Financial Incentives and Physician Treatment Decisions: Evidence from Lower Back Pain*

Yu Ding† Chenyuan Liu‡

January 18, 2021

Click Here for the Latest Version

Abstract

In response to the high cost of health care, the capitated payment model has become more popular in recent years. Under capitation, physicians are compensated a fixed amount per patient regardless of the services generated. We provide new evidence on how the capitation payment model changes physician behavior by studying the treatment of lower back pain, a treatment type that has a lot of leeway for physicians' discretion. We use data from 2003 to 2006 from a large employer-sponsored health insurance claim database, and we leverage capitation variation within the plan and physician to mitigate selection concerns and isolate impacts from other supply-side cost containment strategies. We find that the treatment intensity—mainly from therapy, diagnostic testing, and drugs—of patients under a capitation system is 12% lower than otherwise similar patients in a non-capitated plan. We also find no evidence of increased readmission rates for patients in a capitated plan.

JEL Codes: I13, G22, I11. *Key words:* Capitation, physician behavior, health insurance.

*We thank Alan Sorensen and Justin Sydnor for their guidance in developing the project. We are also grateful to Naoki Aizawa, Ken Hendricks, Danny R. Hughes, John Mullahy, Margie Rosenberg, Joan Schmit, Chris Taber, Nicolas Ziebarth and conference participants at the 2020 World Risk and Insurance Economics Congress for their comments.

†yu.ding@wisc.edu. Department of Economics, University of Wisconsin at Madison

‡liuchy3@sem.tsinghua.edu.cn. Department of Economics, School of Economics and Management, Tsinghua University, Beijing 100084, China. Phone: +86-10-6278-1005.

1 Introduction

Health care spending accounts for a large and increasing share of gross domestic product in the United States. In response, some payers deviate from the traditional fee-for-service payment model and have adopted the capitated payment model. Under capitation, physicians are compensated based on the number of patients they treat rather than the volume of services they prescribe. Payers who have adopted capitation payments claim that it can reduce the use of medical services and the provision of low-value care—two factors contributing to the high cost of health care. For instance, Shrank, Rogstad and Parekh (2019) estimated that the annual cost of overtreatment rose from about $75.7 billion in 2012 to $101.2 billion in 2019. The literature also documents a range of potentially low-value care.[1] Essentially, the capitated payment contract transfers all or part of the financial risk to the physicians, encouraging them to be accountable for the quantity of services they provide. The recently established accountable care organizations in Medicare and private insurers are examples of a capitation payment model.

It can be challenging to assess whether capitated contracts lead to cost savings because there is selection into which providers and payers use capitation. For example, capitation contracts are more common in managed care plans such as health maintenance organizations (HMOs). These plans may attract patients with lower medical needs rather than truly reduce unnecessary care. Similarly, physicians who prescribe less care on average may be more willing to participate in a capitated payment model. Understanding the source of any potential cost differences driven by capitation contracts is essential in evaluating whether such incentives should be implemented more widely.

In this paper, we empirically examine the effects of capitated payment models on physicians' prescribing decisions. The movement toward managed care in the 1990s and early 2000s led to the growing popularity of capitation contracts in many places.[2] This historical movement toward capitation contracts provides an opportunity to study the issue. We focus on the treatment of lower back pain (LBP). The disease is economically significant: about 80% of the US population is affected by lower back pain at some point, and people with this condition spend more than $50 billion annually on treatment. More importantly, the treatment varies greatly across patients and providers (Smith, 2011). For example, using certain diagnostic imaging services is costly, and the benefits are limited (or even harmful for certain patients) (Chou et al., 2009, 2011;

[1] Examples include treatment of marginally-ill patients (Currie and Slusky 2020; Alalouf, Miller and Wherry 2019), use of reperfusion therapy to low-propensity patients (Chandra and Staiger, 2017), intensive post-acute care for marginal patients (Einav, Finkelstein and Mahoney, 2018), and cesarean delivery for low-risk pregnancies (Card, Fenizia and Silver, 2018).

[2] For example, Ho and Pakes (2014) document that 74% of primary care physicians in California were paid under capitation in 2003.

Flynn, Smith and Chou, 2011). However, the use of imaging is intensive: Schwartz et al. (2014) estimated that the 2014 Medicare spending on imaging for unspecified low back pain ranged from \$82 million to \$226 million. These findings raise the question of the efficacy of the treatment, and whether the capitated payment model helps reduce the overuse of low-value services like imaging when treating lower back pain.

The data we use is Truven MarketScan data from 2003 to 2006, a large commercial insurance claim data set based on the working-age US population. We construct and identify more than 60,000 episodes treating lower back pain. For each episode, we identify the primary care physicians, who play a central role in deciding the subsequent care. We directly observe from the data whether these physicians are paid under capitation, and we use this information as the key independent variable. For each episode, we also build a treatment intensity measure based on the weighted sum of the procedures performed. We construct the weights using a hedonic regression, where we regress price on patient age, gender, and year fixed effects using non-capitated contracts for all procedure codes separately. We then predict the average price for each procedure code. By taking this step, price variation is removed from the data, and we can focus only on variation in utilization.

We use a fixed-effects model to control for patient and physician selection into a capitation payment arrangement. First, patients who are treated by a primary care physician under a capitated plan may differ from other patients. To address this selection problem, we control for patient demographic information and chronic conditions generated from past claims. We also identify patients who stay in the same set of plans during the sample period and control for the plan fixed effects. By doing so, we leverage two sources of variation in capitation: from the same plan contracting with multiple providers with different capitation contracts; or from employers switching plans over time (for example, switch from a traditional plan to a managed care plan). This procedure allows us to control for unobserved patient selection into capitated plans. We further control for plan-year fixed effects to use only the first source of variation, thus separating other cost-control strategies from the impact of capitation. Second, we control for physician fixed effects to account for physician selection into capitated contracts. By doing so, we leverage variation in capitation from the same physician treating both capitated and non-capitated patients.

We find that patients treated by capitated physicians experience a moderate reduction in their overall treatment intensity. The overall treatment intensity is 12% lower for patients in a capitated model than for other patients, and results are robust to a range of different specifications. The treatment difference is mainly driven by the utilization of diagnostic testing

(30%), therapy (17%), and drugs (15%). There is almost no difference in the use of surgery.

We also find that the differences in treatment lead to very little difference in readmission rates for lower back pain in subsequent years. For patients in our benchmark sample who can be tracked over the next four years after the end of their episodes, we find that those in a capitated system have a very similar likelihood of having another LBP-related episode within 1 to 4 years. This finding suggests that capitation effectively reduces the use of treatment in lower back pain episodes without causing adverse treatment outcomes.

This paper contributes to the growing literature studying whether the capitated payment model reduces unnecessary care. Researchers have found mixed evidence on the impact of capitated contracts on the cost and quality of health care. Some studies provide evidence that capitation leads to lower costs (Gaynor, Rebitzer and Taylor, 2004; Ho and Pakes, 2014; Andoh-Adjei et al., 2018; Sacks, 2018). In contrast, others show a limited effect of capitation in controlling total health care expenditure or improving health care quality (Altman, Cutler and Zeckhauser, 2003; Duggan, 2004; Kontopantelis et al., 2015; Zhang and Sweetman, 2018). Many studies examine the effects of capitated arrangements using cross-plan or cross-insurer variation (e.g., Altman, Cutler and Zeckhauser 2003; Ho and Pakes 2014; Sacks 2018). The problem with such an approach is that capitation often exists along with other cost-control methods, such as a narrow network, utilization authorization, and selected covered benefits (Glied and Zivin, 2002). We offer new insights by leveraging episode-level variation in capitation and use plan-year fixed effects to separate the effects of capitated contracts from other supply-side cost-control incentives.

This study also contributes to the literature by considering employer-sponsored plans of a large-scale national sample, as opposed to plans only in a specific state (Ho and Pakes, 2014) or only Medicare/Medicaid plans (Duggan, 2004). The recent development in the Medicare bundled payments model, and Accountable Care Organizations (ACO) can be seen as a variation of the capitation payment model, under which physicians are compensated a fixed amount per capita for pre-specified episodes. Literature find evidence of reducing treatment in public insurance programs like Medicare (Einav et al., 2020; Eliason et al., 2020). Our results suggest that private insurance markets, such as employer-sponsored health plans, may also benefit from the capitation model.

More broadly, our work advances the literature exploring physician behaviors and the organization of care. Recent research finds that physicians respond strongly to financial incentives, including volume-based payments (Clemens and Gottlieb, 2014; Jacobson et al., 2013), for example, reimbursement from Medicaid (Alexander and Schnell, 2019) and Medicare (Einav, Finkelstein and Mahoney 2018; Maclean et al. 2018), payments from drug firms (Carey, Lieber and Miller, 2020), physician ownership of practices (Howard, David and Hockenberry, 2017), and episode-based payment (Carroll et al., 2018), etc. Our results indicate that physicians respond to the capitated compensation model by reducing treatment intensity in the case of lower back pain.

The rest of the chapter is organized as follows. Section 2 provides background information on lower back pain and capitation contracts. Section 3 presents the empirical strategy. Results are presented in Section 4. Section 5 concludes.

2 Background

2.1 Lower Back Pain

Lower back pain (LBP) is defined as "pain in the area on the posterior aspect of the body from the lower margin of the twelfth ribs to the lower gluteal folds with or without pain referred into one or both lower limbs that lasts for at least one day" (Deyo, Von Korff and Duhrkoop, 2015). LBP affects most adults, causes disability for some, and is a common reason for seeking healthcare (Deyo, Von Korff and Duhrkoop, 2015). According to the estimation of Luckhaupt et al. (2019), 26.4% of US workers have LBP, 8.1% have frequent and severe LBP, and 5.6% have work-related LBP.

Despite the prevalence of LBP, generally accepted guidelines for diagnosing LBP are absent (Koes et al., 2010). The diagnostic methods include medical history and physical exam, and imaging tests. When combined with clinical evaluations, imaging tests may help diagnose spinal problems. However, imaging tests are not always associated with clinically meaningful benefits, and they can even be harmful. In addition, many imaging tests poorly predict which patients will benefit from surgery (Chou et al., 2011; Goodney et al., 2015). Nevertheless, the utilization of imaging tests is high in the United States. For instance, Schwartz et al. (2014) estimated that the 2014 Medicare spending on imaging (excluding follow-up treatment because of the test results) for unspecified low back pain ranged from $82 million to $226 million.

There is no consensus on the best way to treat LBP either. LBP treatments include medications, noninterventional treatments such as physical therapy and exercise programs, and interventional spine surgeries and procedures. Surgical procedures range from well-established approaches for discectomies and spinal canal decompression to multiple means of addressing segmental fusion using several different approaches, materials, instruments, and indications. However, medical researchers find limited evidence to support the use of many interventional surgical procedures (see Friedly, Standaert and Chan (2010) for a review of the literature.) Meanwhile, the utilization of LBP surgeries continues to increase. For instance, the rate of spinal fusion operations for stenosis increased 67%, from 31.6 per 100,000 Medicare beneficiaries in 2001 to 52.7 per 100,000 Medicare beneficiaries in 2011 (Goodney et al., 2015). Disagreement also exists regarding the benefits of physical therapy, and "international guidelines contain conflicting recommendations for manipulation and exercise therapy" (Koes et al., 2001; Chou et al., 2007). Fritz et al. (2012) and Fritz, Brennan and Hunter (2015) find a large variation among physicians about whether to use and when to use physical therapies.

In summary, due to LBP's proliferation and wide variation in the treatment choices, we concentrate on LBP to examine how the capitation arrangement influences physicians' treatment decisions.

2.2 Capitation

To control health care expenditures, payers may replace a fee-for-service payment model with a capitated payment model by paying physicians based on the number of patients they treat instead of the volume of services they prescribe. Capitation contracts are most common with HMOs and are less common with preferred provider organizations (PPOs) and traditional plans. But even among HMOs, there is a large variation in whether capitation contracts are used. For instance, according to Zuvekas and Cohen (2010), only 15% to 33% of physician office visits for private HMO plan enrollees are under a capitation arrangement. Additionally, the capitation model is more common among primary care doctors and less prevalent among specialists.

The forms of capitation payments can vary. One extreme is the global capitation payment system, which bundles all providers and covers the cost of all services received by patients, including inpatient hospital stays. At the other extreme is a payment that covers only the services provided by the primary care physician or physician group. The latter type is almost always accompanied by "shared risk arrangements," under which a target is set for total spending. Cost savings or overruns relative to the target are shared between the primary care physicians and the insurers. Overall, the capitation payment system deviates from the traditional pay-for-volume model and generates incentives for physicians to share the financial risk of a patient's entire treatment episode.[3]

The capitation payment model appeared in the 1980s and thrived with the proliferation of HMOs. The rate of capitation payment among physicians has decreased since the early 2000s (Zuvekas and Cohen, 2010). Recent years have seen new reforms toward Medicare bundled

[3]See Ho and Pakes (2014) for details about capitation arrangements.

payment model and Accountable Care Organization initiatives; these variations of the capitation idea are intended to create financial incentives for physicians to curb medical expenditures (Friedberg et al., 2015).

3 Empirical Strategy

3.1 Data and Sample

We use data from the Truven MarketScan Commercial Claims and Encounter Data, a large commercial insurance claim data set based on the working-age US population. For each claim record, the data set provides diagnosis and procedure codes and detailed payment information. We directly observe whether a claim is paid under capitation, which allows us to estimate the effects of capitation payment on treatment intensity. Because the Truven MarketScan data also track enrollees over time, we can observe an individual's full medical service use history. We also observe other demographic and socioeconomic characteristics, including age, gender, and employment status. The sample year is from 2003, the first year the capitation measure is reported, to 2006, the last year with enough observations under capitation in the data.

We construct a sample of LBP-related episodes. To build this sample, we first identify a patient's claim encounters with LBP-related diagnoses following the medical literature (see Cherkin et al. (1992*a*) and Appendix Table A.2). We then group these encounters into episodes based on service type and timing. An LBP episode starts from a patient's earliest LBP encounter, followed by subsequent encounters with a time gap shorter than 180 days. An episode ends if there is no additional LBP encounter within 180 days of the last record. Two consecutive LBP encounters that occur more than 180 days apart are designated as two separate episodes. Based on this definition, most patients have one episode during the sample period.[4]

We trim the sample in the following ways to keep the episodes homogeneous and remove non-emergency episodes less likely to be affected by physicians' financial incentives. First, we keep only episodes started with a primary care office visit. This step removes incomplete episodes whose full treatment history is not in our sample. It also removes episodes started with surgical treatment, which follow different treatment strategies compared with non-acute conditions. Second, we exclude episodes with LBP treatment that happened in emergency care or out-of-network providers because we are concerned that physicians face different financial incentives under out-of-network encounters. Third, we remove certain types of patients whose treatment of LBP is different in nature. These include pregnant women (who may have pregnancy-related

[4]We also consider other choices of the time gap, from 90 days to 270 days, and report the robustness of our main results with these alternative definitions of episodes in Appendix Table A.1.

LBP), people with certain severe chronic diseases (whose treatment of LBP will be complicated by other conditions), and people under age 18 or over age 65 (who may have insurance other than provided by their employer).[5] In total, the sample includes 61,369 episodes from 55,620 patients.

3.2 Capitation Measure

We define the key independent variable of capitation based on whether the primary care physician is paid under capitation as recorded in the claim data. We choose primary care physicians because they play a critical role in deciding different treatment options, both directly by prescribing treatment and drugs, or indirectly by referring to other specialists. They are also most frequently targeted by capitation arrangements. Under capitation, insurers often remunerate primary care physicians through fixed monthly payments per patient to cover the cost of patient services, and sometimes reward primary care physicians for savings from the entire episode. Therefore, the capitation arrangement generates a financial incentive for primary care physicians to save on patient treatment.

There are two important caveats about how to interpret our results correctly. First, we do not observe the specific financial terms of the capitation arrangements, so our measure cannot capture the "intensity" of the financial incentives, but rather represents a range of different capitation levels. Second, even though our capitation measure is based on primary care physicians, the arrangement may apply to other downstream specialists, such as radiologists, surgeons, and therapists. We do not use the capitation measure based on these specialists because the claims are not clearly mapped to different specialists. Given that the capitation status of primary care physicians is positively correlated with the capitation status of downstream claims in our data, our capitation measure captures both the capitation status of the primary care physician and the status of the entire episode.

In our sample, patients in the capitated plans are healthier. The sample defined above includes 8,133 capitated LBP episodes and 53,236 non-capitated episodes. In Table 1 we compare the patient individual characteristics of capitated and non-capitated episodes. The patients in capitated plans are slightly younger than their counterparts receiving treatment in a non-capitated system, and they are less likely to have chronic conditions. We also find that patients who receive care in a capitation arrangement are more likely than others to be paid hourly and work part-time.

[5]The chronic conditions we rule out include colorectal cancer, lung cancer, female/male breast cancer, endometrial cancer, prostate cancer, Alzheimer's disease and related disorders or senile dementia, heart failure, acute myocardial infarction, stroke/transient ischemic attack, and hip/pelvic fracture.

Table 1: Summary Statistics of Capitated/Non-capitated Patients, Patient Characteristics

	Capitated		Non-capitated		Difference	
	Mean	SD	Mean	SD	Mean	SE
Demographics						
Female (%)	56.95	49.52	56.21	49.61	0.75	0.59
Age	43.78	10.87	45.00	10.87	-1.22	0.13
Health Status (%)						
Acquired Hypothyroidism	5.86	23.50	7.81	26.84	-1.95	0.29
Anemia	4.30	20.29	5.85	23.47	-1.55	0.25
Cataract	2.61	15.93	3.54	18.49	-0.94	0.19
Obstructive Pulmonary/Bronchiectasis	5.83	23.43	7.39	26.17	-1.57	0.28
Chronic Kidney Disease	1.91	13.67	1.89	13.62	0.02	0.16
Diabetes	8.09	27.27	9.41	29.20	-1.32	0.33
Hyperlipidemia	20.99	40.73	28.79	45.28	-7.80	0.49
Depression	8.63	28.08	9.70	29.60	-1.07	0.34
Hypertension	19.00	39.23	27.08	44.44	-8.08	0.48
Glaucoma	2.71	16.22	3.55	18.51	-0.85	0.20
Ischemic Heart Disease	3.58	18.58	5.26	22.32	-1.68	0.23
Atrial Fibrillation	0.53	7.25	0.80	8.90	-0.27	0.09
Asthma	6.07	23.89	6.01	23.77	0.06	0.28
Benign Prostatic Hyperplasia	1.52	12.25	2.64	16.02	-1.11	0.15
Rheumatoid Arthritis/Osteoarthritis	14.04	34.74	19.07	39.29	-5.03	0.42
Osteoporosis	2.80	16.51	2.91	16.81	-0.11	0.20
Compensation Classification (%)						
Salary Non-union	3.26	17.76	12.25	32.79	-8.99	0.24
Salary Union	0.18	4.29	0.81	8.94	-0.62	0.06
Salary Other	0.06	2.48	1.04	10.13	-0.98	0.05
Hourly Non-union	1.57	12.45	8.73	28.23	-7.16	0.18
Hourly Union	5.02	21.83	7.52	26.36	-2.50	0.27
Hourly Other	0.02	1.57	1.08	10.32	-1.05	0.05
Non-union	3.91	19.38	8.36	27.68	-4.45	0.25
Union	0.15	3.84	1.72	12.98	-1.57	0.07
Unknown	85.82	34.88	58.50	49.27	27.32	0.44
Employment Status (%)						
Active Full Time	19.85	39.89	47.94	49.96	-28.09	0.49
Active Part Time or Seasonal	0.14	3.68	1.35	11.54	-1.22	0.06
Early Retiree	1.62	12.64	5.78	23.34	-4.16	0.17
Medicare Eligible Retiree	0.12	3.50	0.47	6.85	-0.35	0.05
Retiree (status unknown)	0.20	4.43	0.19	4.33	0.01	0.05
COBRA Continuee	0.07	2.72	0.51	7.09	-0.43	0.04
Long-Term Disability	0.02	1.57	0.31	5.59	-0.29	0.03
Surviving Spouse/Depend	0.00	0.00	0.17	4.11	-0.17	0.02
Other/Unknown	77.98	41.44	43.29	49.55	34.69	0.51
Number of Observations	8,133		53,236			

Note: The table shows the summary statistics of patient characteristics for capitated/non-capitated patients separately. SD = standard deviation. SE = standard error.

3.3 Treatment Intensity Measures

We build the measure of treatment intensity of an episode using procedure codes in the claim data. Each procedure code document the specific treatment patient received. An episode often contains hundreds to thousands of procedure codes. To aggregate all procedures at the episode level, we calculate the weighted sum of all the procedures performed in that episode, where the weights are the expected average price of each procedure $\bar{p}_z$:

$$t = \sum_z \bar{p}_z f_z, \tag{1}$$

where f_z is the quantity of each procedure code z, and $\bar{p}_z$ is the weight.

For each medical claim, we observe the transaction price p. The price measure is the actual amount insurers paid to the provider (not list price), including both plan payment and consumer cost-sharing. This price represents the overall resources used for each procedure and captures price variation among insurers and providers. Since our focus is on understanding utilization patterns, we want the treatment intensity measure to reflect only differences in service utilization and not differences in negotiated prices for services across different plans. To eliminate the variation in negotiated prices, we calculate each procedure's average price by regressing the transaction price on the patient's age, gender, chronic conditions, and year fixed effects. We control for these patient characteristics because they might affect the resources used. We use year fixed effects to remove the time trend of medical prices. In this estimation step, we only use the claims from non-capitated claims, because the price is often not accurately reported for capitation contracts. We then predict the price for all claims with that procedure code to get $\bar{p}_z$.

The treatment intensity measure has a bimodal distribution and is highly skewed. Most people receive minimum or no treatment, while some patients receive very intensive treatment. To account for the skewness of the data, we transform the raw treatment intensity measure into log scale using the inverse hyperbolic sine transformation:

$$IHS(t) = log(t + \sqrt{t^2 + 1}).$$

The inverse hyperbolic sine transformation behaves similarly to logarithms transform and preserves zero as zero.

For each episode, we also construct the treatment intensity measure for different types of medical services. We classify LBP-related medical claims into five categories: office visit, diagnostic testing, therapy session, surgeries directly related to LBP treatment, and other surgeries

(Cherkin et al., 1992*b*). We also construct a dummy variable indicating whether each type of service is used at all in an episode. Every observation will have an office visit, but some may not have other services.

Table 2 shows the summary statistics of the outcome variables. The average treatment intensity for all services within an LBP episode for patients in a capitation system is around $421, while that of patients in other types of plans is $584. The average treatment intensity is significantly higher for patients in non-capitated plans for nearly all service categories except for back surgery.

Table 2: Summary Statistics of Capitated/Non-capitated Patients, Outcome

	Capitated		Non-capitated		Difference	
	Mean	SD	Mean	SD	Mean	SE
Treatment Intensity, $						
All Medical Services	420.69	1457.98	584.38	1563.50	-163.69	17.53
Office Visit	151.33	227.22	167.74	227.27	-16.41	2.71
Therapy	47.91	236.53	89.53	376.73	-41.63	3.09
Back Surgery	62.59	469.21	105.40	586.37	-42.81	5.79
Other Surgery	82.83	305.19	114.16	316.16	-31.33	3.65
Diagnostics	70.27	905.26	99.38	956.48	-29.11	10.86
Number of Observations	8,133		53,236			
Drug Usage Intensity, $						
All Drug Usage	213.17	386.50	1163.53	1877.62	-173.33	16.64
Muscle Relaxants	5.00	7.82	23.83	42.05	-2.81	0.35
Opioids	24.84	32.42	290.38	429.66	-7.58	4.04
Number of Observations	7,303		38,503			

Note: The table shows the summary statistics of outcome variables for capitated/non-capitated patients separately. SD = standard deviation. SE = standard error.

For 75% of the episodes in our baseline sample, we observe the related drug claims. For these episodes, we identify LBP-related drug prescriptions and all subsequent refills for these prescriptions. We then construct a similar treatment measure for overall drug use, and the two most common types of drugs: opioids and muscle relaxants. To do so, we group drug claims by a national drug code. We then calculate the average per-day price for each drug in our sample by year. Finally, we multiply the average price by the number of days of supply to determine per-drug spending. The episode-level total drug usage is the sum of the spending on all drugs. This usage measure takes the same price for a specific drug across different plans and insurers and reflects only usage differences, not price differences.

3.4 Regression Model

3.4.1 Baseline Analysis

As noted in Section 3.1, selection is a potentially large problem in our data. As a first way to address the problem, we control for patient characteristics X, including age, gender, employment status, and the existence of chronic conditions.

One important channel through which selection might happen is patients' choice of health insurance plans. For example, capitation contracts are more common in managed care plans, and especially HMO plans, than other plans. However, these plans are different along other dimensions: they may have different demand-side cost-sharing, and impose other cost-control strategies, such as narrower networks, referral restrictions, etc. One concern is that patients in HMOs are different from patients who choose traditional plans. To remove concerns that patient selection of plans confounds the effects of capitation, we control plan fixed effects. Though patients can select different plans based on plan characteristics, the primary care physicians are often assigned by the plan based on zip codes. Besides, the financial arrangement between providers and plans is often unknown to patients. As a result, the capitation differences within the same plan is unlikely to be affected by patient selection.

To construct the plan fixed effects measure, we track individuals' plan choices over the sample period. If plan ID information is not available, we use the unique combination of the insurer, employer, and plan type to impute the plan ID. We then group patients who stay in the same set of plans over time and label them as being in the same "plan." For example, all patients who chose plan A in 2003, switched to plan B in 2004, switched to plan C in 2005, and stayed in plan C in 2006 will be labeled as in the same "plan." This accounted for the fact that when some employers completely change the plan menu over time, there is no active patient selection. Our measure of plan identifier guarantees that comparisons are within patients with the same history of plan choices.

Another source of selection comes from the provider side. Primary care physicians may have different preferences toward capitated arrangements and treatment philosophies, and they may actively select capitation contracts based on their treatment style. For example, physicians who prescribe less intense treatment on average may be more willing to join a capitation contract. To account for this selection channel, we include provider fixed effects in our model. By doing so, we compare capitated and non-capitated patients treated by the same provider.

Our baseline regression models are:

$$y_{it} = \alpha + \beta_1 CAP_{it} + X_{it}\beta_X + \delta_s + \gamma_g + \theta_t + \epsilon_{it}. \tag{2}$$

Here, i is the index for each episode (level of observation), and y_{it} is either the log treatment intensity measure or a dummy variable indicating whether a certain service is used. CAP_{it} is a dummy variable indicating whether the associated primary care physician is under capitation. θ_t indicates year fixed effects. If an episode expands into multiple years, we use the initial visit date to determine the year. δ_s are physician fixed effects and γ_g are plan fixed effects. Controlling for physician fixed effects removes the impact of time-invariant physician characteristics, and plan fixed effects removes the impact of time-invariant factors determining plan choices. Equation (2) controls for both of these effects separately in the same equation. Under the assumption that physician fixed effects are similar across different plans conditional on all other variables, this model controls both physician and patient selection. β_4 represents the treatment effects for similar patients treated by the same physician.

In our baseline model and all subsequent regression models, we cluster the standard errors at the employer or insurer level. The Truven MarketScan data are collected either from employers or insurers. It's reasonable to assume that error terms are likely correlated among episodes from the same employer or insurer.

For robustness check, we also estimate a model controlling for the interactive term of plan and provider fixed effects:

$$y_{it} = \alpha + \beta_2 CAP_{it} + X_{it}\beta_X + \delta_s\gamma_g + \theta_t + \epsilon_{it}. \tag{3}$$

Since there is no variation of capitation status for a plan-provider pair in the same year, β_2 is estimated based on provider-plan pairs changing the capitation arrangement over time. In our sample, 40% observations are thus absorbed by the interactive fixed effects and are not used to estimate β_2. The smaller sample size makes the estimation less precise than our baseline specification, so we do not use it as the benchmark. We will discuss this point in detail in Section 4.

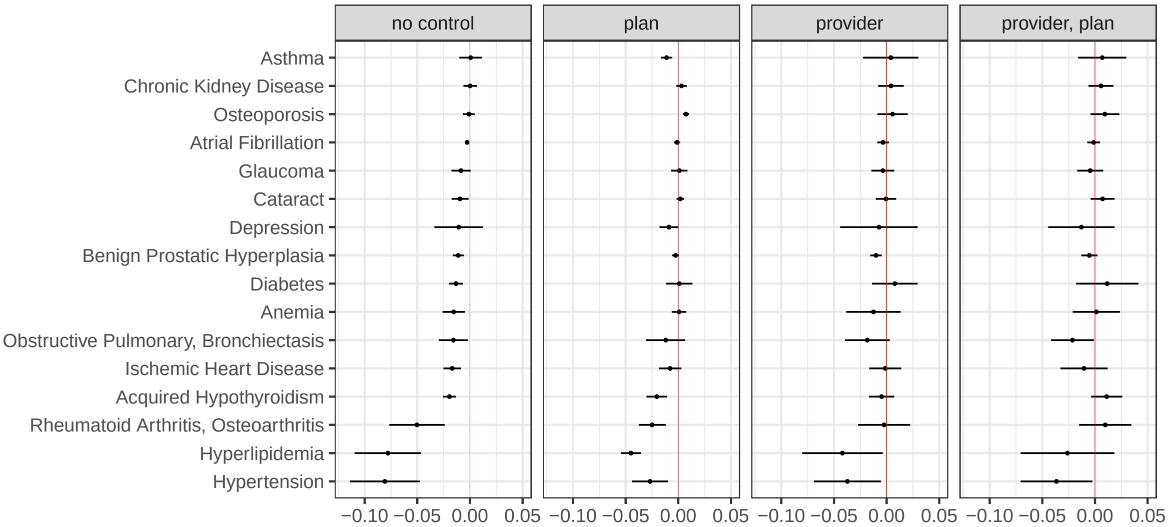
We have done several extra analyses to justify the fixed effects model in reducing selection concerns. First, we show that our fixed effects models can remove selection on observables. Figure 1 offers a comparison of the likelihood of having chronic conditions among patients in capitated and non-capitated plans. The first panel on the left contains the raw mean differences. Patients in a capitation system are less likely to have most of the chronic conditions without controls. Controlling for plan and provider fixed effects reduces the differences to almost zero for almost all chronic conditions. For example, patients under capitation are 7.5% less likely to have high blood fat (hyperlipidemia) than patients in other types of plans under no controls. The estimated difference for hyperlipidemia decreases to 4% once we add plan or provider fixed effects separately, and the difference is 2.5% when we include both plan and provider fixed effects. If unobservables are similar in nature to observables, then our fixed effects model will account for the selection problem.

Another way to assess whether plan and provider fixed effects removed selection concerns is to estimate a model without controlling for individual characteristics:

$$y_{it} = \alpha + \beta_3 CAP_{it} + \delta_s + \gamma_g + \theta_t + \epsilon_{it}. \tag{4}$$

If β_3 is similar to β_1, then the fixed effects model is effective in removing selection concerns.

Figure 1: Chronic Condition Rate Differences between Capitated/Non-Capitated Patients



Note: The dependent variable is whether the patient has a certain chronic condition. The independent variable is whether the patient is under a capitated plan. The four panels from left to right indicate four regression models: no controls; controlling for year, age, and plan fixed effects; controlling for year, age, and provider fixed effects; and controlling for year, age, and plan and provider fixed effects. Lines indicate 95% confidence interval. Standard errors are clustered at the insurer/employer level.

3.4.2 Cross-time and Cross-sectional Variation Decomposition

In our baseline analysis, the impact of capitation is estimated by comparing capitated and non-capitated patients within the same plan and providers. In general, the variation in capitation comes from two sources: the same plan (provider) have different capitation arrangements with different providers (plans) in the same year (cross-sectional variation); the same plan (provider) may change capitation arrangement over time (cross-time variation). To better understand which source drives the overall effects, we decompose the estimated impacts into these two sources.

Decomposition: Plan We first include only plan fixed effects and estimate the following equation:

$$y_{it} = \alpha + \beta_4 CAP_{it} + X_{it}\beta_X + \gamma_g + \theta_t + \epsilon_{it}, \tag{5}$$

We then estimate the impacts of capitation on treatment using only variation in capitation within the same plan-year:

$$y_{it} = \alpha + \beta_{41} CAP_{it} + X_{it}\beta_X + \gamma_{gt} + \epsilon_{it}. \tag{6}$$

Model (6) controls for plan-year fixed effects, so the variation in capitation is from the same plan setting different capitation arrangements with different providers. One benefit of this model is that it also removes the impact of other cost-control methods that vary at the plan level across year. Often capitation happens along with other supply-side cost-control methods, such as utilization authorization and referral restriction. These measures, however, usually vary across plans and are the same within a plan-year. By controlling for plan-year fixed effects, we can hold fixed the variation of other supply-side cost-control measures and identify the net effects of capitation.

Next, we estimate the effects of capitation on treatment intensity using only cross-time change in capitation arrangements:

$$y_{it} = \alpha + \beta_{42} CAP_{gt} + X_{it}\beta_X + \gamma_g + \theta_t + \epsilon_{it}. \tag{7}$$

In model (7), we calculate the average capitation rates within a plan-year, CAP_{gt}, and we use this as the new independent variable. Since we control for plan fixed effects, the coefficient of CAP_{gt} reflects the change in capitation of a specific plan over time.

Decomposition: Provider To understand effects of cross-sectional and cross-time physician variations, we first estimate a model with provider fixed effects:

$$y_{it} = \alpha + \beta_5 CAP_{it} + X_{it}\beta_X + \delta_s + \theta_t + \epsilon_{it}, \tag{8}$$

where y_it is either the log treatment intensity measure $IHS(t)$ or a dummy variable indicating whether a certain service is used, and δ_s are dummy variables for different providers. β_5 captures the difference in treatment decisions for patients in both capitated and non-capitated arrangements who are treated by the same provider.

We further decompose the treatment effects we estimated in equation (8) into the cross-section variation (based on the first source of variation) and cross-time variation (based on the second source of variation) in the following two models:

$$y_{it} = \alpha + \beta_{51} CAP_{it} + X_{it}\beta_X + \delta_{st} + \epsilon_{it}. \tag{9}$$

β_{51} is estimated using the fact that the same physician may enter different contracts with different plans in the same year. Controlling for the provider and year fixed effects will remove the concern that a physician's treatment style is correlated with her or his decision to enter a capitation arrangement, because we compare the treatment within a year. We can also use this model to evaluate whether the provider can differentiate treatment for different patients in the same year. Anecdotal evidence indicates that physicians may not vary their treatment decision among patients with different underlying reimbursement contracts in the same year. To the extent that this is true, a null effect in this model might not indicate that the true treatment effect is zero.

Next, we estimate the impact of capitation on treatment intensity using the fact that the same provider may switch capitation arrangements over time. We estimate the following model:

$$y_{it} = \alpha + \beta_{52} CAP_{st} + X_{it}\beta_X + \delta_s + \theta_t + \epsilon_{it}, \tag{10}$$

where CAP_{st} is the fraction of episodes under capitation for provider s in year t. Model (10) examines how providers' treatment decision change over time when providers move from fewer patients in capitated contracts to more patients in capitated contracts. This specification removes time-invariant physician characteristics correlated with treatment and capitation choice. Under the assumption that there is no change in treatment philosophy that is correlated with the decision to switch between capitation contracts, the model will identify the true treatment effects.

4 Results

4.1 Treatment Intensity of Medical Services

Baseline Table 3 shows the differences in treatment intensity for lower-back pain episodes between capitated and non-capitated patients. The dependent variable is transformed into a log-scale using inverse hyperbolic sine transformation, so the coefficients of "Capitated" are approximately percentage change if the patient switched from a non-capitated physician to a capitated physician. Standard errors are clustered at the insurer (if the insurance company provides the data) or employer level (if the employers provide the data).

Our baseline estimates are presented in column 1. The model controls for both plan and provider fixed effects. The estimate indicates that the patients in a capitation system utilize 12.2% fewer medical resources than patients in non-capitated plans, and the coefficient is significant at the 1% level. Given that the average total expenditure (including insurer payment and patient out-of-pocket spending) of an episode in our sample is around \$563, the treatment intensity differences transform to \$69 expenditure difference per episode (equivalent to \$89 medical expenditure in 2019).

The estimates are robust using different sets of fixed effects. Table 3 column 2 shows the results with interacted provider and plan fixed effects. The point estimate is similar, though because many observations are absorbed by the interactive plan and provider fixed effects, the standard errors are larger. Column 3 shows the results with only provider fixed effects and plan fixed effects, and no individual characteristics. The coefficient has a similar magnitude and significance level as the baseline specification in Column 1. In other words, once controlled for fixed effects, observed individual characteristics are no longer important in explaining the treatment intensity. The result suggests that the fixed effects model is effective in removing selection based on observables. If we believe the unobserved variables are similar to the observed individual characteristics, our baseline specification will account for potential selection through the unobservables. Columns 4 and 5 show the results with plan fixed effects alone and provider fixed effects alone. Both are statistically indifferent from the baseline estimates.

Cross-Section and Cross-Time Variation in Capitation The impact of capitation on treatment intensity in Table 3 are estimated based on both the cross-section and cross-time variation in capitation. Table 4 and Table 5 presents the estimated effects from either source.

First, we break down the estimates into cross-plan variation and cross-time variation. Table 4 Column 1 uses the same specification as column 4 in Table 3, controlling for year fixed effects, patient characteristics, and plan fixed effects. In column 2, we control for plan $\times$ year

Table 3: Regression Results Comparing the Treatment Intensity of All Services for Patients under Capitated and Noncapitated Arrangements

	1	2	3	4	5
Capitated	-0.122***	-0.101	-0.114***	-0.115***	-0.072
	(0.039)	(0.068)	(0.038)	(0.038)	(0.054)
Number of Observations	60,205	36,474	60,205	60,405	61,369
R-squared	0.350	0.401	0.333	0.076	0.323
Prov Fixed Effects	×		×		×
Plan Fixed Effects	×		×	×	
Prov × Plan Fixed Effects		×			
Individual Characteristics	×	×		×	×

Note: The table shows the regression results comparing the treatment intensity of capitated/noncapitated patients. The dependent variable is the inverse hyperbolic sine transformation of treatment intensity of all services. In all models, we control for year fixed effects. The underlying sample is the same across all columns. The "Number of Observations" is different because some observations are absorbed by different fixed effects, and they are excluded from the "Number of Observations" calculation. Standard errors are clustered at the insurer/employer level. $^{*}: p < 0.1,^{**}: p < 0.05,^{***}: p < 0.01$.

fixed effects to identify the impacts of capitation from variation within a year and plan. This specification shows that the treatment intensity of patients treated by capitated physicians is 10.9% lower than that of patients treated by noncapitated physicians, similar to our baseline results. The coefficient is significant at the 1% level. Sometimes a plan may change its network types, other supply-side cost controls, or demand-side cost-sharing attributes across years. For example, a plan may switch from HMO to PPO. However, a health insurance plan typically sets the same supply-side cost-control methods (e.g., utilization authorization, referral restriction, network breadth, etc.) and demand-side cost-sharing within a year. Thus our estimated difference in Column 2 is solely driven by the same plan's capitation arrangement with different providers, holding fixed other potential factors. The results suggest that our baseline estimates are not confounded by other supply-side or demand-side cost control methods.

In Table 4 Column 3, we use the average capitation rate of a plan within a year as the independent variable. This specification uses the plans' variation across years to identify the effect of capitation. The result indicates that a plan with all patients under capitation has 37.2% lower treatment intensity than a plan with no capitation arrangement. In Column 4, we use a dummy variable indicating whether a plan-year has any capitation arrangement. This specification shows that having capitation with some physicians in a plan leads to a 14.9% reduction in the treatment intensity of all services relative to a plan with no capitated patients. Our results show that the reduction in treatment effects results from both differences in capitation rates across plans and the change in capitation within the same plan across years.

Second, we examine the sources of capitation's impacts on treatment intensity from variation within a physician. Table 5 Column 1 shows the results controlling for provider fixed effects. We then break down the estimates using variation in capitation for the same provider within a year. Column 2 shows that, on average, physicians do not vary treatment decisions for patients with different capitation arrangements within a year. The point estimate is much smaller and not statistically different from zero. The results suggest that if a physician has both patients from capitated plans or fee-for-service plans, the physician does not differentiate treatments among these patients. One may wonder whether capitation arrangements in the commercial insurance market have spillover effects on public health insurance programs. Based on our findings that physicians have similar treatment intensity to patients with different capitation arrangements in a given year, it's unlikely the capitation arrangement in commercial insurance programs will have large spillover effects to other insurance programs, not at least in the short term.

The impacts of capitation on treatment intensity mainly come from a change in capitation arrangements across time. In Table 5 Column 3, the independent variable is the average capitation rate for the provider within a year, and the independent variable in Column 4 is whether the provider has any patient under capitation in a year. In either specification, capitation results in about a 10% reduction in treatment intensity.

Table 4: Cross-sectional and Cross-time Plan Capitation Variation on Overall Treatment Intensity

	1	2	3	4
capitated	-0.115***	-0.109***		
	(0.038)	(0.036)		
Average Capitation Rate of Plan-Year			-0.372**	
			(0.178)	
Any Capitated within Plan-Year				-0.149***
				(0.069)
Number of Observations	60,405	59,309	60,405	60,405
R-squared	0.076	0.094	0.076	0.076
Plan Fixed Effects	$\times$		$\times$	$\times$
Plan $\times$ Year Fixed Effects		$\times$		

Note: The table examines the cross-sectional and cross-time variation of plan's capitation status change on the overall treatment intensity. The dependent variable is the inverse hyperbolic sine transformation of treatment intensity of all services. In all models, we control for year fixed effects. The sample size is slightly smaller in column 2 because some observations are absorbed by plan-year fixed effects (while they are not absorbed by plan fixed effects alone). Standard errors are clustered at the insurer/employer level. $^{*}: p < 0.1,^{**}: p < 0.05,^{***}: p < 0.01$.

Treatment Intensity by Types of Services The impacts of capitation on treatment intensity is not equally spread across types of services. Figure 2 presents the impacts of capitation

Table 5: Cross-sectional and Cross-time Provider Variation on Overall Treatment Intensity

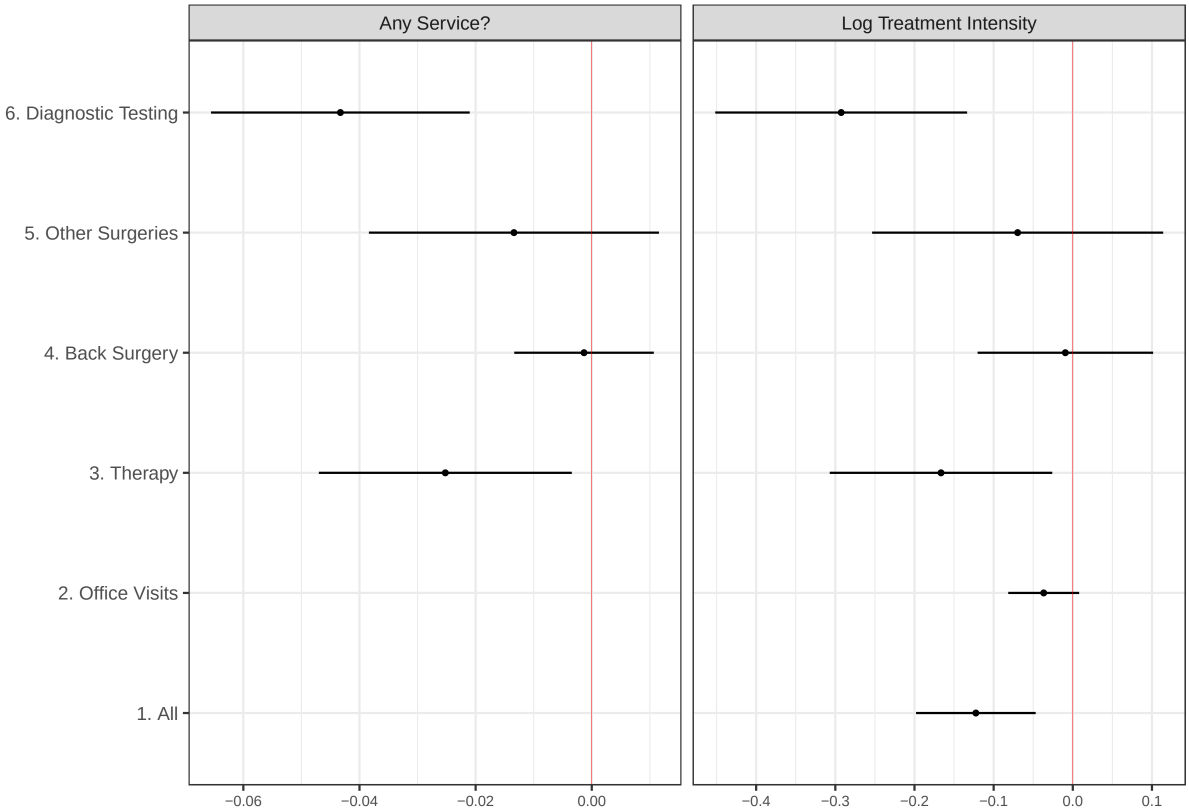
	1	2	3	4
Capitated	-0.072	-0.049		
	(0.054)	(0.049)		
Average Capitation Rate of Prov-Year			-0.101	
			(0.076)	
Any Capitated within Prov-Year				-0.105**
				(0.049)
Number of Observations	61,369	47,133	61,369	61,369
R-squared	0.323	0.368	0.323	0.323
Prov Fixed Effects	×		×	×
Prov × Year Fixed Effects		×		

Note: The table examines the cross-sectional and cross-time variation of provider's capitation status change on the overall treatment intensity. The dependent variable is the inverse hyperbolic sine transformation of treatment intensity of all services. In all models, we control for year fixed effects. The sample size is slightly smaller in column 2 because some observations are absorbed by provider-year fixed effects (while they are not absorbed by provider fixed effects alone). Standard errors are clustered at the insurer/employer level. $^{*}: p < 0.1,^{**}: p < 0.05,^{***}: p < 0.01$.

status on using five major treatment categories: office visits, therapy, back-related surgery, other surgeries, and diagnostic testing. The dependent variable is either a dummy variable indicating any treatment (left panel) or the treatment intensity under log-scale (right panel). We estimate the impacts of capitation on treatment intensity using the same specification and sample as in the baseline analysis. Note that in the left panel, there are no estimates for all service or office visits because all episodes in our sample are started with an office visit to the primary care provider.

Patients under a capitated plan are significantly less likely to have any therapy and diagnosis tests and also have much lower treatment intensity for these services. We estimate that capitated patients are 2.5% less likely to use any therapy and have 15% less therapy treatment overall. The patients under capitated plans are 4% less likely to have diagnostic tests and receive almost 30% less treatment in the diagnostic testing services. In contrast, we do not observe that capitation status significantly influences back surgery or other surgeries. These results may be driven by the fact that surgeries and invasive procedures are used mainly for patients with severe conditions and cannot be easily removed. For more selective services like therapy or diagnostic testing, capitation plays an important role in reducing usage.

Figure 2: Treatment Differences Among Types of Medical Services: Extensive and Intensive Margin



Note: The figure shows the estimated differences in using certain types of services between capitated patients and noncapitated patients. The dependent variable in Panel "Any Service?" is a dummy variable indicating the use of any treatment, while the dependent variable of Panel "Log Treatment Intensity" is the treatment intensity measure under inverse sine hyperbolic transformation (in log scale). The left panel "Any Service?" does not include results for using any service and whether the episode has office visits because all observations in our sample start with a primary care office visit. The y-axis is the coefficient and 95% confidence interval of capitated. All models control for the provider fixed effects, plan fixed effects, and individual characteristics. Standard errors are clustered at the insurer/employer level.

4.2 Drug Utilization

For a subset of episodes, we observe the drug claims and examine capitation's impacts on drug use. Table 6 shows the results with drug utilization. In all models, we control for plan fixed effects, individual characteristics, and year fixed effects. Controlling also for provider fixed effects results in too few observations to estimate the impacts accurately, so we only control for plan fixed effects. We consider the use of any drug, opioids, and muscle relaxants. For each type of drug, we examine both the extensive margin on whether any drugs are used and the intensive margin on the usage intensity.

Table 6 Column 1-3 present the effect of capitation on whether an episode includes any LBP-related drug claims. We find that the patients under capitated plans are 2.2% less likely to use any drugs, 2.3% less likely to use opioids, and 2.6% less likely to use muscle relaxants. Column 4-6 show the impact of capitation on drug use intensity. We find that capitation results in a 14.5% reduction in the use of all drugs. Capitation also reduces the use of opioids and muscle relaxants by 3.5% and 9%, respectively.

Table 6: Regression Results Comparing Drug Usage of All Services for Patients under Capitated and Noncapitated Arrangements

	Any Usage			Treatment Intensity		
	Drug Total	Opioids	Muscle Relaxants	Drug Total	Opioids	Muscle Relaxants
Capitated	-0.022**	-0.023***	-0.026***	-0.145**	-0.035*	-0.090***
	(0.010)	(0.005)	(0.010)	(0.066)	(0.018)	(0.023)
Number of Observations	44,846	44,846	44,846	44,846	44,846	44,846
R-squared	0.130	0.086	0.071	0.149	0.101	0.067
Plan Fixed Effects	×	×	×	×	×	×
Individual Characteristics	×	×	×	×	×	×

Note: The table shows the regression results comparing the drug usage intensity of capitated/non-capitated patients. The dependent variables in each column are: whether the patient used any drug, any opioids or any muscle relaxants during the episode, or the inverse hyperbolic sine transformation of treatment intensity of total drugs, opioids or muscle relaxants. Standard errors are clustered at the insurer/employer level. $^{*}: p < 0.1, ^{**}: p < 0.05, ^{***}: p < 0.01$.

4.3 Placebo Test: Emergency Room Visits

The previous results show that capitation leads to reductions in the intensity of treatment for lower back pain. There are concerns that the estimated impacts of capitation may be confounded by other demand or supply-side factors. To address these concerns, we use the emergency room (ER) visits from patients in our baseline sample as a placebo test. Unlike LBP episodes, ER visits are typically initiated by patients with urgent conditions needing immediate care. The decision to use ER services, especially ER visits for severe medical conditions, should not be affected by providers' financial arrangements.

For the same patients in our baseline analysis, we construct several measures on the ER services utilization. First, we construct a dummy variable indicating whether a patient has at least one ER visit during the episode. The procedure codes related to ER visits also document the severity of the illness and the urgency for care (in five levels). We thus construct a measure for any ER visits or any ER visits with the most severe conditions (Level 5). To reflect the overall utilization intensity, we also calculate the number of days with ER visits for all ER visits and ER visits with the most severe conditions, respectively.

As shown in Table 7, the capitation status of a patient's primary care physician has almost no impact on the patient's utilization of ER services. Patients under a capitated plan or a fee-for-service plan have almost the same likelihood of having any ER visits or severe ER visits and have a very similar number of ER visits. If there is any difference, the patients under capitated plans have slightly more ER days (not statistically different from zero).

Table 7: Placebo Test: Emergency Room Visits

	Any ER Visit		# of ER Visits	
	All ER	Severe ER	All ER	Severe ER
Capitated	-0.007	0.006	0.015	0.024
	(0.009)	(0.004)	(0.013)	(0.034)
Number of Observations	60,205	60,205	60,205	60,205
R-squared	0.263	0.245	0.212	0.253
Plan Fixed Effects	×	×	×	×
Prov Fixed Effects	×	×	×	×
Individual Characteristics	×	×	×	×

Note: The table shows the regression results comparing the emergency room (ER) use of capitated/non-capitated patients. The dependent variable is either whether the patient has any ER visits, any ER visits associated with severe conditions, or the number of these visits. All models use the baseline analysis sample with plan fixed effects, provider fixed effects and control for individual characteristics. Standard errors are clustered at the insurer/employer level. $^{*}: p < 0.1, ^{**}: p < 0.05, ^{***}: p < 0.01$.

4.4 Readmission Rates

A natural question is whether the estimated reduction in treatment intensity for patients under capitated plans represents a reduction in over-treatment or indicates under-treatment of valuable services. The question is important for understanding the nature of capitation incentives and the welfare implication of the capitated payment model. Though we do not entirely know the nature of the reduced services from the claim data, we can indirectly examine the quality of care using patient outcome measures. To do so, we calculate readmission rates for lower back pain in subsequent years to measure the treatment outcome. If the reduction in treatment for capitated patients leads to readmission for lower back pain in subsequent years, then the reduction in services may reflect the under-use of valuable care. On the other hand, if we find that patients in a capitated plan are similar in having subsequent LBP claims, then the estimated reduction in services is likely over-treatment.

We construct readmission measures by tracking patients in our sample over time. We can track about 65% of the baseline sample over the next four years. We then examine whether these patients have any LBP-related diagnosis in the four years after the end of their initial LBP episode. We then run a regression of this readmission measure on their original capitation status, controlling for individual characteristics, plan fixed effects, provider fixed effects, and year fixed effects.

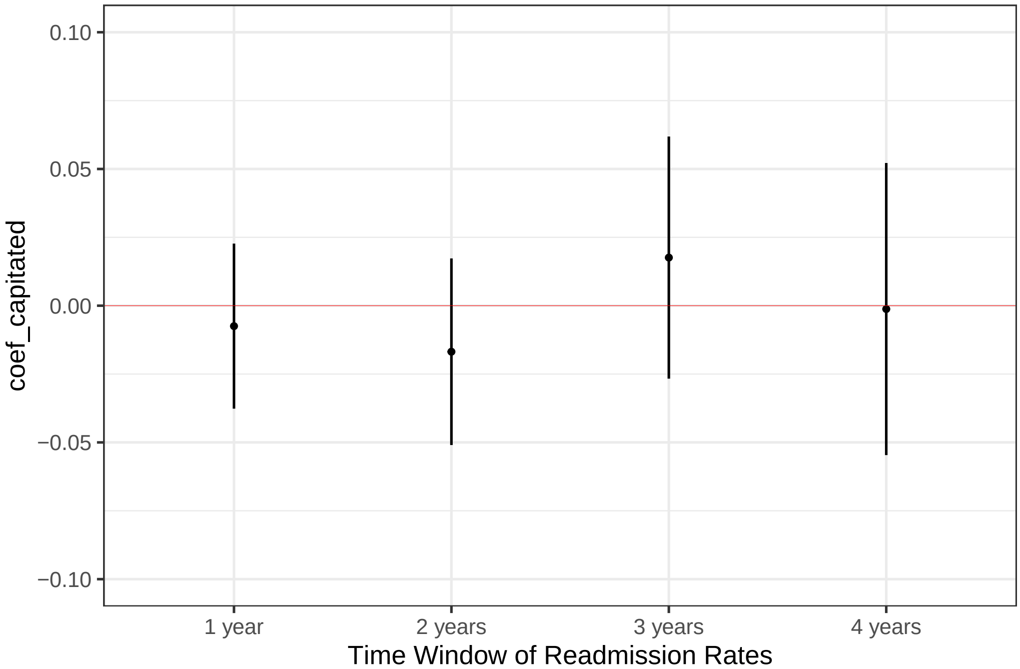
Figure 3 shows the results of our analysis. The x-axis indicates the time since the end of the initial episode, while the y-axis shows the point estimate and 95% confidence interval of the capitation coefficient. On average, patients under a capitated plan are slightly less likely to incur another LBP-related claim within 1, 2, or 4 years and slightly more likely to incur an LBP-related claim in 3 years. All estimates are very close to zero and are statistically insignificant. The results suggest that the reduction in treatment does not lead to higher chances of having the same condition in the future. In treating lower back pain, it seems the capitation payment model only reduces unnecessary care and does not adversely affect patient outcomes. These results are consistent with our previous findings that capitation mainly reduces the use of selective services like diagnostic testing and therapy.

5 Conclusion

This paper explores the effect of capitated payment models on lower back pain treatment in employer-sponsored health insurance plans. We find that patients under capitated plans receive significantly less treatment. The overall treatment intensity is 12.2% lower. Capitation contracts reduce the utilization of therapy, diagnostic imaging, and drugs such as muscle relaxants and opioids but have almost no impact on surgeries. Our identification relies on the panel feature of the claim data. Although patients rarely have multiple LBP episodes, we identify a group of people who enrolled in the same plan over time and control for the plan group fixed effects. We also control for physician fixed effects to further control for selections.

We study patients with lower back pain because medical literature indicates that many of the

Figure 3: Readmission Rates



Notes: The figure shows the differences in readmission rates of lower back pain for capitated and noncapitated patients. The dependent variables are dummy variables indicating whether any lower back pain related claims occur within a certain period after the episode ends. The x-axis is the time since the last day of the episode. The y-axis is the coefficient and 95% confidence interval of capitated. All models control for provider fixed effects, plan fixed effects, and individual characteristics. Standard errors are clustered at the insurer/employer level.

services used to treat this condition have low value. We find that capitation leads to differential treatment for otherwise similar patients. The question then is whether the inefficiency comes from under-treatment of patients under a capitated arrangement or over-treatment of patients in a non-capitated plan. We further our analysis by showing similar readmission rates of LBP conditions in subsequent years for capitated patients and fee-for-service patients. We provide suggestive evidence that in treating lower back pain, the capitation payment model effectively reduces treatment intensity without the cost of patient outcome. Of course, readmission rates are only one of the measures on patient outcome. More detailed data are needed to evaluate whether capitation encourages more efficient use of medical services for lower back pain and other conditions.

In our data, we do not observe the details of the capitation arrangement. Specifically, we do not observe how the incentives are shared among different physicians in a group and whether they face dynamic incentives over time. Our sample period is characterized by the declining popularity of capitation payment contracts. Physicians might not respond to the incentive if they did not stay in this type of contract for the following period, especially if some of the contracts offer dynamic incentives. The current movement toward value-based care may suffer from the same concern. More research is needed to understand the incentives of different capitation contracts and how they affect patients' long-term health status.

Using only within-provider-year variation to estimate the effects of capitation on treatment, we find almost zero treatment effects. This result suggests that physicians do not differentiate care for patients with different insurance plans in the same period. The treatment effects solely come from variations in the average capitation rates a physician faces over time. These findings are consistent with other empirical works providing evidence that treatment is homogeneous within a physician practice in the same year, even though the patients are from both fee-for-service plans and managed care plans (Glied and Zivin, 2002). Most physician groups in the United States have contracts with multiple insurers and face variation in their compensation incentives within the practice. However, the pattern is changing due to the recent trend toward fully integrated systems. For example, providers in some vertically integrated health care systems, such as Kaiser, almost exclusively treat managed care patients from their system. A natural next step is to study whether and to what extent fully integrated systems change physician behaviors relative to a simple capitation model.

References

Alalouf, Mattan, Sarah Miller, and Laura R Wherry. 2019. "What Difference Does a Diagnosis Make? Evidence from Marginal Patients." National Bureau of Economic Research Working Paper 26363. Series: Working Paper Series.

Alexander, Diane, and Molly Schnell. 2019. "The Impacts of Physician Payments on Patient Access, Use, and Health." National Bureau of Economic Research Working Paper 26095. Series: Working Paper Series.

Altman, Daniel, David Cutler, and Richard Zeckhauser. 2003. "Enrollee mix, treatment intensity, and cost in competing indemnity and HMO plans." *Journal of Health Economics*, 22(1): 23–45.

Andoh-Adjei, Francis-Xavier, Bronke Boudewijns, Eric Nsiah-Boateng, Felix Ankomah Asante, Koos van der Velden, and Ernst Spaan. 2018. "Effects of capitation payment on utilization and claims expenditure under National Health Insurance Scheme: a cross-sectional study of three regions in Ghana." *Health Economics Review*, 8(1): 17.

Card, David, Alessandra Fenizia, and David Silver. 2018. "The Health Effects of Cesarean Delivery for Low-Risk First Births." National Bureau of Economic Research w24493.

Carey, Colleen, Ethan M.J. Lieber, and Sarah Miller. 2020. "Drug Firms' Payments and Physicians' Prescribing Behavior in Medicare Part D." National Bureau of Economic Research Working Paper 26751. Series: Working Paper Series.

Carroll, Caitlin, Michael Chernew, A. Mark Fendrick, Joe Thompson, and Sherri Rose. 2018. "Effects of episode-based payment on health care spending and utilization: Evidence from perinatal care in Arkansas." *Journal of Health Economics*, 61: 47–62.

Chandra, Amitabh, and Douglas O Staiger. 2017. "Identifying Sources of Inefficiency in Health Care." National Bureau of Economic Research Working Paper 24035. Series: Working Paper Series.

Cherkin, Daniel C, Richard A Deyo, Ernest Volinn, and John D Loeser. 1992*a*. "Use of the International Classification of Diseases (ICD-9-CM) to identify hospitalizations for mechanical low back problems in administrative databases." *Spine*, 17(7): 817–825.

Cherkin, Daniel C., Richard A. Deyo, Ernest Volinn, and John D. Loeser. 1992*b*. "Use of the International Classification of Diseases (ICD-9-CM) to Identify Hospitalizations for Mechanical Low Back Problems in Administrative Databases:." *Spine*, 17(7): 817–825.

Chou, Roger, Amir Qaseem, Douglas K Owens, and Paul Shekelle. 2011. "Diagnostic imaging for low back pain: advice for high-value health care from the American College of Physicians." *Annals of Internal Medicine*, 154(3): 181–189.

Chou, Roger, Amir Qaseem, Vincenza Snow, Donald Casey, J Thomas Cross, Paul Shekelle, and Douglas K Owens. 2007. "Diagnosis and treatment of low back pain: a joint clinical practice guideline from the American College of Physicians and the American Pain Society." *Annals of Internal Medicine*, 147(7): 478–491.

Chou, Roger, Rongwei Fu, John A Carrino, and Richard A Deyo. 2009. "Imaging strategies for low-back pain: systematic review and meta-analysis." *The Lancet*, 373(9662): 463–472.

Clemens, Jeffrey, and Joshua D. Gottlieb. 2014. "Do Physicians' Financial Incentives Affect Medical Treatment and Patient Health?" 104(4): 1320–1349.

Currie, Janet, and David Slusky. 2020. "Does the Marginal Hospitalization Save Lives? The Case of Respiratory Admissions for the Elderly." National Bureau of Economic Research Working Paper 26618. Series: Working Paper Series.

Deyo, Richard A, Michael Von Korff, and David Duhrkoop. 2015. "Opioids for low back pain." *BMJ*, 350.

Duggan, Mark. 2004. "Does contracting out increase the efficiency of government programs? Evidence from Medicaid HMOs." *Journal of Public Economics*, 88(12): 2549–2572.

Einav, Liran, Amy Finkelstein, and Neale Mahoney. 2018. "Provider Incentives and Healthcare Costs: Evidence From Long-Term Care Hospitals." *Econometrica*, 86(6): 2161–2219.

Einav, Liran, Amy Finkelstein, Yunan Ji, and Neale Mahoney. 2020. "Voluntary Regulation: Evidence from Medicare Payment Reform." National Bureau of Economic Research w27223, Cambridge, MA.

Eliason, Paul J, Benjamin Heebsh, Riley J League, Ryan C McDevitt, and James W Roberts. 2020. "The Effect of Bundled Payments on Provider Behavior and Patient Outcomes." 84. Working Paper.

Flynn, Timothy W., Britt Smith, and Roger Chou. 2011. "Appropriate Use of Diagnostic Imaging in Low Back Pain: A Reminder That Unnecessary Imaging May Do as Much Harm as Good." *Journal of Orthopaedic & Sports Physical Therapy*, 41(11): 838–846.

Friedberg, Mark W, Peggy G Chen, Chapin White, Olivia Jung, Laura Raaen, Samuel Hirshman, Emily Hoch, Clare Stevens, Paul B Ginsburg, Lawrence P Casalino, et al. 2015. "Effects of health care payment models on physician practice in the United States." *RAND Health Quarterly*, 5(1).

Friedly, Janna, Christopher Standaert, and Leighton Chan. 2010. "Epidemiology of spine care: the back pain dilemma." *Physical Medicine and Rehabilitation Clinics*, 21(4): 659–677.

Fritz, Julie M, Gerard P Brennan, and Stephen J Hunter. 2015. "Physical therapy or advanced imaging as first management strategy following a new consultation for low back pain in primary care: associations with future health care utilization and charges." *Health Services Research*, 50(6): 1927–1940.

Fritz, Julie M, John D Childs, Robert S Wainner, and Timothy W Flynn. 2012. "Primary care referral of patients with low back pain to physical therapy: impact on future health care utilization and costs." *Spine*, 37(25): 2114–2121.

Gaynor, Martin, James B. Rebitzer, and Lowell J. Taylor. 2004. "Physician Incentives in Health Maintenance Organizations." *Journal of Political Economy*, 112(4): 915–931.

Glied, Sherry, and Joshua Graff Zivin. 2002. "How do doctors behave when some (but not all) of their patients are in managed care?" *Journal of Health Economics*, 21(2): 337–353.

Goodney, Philip R, Nino Dzebisashvili, David C Goodman, and Kristen K Bronner. 2015. "Variation in the care of surgical conditions." *The Dartmouth Institute*.

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

Howard, David H., Guy David, and Jason Hockenberry. 2017. "Selective Hearing: Physician-Ownership and Physicians' Response to New Evidence." *Journal of Economics & Management Strategy*, 26(1): 152–168. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/jems.12178.

Jacobson, Mireille, Tom Y Chang, Joseph P Newhouse, and M.D. Earle, Craig C. 2013. "Physician Agency and Competition: Evidence from a Major Change to Medicare Chemotherapy Reimbursement Policy." National Bureau of Economic Research Working Paper 19247. Series: Working Paper Series.

Koes, Bart W, Maurits Van Tulder, Chung-Wei Christine Lin, Luciana G Macedo, James McAuley, and Chris Maher. 2010. "An updated overview of clinical guidelines for the management of non-specific low back pain in primary care." *European Spine Journal*, 19(12): 2075–2094.

Koes, Bart W, Maurits W van Tulder, Raymond Ostelo, A Kim Burton, and Gordon Waddell. 2001. "Clinical guidelines for the management of low back pain in primary care: an international comparison." *Spine*, 26(22): 2504–2513.

Kontopantelis, Evangelos, David A Springate, Mark Ashworth, Roger T Webb, Iain E Buchan, and Tim Doran. 2015. "Investigating the relationship between quality of primary care and premature mortality in England: a spatial whole-population study." *Bmj*, 350.

Luckhaupt, Sara E, James M Dahlhamer, Gabriella T Gonzales, Ming-Lun Lu, Matthew Groenewold, Marie Haring Sweeney, and Brian W Ward. 2019. "Prevalence, Recognition of Work-Relatedness, and Effect on Work of Low Back Pain Among US Workers." *Annals of Internal Medicine*, 171(4): 301–304.

Maclean, Johanna Catherine, Chandler McClellan, Michael F Pesko, and Daniel Polsky. 2018. "Reimbursement Rates for Primary Care Services: Evidence of Spillover Effects to Behavioral Health." National Bureau of Economic Research Working Paper 24805. Series: Working Paper Series.

Sacks, Daniel W. 2018. "Why do HMOs spend less? Patient selection, physician price sensitivity, and prices." *Journal of Public Economics*, 168: 146–161.

Schwartz, Aaron L, Bruce E Landon, Adam G Elshaug, Michael E Chernew, and J Michael McWilliams. 2014. "Measuring low-value care in Medicare." *JAMA Internal Medicine*, 174(7): 1067–1076.

Shrank, William H., Teresa L. Rogstad, and Natasha Parekh. 2019. "Waste in the US Health Care System: Estimated Costs and Potential for Savings." *JAMA*.

Smith, R. 2011. "Dartmouth Atlas of Health Care." *BMJ*, 342(mar23 4): d1756–d1756.

Zhang, Xue, and Arthur Sweetman. 2018. "Blended capitation and incentives: Fee codes inside and outside the capitated basket." *Journal of Health Economics*, 60: 16–29.

Zuvekas, Samuel H, and Joel W Cohen. 2010. "Paying physicians by capitation: is the past now prologue?" *Health Affairs*, 29(9): 1661–1666.

Appendix

Table A.1: Robustness: Alternative Episode Window Definition

	90-day	180-day	270-day
Capitated	-0.079***	-0.122***	-0.119***
	(0.030)	(0.039)	(0.040)
Number of Observations	67,162	60,205	56,964
R-squared	0.332	0.350	0.358
Prov Fixed Effects	×	×	×
Plan Fixed Effects	×	×	×
Individual Characteristics	×	×	×

Note: The table shows the regression results comparing the treatment intensity of capitated/non-capitated patients using different definition of episode window. Each observation is an episode. For a patient, an LBP episode starts from his/her earliest LBP encounter, followed by subsequent encounters with a time gap shorter than 90, 180 or 270 days (column 1 - 3 respectively). An episode ends if there is no additional LBP encounters within 90/180/270 days of the last record. Two consecutive LBP encounters with larger than 90/180/270-day gaps are designated to two separate episodes. The dependent variable is the inverse hyperbolic sine transformation of treatment intensity of all services. All regressions control for provider fixed effects, plan fixed effects and individual characteristics. Standard errors are clustered at the insurer/employer level.

Table A.2: Diagnoses for Lower Back Pain (Cherkin et al., 1992*a*)

ICD-9 Code(s)	Diagnosis
721.3	Lumbosacral spondylosis without myelopathy
721.42	Spondylogenic compression of lumbar spinal cord
721.9	Spondylosis of unspecified site without myelopathy
721.91	Spondylogenic compression of spinal cord, not specified
722.1	Displacement of thoracic or lumbar disc without myelopathy
722.1	Displacement of lumbar disc without myelopathy
722.2	Displacement of unspecified disc without myelopathy
722.52	Degeneration of lumbar or lumbosacral disc
722.6	Degeneration of disc, site unspecified
722.7	Disc disorder with myelopathy, site unspecified
722.73	Lumbar disc disorder with myelopathy
722.8	Postlaminectomy syndrome, unspecified region
722.83	Postlaminectomy syndrome, lumbar
722.9	Other and unspecified disc disorder, site unspecified
722.93	Other and unspecified lumbar disc disorder
724	Spinal stenosis, unspecified site (not cervical)
724.02	Lumbar stenosis
724.09	Spinal stenosis, other
724.2	Lumbago
724.3	Sciatica
724.4	Thoracic or lumbosacral neuritis or radiculitis, unspecified
724.5	Backache, unspecified
724.6	Disorders of sacrum (including lumbosacral joint instability)
724.8	Other symptoms referable to back
724.9	Other unspecified back disorders
738.4	Acquired spondylolisthesis
739.3	Nonallopathic lesions, lumbar region
739.4	Nonallopathic lesions, sacral region
756.11	Spondylolysis, lumbosacral region
756.12	Spondylolisthesis
847.2	Sprains and strains, lumbar
847.3	Sprains and strains, sacral
847.9	Sprains and strains, unspecified region
307.89	Psychogenic backache
721.5-8	Unique or unusual forms of spondylosis
722.30	Schmorl's nodes, unspecified region
722.32	Lumbar Schmorl's nodes
737.10-737.30	Idiopathic scoliosis
738.5	Other acquired deformity of back or spine
756.10	Anomaly of spine, unspecified
756.13-756.19	Various congenital anomalies
805.4	Lumbar fracture
805.6	Sacral or coccygeal fracture
805.8	Vertebral fracture of unspecified site
846.0-9	Sprains and strains, sacroiliac
996.4	Mechanical complication of internal orthopedic device, implant and graft

Note: This table exhibites the ICD-9 diagnosis codes related to LBP (Cherkin et al., 1992*a*). refers to diagnoses applicable only to nonsurgical cases.