November 17. <http://khn.org/news/new-health-plans-offer-discounts-for-diabetes-care>.
- Azevedo, Eduardo M., and Daniel Gottlieb.** 2017. "Perfect Competition in Markets with Adverse Selection." *Econometrica* 85 (1): 67–105.
- Bhattacharya, Jay, and M. Kate Bundorf.** 2009. "The Incidence of the Healthcare Costs of Obesity." *Journal of Health Economics* 28 (3): 649–58.
- 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, edited by Mark V. Pauly, Thomas G. McGuire, and Pedro P. Barros, 691–762. Amsterdam: North-Holland.
- 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): 3335–64.
- Cardon, James H., and Igal Hendel.** 2001. "Asymmetric Information in Health Insurance: Evidence from the National Medical Expenditure Survey." *RAND Journal of Economics* 32 (3): 408–27.
- Carey, Colleen.** 2017. "Technological Change and Risk Adjustment: Benefit Design Incentives in Medicare Part D." *American Economic Journal: Economic Policy* 9 (1): 38–73.
- Carey, Colleen.** 2018. "Sharing the Burden of Subsidization: Evidence on Pass-Through from a Payment Revision in Medicare Part D." Unpublished.
- Centers for Medicare and Medicaid Services.** 2015. *Summary Report on Transitional Reinsurance Payments and Permanent Risk Adjustment Transfers for the 2014 Benefit Year*. Washington, DC: 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.** 2017. "Do Health Insurers Innovate? Evidence from the Anatomy of Physician Payments." *Journal of Health Economics* 55 (1): 153–67.
- Cosby, Steven G.** 2016. *March 31, 2016, HHS-Operated Risk Adjustment Methodology Meeting—Discussion Paper*. Washington, DC: Department of Health and Human Services.
- Decarolis, Francesco, and Andrea Guglielmo.** 2017. "Insurers' Response to Selection Risk: Evidence from Medicare Enrollment Reforms." *Journal of Health Economics* 56: 383–96.
- DeLeire, Thomas, Andre Chappel, Kenneth Finegold, and Emily Gee.** 2017. "Do Individuals Respond to Cost-Sharing Subsidies in Their Selections of Marketplace Health Insurance Plans?" *Journal of Health Economics* 56: 71–86.
- 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, Amy Finkelstein, Raymond Kluender, and Paul Schrimpf.** 2016. "Beyond Statistics: The Economic Content of Risk Scores." *American Economic Journal: Applied Economics* 8 (2): 195–224.
- Einav, Liran, Amy Finkelstein, and Maria Polyakova.** 2018. "Private Provision of Social Insurance: Drug-Specific Price Elasticities and Cost Sharing in Medicare Part D." *American Economic Journal: Economic Policy* 10 (3): 122–53.
- Ellis, Randall P., and Thomas G. McGuire.** 2007. "Predictability and Predictiveness in Health Care Spending." *Journal of Health Economics* 26 (1): 25–48.
- Fang, Hanming, and Alessandro Gavazza.** 2011. "Dynamic Inefficiencies in an Employment-Based Health Insurance System: Theory and Evidence." *American Economic Review* 101 (7): 3047–77.
- Feldstein, Martin S.** 1973. "The Welfare Loss of Excess Health Insurance." *Journal of Political Economy* 81 (2): 251–80.
- Frank, Richard G., Jacob Glazer, and Thomas G. McGuire.** 2000. "Measuring Adverse Selection in Managed Health Care." *Journal of Health Economics* 19 (6): 829–54.
- Geruso, Michael, and Timothy J. Layton.** 2017. "Selection in Health Insurance Markets and Its Policy Remedies." *Journal of Economic Perspectives* 31 (4): 23–50.
- Geruso, Michael, and Timothy Layton.** Forthcoming. "Upcoding: Evidence from Medicare on Squishy Risk Adjustment." *Journal of Political Economy*.
- Geruso, Michael, Timothy Layton, and Daniel Prinz.** 2019. "Screening in Contract Design: Evidence from the ACA Health Insurance Exchanges: Dataset." *American Economic Journal: Economic Policy*. <https://doi.org/10.1257/pol.20170014>.
- 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.

- 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 and Policy* 15 (2): 1558–95.
- 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): 1055–71.
- 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): 1030–66.
- Han, Tony, and Kurt Lavetti.** 2017. "Does Part D Abet Advantageous Selection in Medicare Advantage?" *Journal of Health Economics* 56: 368–82.
- Handel, Ben, Igal Hendel, and Michael D. Whinston.** 2015. "Equilibria in Health Exchanges: Adverse Selection versus Reclassification Risk." *Econometrica* 83 (4): 1261–1313.
- Heiss, Florian, Adam Leive, Daniel McFadden, and Joachim Winter.** 2013. "Plan Selection in Medicare Part D: Evidence from Administrative Data." *Journal of Health Economics* 32 (6): 1325–44.
- Ho, Kate, Joseph Hogan, and Fiona Scott Morton.** 2017. "The Impact of Consumer Inattention on Insurer Pricing in the Medicare Part D Program." *RAND Journal of Economics* 48 (4): 877–905.
- 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.
- Kuziemko, Ilyana, Katherine Meckel, and Maya Rossin-Slater.** 2018. "Does Managed Care Widen Infant Health Disparities? Evidence from Texas Medicaid." *American Economic Journal: Economic Policy* 10 (3): 255–83.
- Lavetti, Kurt, and Kosali Simon.** 2018. "Strategic Formulary Design in Medicare Part D Plans." *American Economic Journal: Economic Policy* 10 (3): 154–92.
- Layton, Timothy J., Randall P. Ellis, Thomas G. McGuire, and Richard van Kleef.** 2017. "Measuring Efficiency of Health Plan Payment Systems in Managed Competition Health Insurance Markets." *Journal of Health Economics* 56 (1): 237–55.
- Layton, Timothy J., and Thomas G. McGuire.** 2017. "Marketplace Plan Payment Options for Dealing with High-Cost Enrollees." *American Journal of Health Economics* 3 (2): 165–91.
- 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): 1736–44.
- Pauly, Mark V.** 1968. "The Economics of Moral Hazard: Comment." *American Economic Review* 58 (3): 531–37.
- 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): 629–49.
- Shepard, Mark.** 2016. "Hospital Network Competition and Adverse Selection: Evidence from the Massachusetts Health Insurance Exchange." National Bureau of Economic Research (NBER) 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. "Externalities and Benefit Design in Health Insurance." National Bureau of Economic Research (NBER) 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, edited by Anthony J. Culyer and Joseph P. Newhouse, 755–845. Amsterdam: North-Holland.
- Veiga, André, and E. Glen Weyl.** 2016. "Product Design in Selection Markets." *Quarterly Journal of Economics* 131 (2): 1007–56.
- 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.