The Role of Insurers in Health Care Spending and Production: Evidence from Utah

Benjamin Handel, Jonathan Holmes, Jonathan Kolstad, and Kurt Lavetti *

November 18, 2025

Abstract

Private health insurers play a central role in determining the cost and quantity of health care in the United States. Despite this centrality, there has been limited empirical work studying how insurers differentially affect spending and consumption, especially for expensive and pervasive chronic conditions. We use hundreds of natural experiments involving employers switching their primary health insurer, together with a movers design, to estimate these causal effects. We find meaningful differences in total cost and prices for medical and pharmaceutical spending. Within-person changes in spending caused by forced reassignment across pairs of insurers, holding fixed plan generosity, are as large as 30% of total medical spending and 37% of drug spending. We also find substantial dispersion in insurer causal effects on health care spending and quantities for patients with diabetes, hypertension, and other chronic conditions. We find strong evidence for drug offsets as insurers who causally increase drug utilization also reduce medical costs and quantities overall and by condition. Finally, we find that employers do not select plans that reduce costs for their employees based on the match between employee conditions and estimated plan treatment effects, leaving significant cost savings on the table.

*Handel: Berkeley, Evans Hall, Berkeley, CA 94720, handel@berkeley.edu. Holmes: University of Ottawa, jonathan.holmes@uottawa.ca. Kolstad: Berkeley, Haas School, Berkeley, CA 94720, jkolstad@berkeley.edu. Lavetti: Ohio State University, lavetti.1@osu.edu. We thank Michael Nguyen-Mason and Seung Sung for invaluable research assistance; Dr. Norman Thurston and the Utah Office of Health Care Statistics for providing data; and seminar participants at Bates White, the Berkeley-Stanford IO Fest, CalTech, Duke, the Chicago Fed, Columbia, Cornell, CRA, Washington University, UC Irvine, UNC, MIT, and the University of Chicago Harris. This research was supported by funding from the Robert Wood Johnson Foundation (72674) and the Ohio State University Institute for Population Research through a grant from the Eunice Kennedy Shriver National Institute for Child Health and Human Development of the National Institutes of Health (P2C-HD058484).

1 Introduction

A defining feature of the US health care system is its reliance on, and great effort to facilitate, market-based, private provision of health insurance. In contrast, many other high-income countries rely on some form of government provision of health insurance. A central argument behind the US choice of insurance model is that competition between private insurers does more than just improve the efficiency of financial risk transfer —it adds value for consumers in the production of health itself (Enthoven, Garber and Singer, 2001). Greater competition between insurers is presumed to create value by forcing insurers to innovate in ways that improve the quality of care, the efficiency of care delivery, or product variety.

However, it remains an open question whether managed competition delivers on these goals or simply transfers rents to private insurers through inefficient differentiation that attracts partially informed consumers. Such rent transfers via private markets may not be costless. A complex web of regulations is necessary to contend with fundamental issues in insurance markets, such as adverse selection (Akerlof, 1970; Rothschild and Stiglitz, 1976), and insurers can extract rents through unintended responses to these regulations. Prior research has made significant strides studying one side of this tradeoff, regulations and insurer responses to them, but has made little progress studying the other side— value generated by differentiated private insurers.

Thus, despite the centrality of private insurer productivity and innovation to the design of the U.S. healthcare system, there is limited causal evidence on what insurers do differently from one another and, subsequently, whether customers choose insurers based on that differentiation (Handel and Kolstad, 2022). Of particular interest is the employer-sponsored health insurance (ESHI) market, which is the health insurance source for the near majority of Americans (48.7% in 2022 according to the Kaiser Family Foundation). A key research challenge, especially in employer-sponsored markets, is that health plans are rarely exogenously assigned.

In this paper, we leverage hundreds of natural experiments in which employers switch insurers to estimate the distribution of insurer treatment effects on aggregate and condition-specific healthcare spending and utilization. We show that assignment to different insurers drives meaningful differences in the quantities, prices, and composition of healthcare delivered to patients. We estimate insurer-specific spending differences and link them to insurer strategies such as quantity rationing, substitution between drugs and medical care, insurer price bargaining, provider network construction, and patient steering within insurer networks. Finally, we assess whether employers, the key intermediary purchaser of insurance in this market, select plans in ways that respond to the meaningful causal cost differences that insurers exhibit in treating patients with different medical conditions.

To do this, we use the Utah All-Payer Claims Data (APCD) from 2013-2015. Utah provides an appealing context to study managed competition, with six private insurers each holding at least a 5% share of the employer-sponsored market. These private insurers are also structurally diverse, including a vertically integrated insurer system (SelectHealth), a public sector non-profit insurer that competes with private insurers in offering coverage to public agencies (Public Employees Health Plan, or PEHP), a nonprofit membership cooperative created under the Affordable Care Act (Arches), and several large national insurers (United, Aetna, Cigna, Regence and Humana). Utah's relatively permissive regulatory environment for insurance market competition suggests insurers have substantial leeway to differentiate.

We analyze private employer-provided insurance enrollees and their dependents using comprehensive data on medical claims, demographics, and insurer identity. A unique feature of the data is that it contains employer and plan IDs. This allows us to study a large set of natural experiments in which employers switch insurers, forcing their employees and dependents to switch if they wish to continue receiving health benefits from the employer. From the perspective of an individual consumer, this forced switch provides exogenous variation to study the impact of insurer assignment on healthcare utilization and spending. To reduce complications related to plan selection, our main analysis focuses on cases where employers offer only one plan before and after the change in insurers, though we relax this restriction in certain analyses. We observe about 50,000 consumers forced to switch insurers, depending on the sample selection criteria.

We evaluate this natural experimental variation using an estimation framework that leverages within-person moves across insurers. The empirical model combines features of event studies with the two-way fixed effects decomposition framework developed by Abowd, Kramarz and Margolis (1999) (and more recent work in health by, e.g., Finkelstein, Gentzkow and Williams (2016)). This model specification allows us to recover treatment effects of insurer assignment under a wide range of potentially nonrandom sorting in labor markets and nonrandom employer selection of insurance plans. The event study component is incorporated to account for short-term systematic disruption effects associated with insurance switches, about which we show new evidence, while identifying the persistent impact of insurers on key outcomes of interest. Our baseline specification controls for plan financial benefit design characteristics (e.g., deductibles, coinsurance) so that the estimated insurer effects exclude the impacts of differences in plan design due to cost-sharing parameters.

We find substantial heterogeneity in estimated insurer treatment effects on both medical and pharmaceutical spending. For medical spending, within-person changes caused by forced reassignment across pairs of insurers, holding fixed plan generosity, are as large as 30% of total medical spending. The largest pairwise difference is between Educators Mutual (EMI), a local insurer, and PEHP, the non-profit providing coverage to public agencies: assignment to EMI reduces spending by 30% (approx. $685 per year) relative to PEHP, conditional on individual fixed effects, demographics, and plan cost-sharing. The range of treatment effects is even larger for pharmaceutical spending, spanning 37% of mean prescription drug spending. In contrast to their treatment effects on medical spending, EMI enrollees have the largest positive effect on drug spending, and PEHP have the largest negative effect. Across the range of plans, many of the pairwise comparisons of insurer treatment effects on medical and drug spending are statistically and economically significant, indicating meaningful causal impacts of insurer assignment.

Combining these effects, across the eight focal insurers that we study, we find a clear negative correlation between the treatment effects on medical and drug spending, indicating that insurers create structures that systematically shift the composition of healthcare delivery, with tradeoffs between medical and drug care. This suggests that managed competition has led to an equilibrium in which insurers are substantially differentiated from each other, and they play an important role in affecting the composition and quantity of medical and drug care provided to patients.

We next examine how prices contribute to spending differences across insurers. We estimate insurer-specific price indices by measuring the average markup charged by healthcare providers to each insurer, controlling for the mix of procedures performed.[1] This approach isolates pure price effects by netting out variation in service quantity and intensity, allowing us to compare prices across insurers while averaging over all contracted providers.

As with total spending, we find a meaningful range of insurer causal effects on prices. We decompose these price effects into three mutually exclusive and exhaustive components. The first is a provider network design effect, which quantifies whether an insurer's network includes providers who are systematically more expensive across all insurers. The second is an insurer bargaining effect, which quantifies whether an insurer negotiates lower prices for the same providers-service pair relative to other insurers. The third is a network steering effect, which quantifies whether insurers can achieve lower prices by steering consumers towards relatively cheaper providers within their established network.

[1] This method is inspired by Relative Value Units from Medicare, but uses privately negotiated prices instead of the Medicare RVU schedule to estimate the relative price of different procedures.

We find substantial heterogeneity in bargaining effects for medical care, suggesting that insurers pay very different prices for the same provider-service pairs. The range of insurer bargaining effects spans 28% of the mean overall price level, with significant differences in many pairwise comparisons. For example, Arches, a non-profit co-op, paid 28% higher bargained prices than EMI, a significant regional carrier focusing on employer markets.[2] Even among national carriers, we find meaningful variation—Cigna paid 12% higher bargained prices than Aetna for the same services.

Network steering generates smaller but still meaningful price differences, spanning about 10% of mean prices. Some insurers appear to use steering as a substitute for bargaining power—for example, Cigna achieves approximately 8% lower prices than Aetna for the same procedures by steering patients to cheaper in-network doctors. While bargaining and steering effects show substantial variation, the network design component of prices is very similar across insurers. This suggests insurers have similar mixes of systematically high and low-cost providers, though the specific providers in each insurer's network differ substantively.

It is clear that insurers are meaningfully differentiated and play important roles in affecting the overall delivery and cost of medical care. A natural question is how the specific care patients receive differs across insurers. We study this question in three parts: first, we examine preventive and high-value care broadly, then analyze prescription drug utilization patterns, and finally investigate treatment patterns for patients with specific chronic conditions.

We find meaningful variation in preventative care utilization across insurers, with treatment effects spanning 10% of average utilization levels for services like preventative office visits and vaccinations.[3] These differences suggest insurers' strategies materially affect the use of basic preventive services.

Similarly, we find large differences across insurers in the use of high-value prescription drugs that can prevent unnecessary hospitalizations. We use the drug classification system developed by Chandra, Gruber and McKnight (2010), which designates drugs as "acute" or "chronic" care drugs if failure to take the drug substantially increases the probability of an adverse health event within 2 or 12 months, respectively. The effects of insurer assignment on the utilization of these drugs are substantial. For example, Arches increases the use of both acute and chronic care drugs by at least 30% relative to PEHP. PEHP, in turn,

[2] Arches went bankrupt in 2015.

[3] We also find substantial differences in the use of low-value care, as defined by Charlesworth et al. (2016), although we lack statistical power to reject the equality of treatment effects for these relatively uncommon procedures.

has substantially higher use of inpatient hospital care for patients with chronic conditions, suggesting consequences of lower drug utilization.

To better understand disease-specific heterogeneity in how insurers manage healthcare delivery for high-need patients, we analyze treatment patterns for five chronic conditions: (i) diabetes, (ii) hypertension, (iii) asthma, (iv) lower back pain, and (v) bronchitis.

We find large condition-specific differences in insurer effects. Insurer effects on medical spending span about \$3,600 per patient per year for diabetes and nearly \$2,000 per patient-year for hypertension. These dramatic differences raise important questions about the mechanisms driving this large variation. We decompose differences in spending and care quantities into mutually exclusive categories: inpatient, outpatient, physician services, and drugs. We find that spending differences for diabetics are largely driven by drug and inpatient spending, with offsetting effects between these categories. However, there is no single pattern that fully explains insurer heterogeneity, nor is it easily explained by structural differences in insurer organizations. On average, insurer-condition spending effects tend to amplify the base heterogeneity in insurer spending effects, though there is meaningful heterogeneity across conditions within insurers, and vice versa.

Insurer differences in treatment effects on drug utilization for diabetics are particularly salient and informative. We show that when diabetic patients are forced to switch insurers there are dramatic shifts in the type of insulin they consume. For example, Aetna and EMI patients, respectively, use 105% and 144% more insulin lispro and 68% and 79% less insulin aspart. These effects provide a useful datapoint showing that forcibly reassigning patients to different insurers substantially alters the practice of medicine they receive.

Broadly, our results highlight that in the US health care system, health insurers do substantially more than provide risk protection for consumers. Insurers play important roles in affecting the quantity, composition, and cost of medical and drug care. Given the large differences in insurer treatment effects overall and for specific chronic conditions, employers who care about controlling costs have strong incentives to consider the types of conditions their employees have when choosing insurers. However, it is an open question whether employers understand that these differences exist and respond to them when selecting plans. This question is of central importance to the functioning of a health insurance system that depends on managed competition. If employers do not respond through their selection, this mutes upstream incentives for insurers to innovate and achieve cost reductions in the marketplace, particularly when these cost reductions are passed through to end users.

To assess this, we quantify employer-level heterogeneity in chronic condition prevalence and use the estimated insurer-condition treatment effects to predict total costs for each employer-insurer pair. We show that there is substantial scope for cost savings: if employers chose their cost-minimizing plans, average medical spending would decrease by 15%. Despite this large potential for savings, the observed plan choices of employers are indistinguishable from random choices (weighted for insurer market shares). Although this back-of-the-envelope calculation suggests potentially large unrealized cost savings, much of the differences in cost are explained by quantities of care rather than prices. More research is needed to assess whether the cost reductions that could be achieved are dynamically efficient or myopic savings. At a minimum, the results show that either (i) employers are unaware of insurer treatment effects on cost, so don't select on them, or (ii) employers know about these treatment effects but don't want to reduce healthcare utilization to achieve cost reductions.

Our study of aggregate healthcare costs and utilization builds on a broad literature, though there is limited evidence on the causal role of insurers in care delivery or differentiation in insurer strategies. Recent work by Abaluck et al. (2021) shows meaningful differentiation between private insurers in Medicare Advantage on mortality, and finds that patients don't select plans based on these differences. While their work establishes the importance of insurer heterogeneity, we focus on identifying specific mechanisms through which insurers may achieve differentiation. Several recent articles study insurer-specific production in Medicaid in cases where consumers are randomized into plans. Geruso, Layton and Wallace (2020) and Wallace (2023) show that private Medicaid plans have meaningful causal effects on medical spending and that these spending changes are achieved through quantity changes (for low- and high-value care) rather than price changes. Notably, as in Abaluck et al. (2021), consumers do not seem to strongly select plans based on their underlying care differentiation. Agafiev Macambira et al. (2022) also studies the causal effects of privatized vs. non-privatized implementation of Medicaid in Louisiana, and finds that privatization led to increases in prior authorization and reductions in drug utilization in a sample of primarily children.

Our work also builds on the literature that considers elements of insurer strategy (e.g., Cutler, McClellan and Newhouse (2000), Gruber and McKnight (2016), Lavetti and Simon (2018), Curto et al. (2019), Cooper et al. (2019), Starc and Town (2019), Brot-Goldberg et al. (2021), Tilipman (2021), Dunn et al. (2023), and Brot-Goldberg et al. (2023)). Notable among these for our paper are Lavetti and Simon (2018) and Starc and Town (2019), who find that Medicare Advantage insurers strategically use increased drug quantities to reduce medical spending when they are responsible for both medical and drug spending. There are also studies that leverage employer changes to study specific kinds of health plan effects (e.g., Brot-Goldberg et al. (2017) and Einav, Finkelstein and Cullen (2010)). This set of papers typically only studies a single employer changing plans rather than all changes in a large region.

2 Data

Our primary data source is the Utah All-Payer Claims Dataset (APCD) from 2013-2015. The APCD is an administrative database constructed by the Utah Department of Health and Human Services. Commercial insurers are required by Utah state law[4] to submit administrative enrollment records and insurance claims to the state Department of Health and Human Services, and the APCD is constructed from these files.[5]

The APCD consists of four main components submitted to the state by insurers following standardized submission guidelines. The first component is a person-month eligibility file, which includes every individual enrolled in a commercial insurance plan each month. This file contains demographic information about each individual (age, race, gender, location), the relationships between individuals who are dependents on another person's plan, and details about sources of coverage. Records for employer-sponsored plans include employer IDs and plan IDs for different plans offered by the employer. Other key components for our analyses include an (anonymized) individual ID, insurer name, group policy sizes, whether the plan is fully insured or self-insured, and whether the plan is linked to a high deductible health savings account.

The second and third components are medical and prescription drug claims. Both claims databases contain charged amounts, negotiated (allowed) amounts, amounts payed by insurers, member liabilities, copayment and/or coinsurance amounts, deductible amounts, and provider (or prescribing physician) identifiers. All amounts are broken out by line-item whenever applicable. The medical claims also contain service codes, dates, and diagnoses, which are reported for inpatient, outpatient, and professional services following similar conventions as Medicare claims. Prescribed drug claims also include drug codes, fill dates, quantities, refills, days supplied, dispensing fees, and pharmacy IDs and locations.

Finally, the fourth component provides detailed information about providers. This includes National Provider Identification (NPI) numbers, provider group IDs, specialties, site

[4]See Utah Health Data Authority Act, Ch 33a

[5]Individuals covered by self-insured employer plans administered through a commercial insurer are included in the data, as they are treated as enrollees of the commercial insurer.

of care information, and the location at which patient care was provided.

In this paper, we focus on the employer-provided segment of the private insurance market in Utah, which makes up approximately 67% of all person-month observations in the APCD. Table 2 presents summary statistics for our two main analysis samples, described below, and the population of employer-sponsored insurance (ESI) plans in the state.[6]

The overall ESI market is split roughly evenly between HMO plans (34%), PPO plans (29%), and POS plans (31%). About one-fourth of plans are high-deductible plans with an associated health savings account.

Utah has a vibrant health insurance market, with a large number of insurers. This profile of insurers makes the Utah context especially well-suited for our study: the market is less concentrated than in most states, with many of the different kinds of insurers we see around the country simultaneously represented.

The largest insurer is Select Health with a 39% share of the employer-sponsored market among in-sample insurers.[7] Select is vertically integrated with Intermountain, a well-regarded health care delivery system, but the relationship is not exclusive, and other insurers also contract with Intermountain. There are three large national carriers in our sample, including Aetna (22%), Cigna (11%), and Humana (1%). The regional BCBS carrier, Regence, has a 13% share.

In addition to these carriers, there are several additional carriers with different models. The Public Employees Health Program (PEHP) is a non-profit plan for public employees in different sectors (e.g., school districts can contract with PEHP). This plan has 12% of the ESI market. EMI Health is a regional non-profit focused on the employer market segment, covering 2% of the market. Finally, our study period also saw the entry of a new cooperative insurer created under the ACA, Arches Health Plan, which quickly attracted about 3% of the overall market, with roughly half of its enrollees in the ESI segment. Arches abruptly went bankrupt and shut down just after our study period (in 2016).

2.1 Sample Construction

Our research design is based on natural experiments in which individuals switch health insurance plans. To construct the analysis samples, we begin with all Utah residents who

[6]The fifth columns of Tables 2 and 3 include all plans sold in the small group and large group markets. This includes small firms, large firms, and plans purchased by state and local governments.

[7]Note: Summary statistics in Table 2 exclude several insurers that do not meet our data quality standards. The largest of these insurers is UnitedHealthcare, which accounts for approximately 20% of the employer-sponsored market in Utah.

received health insurance from an employer in Utah. This sample includes dependents and spouses who share the same insurance plan as the focal employee.

From this data, we construct four samples. The *forced switchers sample* contains individuals who switched insurance plans because the employer of the focal worker changed insurance plans. The *job movers sample* contains individuals whose insurance plans changed because the focal employee switched jobs.[8] For each of these groups, we also construct a sample of control individuals who did not change insurance companies. For the forced switcher sample, the matched controls are randomly chosen individuals who did not change jobs, and whose employer did not switch insurers. The control group for the job movers sample includes individuals who switched firms but did not change insurance companies.

Table 1: Sample Restrictions

	Forced Switchers		Job Movers	
	Across Insurers	Matched Controls	Across Insurers	Within Insurer
All plan switches	165029		149312	54446
In-sample insurers	103780		75791	39430
No employee plan choice	81416		58171	17755
Balanced panel	60291		15989	5864
Covers prescription drugs	50525	126316	13931	5453

This table reports the number of individuals in the analysis sample. An individual is included in columns 1 or 3 if they have at least one switching "event" that qualifies according to the sample restrictions. Matched control observations in column 2 are randomly selected using a procedure defined in the text, and observations in column 4 include individuals who switched employers but not insurers. The first row, "all employer switches," includes all persons who have switched plans in the employer market at least once (columns 1, 3) and individuals who have switched jobs but not plans (column 4). The second row restricts to in-sample insurers that meet our data quality standards. In the third row, we restrict the sample to cases when subscribers had only one option at their current employer before and after switching insurance plans. In the fourth row, we require all enrollees to have at least six months of coverage before and after their switch and drop all those individuals who switch from subscriber to dependent or vice versa. We also exclude individuals with more than two months of non-insurance between their origin and destination insurers. In the fifth row, we restrict to insurance plans that cover prescription drugs. The matched control sample is required to fulfill the same sample restrictions as the other three columns.

[8] Since we can observe all covered employees at each firm, we can distinguish job changes by individual workers from changes in employer IDs that affect all workers, for example, due to a merger of firms. We exclude from the job movers sample sets of coworkers who change jobs in large, correlated blocks to avoid this latter type of variation.

To construct these four samples, we begin with a dataset of all cases where a person switches from one employer-sponsored insurance plan to another. To this data, we add a randomly selected group of matched controls, which are cases when a person keeps the same insurance plan and does not switch employers. These matched controls or "placebo events" are selected using a random sampling procedure designed to match the characteristics of the forced switcher sample according to the following variables: date of switch, health insurance company, age, gender, and member relationship to the focal employee. This process leaves us approximately 165,000 candidate forced switchers and 149,000 job movers, as shown in the first row of Table 1.

To construct our analysis dataset, we first restrict the sample to the 8 insurers that have meaningful market shares. Our analysis sample includes three large private insurers (Aetna, Cigna, and Humana), two regional non-profit insurers (Regence, Educators Mutual/EMI), a non-profit HMO (SelectHealth), an ACA co-op (Arches), and a government-affiliated plan for public employees called the Public Employees Health Program (PEHP).[9] UnitedHealthcare is a major insurer that operates in Utah, but we exclude them from the analyses because of a known data quality problem in our study years that caused some United claims to be omitted from the APCD. We also exclude several smaller insurers that only cover a small number of employers.

To avoid the impact of changes in plan menus on changes in selection across plans within the same employer, we restrict our main analysis sample to focal workers (and dependents) whose employers offered only one health insurance plan before and after the plan change (shown in row 3 of Table 1). The menu of employer plan options at each firm is inferred using group policy IDs and insurer information.[10] In row 4, we further restrict to events where we observe at least six months before and after the insurance switch. We also exclude individuals whose coverage dependence status changed. This eliminates, for example, households who switch coverage from one spouse's plan to another. We also drop any individuals and their dependents if there is a gap in insurance coverage between switches that exceeds two months. Finally, in row 5 we drop any insurance plan that does not cover prescription drugs.

Table 2 reports summary statistics on insurers and plan types for the main analysis samples. The forced switcher sample includes 50,525 individuals who were forced to switch insurers because their employer switched coverage providers.[11]There are 883 distinct em-

[9]PEHP only covers public employees, but we can include them in the sample because our sample of ESI plans includes public organizations (eg: school boards and local governments)

[10]Since we can only observe plans that at least one employee chooses, our inferred menu of plan options excludes potential options that no employees ever choose.

[11]The table reports the number of individuals who switch. In our estimating equation, the unit of ob-

Table 2: Utah Health Insurance Market

	Forced Switchers Across Insurers	Forced Switchers Matched Controls	Job Movers Across Insurers	Job Movers Within Insurer	All Employer Plans
Insurer:					
Aetna	0.27	0.19	0.22	0.08	0.22
Cigna	0.10	0.12	0.09	0.05	0.11
EMI	0.04	0.01	0.02	†	0.02
Arches	0.04	0.00	0.01	†	0.00
Humana	0.03	0.03	0.04	0.01	0.01
PEHP	0.07	0.05	0.11	0.01	0.12
Regence	0.19	0.21	0.18	0.10	0.13
Select	0.26	0.39	0.33	0.75	0.39
Plan Type:					
HMO	0.30	0.39	0.34	0.75	0.41
PPO	0.31	0.29	0.31	0.12	0.31
POS	0.30	0.20	0.24	0.07	0.20
Other Plan Type	0.09	0.12	0.10	0.05	0.08
High Deductible	0.30	0.31	0.34	0.35	0.34
Observations					
Person-Months	1383766	3390147	362569	143458	24339175
Individuals	50525	126151	13931	5453	1202884
Subscribers	18972	48450	4510	1742	450998
Employers	883	9138	3537	1763	17806
Average Firm Size	308.6	221.2	1153.9	160.4	2048.8

This table reports the market shares of major insurers and the types of plans sold in our primary analysis sample, compared to plans sold by in-sample insurers in the overall employer-sponsored market in Utah. Market shares are reported separately for the forced switcher and job mover samples and all employer-sponsored plans. A HMO is a Health Maintenance Organization, a PPO is a Preferred Provider Organization, a POS is a Point of Service Plan, and "Other" includes plans that do not fall into these categories. The number of employers counts all firms both before and after a switch, so two different firms employ each person in the job movers sample.
† Small cell suppressed for privacy.

ployers, 18,972 focal workers (or subscribers), 31,553 dependents, and a total of 1,383,766 person-month observations. The average employer size (a worker-month level average) is 309. The matched control sample, containing people who did not switch insurers, contains 126,316 individuals.

Our job movers sample includes 13,931 individuals whose coverage switched because the focal worker changed employers. The job movers sample contains 4,510 focal workers, 9,421 dependants, and 362,569 total person-month observations.[12] The control sample of job movers, in Column 4, contains workers 1,742 workers (and 3,711 dependents) who changed jobs but remained covered by the same insurer before and after the switch.

As shown in Table 2, the plan types in both analysis samples are similar to the overall ESI market. The job movers sample has a higher share of individuals covered by SelectHealth, the largest insurer, which primarily sells HMO plans.

2.2 Sample Demographics

Table 3 provides detailed summary statistics on the two analysis samples and the overall ESI market in Utah. The job movers sample is slightly younger than the forced switcher sample and the ESI market, with an average age of 24. This is explained in part by the higher share of dependent children in the job mover sample. Given this younger sample, the average monthly spending on medical care and pharmaceuticals is lower in the job mover sample, around \$240 per month compared to the ESI average of \$266 and forced switcher average of \$275.

About 67% of forced switcher plan-months (which includes both pre and post-switch plans) have annual deductibles in excess of \$1250, compared to 55% of job mover plan-months, and 53% of the ESI market overall. Individuals in the analysis samples also pay a slightly higher share of non-deductible claims out-of-pocket.

2.3 Flows Into and Out of Insurers

Since the focus of our analyses is on plan switchers, we show summary statistics on the rate of plan transitions by insurer in Table 4. Each cell in the transition matrix reports the

servation for this dataset is a switch rather than a person, so an individual who switches insurers multiple times could appear in the data multiple times. It is common for people to switch insurers multiple times, but very few people have more than one switch that satisfies all of our sample criteria. Our final sample includes 52,466 forced switch events across insurers, and 14,095 job move events across insurers.

[12] There are a very small number of individuals who have non-overlapping event windows for both a forced switch and a job move.

Table 3: Analysis Sample Summary Statistics

	Forced Switchers		Job Movers		All
	Across Insurers	Matched Controls	Across Insurers	Within Insurer	Employer Plans
Demographics:					
Age in Jan 2013	28.3	28.7	24.1	24.2	28.1
Fraction Male	0.51	0.51	0.51	0.51	0.51
Member Type:					
Subscriber/Employee	0.37	0.38	0.32	0.32	0.36
Dependent Spouse	0.19	0.20	0.17	0.17	0.20
Dependent Child	0.31	0.37	0.43	0.49	0.40
Other Dependent	0.12	0.049	0.081	0.018	0.052
Deductible:					
No Deductible	0.0061	0.024	0.016	0.013	0.029
1-500	0.074	0.13	0.12	0.12	0.13
501-1249	0.22	0.22	0.24	0.21	0.26
Over 1250	0.67	0.58	0.55	0.59	0.54
Copay/Coinsurance:					
$< 1\%$	0.055	0.050	0.050	0.055	0.030
$1\% - 9\%$	0.38	0.37	0.29	0.39	0.40
$10\% - 19\%$	0.50	0.51	0.55	0.47	0.52
$>= 20\%$	0.056	0.065	0.087	0.074	0.040
Probability of Claim:					
Medical	0.25	0.25	0.24	0.26	0.26
Pharmacy	0.28	0.28	0.25	0.25	0.24
Any Claim	0.39	0.39	0.36	0.37	0.37
Monthly Spending:					
Medical	223.1	228.9	199.8	225.5	243.5
Prescription Drug	51.6	51.9	40.3	39.4	47.0
Total	274.7	280.8	240.2	264.9	290.5
Observations					
Person-Months	1383766	3390147	362569	143458	24339175
Individuals	50525	126151	13931	5453	1202884

This table reports the market shares of major insurers and the types of plans sold in our primary analysis sample, compared to plans sold by in-sample insurers in the overall employer-sponsored market in Utah. Market shares are reported separately for the forced switcher and job mover samples and all employer-sponsored plans. Copay/coinsurance rates report the fraction of health spending not subject to a deductible that is paid by consumers as either co-pays or as coinsurance. Deductible and copay/coinsurance rates are estimated based on the utilization of other enrollees in the same plan.

Table 4: Insurer Transition Matrix

Initial Insurer	Final Insurer Aet.	Arch.	Cig.	EMI	Hum.	PEHP	Reg.	Sel.	Total
Aetna	-	2.03	6.38	3.09	0.56	2.30	9.00	12.62	35.98
Arches	0.10	-	†	†	†	†	†	†	0.13
Cigna	3.61	0.62	-	0.03	0.04	0.20	3.10	1.32	8.92
EMI	†	†	†	-	†	†	†	†	†
Humana	0.97	0.31	0.49	0.05	-	0.01	0.40	2.91	5.15
PEHP	0.53	1.43	0.08	0.03	†	-	0.24	3.98	6.28
Regence	4.54	0.28	1.82	0.33	0.39	1.08	-	10.33	18.77
Select	3.20	3.67	3.87	5.26	0.33	4.29	4.13	-	24.75
Total	12.96	8.34	12.66	8.78	1.32	7.89	16.87	31.18	100.00
	(8623)	(5550)	(8427)	(5846)	(878)	(5254)	(11229)	(20754)	(66561)

This table represents the transition probabilities of all analysis sample observations, including forced switchers and job movers, but excluding all control observations. Each cell is a percentage of the overall analysis sample, and total sample sizes are reported in parentheses below each column.
† Small cell suppressed for privacy.

share of the combined forced switcher and job mover samples associated with that switch, not including control observations. For example, the first row indicates that 2.03% of the analysis sample switched from an Aetna plan to an Arches plan.

The transition matrix describes a dynamic and interconnected insurance market. Most insurers are linked to each other through both leavers and joiners. The clear exceptions are Arches and EMI. Arches is a new co-operative insurer established in 2014, and so was not old enough to have a meaningful number of individuals and firms switching out of Arches before the end of our sample period in June 2015. Similarly, EMI is a small insurance company gaining market share over this period. Since our analysis strategy benefits from a dense network of interconnected firms, we perform robustness checks in Appendix Figure A9 in which we exclude these growing insurers.

3 Empirical Approach

Our empirical approach uses the natural experimental variation from insurance switchers to estimate the causal effect of insurer assignment on healthcare spending. Our estimating equation combines features of event study models with the statistical decomposition framework developed by Abowd, Kramarz, and Margolis (AKM, 1999). Our primary estimating equation is:

$$
\begin{aligned}
y_{ist} &= \psi_{J(i,t)} + \rho AV_{P(i,t)} + \sum_{k \neq -2} \left[\gamma_k^{FS} * FS_{is} + \gamma_k^{JMA} * JMA_{is} \right. \\
&+ \left. \gamma_k^{JMW} * JMW_{is} \right] + \tau_t + \theta_{is} + \epsilon_{ist} \qquad (1)
\end{aligned}
$$

where y_{ist} is the spending of person i in switch event s in month t. Although the vast majority of people in the sample only switch once, we allow the switch event to vary within person for those who switch multiple times.[13]

Our primary outcome of interest, $\psi_{J(i,t)}$, is the effect of being assigned to insurer J on the average spending of enrollees, where the subscript $J(i,t)$ is an index function that maps to insurer J in whose plan person i is enrolled in month t. These insurer effects reflect several dimensions upon which insurers differentially affect healthcare spending, including the design of provider networks, incentive structures with providers (such as pay for performance), and bargaining in price negotiations.

We remove the impacts of plan design features by including a vector $AV_{P(i,t)}$, which contains a set of indicator variables for categories of plan deductible levels ($0, $1-$500, $501-$1249, $1250+, and unknown), and indicators for categories of the realized leave-one-out actuarial value of the post-deductible portion of the plan. These categorical controls (<1%, 1-9.9%, 10-19.9%, ≥20%) quantify the average share of post-deductible claims that are paid out-of-pocket by consumers.[14] ρ captures the effect of cost-sharing plan design features (or "demand side" factors) on spending.

The unit of observation is a switch, when a person changes from one insurer to another. To measure insurer treatment effects, we only include observations for the insurers immediately prior to and immediately after the switch, and observations for any other insurer are dropped. A small number of individuals have more than one switch that satisfies our other sample restrictions (for example, a minimum of 6 months without changing insurers before and after the switch). In these cases, we include both switches as separate events, resulting in a total of 52,466 forced switcher events and 14,095 job mover events. For our matched controls, we require individuals to not have changed insurance companies for at least 12 months.

[13]For each switch s, the analysis data includes only observations where the person i has insurance coverage from the insurer immediately before or after switch s.

[14]For each plan, we measure cost-sharing characteristics for medical spending, pharmaceutical spending, and total (combined) spending. When estimating section 3, actuarial value controls are chosen to match the dependent variable. So when medical spending is the dependent variable, we use medical deductible and coinsurance rate as controls. When pharmaceutical spending is the dependent variable, we use pharmaceutical actuarial value and deductible as controls.

We control for fixed person-by-switch-event effects θ_{is}, which nest person effects and therefore capture static across-person differences in spending. Including switch effects rather than person effects allows spending levels to differ arbitrarily within person across switch events. τ_t are calendar month effects, and ε_{it} is an error term that we discuss below.

The remaining sets of variables are different forms of event time effects. If the effects on healthcare spending caused by forced switches across insurers were instantaneous and static, then the AKM decomposition framework would allow us to use the natural experimental variation caused by switches to infer the distribution of insurer treatment effects, $\psi_{J(i,t)}$. However, we will show new empirical evidence on the dynamics of healthcare spending around the time of insurance switches that suggests there are important time-varying disruption effects caused by insurance switches. Because the rate of switching between pairs of insurers is not perfectly balanced across all insurers and periods, failing to account for these disruption effects would bias the estimates of insurer treatment effects, $\psi_{J(i,t)}$.

To address this issue, we control for the disruption effects by including event-time indicators in Equation 1. We include separate indicators for three (mutually exclusive) types of switchers: Forced switchers (FS, those who change insurers without changing jobs), access-insurer job movers (JMA), and within-insurer job movers (JMW). Event-time dummies for within-insurer job movers allow us to test for any disruption effects from individuals who change jobs but keep the same insurer. Across-insurer job movers allow us to test for disruption effects from individuals who change both jobs and insurers. Finally, the forced switcher sample allows us to test for disruption effects of changing insurers without changing jobs.

The inclusion of the event time controls in the model allows for arbitrary differences in average spending across months in the time window surrounding a change in plans, while allowing for unrestricted correlation between event time effects, the unobserved person or switch-level component, θ_{is}, and the unobserved insurer component, $\psi_{J(i,t)}$. The γ_k^{FS}, γ_k^{JMA}, and γ_k^{JMW} coefficients are each normalized to zero in the second to last month of coverage prior to the plan switch because there is slight evidence of a pre-switch increase in pharmaceutical spending in month $t-1$, consistent with stockpiling, as well as a slight pre-switch decrease in medical spending. Normalizing $\gamma_{-2} = 0$ facilitates visual inspection of the estimates in the absence of these final month anticipatory effects.

Recent literature has documented issues with two-way fixed effects models because of the collinearity between time dummies (τ_t) and event-time dummies (γ).[15] To separately identify these two effects, we include a large control group of individuals who do not switch

[15]See Roth et al. (2023) for a review of the literature on differences in differences estimation.

health insurance companies over a long period, selected to match the characteristics of the population of forced switchers (see section 2.1 for details). For all event-time dummies, we bin endpoints after 12 months, both pre- and post-event. As a further robustness check, in Appendix Section A.2 we use a two-step process which completely avoids the issue of negative weights by de-trending the dependent variable as a first step, and then estimating the equation on the seasonally adjusted variables without τ_t as a second step. We find results that are nearly identical to our main results.

In all figures, we estimate standard errors using the method described in appendix section A.1 to test whether any given insurer is an outlier relative to the Utah average. We cluster standard errors at the person level. To reduce the impact of outliers, we winsorize all health spending values at the 99.9th percentile in our main results.[16]

3.1 Model Assumptions, Identification, and Interpretation

Model 1 allows for many different forms of nonrandom sorting, including sorting of workers into firms in the labor market, or sorting of employers into insurance plans. For example, it allows for the possibility that a worker with a chronically sick child chooses to work at a firm because they offer generous coverage. This type of sorting is simply a correlation between $AV_{P(i,t)}$ and θ_{is}, and does not violate the exogeneity conditions. Similarly, since the model allows unobserved person-level effects in θ_{is} to be correlated with insurer effects and plan design features, the model can also accommodate arbitrary employer-level selection across insurers, or selection into plans of different generosity. This type of employer-level sorting can always be expressed as an aggregation of the individual sorting patterns of the workers at that employer. Therefore, if a firm has abnormally high-spending workers and reduces costs by offering less generous coverage, this type of selection does not violate the exogeneity conditions.

A further potential concern is unmodelled trends in the error term, such as individual health shocks, that could systematically precipitate switches into or out of a particular insurer. To address this, we rely only on plan switches where employers changed the plan they offer from a single option to a different single option. An advantage of using an employer-based switching design is that the timing of individual health shocks is unlikely to be related to the timing of an employer's decision to change plans. This is especially true at medium and large employers, where the relative costs of any particular worker are smaller and unlikely

[16]Winsorization reduces the extent to which results are driven by extremely rare but expensive health events. Appendix Figure A12 shows results from our core specification without winzorization.

to affect firm level plan decisions.[17]

Interpretation of the parameters of interest in Equation 1 depends upon several orthogonality conditions. The key orthogonality conditions are:[18]

$$E\left[av_k\varepsilon\right]=0 \;\; \forall k, \;\; E\left[\theta_{is}\varepsilon\right]=0 \;\; \forall i, \;\; E\left[\psi_j\varepsilon\right]=0 \;\; \forall j \tag{2}$$

where here av_k denotes the kth element of $AV_{P(i,t)}$, θ_{is} is the isth column of the design matrix of indicators for each person-by-switch event, and ψ_j is the jth column of the design matrix of indicators for insurers.

To clarify these conditions, we consider the error term ε as representing the sum of three random effects: an individual-switch-insurer match effect $\Phi_{s(i)J(i,t)}$, a random walk process ξ_{it}, and transitory error ϵ_{it}:

$$\varepsilon_{it}=\Phi_{s(i)J(i,t)}+\xi_{it}+\epsilon_{it} \tag{3}$$

The match effect $\Phi_{s(i)J(i,t)}$ is assumed to be mean zero for every individual-by-switch and every insurer j in the sample. This term represents heterogeneity across individuals in the effect of insurer j on spending, including, for example, condition-specific innovations made by insurers like disease management tools for chronic conditions. To the extent insurers share the same common innovations, the impacts of these innovations on spending would be unchanged when individuals switch insurers, and therefore the effects of common innovations are not included in $\Phi_{s(i)J(i,t)}$ but are contained in θ_{is} and/or other model parameters. Because $\Phi_{s(i)J(i,t)}$ is mean zero, it measures only the systematic component of heterogeneity in spending effects across insurers and across individuals.

The random walk process ξ_{it} is assumed to have zero mean for each individual, capturing idiosyncratic drift in the person-specific component of health spending.[19] The transitory error ϵ_{it} captures other shocks to health spending, and is assumed to be mean zero for each person.[20] A composite health spending error term of this form is consistent with specification tests by French and Jones (2004), who find that log health spending is well-represented by the sum of a very highly persistent AR(1) process and a white noise process. Equation 3

[17]A violation of this assumption implies that decision makers at the firm (e.g. Human Resources) would make firm level insurance decisions based on specific individual's deviations in health in a specific year. This seems difficult to implement and unlikely in practice.

[18]Since τ is not a parameter of interest, and all of the parameters of interest are static, it is not crucial that τ be orthogonal to ε. For example, if a contagious disease outbreak caused temporally correlated increases in ξ_{it}, whether this variation loads onto τ or ξ_{it} is not material for interpreting ρ, θ, or ψ.

[19]Since the model includes both calendar and event time effects, ξ_{it} does not include systematic drift in spending, such as drift due to aging or healthcare inflation.

[20]See Card, Heining, and Kline (2013) for a similar model of earnings processes.

contains a perfectly persistent AR(1) process (ξ_{it}), a white noise process (ϵ_{it}), and a match component.[21]

The first set of orthogonality conditions $E\left[av_k\varepsilon\right]=0\ \ \forall k$ says that the composite error cannot be correlated with changes in av_k. Since variation in av_k is caused by changes in plan design, this condition requires that one individual's realization of ε cannot affect the actuarial value of a plan. Note that when calculating av_k we always use leave-one-out estimates so there is no mechanical effect between a person's spending and the actuarial value of the plan in which they are enrolled. One example of a violation of this condition could be if the high cost of an extremely sick enrollee causes a small employer to change the menu of plan options for all workers at a firm. This type of violation is less likely to happen at medium to large firms, or at firms that pay premiums that are community-rated, as in the small business exchange.

The second set of key orthogonality conditions, $E\left[\theta_{is}\varepsilon\right]=0$, are directly implied by the assumptions that $\Phi_{s(i)J(i,t)}$, ξ_{it}, and ϵ_{it} are mean zero for each person.

The third set of conditions is $E\left[\psi_j\varepsilon\right]=0\ \ \forall j$, which says that insurer assignment is conditionally exogenous. Common forms of insurance selection, like employers choosing plans while taking into account the health status of their workers, do not violate this condition because the model allows unobserved person-by-switch effects to be arbitrarily correlated with insurer identity and plan characteristics. But one potential way in which this condition could be violated is if employers select insurers based in part on $\Phi_{s(i)J(i,t)}$. This type of idiosyncratic complementarity between the health conditions of workers and heterogeneity in the treatment effects of insurers on spending could create problematic correlations between ψ_j, θ_i, and $\Phi_{s(i)J(i,t)}$.[22] We show in Section 5 that such sorting does not occur in practice, despite the potential for employers to substantially reduce costs by sorting into plans based on $\Phi_{s(i)J(i,t)}$. If this type of sorting were to occur, it would also imply that switches into an insurer should have asymmetric effects on spending relative to switches out of the same insurer, conditional on person effects and other covariates. We evaluate this directly by re-estimating a model similar to Equation 1 that includes ordered pair effects for each directional combination of insurer switches. As we discuss in Section 4.6 and show in Appendix Figure A8, insurer spending effects are fairly symmetric, suggesting that the conditional exogeneity assumptions appear reasonable.

[21] In several specifications, French and Jones (2004) fail to reject the hypothesis that the AR(1) parameter is 1.

[22] In the job mover sample this sorting could occur for similar reasons if workers select jobs on the basis of idiosyncratic complementarities between individual health conditions and insurer effects.

A different form of violation of the third restriction could occur if the random walk component, ξ_{it}, is predictive of changes in insurers. For example, if adverse health shocks cause workers to change insurance coverage in systematic ways, this would violate the conditional exogeneity of insurer assignment. Since we restrict the sample to workers employed at firms that only offer one plan option, this type of worker-initiated sorting would require workers to change jobs in order to obtain different insurance coverage. A similar violation could occur if firms change the insurance plan offered to workers because of changes in the firm-level expected value of ξ_{it}. Such a violation is also testable because it implies systematic upward drift in the error term preceding changes in insurance coverage, which would be observable as a pre-trend in the event time parameters. As we show in Figure 1, there is no such evidence of a pre-trend. For the same reason, this evidence is also suggestive that ϵ_{it} and ψ_j are not conditionally correlated.

Overall, these diagnostic tests are consistent with the orthogonality conditions in Equation 1 holding. Under these conditions, we can interpret $\psi_{J(i,t)}$ as the treatment effect on spending of reassigning person i to insurer j.

4 Insurer Treatment Effect Estimates

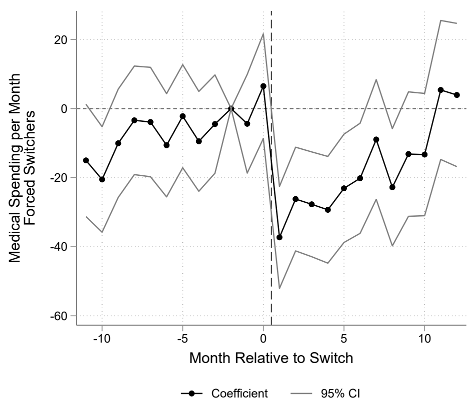
In this section we present our primary empirical results. We begin by showing new evidence on the dynamics of healthcare spending patterns around the time of insurance switches. We then present the main estimates of insurer treatment effects on aggregate spending levels. We then decompose the impacts of prices from quantities in driving insurer spending differences, and further identify the mechanisms by which insurers might affect prices by estimating the contributions of price bargaining, sorting, and network design channels. Finally, we assess variation in the provision of high- and low-value care, overall and for specific chronic conditions of interest, providing estimates of drivers of spending differences as well as measures of clinical quality.

4.1 Healthcare Spending Dynamics around Plan Changes

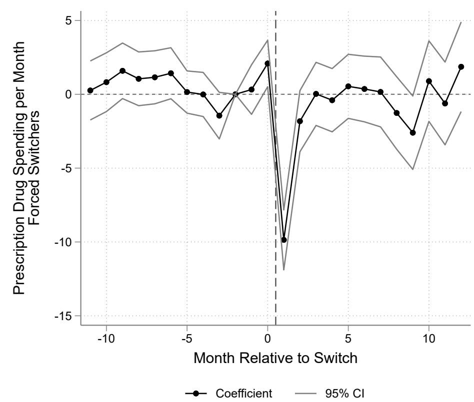
In Figure 1 we show estimates of $\widehat{\gamma_k^{FS}}$, $\widehat{\gamma_k^{JMA}}$, and $\widehat{\gamma_k^{JMW}}$ from Equation 1, separately for medical and drug spending and for the forced switcher and job mover samples. While these disruption effects are not the main focus of our analyses, they are important for demonstrating the need to combine an event study framework within the AKM decomposition model. This is also the first, to our knowledge, evidence of the impacts of insurance plan switches on disruptions in the continuity of healthcare, and it suggests an important avenue for future research on the consequences of these disruptions for health outcomes.

Figure 1: Disruption Effects of Plan Changes on Health Spending

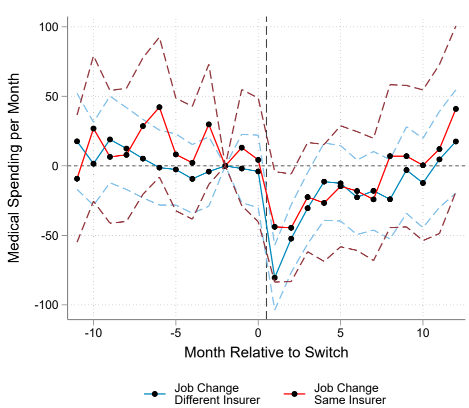
(a) Medical Spending, Forced Switcher Sample



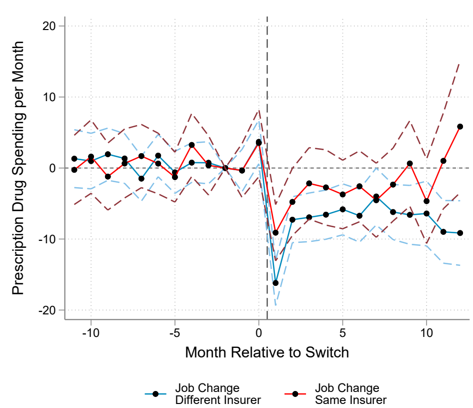
(b) Drug Spending, Forced Switcher Sample



(c) Medical Spending, Job Mover Sample



(d) Drug Spending, Job Mover Sample



Notes: This figure reports event-study coefficients γ_k from Equation 1, separately estimated using the sample of forced switchers and job movers. The dependent variable is total monthly medical spending (panels a and c) and total monthly drug spending (panels b and d). In panels c and d the blue line shows γ_k for workers who changed jobs and switched insurers, and the red line shows estimates for workers who changed jobs and insurance plans but maintained the same insurer.

As the figure shows, there are substantial declines in both medical and drug spending shortly after switches across different insurers, as well as after job switches in which workers change plans but keep the same insurer (the red lines in Figures 1c and 1d.).[23] In the first month following a forced insurer switch, (within-worker) medical and drug spending fall by

[23]In Appendix Figure A1 we also show estimates of $\widehat{\gamma_k}$ from models in which the dependent variable is a binary measure of any medical or drug utilization.

nearly 20% and 25%, respectively. Drug spending rebounds to the pre-change level quickly, by around month 3 after the change. However, medical spending remains lower on average for 10 months after the switch, potentially driven by the need for people to find physicians in their new insurance network.

The inclusion of the γ_k terms in Equation 1 allows us to separate the persistent treatment effects of insurer assignment from these temporary disruption effects. Separating these effects can be important for identifying ψ_j if switches across pairs of insurers are imbalanced, so that some insurers are overrepresented in the post-switch periods.

4.2 Impact of Insurer Assignment on Aggregate Spending

Figure 2 presents estimates of our primary parameters of interest from Equation 1, insurer treatment effects $\widehat{\psi_j}$. For both medical and drug spending, we find substantial heterogeneity in treatment effects across insurers.

We estimate that within-person changes in medical spending caused by forced reassignment across pairs of insurers, holding fixed plan generosity, are as large as 30% of average medical spending.[24] Many of the pairwise comparisons between insurer treatment effects on medical spending are statistically significantly different from each other and imply economically meaningful causal spending impacts.

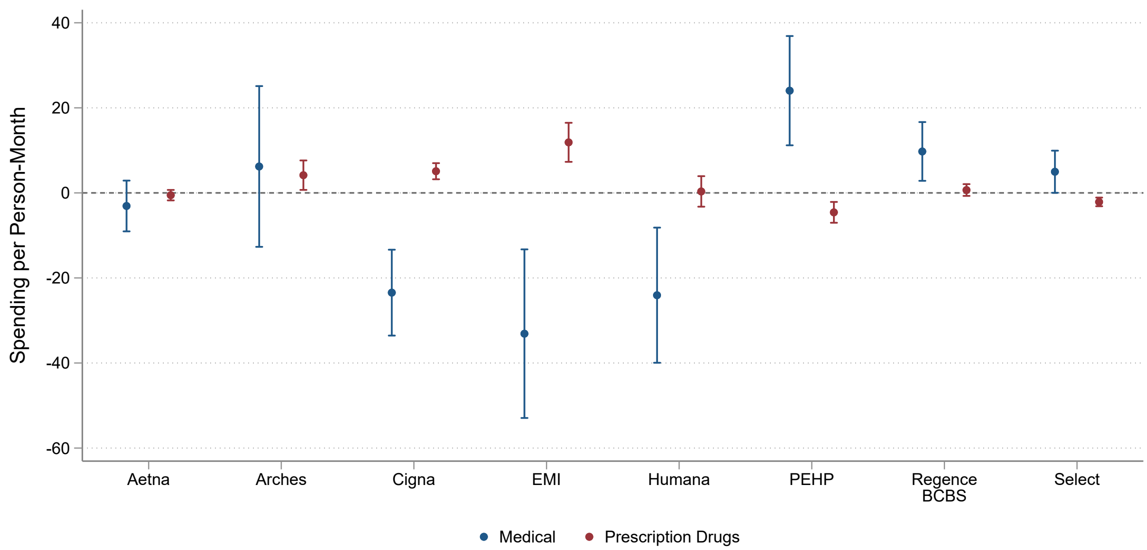
The Public Employee Health Plan (PEHP) has the largest positive effect on medical spending, increasing expenditures by $24 per month, representing a 13% increase relative to the enrollee-weighted average insurer effect (which is normalized to zero). At the other extreme, EMI Health, a regional insurer, reduces spending by $33 per month (-17% of average monthly spending). Among national carriers, Cigna and Humana most effectively reduce medical spending, with treatment effects of -$23 (-12%) and -$24 (-13%) respectively. SelectHealth and Aetna both have treatment effects near the weighted average insurer effect, while Regence Blue Cross increases spending by $11 (5%) per month.

There is an even larger range of relative treatment effects on drug spending, which span $16 per month, which is 37% of mean drug spending.[25] At the edges of this range, EMI Health causes a $12 (27%) increase in monthly drug spending while PEHP causes a $5 (-

[24]Percentages are all calculated relative to average winsorized monthly spending, which is $190 per month for medical, and $44 per month for prescription drugs.

[25]We note that the drug spending we study is total drug spending recorded in claims data, which does not account for drug rebates delivered by manufacturers to pharmacy benefit managers and, potentially, passed through to insurers to some extent. The typical spread in rebates across insurers of these rebates is smaller than the treatment effects estimated here (Plummer et al. (2022)). Also, we find that most of these changes are due to quantity changes rather than price changes.

Figure 2: Spending Effects $\psi_{J(i,t)}$



Notes: This figure reports insurer fixed effects from equation 1 in our full analysis sample, which includes forced switchers and job movers. Results from two separate regressions are reported, with the dependent variable being monthly medical spending (in blue) and prescription drug spending (in red). In each case, we winsorize the dependent variable at the 99.9th percentile. Coefficients report the difference in spending between each insurer and the weighted average across all insurers, weighted by the number of enrollees. Error bars represent the 95% confidence interval for each coefficient to test the hypothesis that the coefficient is different from the weighted average of all insurers (the dotted line).

10%) decrease. Cigna increases drug spending by \$5 (12%) per month, while SelectHealth decreases spending by about \$2 per month. Regence, Aetna, and Humana have tightly estimated effects around that of the average insurer, while Arches has a wider confidence band around zero.

Comparing across the eight insurers we study, we find a clear negative correlation between the treatment effects on medical and drug spending, indicating that insurer strategies alter the within-person substitution patterns between these categories of care, rather than just causing upward or downward shifts in spending for all types of care. Cigna and EMI have large negative effects on medical spending but positive effects on drug spending, while the reverse is true for PEHP and SelectHealth. Humana has large negative effects on medical spending but no effect on drug spending, while Regence has a large positive effect on medical spending and no effect on drug spending. This evidence for drug offsets in insurer strategies is consistent with prior work studying privatized Medicare by Starc and Town (2019) and Lavetti and Simon (2018).

We perform symmetry analyses to evaluate whether switching from one insurer to another leads to an equal and opposite-sized treatment effect relative to switching in the reverse direction. We find evidence for close to symmetric responses. In Section 4.6 we discuss this analysis in more depth and investigate alternative specifications, including Poisson and two-part models.

4.3 Insurer Effects on Prices and Quantities

In order to understand whether insurer effects on total spending are due to differences in the prices paid for services or differences in the quantity of care provided, we perform a price-quantity decomposition of the overall spending treatment effects. There is some subtlety in doing this for total medical spending (or other categories that have many possible treatments), since, in addition to doing more or less of a given procedure, or having higher or lower prices for a given procedure, insurers may also induce substitution from higher to lower resource procedures to treat the same condition.

We construct an index that is similar in spirit to a Relative Value Unit (RVU) that accounts for quantity changes while also allowing for procedure substitution, following the methodology of Hsiao et al. (1988).[26] Our method uses a data-driven approach to identify the value of a procedure: It attributes value based on the average price that all insurers are willing to pay. To do this, we estimate a typical cost for each procedure using the following decomposition model:

$$p_{jdx} = \alpha_{jd} + \delta_x + \eta_{jdx} \tag{4}$$

where p_{jdx} is the log cost for procedure x performed by provider d and paid for by insurer j. The estimate $\widehat{\delta_x}$ serves as an index proportional to the typical cost for a given procedure x. We estimate separate regressions for inpatient, outpatient, professional, and prescription drugs. In each case, we estimate equation 4 on the full sample of all insurance claims in our sample (including claims for individuals who do not meet our other sample restrictions). We then construct a quantity index as:

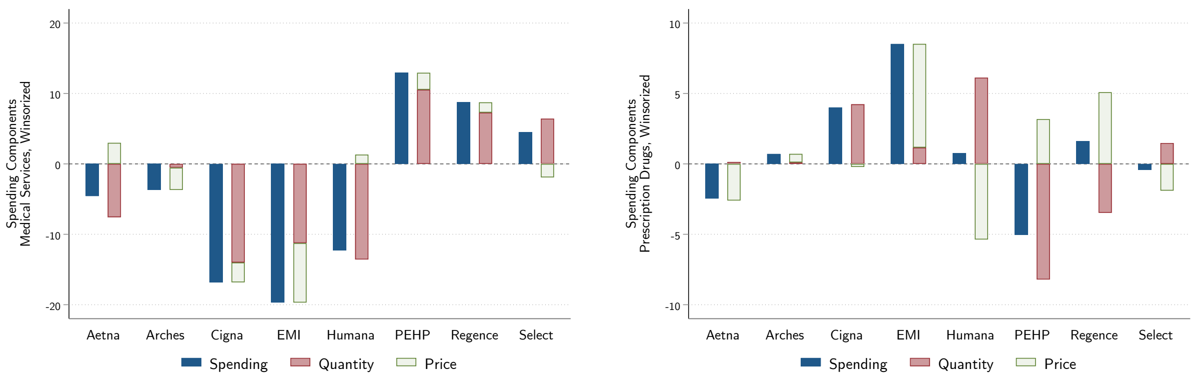
$$Q_{jx} = \Sigma_d W_{dj}^x exp\left\{\hat{\delta_x}\right\} \tag{5}$$

where W_{dj}^x represents the frequency of procedure x performed by doctor d for insurer j. Q_{jx} is

[26] This method is inspired by Relative Value Units from Medicare, but uses privately negotiated prices instead of the Medicare RVU schedule to estimate the relative price of different procedures.

Figure 3: Prices vs. Quantities, $\psi_{J(i,t)}$
Medical Spending and Pharmaceutical Spending

(a) Medical Spending (b) Prescription Drug Spending



Notes: This figure reports insurer fixed effects from equation 1 for three separate regressions. The top panel presents results for medical spending and the bottom panel presents results for drug spending.The dependent variables are total spending (in blue), total utilization (in red), and total price (in green). See text for how we constructed these three measures. The sample in this regression is a combination of forced switchers and job movers.

the associated quantity for treatment x and insurer j. We then compute the total aggregated quantity of medical care for insurer j by summing over Q_{jx} such that $Q_j = \Sigma_x Q_{jx}$. This calculation is done in each quarter to calculate Q_{jt}. We then define the price paid for care by insurer j as:

$$P_{jt} = \frac{TS_{jt}}{Q_{jt}}$$

We apply these definitions throughout our analyses, including for subsamples (e.g., patients with diabetes) and treatment categories (e.g., inpatient care).

Figure 3 presents the decomposition of insurer effects into price and quantity components for both medical and drug spending. As with total spending, we find substantial heterogeneity in insurer price effects. Humana, PEHP, and Aetna pay higher prices for medical care than the typical insurer (enrollee-weighted,) while Cigna, EMI, and Regence pay lower prices.

To better understand the mechanisms driving price variation across insurers, we decompose price differences into three mutually exclusive and exhaustive components:

1. **Network Design:** whether an insurer's network includes providers who are systematically more or less expensive across all insurers.

2. **Price Bargaining:** whether an insurer pays different prices than other insurers for

the same provider-service pairs (similar in spirit to the work in Cooper et al. (2019)).

3. **Network Steering:** whether an insurer successfully steers patients to lower-price providers within their network.

We estimate each component of the decomposition for each procedure x and then aggregate price effects to the insurer level by summing over procedures. For parsimony, we suppress this summation over x in the notation below and express the price decomposition for a given x as:

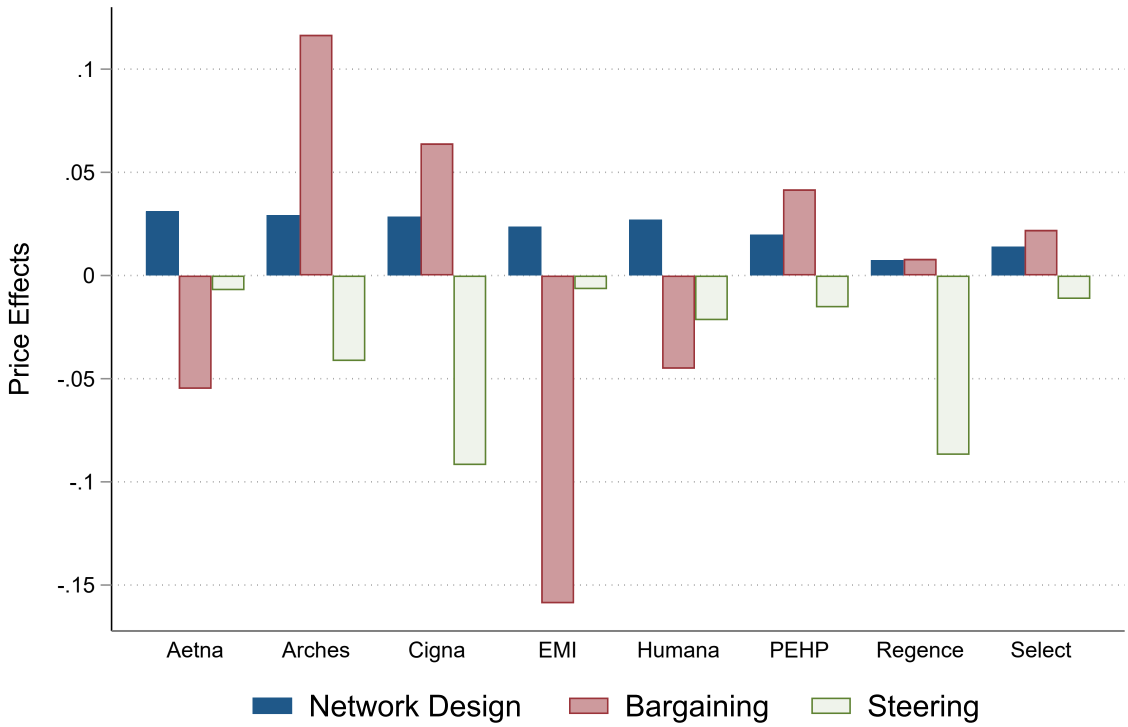
$$
\begin{aligned}
P_j = & \sum_d \left\{ 1[d \in N_j]\overline{\alpha_d}\frac{\overline{S_d}}{\sum_d \overline{S_d} 1[d \in N_j]} \right\} && \text{(Network Design)} \\
& + \sum_d \left\{ 1[d \in N_j](\alpha_{jd} - \overline{\alpha_d})\frac{\overline{S_d}}{\sum_d \overline{S_d} 1[d \in N_j]} \right\} && \text{(Bargaining)} \\
& + \sum_d 1[d \in N_j]\alpha_{jd} \left(S_{jd} - \frac{\overline{S_d}}{\sum_d \overline{S_d} 1[d \in N_j]} \right) && \text{(Steering)}
\end{aligned}
$$

where N_j represents insurer j's network, $\overline{\alpha_d}$ is the average price paid to provider d across all insurers, and α_{jd} is the average price paid by insurer j to provider d. Similarly, $\overline{S_d}$ is the average market share of provider d across all insurers, and S_{jd} is the market share for provider d within insurer j.[27]

Figure 4 presents results from this price effects decomposition. There are several key patterns. First, we find substantial heterogeneity in bargaining effects for medical care, suggesting that insurers pay very different prices for the same provider-service pairs. The range of insurer bargaining effects spans 28% of the mean overall price level, with significant differences in many pairwise comparisons. For example, Arches, a non-profit co-op, paid 28% higher bargained prices than EMI, a significant regional carrier focusing on employer markets. Cigna, a major national carrier, paid 12% higher bargained prices than Aetna and 10% higher prices than Humana, two other major national carriers. PEHP, SelectHealth, and Regence paid only slightly higher bargained prices than other insurers.

[27] In our empirical implementation, since we don't observe exact provider networks (e.g., from a list) we derive these networks from claims data by leveraging fields that indicate whether providers are in the network or not. If a very high proportion of a provider's claims with a given insurer are recorded as in-network, we count that provider as being in-network. If we vary the threshold for this definition of in-network, it impacts which providers are counted as in or out of network, though any errors in classification are naturally for providers with a smaller, rather than larger, presence.

Figure 4: Price Effect Decomposition: Medical Spending



Notes: This figure reports the results of our price effect decomposition into (i) network inclusion (ii) price bargaining and (iii) within-network steering. See the text for formal definitions of these three measures. The sample includes all fee-for-service claims for employer-sponsored plans.

We estimate that network steering effects span a range of about 10% of mean prices. Cigna and Regence achieve approximately 8% lower prices than Aetna for the same procedures by steering patients to cheaper in-network doctors. Since Cigna and Regence have meaningfully higher bargained rates than Aetna, this suggests that policies to steer patients within networks may exist as a substitute for price bargaining. Most other insurers have limited steering effects, though Arches generates modest savings through steering.

Finally, we estimate that the network inclusion component is very similar across insurers, suggesting that insurers have similar mixes of systematically high and low cost providers. This aggregate similarity masks meaningful differences in the composition of specific providers included in each insurer's network. That is, the rates for specific providers and specific classes of providers do differ substantively across insurers.

4.4 Impacts on Value of Care Provided

It is clear that insurers are meaningfully differentiated from one another and play important roles in affecting the overall delivery and cost of medical care. A natural question is how these differences across insurers manifest in the provision of specific types of care. We first assess differences across broad categories of care (e.g., preventive care, specific drug classes) and then investigate treatment patterns for patients with common chronic conditions, with a detailed case study of patients with diabetes.

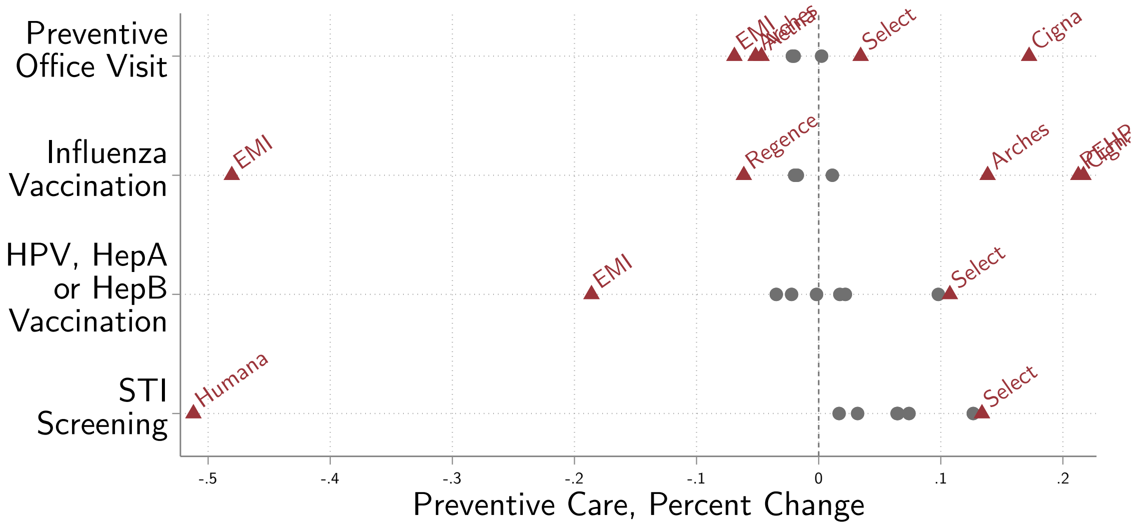
The results from these analyses matter for several reasons. First, we are intrinsically interested in understanding whether the variation in spending that we document stems from differences in clinically validated high-value versus low-value care. Second, assessing these differences provides insight into health outcomes insofar as outcomes depend on evidence-based care inputs. This is particularly relevant in our relatively young and commercially insured population where mortality and other traditional measures of health outcomes are less likely to vary in informative ways.[28]

Figure 5 presents estimates of insurer effects on four types of preventive care that are widely considered to be high-value (and often under-used): (i) preventive office visits (ii) influenza vaccines (iii) HPV / HepA / HepB vaccinations, and (iv) screening for sexually transmitted infections. We find meaningful causal differences in preventative care utilization across insurers, with treatment effects spanning more than 20% of average utilization levels for all four measures. For example, Cigna has large positive effects on preventive office visits and flu vaccinations (20% above the mean), while SelectHealth has large (10-15%) positive impacts on HPV / HepA / HepB vaccinations and STI screenings. In contrast, EMI and Humana have large negative impacts on flu vaccinations (-48%) and STI screening (-52%) respectively. The figure also illustrates other smaller-scale but economically meaningful differences in insurer effects on these preventive utilization measures.

We also find substantial variation in the use of high-value prescription drugs that can prevent unnecessary hospitalizations. We use the drug classification system developed by Chandra, Gruber and McKnight (2010), which categorizes drugs as "acute" or "chronic" care drugs if non-adherence substantially increases the probability of an adverse health event

[28]Focusing on care inputs also allows us to address the relatively short time frame for observation. In both cases, mortality differences that an insurer might impact are unlikely to manifest while an individual is enrolled in a specific plan, or in private ESHI in general, given they primarily occur in older age (see e.g. Danesh et al. (2024) for the role of chronic disease in prime-age versus old-age mortality). Therefore, we view measures of clinically validated high-value care as good measures of clinical quality, arguably better than short-run outcomes.

Figure 5: Preventive Care by Insurer



Notes: This figure reports insurer fixed effects from equation 1, separately estimated for 4 different examples of preventive care as defined in the Affordable Care Act. The dependent variable is the probability of observing each type of preventive care in a given month. We then standardize each dependent variable by dividing each coefficient by the overall mean of that dependent variable in the entire sample, which we report as a percentage change. Fixed effects are normalized to have the same mean outcome as the overall population. Insurer fixed effects that are statistically different from the mean at the 5% level are identified in red.

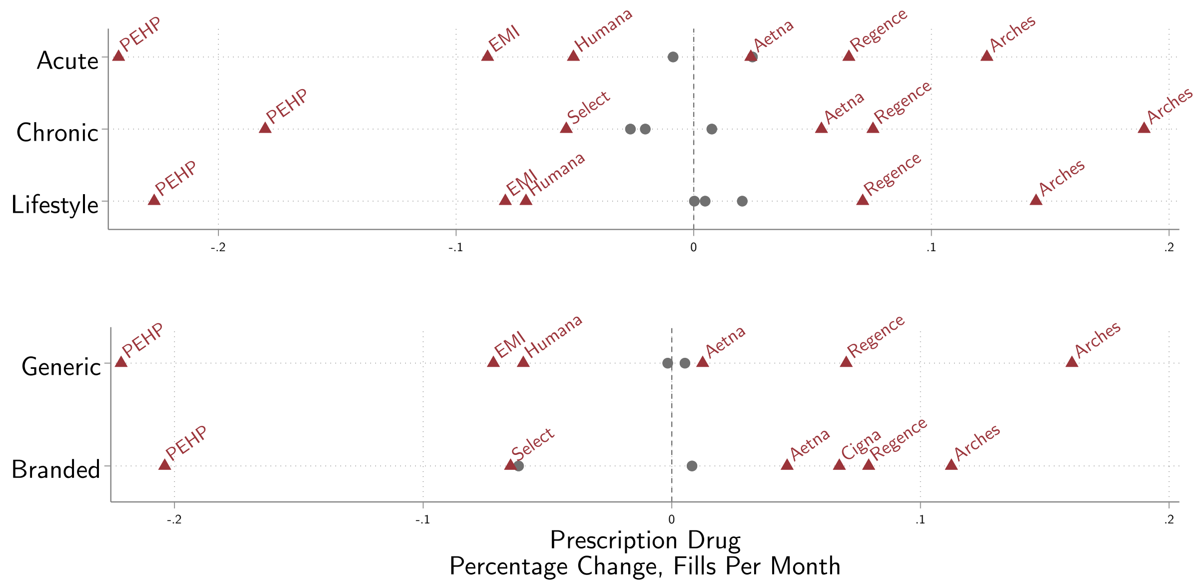
within 2 or 12 months, respectively.

Figure 6 presents our estimates of insurer effects on these drug categories. We find substantial heterogeneity in the effect of insurer assignment on the use of these drugs. Arches increases the use of both acute and chronic care drugs by at least 30% relative to PEHP. In turn, PEHP, as we will show later, has substantially higher use of inpatient hospital care for patients with chronic conditions like diabetes and hypertension.

Most insurers have meaningful causal effects on drug use across acute, chronic, and lifestyle drugs, though these effects are highly correlated across the three categories. This pattern suggests insurers generally pursue strategies that affect overall drug use rather than policies that target specific categories of drugs. For example, in addition to the PEHP-Arches contrast noted above, Regence has positive impacts relative to baseline across all three categories, while EMI has negative impacts on all three.

These patterns broadly persist in results comparing the use of branded versus generic drugs. Arches, Regence, Aetna, and Cigna cause increases in utilization of both branded and generic drugs, while EMI and PEHP reduce utilization of both. Notably, SelectHealth

Figure 6: Causal Effect on Drug Utilization by Drug Type



Notes: This figure reports insurer fixed effects from equation 1, separately estimated for generic vs branded drugs, as well as drug categories (Acute, Chronic, Lifestyle) defined by Chandra, Gruber and McKnight (2010). The dependent variable is the total spending on each type of drug in each month, and fixed effects are normalized by dividing by the average monthly spending to report percentages. Acute drugs, if not taken, are likely to lead to a near-term hospital admission. Chronic drugs, if not taken, may lead to a hospital visit within the year. Lifestyle drugs are those that are unlikely to result in an adverse health event if not taken. Insurer fixed effects that are statistically different from the mean at the 5% level are identified in red.

is an interesting exception to this pattern, with a negative (8%) effect on branded drug use no effect on generic utilization.

In Appendix Figure A2, we investigate insurer treatment effects for 13 different procedures identified as examples of low-value care in the medical literature by Charlesworth et al. (2016). The figure reports 104 estimates, one for each insurer-procedure pair. While the point estimates show meaningful dispersion, we lack the statistical power to detect reasonably large effects for these relatively uncommon procedures. Only 4 out of the 104 estimated treatment effects are statistically different than the mean.

4.5 Insurer Heterogeneity in Chronic Disease Treatment

A key potential differentiator between insurers is their approach to managing the treatment of patients with chronic conditions. Insurers often emphasize chronic disease management treatment strategies in their internal cost-reduction efforts, as well as in marketing aimed at consumers and employers. To this end, we investigate insurer-by-disease heterogeneity in treatment effects to understand how insurers affect healthcare delivery for high-need patients with five important chronic conditions: diabetes, hypertension, asthma, lower back pain, and bronchitis. We conduct an in-depth case study for patients with diabetes and provide evidence of meaningful effects at more aggregate levels for the other four conditions.

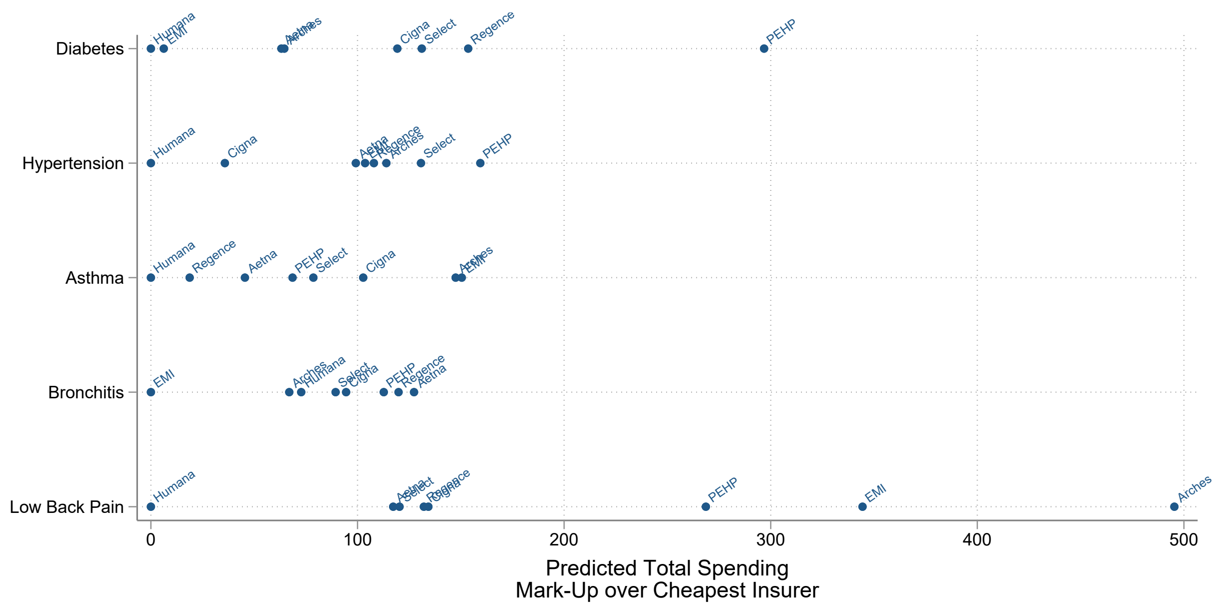
Figure 7 presents estimates of insurer-specific effects on total spending (medical plus drug) for each of these five chronic conditions. The estimates depicted in the figure come from a version of equation 1 that includes separate fixed effects for each insurer by condition pair. To facilitate comparisons, we use the estimates from this model to construct insurer by condition markups relative to the lowest cost insurer for each condition, and present these markups in Figure 7.

The figure shows compelling evidence of insurer-specific dispersion in spending by condition. For diabetes, insurer treatment effects span nearly \$300 per month (\$3,600 annually), with PEHP generating the highest spending and Humana the lowest. EMI, Aetna, and Arches all substantially reduce spending by over \$2,500 per year relative to PEHP. The highest and lowest estimated treatment effects are statistically significantly different from each other. The results for patients with lower back pain also imply large dispersion, though the ordering of insurers is different across conditions. For example, EMI and Arches are the highest-spending insurers for lower back pain, with relative treatment effects of about \$4000 and \$5900 per year, respectively, despite being among the lower-spending insurers for diabetes. For the other three conditions (hypertension, asthma, and bronchitis), the dispersion and magnitude of effects are broadly similar. Again, the highest and lowest estimated treatment effects for each condition are statistically significantly different from each other.

There are some noteworthy insurer-level patterns visible when comparing across conditions. Humana is the lowest-spending insurer for four of the five conditions and is in the lower half of spending for the fifth. In contrast to Humana's consistency, EMI and Arches both have very high condition-specific variation, with the lowest two treatment effects for bronchitis but the highest treatment effects for asthma and lower back pain. Cigna, SelectHealth, and Aetna are generally near the median treatment effect for most conditions.

This heterogeneity in total spending effects has substantial implications for insurer, employer, and consumer costs. A natural next step is to unpack the channels that drive these differences. We conduct a detailed investigation of diabetes in this section, since it is the most common of these chronic conditions, and diabetic patients have the highest average spending levels among the five conditions. In Appendix Figures A5 and A7, we also report

Figure 7: Average Total Spending Markups by Condition



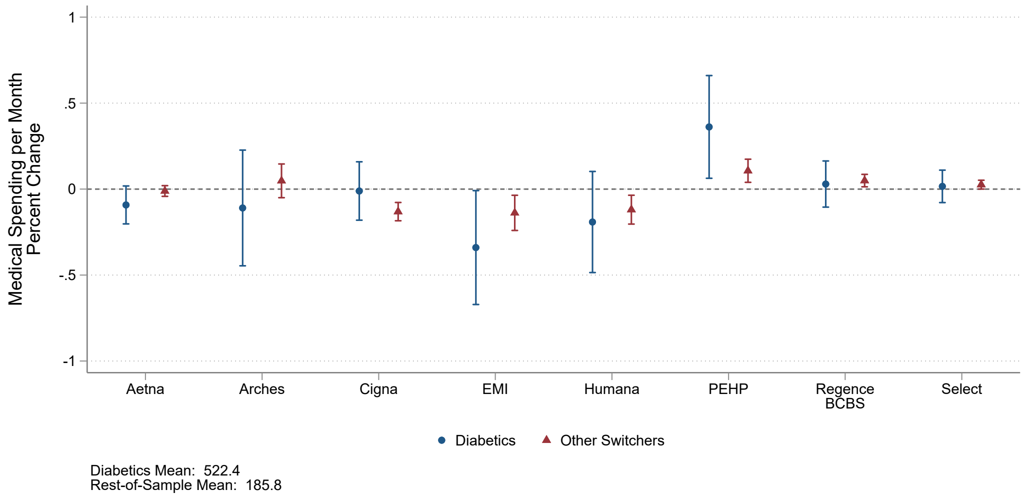
Notes: This figure reports results from a modified version of equation 1 with separate fixed effects for each insurer by condition. The dependent variable is monthly total spending winsorized at the 99.9th percentile, and the sample is a combination of forced switchers and job movers. The Figure plots predicted spending by insurer-condition normalized to show differences in spending relative to the cheapest insurer. The normalization is done separately by condition.

estimates from a detailed investigation of spending effects for patients with hypertension.

Figure 8 presents the estimated insurer treatment effects for patients with diabetes. The chart shows the same general patterns as Figure 7 for diabetes, but here treatment effects are expressed as percentage differences relative to the enrollment-weighted average of all insurers (which is normalized to zero). The figure also reports insurer average treatment effects for all non-diabetic patients, along with standard errors. This perspective shows that the PEHP treatment effect for diabetes is statistically different from those of all other insurers, while only a few of the other pairwise comparisons are statistically different from one another. A pattern in this figure is that insurer treatment effects for diabetic patients typically have the same sign as treatment effects for the same insurer for non-diabetic patients, suggesting that condition-specific heterogeneity may follow (or amplify) baseline insurer spending heterogeneity.

We further decompose these differences in spending for enrollees with diabetes into mutually exclusive categories (inpatient, outpatient, physician fees, drug use), and into price and quantity components using the same procedure described in Section 4.3. We find that the largest dispersion in insurer treatment effects on spending comes from inpatient care,

Figure 8: Insurer Treatment Effects on Total Spending for Patients with Diabetes



Notes: This figure reports insurer fixed effects from a modified version of equation 1 where we estimate separate fixed effects for the population of consumers who have ever had a diagnosis of diabetes, and those who have not. The dependent variable is monthly total spending. We winzorize the dependent variable at the 99.9th percentile, and then divide it by the average spending in either the diabetes or the non-diabetes sample, depending on whether the person has diabetes. This allows us to report percentage effects on spending. Coefficients report the difference in spending between each insurer and the weighted average of spending across all insurers in the diabetes and non-diabetes samples, respectively. Error bars represent the 95% confidence interval to test the hypothesis that the coefficient is different average of all insurers for the diabetes and non-diabetes samples, respectively.

for which treatment effects span a range of nearly \$2,100 annually, or 144% of mean annual inpatient spending per capita. Of this dispersion, about 57% is explained by differences in quantities, and the remaining 43% is attributed to differences in prices. The dispersion in treatment effects for drug and professional spending is also sizable, \$1,230 (41% of mean) and \$1,014 (45% of mean) per person-year, respectively, and these differences are almost entirely explained by quantities. These results are reported in Appendix Table A3.

These substantive differences in diabetes spending, overall and across categories, have direct implications for patient care and patient health. We illustrate this for patients with diabetes by highlighting how, when diabetic patients are forced to switch across insurers there are dramatic shifts in the type of drugs they consume. Table 5 reports insurer treatment effects for five common diabetes drugs, four of which are types of insulin, and the fifth (Sitaglipin) being a drug used in combination with insulin and metformin to improve glycemic control for type-2 diabetics.

The table shows clear and quite large statistically significant differences in insurer treat-

Table 5: Utilization Impacts: Diabetes Drugs

	Insulin Glargine	Insulin Aspart	Insulin Lispro	Insulin Human	Sitagliptin Phosphate
Aetna	-30%*	-68%*	105%*	16%	-49%*
Arches	-98%	-63%*	146%*	34%	-7%
Cigna	22%	-8%	30%*	84%*	25%
EMI	3%	-79%*	144%*	-2%	1%
Humana	140%*	43%	7%	102%	126%
PEHP	-1%	53%*	-93%*	-18%	-14%
Regence	-18%*	-53%*	10%	12%	35%
Select	15%*	80%*	-94%*	-57%*	-3%

Notes: This figure reports insurer treatment effects from equation 1 for total utilization of 5 common diabetes drugs. After estimating coefficients, we divide by the mean utilization in the analysis sample of switchers to measure percentage effects. Fixed effects in both groups are normalized to have an enrollment-weighted mean of zero.

ment effects on use for all five of these drugs. For example, SelectHealth enrollees use 80% more insulin aspart than the average insurer, while they use 94% less insulin lispro and 57% less human insulin than typical. PEHP has a similar pattern, increasing aspart use by 53% and decreasing lispro use by 93%. Conversely, EMI, Aetna, and Arches all show large (63-79%) decreases in insulin aspart and big increases in insulin lispro (105-146%) relative to the average insurer. Cigna has smaller effects for those two drugs, but a huge 84% increase in human insulin.

Most insurers use more of some of these drugs and less of others. However, Humana is an exception, with positive point estimates across all five drugs, with a statistically significant 140% increase in insulin glargine relative to baseline. This is noteworthy because, despite the high utilization of drugs to manage diabetes, Humana has the lowest total effect on spending for diabetics because the high rates of insulin use are offset by lower inpatient and professional spending for diabetic patients. This is another example of insurers choosing strategies that take drug and medical spending offsets into account to manage costs.

Taken together, these results show substantial insurer-specific impacts on both the composition and levels of health care received by patients with diabetes. It is important to note that our estimated spending and quantity effects are short-run (one-year) estimates. It is entirely possible that having lower care utilization in the short run increases subsequent care and spending needs beyond our analysis period.

4.6 Specification Analysis and Robustness

In this section we discuss a range of analyses to either support our main specification choice or assess the robustness of our results to specification assumptions.

4.6.1 Treatment Effect Symmetry

As discussed in Section 3.1, one assumption in Equation 1 is that insurer treatment effects are symmetric when individuals switch into or out of the insurer. This assumption could be too strong if, for example, insurers have specific strategies that relate to the onboarding of new customers.

To evaluate this assumption, we estimate a version of Equation 1 that replaces insurer effects with directional pairwise combinations of all insurer switches. Specifically, if t_0 is the period just prior to a switch and t_1 is the subsequent period, we define $\omega_{j,k,t} = 1\{(J_{i,t_o} = j)\} \times 1\{(J_{i,t_1} = k)\} \times 1\{(t \geq t_1)\}$, which captures the effects of switching from insurer j to k in period t_1.[29] We then estimate the following equation:

$$\begin{aligned} y_{ist} &= \omega_{j,k,t} + \rho AV_{P(i,t)} + \sum_{k \geq -11}^{k \leq 12} \left[\gamma_k^{FS} * FS_{is} + \gamma_k^{JMA} * JMA_{is} \right. \\ &+ \left. \gamma_k^{JMW} * JMW_{is} \right] + \tau_t + \theta_{is} + \epsilon_{ist} \end{aligned} \tag{6}$$

Appendix Figure A8 presents estimates from Equation 6. The figure shows that the effects on medical spending caused by switches into and out of each insurer are broadly symmetric.

4.6.2 Sensitivity to Functional Form and Sample Inclusion

Our baseline specification in Equation 1 is a linear model of spending. In Appendix Figure A11 we evaluate this functional form assumption by showing results from a Poisson model. The estimated insurer treatment effects from the Poisson model have the same qualitative and quantitative effects as our baseline results, although the scale of the dependent variable is measured in log points rather than levels in the Poisson estimates.

The dependent variable in Equation 1, health spending, is also winsorized at the 99.9th percentile so that our results are not unduly driven by extreme outliers. In Appendix Fig-

[29] When we look at every pairwise combination of insurers as a separate experiment, then the disruption effects are no longer identified. To allow identification, we assume that disruption effects only occur for at most 12 months before or after the switch.

ure A12 we re-estimate the model using non-winsorized spending, and show that the estimates are very similar in magnitude to our baseline estimates in Figure 2, though the outlier values increase standard errors when the dependent variable is not winsorized.

We also estimate a quantile model of medical spending in Appendix Figure A3. This analysis examines, for example, whether insurers that encourage patients to visit their doctors more frequently are also more likely to experience high outlier spending events. We show that insurer treatment effects on spending are ordinally similar across quantiles of the medical spending distribution. That is, insurers that are more effective at constraining very high spending levels ($> \$10,000$ per month) are also less likely to have spending above other quantiles of the spending distribution ($> \$0$, $> \$100$, and $> \$1,000$ per month), suggesting that heterogeneity in insurer strategies does not primarily operate through targeting high versus low utilization enrollees on average.

In Appendix Figure A10, we show results with and without controls, and find that controls do not have a major impact on the results. In Appendix Figure A9 we examine a specification that restricts the sample to larger, established insurers. Again, we find nearly identical results to our main estimates. Finally, in Appendix A.2 we consider a different specification that corrects for seasonal trends prior to estimating insurer effects and disruption effects. We find that this specification has little impact on our estimates.

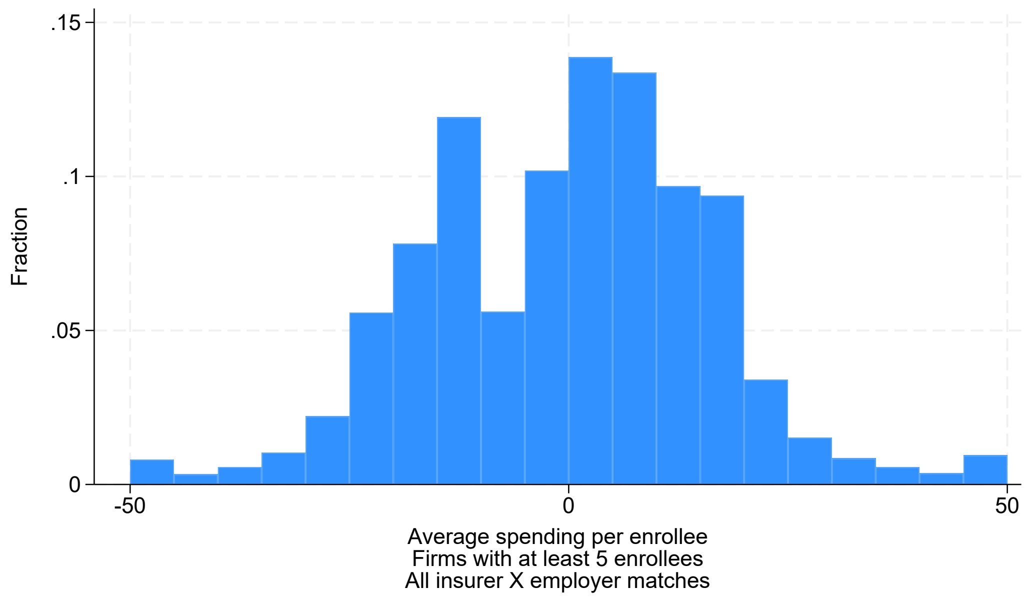
5 Employer Choice

We find broad evidence that insurers have meaningfully distinct treatment effects on costs and quantities for patients with chronic conditions. To the extent that there are firm-level differences in the prevalence of different chronic conditions among employees (and their dependents), the cost-minimizing insurer may differ across employers. For example, the estimates in Section 4.5 suggest that an employer with an above-average share of diabetic workers may minimize costs by choosing Aetna, while one with an above-average share of hypertensive employees may reduce costs by choosing Cigna. In this section, we assess how employers select insurers based on these insurer-by-condition treatment effects.

We also decompose firm-level differences in spending impacts across insurers into quantity and price components. Of course, if reduced costs or quantities are indicative of lower quality care, it is not obvious that employers *should* choose the lowest cost insurer. Thus, we interpret the evidence in this section as descriptive of employer choice behavior in response to potential financial savings, rather than as prescriptive.[30]

To do this, we begin by estimating a version of equation 1 that includes separate fixed effects for each insurer by condition pair (as in Section 4.5). We then aggregate the estimated insurer-by-condition treatment effects across all employees (and dependents) in each firm to calculate firm-by-insurer treatment effects for all combinations of firms and insurers, including counterfactual combinations. The estimates therefore capture a combination of differences across employers in the prevalence of each chronic condition, along with heterogeneity across insurers in the effects of each chronic condition on total spending.

Figure 9: Distribution of Employer-by-Insurer Treatment Effects on Total Spending



Notes: This figure presents estimates of the distribution of insurer treatment effects on total monthly per capita spending for all combinations of employers and insurers in the sample. The sample is restricted to employers with at least 5 enrollees. Treatment effects are top-coded at 50 and bottom-coded to -50.

Figure 9 depicts the distribution of these estimated treatment effects on total spending. The estimates in the figure are scaled to depict the distribution of treatment effects on spending per enrollee per month, with the average treatment effect normalized to zero. The figure shows meaningful dispersion in per capita monthly spending: a one standard deviation change in average employer-by-insurer treatment effects on monthly per capita spending is \$21, or about 8% of average spending levels.

[30]We also note that, in a competitive marketplace, if one insurer is much cheaper for a given employer, they may choose to charge higher markups to that employer, which we do not observe.

These meaningful differences in spending effects across insurers motivate our analysis of employer choices of insurers, which aims to understand whether employers respond to these differences in costs when choosing plans. Table 6 presents the results for this employer choice analysis. In the first row, we report mean observed per capita monthly spending in the analysis sample (column 1) and the full population of observed ESI plans (column 2). In row two, we report simulated counterfactual estimates of mean per capita spending if employers were randomly assigned to insurers. In the analysis sample (column 1), observed average spending is equal to 99.2% of spending under random insurer assignment (241.9/243.9). This suggests that, in terms of projected costs for enrollees, employers are doing no better (or no different) than choosing randomly. This finding also holds for estimated quantity effects, shown in columns 3 and 4.

Table 6: Predicted Spending Under Different Allocations of Enrollees to Insurers

	Total Spending		Utilization	
	Analysis Sample	Full Sample	Analysis Sample	Full Sample
Actual Insurer	241.9	241.5 (100%)	164.1	164.3 (100%)
Random Insurer	243.9	239.7	165.5	163.5
Indiv Cost Min	198.1	197.1 (82%)	142.7	142.1 (86%)
Employer Cost Min	214.0	215.0 (89%)	151.1	152.3 (93%)
Person-Months	1,739,590	24,339,175	1,739,590	24,339,175

Notes: This table reports estimates of actual and simulated counterfactual total monthly spending (columns 1-2) and utilization (columns 3-4) under four different scenarios. The first row reports mean observed spending and utilization in the analysis sample (column 1) and the full sample of individuals with employer-sponsored coverage in the APCD (column 2). In row 2, we randomly allocate each person to an insurer and predict spending based on the individual's medical conditions. In row 3, we assign each person to their cost-minimizing insurer given their health conditions. In row 4, we assign each employer to the lowest-cost insurer given the health conditions of all employees at the firm, with the restriction that each employer can only choose a single insurer. In all cases, we winsorize monthly spending at the 99.9th percentile.

In row three, we report average counterfactual spending if each individual employee was assigned to the cost-minimizing insurer given their set of chronic conditions. We estimate that this would reduce spending by 18.1% on average relative to observed insurer choices in row 1 (from \$241.9 to \$198.1). Notably, column 3 shows that about 72% of this cost-savings is explained by reductions in the estimated quantity of care, which falls by 13.0% relative to row 1.

Finally, in row four, we report estimates from a similar counterfactual exercise that constrains each employer to offer only one cost-minimizing plan to all of its employees. We estimate that employers could reduce healthcare spending by 11.5% (from \$241.9 to \$214.0) if they chose the cost-minimizing insurer given the health conditions of their employees. Again, column 3 suggests that about 69% of this cost-savings is explained by reductions in quantity. Thus, it is not obvious that employers are making a mistake when they don't select the cost-minimizing insurer: instead, it could be that they are unwilling to trade off lower spending for reduced quantities. It is also worth noting that these estimates are short-run effects, and reductions in quantity may lead to dynamic changes in long-run spending. A full welfare assessment of employer choices would require a dynamic analysis of the value of reductions in spending and quantities of care, which is beyond the scope of our analyses in this paper.

6 Conclusion

A core tenet underlying the managed competition paradigm is that insurers will compete with one another through product innovations that appeal to consumers (Enthoven, Garber and Singer, 2001). Given these innovations, insurers compete on price and pass through the surplus generated by innovations to consumers. Whether, in practice, insurers offer meaningfully different products is an empirical question that is critical to understanding policy and market design for the U.S. health care system.

In this paper, we aim to fill a gap in the literature by quantifying a wide range of insurer treatment effects on health care in the private employer-provided insurance sector, which covers the majority of the insured population of the United States. We ask whether (i) insurers have meaningful causal impacts on spending and utilization, and (ii) what specific forms that differentiation takes. To do this, we leverage all-payer claims data from Utah that contain individual-level claims for virtually the entire private employer-sponsored insurance market, along with information about employers that provide coverage and insurers. Using a movers-design approach that blends event study with two-way fixed effects models, while controlling for differences in plan design, we estimate each insurer's effect on a variety of outcomes of interest. Our study examines hundreds of employer-driven insurance switches across various insurer pairings.

We find meaningful variation across insurers in aggregate spending measures but also in different strategies, including the degree of offset between medical and pharmaceutical spending, approaches to price bargaining and network design, and how specific chronic conditions are treated. We find pairwise differences in medical cost that, holding fixed plan generosity, change medical spending by as much as 32%. Impacts of insurer assignment on pharmaceutical spending are even larger, spanning 44% of the mean spending level for drugs. We also find that insurers pay meaningfully different prices for the same services. These bargaining effects span 28% of the mean overall price level. We further quantify network steering effects — the ability of a plan to move care within the existing network to lower-cost providers — and find effects that are smaller, though material, of approximately 10% of the mean price level. Aggregate price bargaining and steering are negatively correlated at the insurer level, suggesting they are substitutes in insurer strategies to reduce cost. Finally, we find that aggregate network design differences are minimal. All insurers include similar numbers of higher- and lower- cost providers. We also find a negative relationship between medical and pharmaceutical spending at the plan level. Insurers who consistently spend more on drugs spend less on medical care, and tend to have lower aggregate spending as well.

We explore the specific strategies employed by insurers for a set of five specific chronic conditions: (i) diabetes, (ii) hypertension, (iii) asthma, (iv) lower back pain, and (v) bronchitis. We find even larger variation across insurers for these specific conditions. For example, annual spending on patients with diabetes varies across insurers by a range of nearly $3,600, and the insurer variation in spending for hypertension spans about $2,000. In a case study focused on diabetes, we find important differences in aggregate spending that are largely driven by trade-offs between higher (lower) drug and lower (higher) inpatient spending. At a more granular level, we also find that switching between insurers causally impacts the composition of drugs consumed by patients.

Taken together, our results provide the first large-scale evidence of insurer differentiation in aggregate outcomes, as well as direct impacts on care. We demonstrate that insurers play a role beyond simply providing risk protection and have important impacts on aggregate utilization and condition-specific provision of care.

Our final analysis studies a key tenet of managed competition: effective market signals. The market for employer-sponsored health insurance relies on employer intermediaries (often the Human Resources departments of firms) to determine plan offerings, set price incentives facing employees (i.e. premiums and subsidies) and choose myriad aspects of education and experience that affect plan choices. Whether employers' choices of plan offerings reward innovative insurers is an empirical question. To study this, we use our estimates to explore the degree to which the causal variation in cost we find actually impacts insurance demand by employers. We estimate the potential cost savings on an employer-by-employer basis for each plan based on our estimated causal effects by condition and the underlying employee population of each firm. We find that the potential cost savings are large, approximately

11.5% of total health care spending. Nevertheless, we find that actual plan offerings are nearly indistinguishable from random choice (on a market-weighted basis).

Our results underscore the important role of firm-specific strategies taken by private insurers in the functioning of the U.S. health care system. We develop and implement an approach to recovering key empirical estimates in considering policy and market design in the managed competition model of health insurance provision. We see this work as a first step in studying the central role played by competition in the insurance market in the U.S.. Future work that directly links the magnitude and form of competition to insurer differentiation and pricing would build on these results to provide more direct evidence on optimal policy design. Our findings underscore the need for such work given the large variation we find. Ultimately, being able to compare the costs of regulatory and market design tools to combat adverse selection in insurance markets with the benefits of meaningful insurer differentiation and innovation is one of the key questions in health economics and policy.

References

Abaluck, Jason, Mauricio Caceres Bravo, Peter Hull, and Amanda Starc. 2021. "Mortality Effects and Choice Across Private Health Insurance Plans*." *The Quarterly Journal of Economics*, 136(3): 1557–1610.

Abowd, John M., Francis Kramarz, and David N. Margolis. 1999. "High Wage Workers and High Wage Firms." *Econometrica*, 67(2): 251–333.

Agafiev Macambira, Danil, Michael Geruso, Anthony Lollo, Chima D Ndumele, and Jacob Wallace. 2022. "The Private Provision of Public Services: Evidence from Random Assignment in Medicaid." National Bureau of Economic Research Working Paper 30390.

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

Borusyak, Kirill, Xavier Jaravel, and Jann Spiess. 2024. "Revisiting Event-Study Designs: Robust and Efficient Estimation." *The Review of Economic Studies*, 91(6): 3253–3285.

Brot-Goldberg, Zarek C, Amitabh Chandra, Benjamin R Handel, and Jonathan T Kolstad. 2017. "What Does a Deductible Do? The Impact of Cost-Sharing on Health Care Prices, Quantities, and Spending Dynamics." *The Quarterly Journal of Economics*, 132(3): 1261–1318.

Brot-Goldberg, Zarek C, Samantha Burn, Timothy Layton, and Boris Vabson. 2023. "Rationing Medicine Through Bureaucracy: Authorization Restrictions in Medicare." National Bureau of Economic Research Working Paper 30878.

Brot-Goldberg, Zarek, Timothy Layton, Boris Vabson, and Adelina Wang. 2021. "The Behavioral Foundations of Default Effects: Theory and Evidence from Medicare Part D." University of Chicago Working Paper.

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

Charlesworth, Christina J., Thomas H. A. Meath, Aaron L. Schwartz, and K. John McConnell. 2016. "Comparison of Low-Value Care in Medicaid vs Commercially Insured Populations." *JAMA Internal Medicine*, 176(7): 998–1004.

Cooper, Zack, Stuart V Craig, Martin Gaynor, and John Van Reenen. 2019. "The Price Ain't Right? Hospital Prices and Health Spending on the Privately Insured." *The Quarterly Journal of Economics*, 134(1): 51–107.

Correia, Sergio, Paulo Guimarães, and Thomas Zylkin. 2020. "Fast Poisson estimation with high-dimensional fixed effects." *The Stata Journal*, 20(1): 95–115.

Curto, Vilsa, Liran Einav, Amy Finkelstein, Jonathan Levin, and Jay Bhattacharya. 2019. "Health Care Spending and Utilization in Public and Private Medicare." *American Economic Journal: Applied Economics*, 11(2): 302–32.

Cutler, David, Mark McClellan, and Joseph Newhouse. 2000. "How Does Managed Care Do It?" *RAND Journal of Economics*, 31(3): 526–48.

Danesh, Kaveh, Jonathan T Kolstad, Johannes Spinnewijn, and William D Parker. 2024. "The Chronic Disease Index: Analyzing Health Inequalities Over the Lifecycle." National Bureau of Economic Research.

de Chaisemartin, Clément, and Xavier D'Haultfœuille. 2022. "Two-way fixed effects and differences-in-differences with heterogeneous treatment effects: a survey." *The Econometrics Journal*, 26(3): C1–C30.

Dunn, Abe, Joshua D Gottlieb, Adam Hale Shapiro, Daniel J Sonnenstuhl, and Pietro Tebaldi. 2023. "A Denial a Day Keeps the Doctor Away*." *The Quarterly Journal of Economics*, qjad035.

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

Enthoven, Alain, Alan Garber, and Sara Singer. 2001. "Near-Universal Coverage Through Health Plan Competition: An Insurance Exchange Approach." *Covering America: Real Remedies for the Uninsured, Editors J. Meyer and E. Wicks*, Washington DC: 155–172.

Finkelstein, Amy, Matthew Gentzkow, and Heidi Williams. 2016. "Sources of Geographic Variation in Health Care: Evidence From Patient Migration*." *The Quarterly Journal of Economics*, 131(4): 1681–1726.

Geruso, Michael, Tim Layton, and Jacob Wallace. 2020. "Are All Managed Care Plans Created Equal? Evidence from Random Plan Assignment in New York Medicaid Managed Care." *NBER Working Paper 27762*.

Gruber, Jonathan, and Robin McKnight. 2016. "Controlling Health Care Costs through Limited Network Insurance Plans: Evidence from Massachusetts State Employees." *American Economic Journal: Economic Policy*, 8(2): 219–250.

Handel, Benjamin, and Jonathan Kolstad. 2022. "The Affordable Care Act After a Decade: The Industrial Organization of the Insurance Exchanges." *Annual Review of Economics*, 14: 287–312.

Hsiao, William C, Peter Braun, Daniel Dunn, and Edmund R Becker. 1988. "Resource-based relative values: an overview." *Jama*, 260(16): 2347–2353.

Lavetti, Kurt, and Kosali Simon. 2018. "Strategic Formulary Design in Medicare Part D Plans." *American Economic Journal: Economic Policy*, 10(3): 154–92.

Mihaly, Kata, Daniel F. McCaffrey, J. R. Lockwood, and Tim R. Sass. 2010. "Centering and Reference Groups for Estimates of Fixed Effects: Modifications to Felsdvreg." *The Stata Journal*, 10(1): 82–103.

Plummer, Elizabeth, Mariana P. Socal, Jeromie M. Ballreich, Gerard F. Anderson, and Ge Bai. 2022. "Trends of Prescription Drug Manufacturer Rebates in Commercial Health Insurance Plans, 2015-2019." *JAMA Health Forum*, 3(5): e220888–e220888.

Roth, Jonathan, Pedro H.C. Sant'Anna, Alyssa Bilinski, and John Poe. 2023. "What's trending in difference-in-differences? A synthesis of the recent econometrics literature." *Journal of Econometrics*, 235(2): 2218–2244.

Rothschild, Michael, and Joseph Stiglitz. 1976. "Equilibrium in Competitive Insurance Markets: An Essay on the Economics of Imperfect Information." *The Quarterly Journal of Economics*, 90(4): 630–649.

Starc, Amanda, and Robert J Town. 2019. "Externalities and Benefit Design in Health Insurance." *The Review of Economic Studies*, 87(6): 2827–2858.

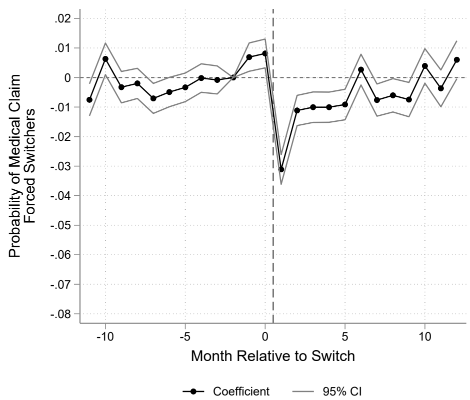
Tilipman, Nicholas. 2021. "Employer Incentives and Distortions in Health Insurance Design: Implications for Welfare and Costs." *Working paper*.

Wallace, Jacob. 2023. "What does a provider network do? Evidence from random assignment in Medicaid managed care." *American Economic Journal: Economic Policy*, 15(1): 473–509.

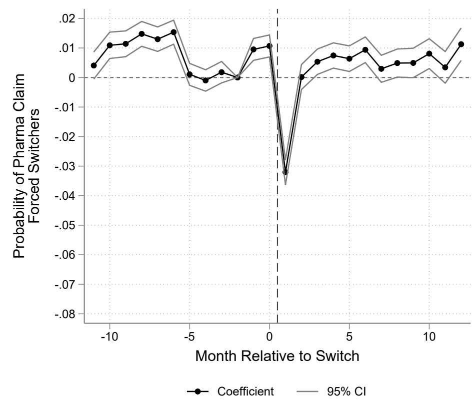
A Appendix

Figure A1: Disruption Effects of Plan Changes on Extensive-Margin Use

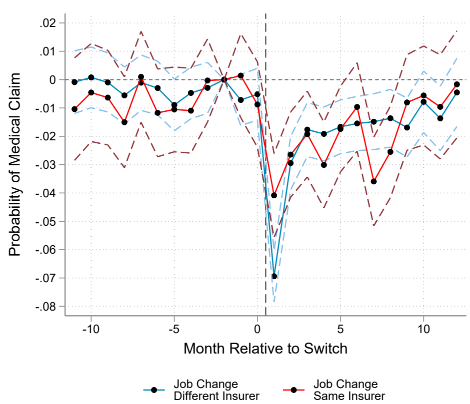
(a) Any Medical Care Use, Forced Switcher Sample



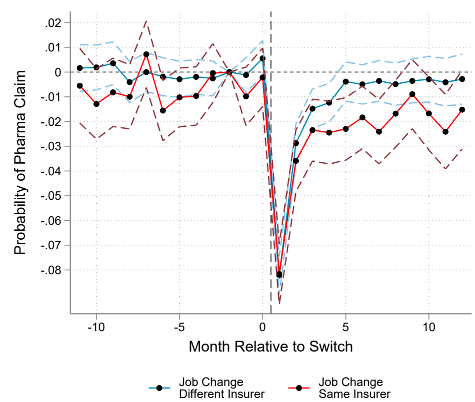
(b) Any Drug Use, Forced Switcher Sample



(c) Any Medical Care Use, Job Mover Sample

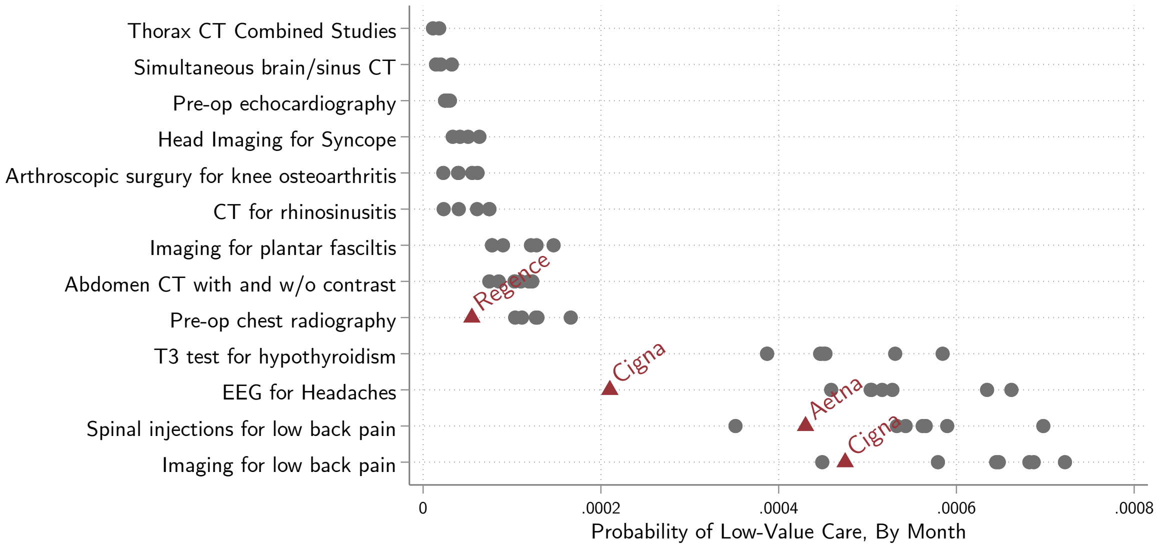


(d) Any Drug Use, Job Mover Sample



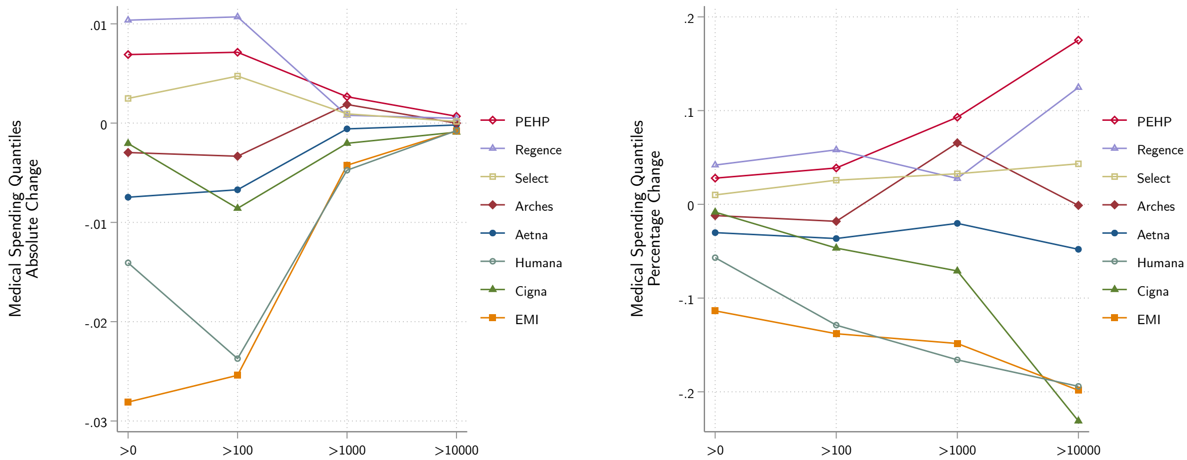
Notes: This figure reports event-study coefficients γ_k from Equation 1, separately estimated using the sample of forced switchers and job movers. The dependent variable is a binary indicator for any medical care use (panels a and c) and any drug use (panels b and d). In panels c and d the blue line shows γ_k for workers who changed jobs and switched insurers, and the red line shows estimates for workers who changed jobs and insurance plans but maintained the same insurer.

Figure A2: Low Value Care Utilization by Insurer



Notes: This figure reports insurer fixed effects from equation 1, separately estimated for 13 different examples of low-value care taken from Charlesworth et al. (2016). The dependent variable is the probability of observing each type of low-value care in one month. Fixed effects are normalized to have the same mean outcome as the overall population. Insurer fixed effects that are statistically different from the mean at the 5% level are identified in red. Each regression is estimated on the full sample, but we only present point estimates that meet the minimum cell size restriction (at least 10 cases of an insurer performing each type of low-value care).

Figure A3: Effects on Quantiles of Health Spending

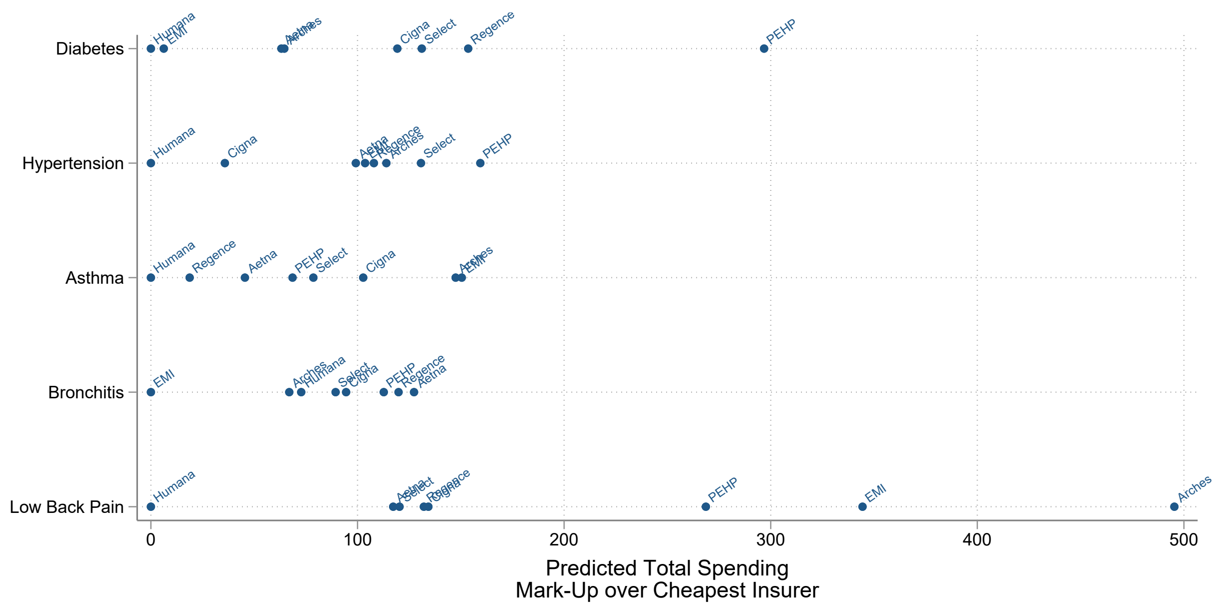


(a) Quantile Effects

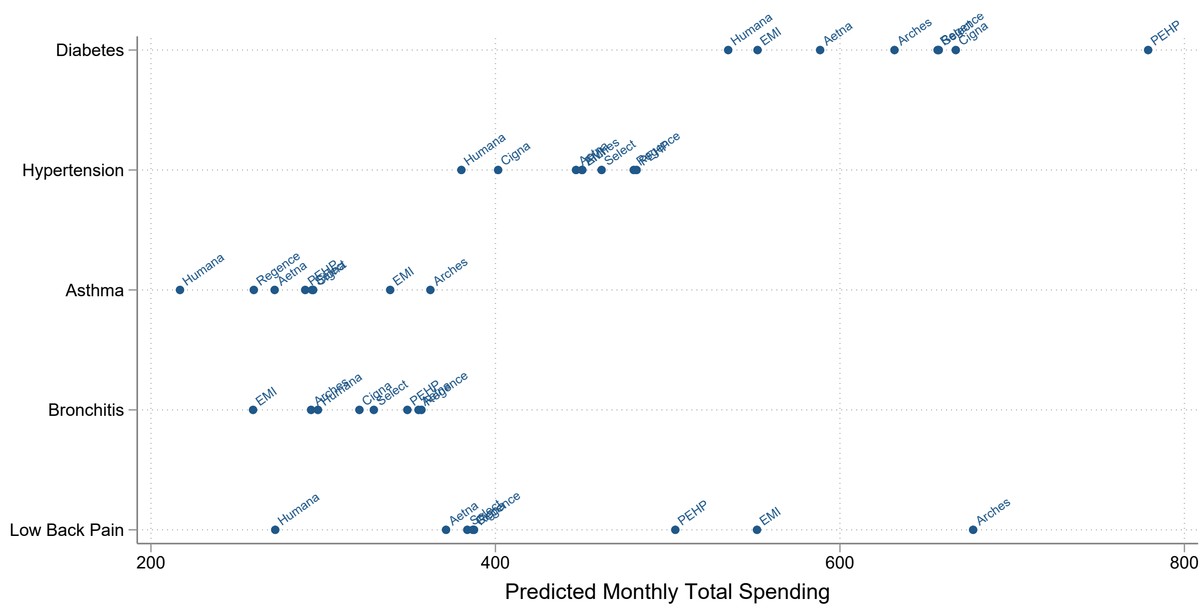
(b) Quantile Effects (Percent Change)

Notes: This figure reports insurer fixed effects from equation 1, separately estimated for quantiles of monthly medical health spending. In the left panel, the dependent variable is the probability of observing spending of at least x in a month, where x is listed on the x-axis. In the right panel, we normalize the coefficients by dividing by the average probability of spending at least x to measure percentage effects.

Figure A4: Average spending for insurers across 5 health conditions



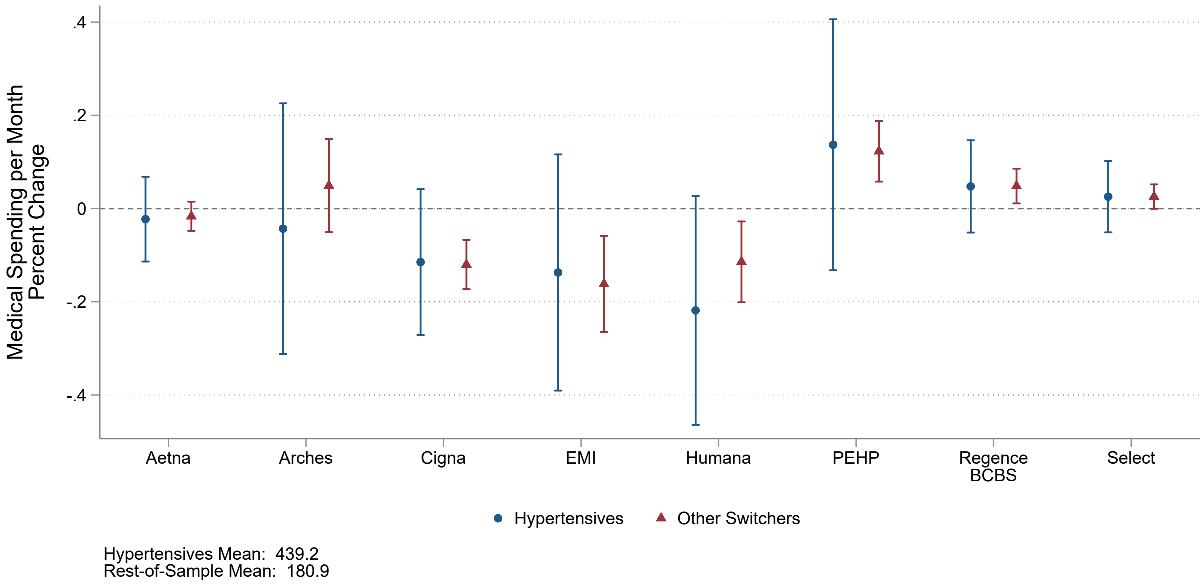
(a) Spending markups



(b) Mean spending

Notes: This figure reports results from a modified version of equation 1 with separate fixed effects for each insurer and each of 5 different conditions. The dependent variable is monthly total spending winsorized at the 99.9th percentile, and the sample is a combination of forced switchers and job movers. We then plot predicted spending by insurer for each of the 5 different conditions. Both panels (a) and (b) report the same data, but panel (B) is normalized to show average spending and panel (A) is normalized to show total possible savings relative to the cheapest insurer (by condition).

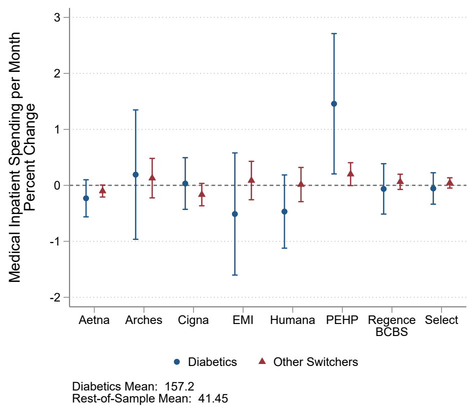
Figure A5: Insurer Treatment Effects on Total Spending: Patients with Hypertension



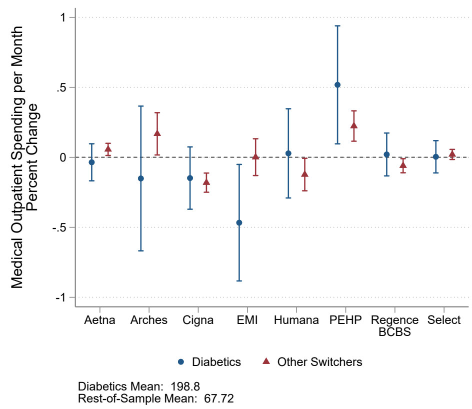
Notes: This figure replicates Figure 8, except we look at individuals with hypertension instead of individuals with diabetes.

Figure A6: Monthly Spending Effects by Spending Category: Diabetes Sample

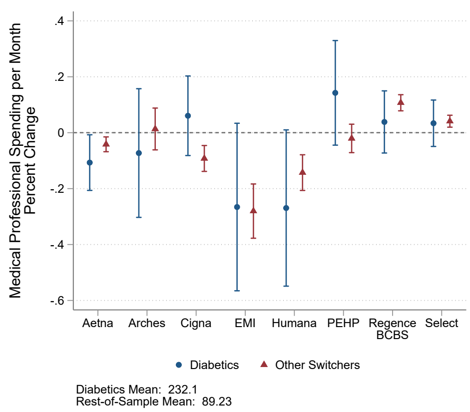
(a) Inpatient Spending



(b) Outpatient Spending



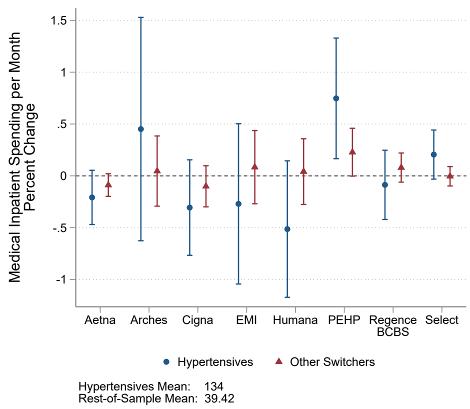
(c) Professional Services Spending



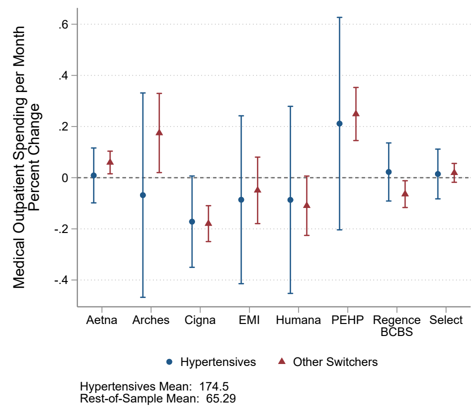
Notes: This figure replicates Figure 8 with inpatient, outpatient, and professional services spending as the dependent variable in panels (a), (b), and (c).

Figure A7: Monthly Spending Effects by Spending Category: Hypertension Sample

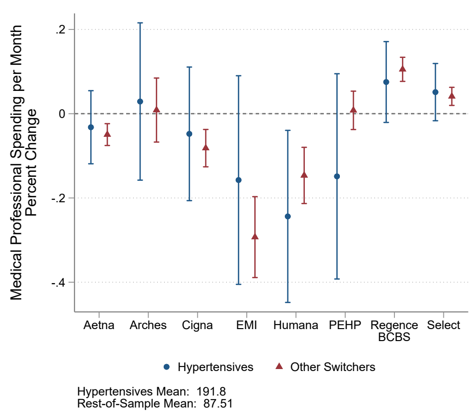
(a) Inpatient Spending



(b) Outpatient Spending



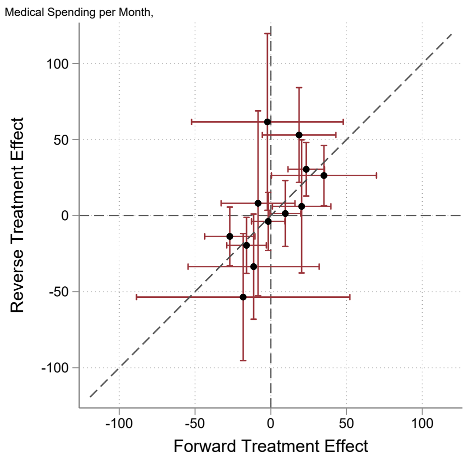
(c) Professional Services Spending



Notes: This figure replicates Appendix Figure A5 with inpatient, outpatient, and professional services spending as the dependent variable in panels (a), (b), and (c).

Figure A8: Insurer Treatment Effect Symmetry

(a) No Controls



(b) All Controls

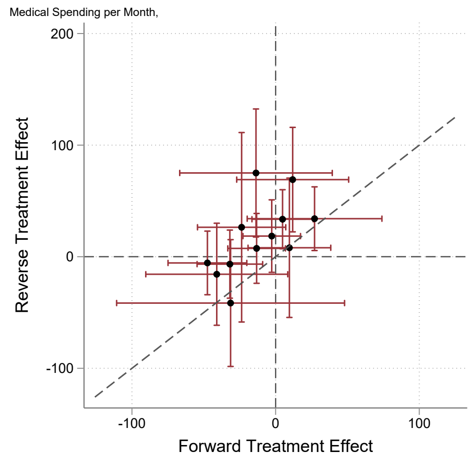
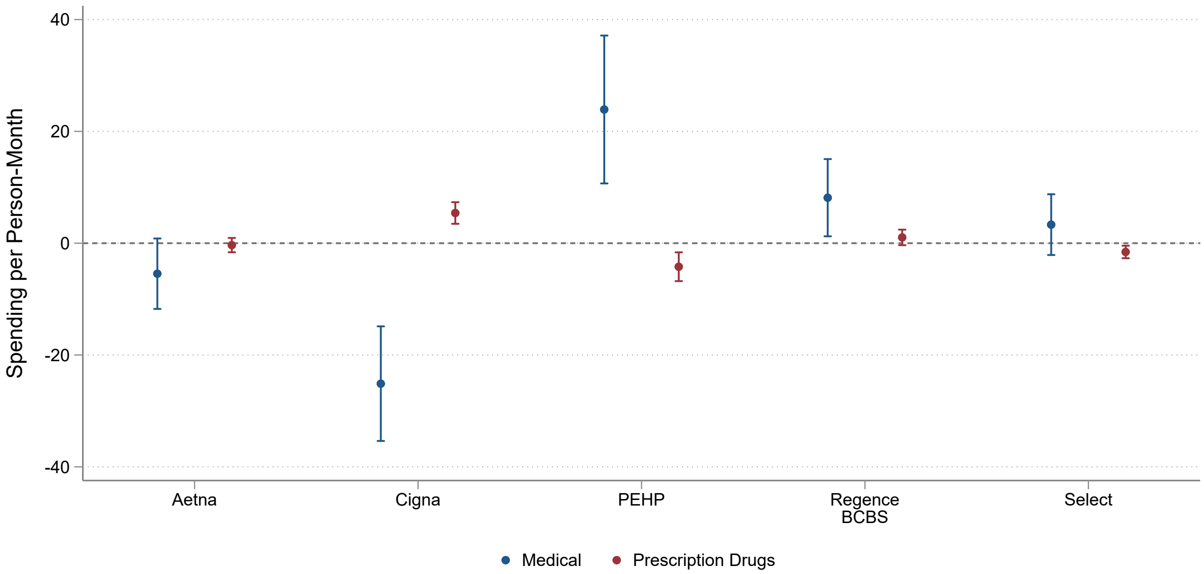
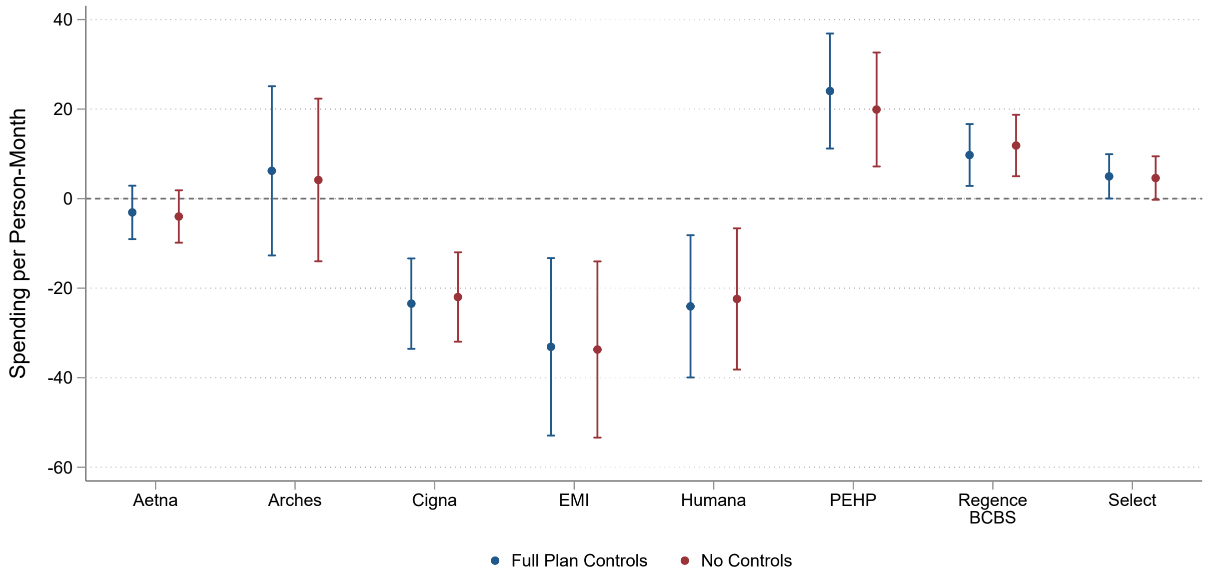


Figure A9: Core Results Excluding Small Growing Insurers



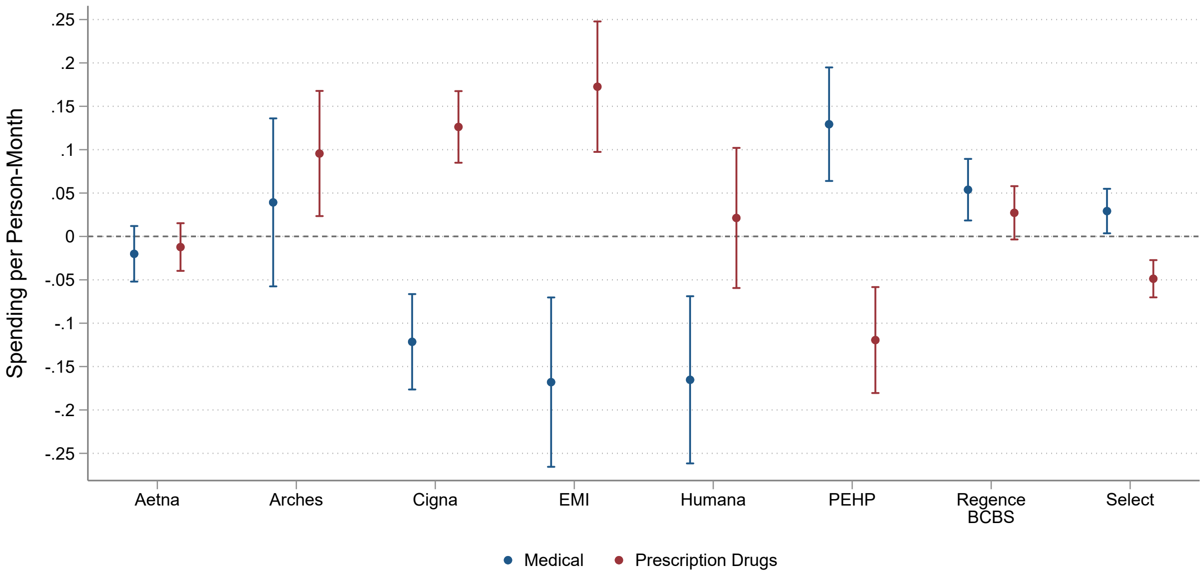
Notes: This figure replicates Figure 2 except we exclude the small insurers Arches, Humana, and EMI from the sample.

Figure A10: Core Results With and Without Controls



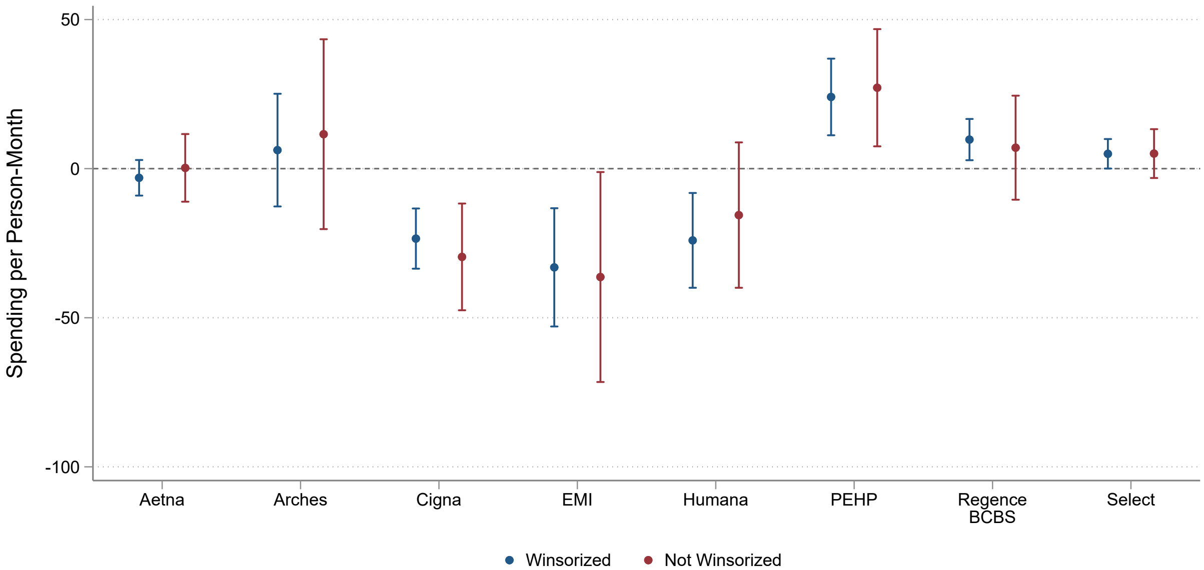
Notes: This figure replicates Figure 2 with and without controls. Results in blue replicate Figure 2 exactly. Results in red estimate an identical regression that does not include plan actuarial value controls $AV_{P(i,t)}$, but is otherwise identical. The dependent variable is total medical spending.

Figure A11: Core Results - Poisson Regression



Notes: Notes: This figure replicates Figure 2 using Poisson regression instead of OLS. We estimate a poisson regression using the ppmlhdfe command in Stata Correia, Guimarães and Zylkin (2020). The dependent variable is total medical spending, and the sample is a combination of forced switchers and job movers.

Figure A12: Core Results - Effect of Winsorization

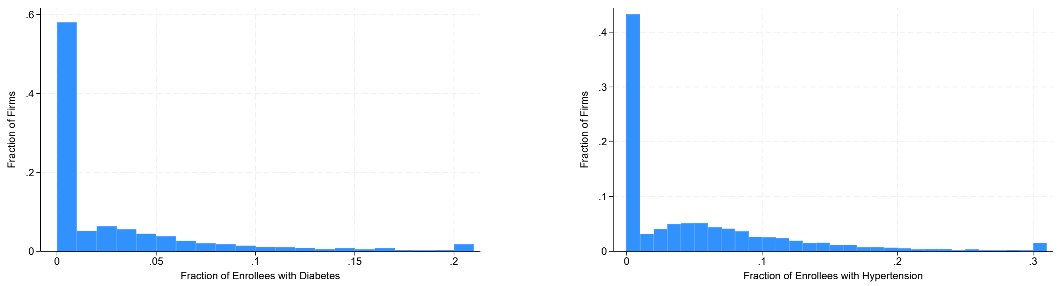


Notes: This figure replicates Figure 2 with and without winzorization. Results in blue replicate Figure 2 exactly. Results in blue report results without winzorizing the dependent variable. In both cases, the dependent variable is total medical spending, and the sample is a combination of forced switchers and job movers.

Figure A13: Distribution of Diabetes and Hypertension

(a) Diabetes

(b) Hypertension



Notes: These figures report the fraction of firms by disease burden (diabetes and hypertension, respectively). The sample is limited to firms with at least 5 employees. Hypertension is top-coded at 30% and diabetes at 20%.

Table A1: Medical Spending Variance Decomposition

	Forced Switcher Sample		Job Mover Sample	
	Component	Variance Share	Component	Variance Share
Medical Spending per Month	1.62e+06	100%	1.44e+06	100%
Person Effects (θ_i)	248589.65	15.3%	185379.22	12.9%
Insurer Effects ($\psi_{J(i,t)}$)	170.54	0.0%	167.72	0.0%
Plan Design ($\rho AV_P(i,t)$)	1151.13	0.1%	1035.93	0.1%
Month Effects (τ_t)	199.40	0.0%	192.77	0.0%
Residual (ε_{it})	1.37e+06	84.4%	1.25e+06	87.0%
Residual Match Component ($\Phi_{iJ(i,t)}$)	2.90	0.0%	11.20	0.0%
Cov($AV_P(i,t),\ \theta_i$)	-321.52	-0.0%	54.09	0.0%
Cov($AV_P(i,t),\ \psi_{J(i,t)}$)	-17.85	-0.0%	33.82	0.0%
Cov($\theta_i,\ \psi_{J(i,t)}$)	-10.72	-0.0%	-67.37	-0.0%

Notes: This table reports variance components from Equation 1, separately estimated using a sample of forced switchers and job movers. The dependent variable is total monthly medical spending.

Table A2: Drug Spending Variance Decomposition

	Forced Switcher Sample		Job Mover Sample	
	Component	Variance Share	Component	Variance Share
Prescription Drug Spending per Month	56091.48	100%	40167.46	100%
Person Effects (θ_i)	31644.05	56.4%	21193.73	52.8%
Insurer Effects ($\psi_{J(i,t)}$)	13.00	0.0%	7.48	0.0%
Plan Design ($\rho AV_P(i,t)$)	1.82	0.0%	10.71	0.0%
Month Effects (τ_t)	32.16	0.1%	13.36	0.0%
Residual (ε_{it})	24436.28	43.6%	18942.40	47.2%
Residual Match Component ($\Phi_{iJ(i,t)}$)	1.65	0.0%	6.76	0.0%
Cov($AV_P(i,t),\ \theta_i$)	0.28	0.0%	2.53	0.0%
Cov($AV_P(i,t),\ \psi_{J(i,t)}$)	-0.34	-0.0%	-0.35	-0.0%
Cov($\theta_i,\ \psi_{J(i,t)}$)	-13.19	-0.0%	0.69	0.0%

Notes: This table reports variance components from equation 1, separately estimated using a sample of forced switchers and job movers. The dependent variable is total monthly prescription drug spending.

Table A3: Yearly Insurer Differences for Patients with Diabetes

	Medical	Drug	Inpatient	Outpatient	Profess.
Total Spending Effects					
Aetna	-489*	-326*	-325	-41	-244*
Arches	-500	904*	-715	22	-208
Cigna	133	329*	-17	-41	105
EMI	-1064	831*	-950	33	-793*
Humana	-1095	-282	-883	22	-681
PEHP	1128	-294	1144*	318*	221
Regence BC	97	-124	212	-103*	152
Select	146	133*	13	35	95
Person-Yr Avg.	3933	3012	1454	620	2247
Quantity / RVU Effects					
	Medical	Drug	Inpatient	Outpatient	Profess.
Aetna	-407	44	-115	-83*	-134
Arches	-399	623*	-402	-85	-74
Cigna	-20	430*	-25	-65	28
EMI	-1119*	229	-807*	41	-652*
Humana	-884	507	-532	-67	-519
PEHP	699	-747*	390	166	275
Regence BC	89	-687*	140	-111*	50
Select	235	285*	35	115*	81
Person-Yr Avg.	3551	3311	1157	514	2099

Notes: This figure reports results from 10 regressions that each replicate the specification from Figure 8 with different dependent variables. We only show coefficients for the subgroup that has a diagnosis of diabetes. Fixed effects in both groups are normalized to have a mean of zero. In Panel (A), the dependent variable is total spending by category (medical, prescription drug, inpatient, outpatient, and professional). When calculating total spending, we only consider procedures that are also in panel B because we are able to calculate their predicted price. In Panel (B), the dependent variable is total utilization by category, where utilization is defined as the total amount that a generic Utah-based health insurer would spend on a given set of claims, using the methods in Section 4.1 to estimate prices. All variables are winsorized at the 99.9th percentile. Estimates that are statistically different from the mean across all insurers at the 95% level are starred.

Table A4: Yearly Insurer Differences for Patients with Hypertension

Total Spending Effects					
	Medical	Drug	Inpatient	Outpatient	Profess.
Aetna	-91	-150*	-227	11	-82
Arches	-246	252	-263	145	-62
Cigna	-557*	195*	-406	-85	-199*
EMI	-476	573*	-418	100	-424*
Humana	-722	-128	-666	5	-345*
PEHP	195	-133	687*	134	-250
Regence BC	227	-60	79	-87*	218*
Select	177	48	195	46	97
Person-Yr Avg.	3340	1354	1247	564	1895
Quantity / RVU Effects					
	Medical	Drug	Inpatient	Outpatient	Profess.
Aetna	-124	-3	-89	-37	-49
Arches	-114	134	-137	25	82
Cigna	-519*	240*	-421*	-59	-152
EMI	-325	179	-534	47	-219
Humana	-716*	186	-523	-125	-213
PEHP	67	-314*	446*	24	-67
Regence BC	214	-236*	178	-79*	124*
Select	204	113*	87	103*	52
Person-Yr Avg.	3069	1409	1024	455	1795

Notes: This figure reports results from 10 regressions that each replicate the specification from Figure A5 with different dependent variables. We only show coefficients for the subgroup that has a diagnosis of hypertension. The population includes all individuals who have ever had a diagnosis of hypertension. Fixed effects in both groups are normalized to have a mean of zero. In Panel (A), the dependent variable is total spending by category (medical, prescription drug, inpatient, outpatient, and Professional). When calculating total spending, we only procedures that are also in panel B because we are able to calculate their predicted price. In Panel (B), the dependent variable is total utilization by category, where utilization is defined as the total amount that a generic Utah-based health insurer would spend on a given set of claims, using the methods in Section 4.1 to estimate prices. All variables are winsorized at the 99.9th percentile. Estimates that are statistically different from the mean across all insurers at the 95% level are starred.

A.1 Standard Errors

In this section, we describe our method of constructing standard errors for insurer effects ($\psi_{J(i,t)}$) in equation 1. This method allows us to estimate standard errors for each insurer separately without having to declare one of our insurers as the reference or "leave-out" insurer. To our knowledge, this method of constructing standard errors is novel, and could be applied in any regression where a researcher wishes to compare different categories or brands.[31]

We begin with the standard fixed effects model:

$$y_{it} = x'_{it}\beta + \psi_{J(i,t)} + \epsilon_{i,t} \tag{8}$$

In equation 8, person i is assigned to insurer brand $J(i,t)$ in each period t, and x'_{it} is a vector of controls which could include individual-level fixed effects. Our interest is to compare brand-level fixed effects $\hat{\psi}_{J(i,t)}$ to test which of the k insurer brands are outliers relative to the market average. Equation 8 also has a matrix form listed in equation 9. In this equation, $\mathbf{F} = \{\mathbf{d_1}, \mathbf{d_2}, ..., \mathbf{d_k}\}$ is a matrix of dummy variables where each variable d_j is equal to 1 if $J(i,t) = j$ and zero otherwise.

$$\mathbf{Y} = \mathbf{X}\beta + \mathbf{F}\psi + \epsilon \tag{9}$$

If $(\mathbf{X}, \mathbf{F})$ is full column rank, we can estimate equation 9 directly using least squares. Then, $\hat{\psi}_j = \frac{1}{N_j}\sum \tilde{y}_{it}1\{J(i,t) = j\}$, where N_j is the number of observations in group j, $\tilde{y}_{it} = y_{it} - x'_{it}\beta$. In other words, $\hat{\psi}$ is simply the average residualized value of the dependent variable in each group.

However, in many cases $(\mathbf{X}, \mathbf{F})$ will be over-parameterized and cannot be estimated without dropping columns of X or F. For example, this is true whenever X contains the constant term or a second set of fixed effects. To combat perfect multicollinearity from an over-

[31] A related method is used to construct the Stata command felsdvregdm (Mihaly et al., 2010). Felsdvregem uses a similar way if constructing F except it replaces $F_j = -1$ if $d_1 = 1$, whereas our approach uses $F_j = -\frac{N_j}{N_1}$. Using the felsdvregdm yields group effects that are normalized by comparing across the average of all fixed effects as shown here:

$$\hat{\psi}_j = \frac{1}{N_j}\sum 1\{J(i,t) = j\}\tilde{y}_{it} - \frac{1}{N}\sum_{k\in J}\frac{1}{N_k}\sum 1\{J(i,t) = k\}\tilde{y}_{it} \tag{7}$$

Our method collapses to felsdvregdm in cases where all groups have an equal number of observations. In cases with different group sizes, the felsdvregdm method gives small groups a much larger weight when calculating the overall reference group. This means that the average of the reference group is more sensitive to outliers.

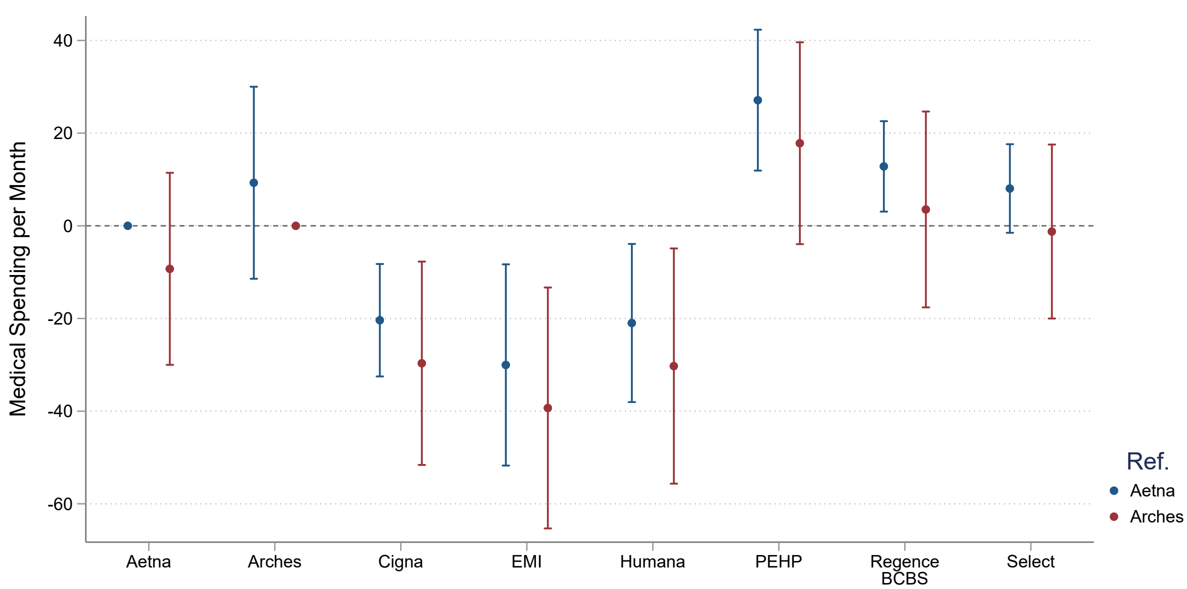
parameterized model, the default approach is to drop a column of $\mathbf{F}$. Define $\mathbf{F}_1 = \{d_2, ..., d_k\}$ to be the matrix where we drop the first column of $\mathbf{F}$. We show the estimate of $\hat{\psi}_j$ from estimating equation 9 using F_1 instead of F in equation 10.

$$\hat{\psi}_j = \frac{1}{N_j} \sum 1\{J(i,t) = j\} \tilde{y}_{it} - \frac{1}{N_1} \sum 1\{J(i,t) = 1\} \tilde{y}_{it} \tag{10}$$

This yields the familiar result that dropping the first column of F identifies the reference group as group 1, and $\hat{\psi}_j$ can be interpreted as the difference in the average residualized outcome $\tilde{y}$ between the focal group j and the reference group 1. In this specification, the choice of leave-out group has a material impact on the estimated variance.

In Appendix Figure A14, we show that the choice of leave-out group has a major effect on standard errors in our case. This figure shows results from estimating equation 1 dropping either Aetna or Arches as the leave-out group. Standard errors are larger when Arches is the leave-out group, sometimes significantly so. For example, the standard errors for SelectHealth are approximately twice as large when Arches is the leave-out group relative to Aetna. This is explained by the fact that Aetna is a much larger insurer, meaning the contrast between SelectHealth and Aetna is more precisely estimated than the difference between SelectHealth and Arches.

Figure A14: Insurer Treatment Effects Using Two Different Leave-Out Groups



Notes: This figure reports insurer fixed effects from a modified version of equation 1, estimated with the reference group being either Arches or Aetna. The dependent variable is total medical spending, and the sample is a combination of forced switchers and job movers.

In this paper, we define another parameterization of $\mathbf{F}$ which we denote $\tilde{\mathbf{F}}$. To construct $\tilde{\mathbf{F}}$, we first drop the first column of $\mathbf{F}$. We construct each other column of $\tilde{\mathbf{F}}$ as follows:

$$\tilde{F}_j = \begin{cases} 1 & if\, d_j = 1 \\ -\frac{N_j}{N_1} & if\, d_1 = 1 \\ 0 & otherwise \end{cases} \tag{11}$$

When combined with the constant, $\tilde{\mathbf{F}}$ spans the same space as $\mathbf{F}$, and so it does not affect the estimates of β or any other coefficients. Using $\tilde{\mathbf{F}}$ in equation 9 leads estimates of the brand effects to be:

$$\hat{\psi}_j = \frac{1}{N_j} \sum 1\{J(i,t) = j\} \tilde{y}_{it} - \frac{1}{N} \sum \tilde{y}_{it} \tag{12}$$

Using this approach, estimated fixed effects are compared not to a reference group, but to the overall mean outcome across all observations. Using this estimator has important advantages over choosing a reference group. First, each fixed effect is compared to the global mean rather than an arbitrary fixed effect. A statistically significant observation is an outlier relative to the market average, which is exactly the outcome of interest in this case. Second, both estimates of fixed effects $\hat{\psi}$ and the standard errors of these estimates are not sensitive to which group was dropped. Finally, because the fixed effects estimates are not sensitive to which column of F was dropped, it is possible to estimate all fixed effects separately by running the same regression twice and dropping different fixed effects. We do this in our core results, which report all the brand fixed effects together on the same graph.

A.2 Two-Step Estimator

Equation 1 incorporates elements of an event study to estimate disruption effects, combined with the decomposition framework developed by Abowd, Kramarz and Margolis (1999) to show variation across insurers. To ensure that we are not subject to the critiques from the recent literature on event studies (for two recent reviews, see Roth et al. (2023); and de Chaisemartin and D'Haultfœuille (2022)), we take steps to ensure that our model is identified and does not suffer from bias caused by the issue of negative weights.

To ensure identification, we include a large control group that never switches insurance companies or switches jobs. This control group identifies the overall time trend and seasonality patterns. We also bin the end-points for event-time dummies after 12 months.

Together, these two factors ensure that our model does not suffer from the identification issues described in Borusyak, Jaravel and Spiess (2024).

We estimate this model using ordinary least squares (OLS), including dummy variables to identify person, time, and disruption fixed effects. Under the additional assumption of homogenous treatment effects, OLS yields consistent estimates of all coefficients. Given that our panel length is just 2.5 years, we believe it is reasonable to assume that the disruption caused by switching insurance companies does not vary a lot over time.

However, if disruption effects varied strongly across switches with different event dates, recent work in event-study estimation has shown that estimates of these disruption effects could be biased because OLS puts some weight on "forbidden comparisons" (such as using a "treated group" as a control group). To confirm that this is not an issue in our case, we estimate a version of the model which is not subject to these concerns, and find nearly identical results.

Our alternative model is a two-step model. In the first step, we restrict only to our synthetic controls in Section 2 to construct a measure of healthcare spending that is purged of the pure time-based component, which includes general inflation in healthcare prices, technological change, and seasonality in utilization:

$$y'_{it} = \tau_t + y_{it}$$

In this equation, y'_{it} is the total healthcare spending of person i in month t, and τ_t is a set of month-by-year effects that account for any changes in average spending over time. We estimate this equation using only observations in our pure controls, who never switch insurance companies. We then use y_{it} as the main outcome variable in our estimating equation. The second step of this two-step process is:

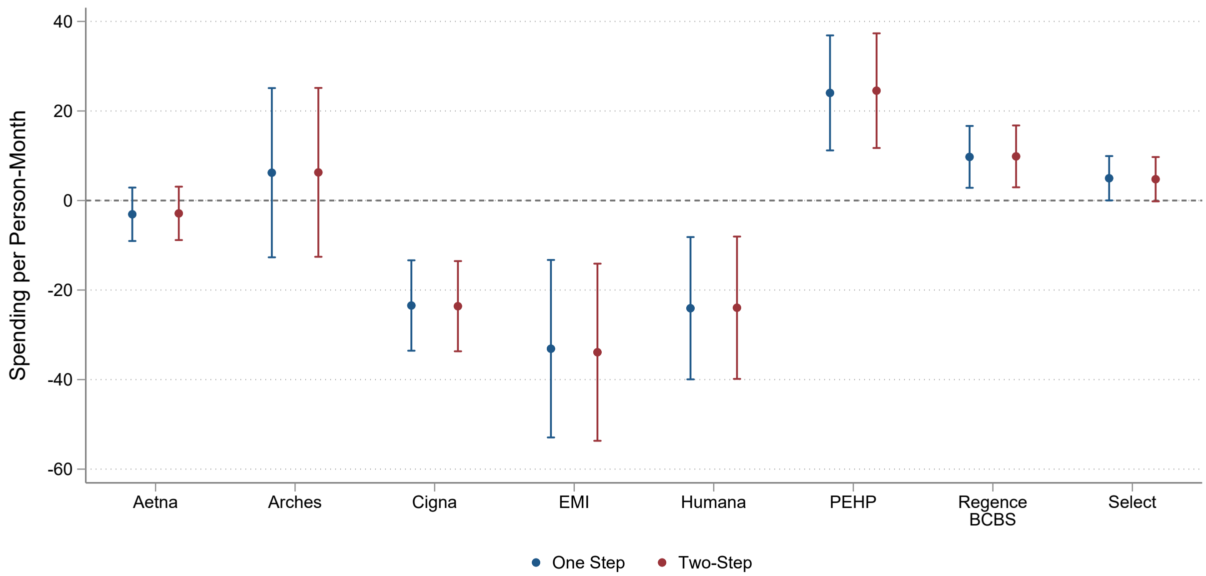
$$\begin{aligned} y_{it} &= \rho AV_{P(i,t)} + \theta_i + \psi_{J(i,t)} + \sum_{k \neq -2} \gamma_k * \mathbb{I}\left[J(i,pre) = J(i,post)\right] \\ &+ \sum_{k \neq -2} \gamma'_k * \mathbb{I}\left[J(i,pre) \neq J(i,post)\right] + \varepsilon_{it} \end{aligned} \tag{13}$$

This specification is identical to the specification in equation 1, except that it does not include month/year controls. Instead, the two-step regression explicitly identifies time trends on the basis of only the control group.[32] Doing this solves the problem of forbidden com-

[32]Note: In results not shown, we have also estimated equation 13 using a dependent variable that has not

parisons because treated observations are no longer implicitly being used to identify control observations.

Figure A15: Insurer Treatment Effects: One Step vs. Two Step Estimator



Notes: This figure compares one-step and two-step estimators of insurer effects $\psi_{J(i,t)}$. The one-step estimator is Equation 1, and the two-step estimator is Equation 13. In both cases, the dependent variable is total monthly medical spending.
Notes: This figure compares one-step and two-step estimators of insurer effects $\psi_{J(i,t)}$. The one-step estimator is Equation 1, and the two-step estimator is Equation 13. In both cases, the dependent variable is total monthly medical spending.

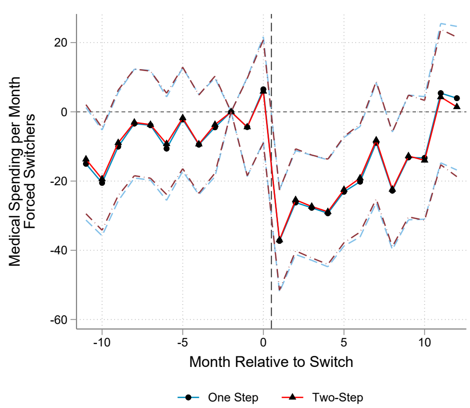
We show results estimating the regression in equation 13 in Appendix Figures A15 and A16. In Appendix Figure A15, we show insurer effects $\Psi_{J(i,t)}$, comparing our main one-step specification to the two-step estimator described in this appendix section. Both point estimates and standard errors are unaffected. In Appendix Figure A16, we compare estimates of the disruption effect from our core 1-stop estimator to the 2-step approach. They are also nearly identical.

From this exercise, we conclude that our results are not being driven by forbidden comparisons. Thus, the event-time dummies are the weighted average disruption effect of switching health insurance companies across all insurers in the data.

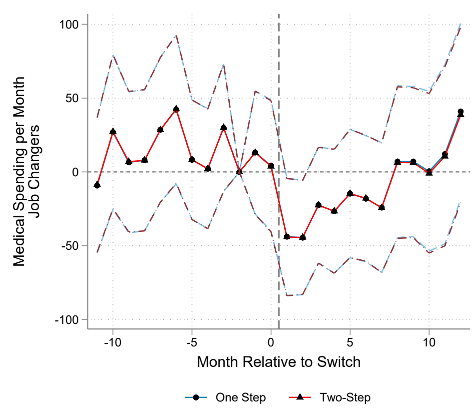
been de-trended. This version assumes that seasonality does not matter. This approach also gives nearly identical results, suggesting the overall seasonal pattern of health insurance spending does not matter over our short panel.

Figure A16: Disruption Effects in Medical Spending: One Step vs. Two Step Estimator

(a) Forced Switcher Sample



(b) Job Mover Sample



This figure compares one-step and two-step estimators of event-study coefficients γ_k. The One-step estimator is Equation 1, and the two-step estimator is Equation 13. We show results for the forced switcher sample in panel (a), and results for job movers who switch insurers in panel (b). The dependent variable is total monthly medical spending.