Do Payment Contracts in Healthcare Influence Medical Treatment Decisions?*

Cici McNamara
Georgia Institute of Technology

Natalia Serna
Stanford University

January 21, 2025

Abstract

We estimate the plausibly causal effect of fee-for-service (FFS) contracts relative to less retrospective contracts between health insurers and healthcare providers on treatment decisions for deliveries. We find that covering c-sections under FFS increases the c-section rate by 27% while covering vaginal deliveries under FFS has no impact on procedure choice. Effects are explained mainly by payment retrospectiveness rather than payment amount. Impacts of FFS on treatment decisions do not vary by pregnancy risk, but also have no impact on maternal or infant health outcomes. Results constitute evidence of substantial provider moral hazard generated by retrospective contracts.

Keywords: Moral hazard, Cesarean section, Fee-for-service, Capitation.

JEL codes: I11, I13, I18.

---

*McNamara: cmcnamara30@gatech.edu. Serna: nserna@stanford.edu.

We are deeply grateful to the Colombian Ministry of Health for providing the data. We thank Liran Einav, Grant Miller, Corina Mommaerts, Ciaran Phibbs, Maya Rossin-Slater, Heather Royer, Dan Sacks, David Silver, Alan Sorensen, and Ashley Swanson for their useful comments. We also thank participants to the 2024 Health Economics and Policy Innovation Collaborative mini-conference and the 2024 ASHEcon conference, and seminar attendees at UCSB. Our findings do not represent the views of any involved institution. All errors are our own.

1 Introduction

A fundamental question in health economics is whether contracts between insurers and providers affect medical treatment decisions. The most common contracts across different health systems are fee-for-service (FFS) and capitation, which lie on opposite ends of the spectrum of payment retrospectiveness. Other contracts such as bundled payments have also emerged recently, which are less than perfectly prospective or retrospective (Einav et al., 2022; Agarwal et al., 2020; Press et al., 2016). Most of the literature to date documents a correlation between FFS contracts and health care costs (e.g., Baker, 1999, 1997) and establishes incentives that insurers and providers face under each contract (e.g., Acquatella, 2021; Ho and Pakes, 2014a; Ellis and McGuire, 1986). However, evidence on the causal impacts of FFS contracts on treatment decisions is scarce because of a lack of comprehensive data on insurer-provider contracts and limited sources of exogenous variation in these bilateral contracting decisions.

In this paper, we estimate the aforementioned causal effect in the setting of the Colombian health care system, where multiple insurers negotiate with multiple providers over contracts for every health service. We focus on the impact of FFS relative to less retrospective contracts like capitation and bundled payments (CAP) on the decision to provide a cesarean section (c-section). In our setting, insurers and providers may cover c-sections under a different contract than vaginal deliveries, resulting in contract variation within insurer-provider pair. Contracts for each service also vary across types of women –

defined by age group and pregnancy risk – as well as over time. Using data on deliveries performed between 2013 and 2015, we show that variation over time in contracting decisions is unrelated to a wide range of insurer, provider, and woman characteristics.

We take advantage of this quasi-random variation in contracts over time to estimate the plausibly causal effect of FFS contracts relative to CAP on c-section rates and other outcomes of interest.[1] We regress different outcomes on an indicator for whether c-sections or whether vaginal deliveries are covered under FFS. Our findings show that the c-section rate increases 27% when c-sections are covered under FFS but contracts for vaginal delivery have no effect on the c-section rate. These results are consistent with the predictions of a model in which insurer-provider pairs may cover different procedures under different contract types and where providers choose procedures to maximize profits. In this model, the impact of FFS is asymmetric across c-sections and vaginal deliveries due to differences in their prices and marginal costs.

We do not find that the impact of FFS varies according to pregnancy risk, which is particularly worrisome for high-risk pregnancies for whom one procedure is likely to be medically recommended. We do find that responsiveness to contract type is stronger among small providers, where incentives between physicians and managers may be better aligned.

Our findings also show that providers are responsive to contract type even when controlling for the payment amount from insurers. This demonstrates that, conditional on payment amount, providers respond to the incentives gen-

[1] The c-section rate is the fraction of deliveries performed by c-section within an insurer, provider, type of woman, and year.

erated by whether payments are made prospectively or retrospectively. Furthermore, there is no evidence that CAP contracts generate risk selection incentives on the intensive margin; that is, in the style of Brown et al. (2014), there is zero correlation between ex-ante health care spending and contracting decisions. Hence, the main mechanism at play is moral hazard. We additionally estimate whether rates of severe maternal morbidity after birth and infants' birth weights and 1-minute APGAR scores differ depending on the delivery procedure contract.[2] We find no evidence that the use of FFS affects these measures of maternal and infant health.

Our paper contributes to a long line of literature dating back to the 1980s studying provider moral hazard under different payment schemes (e.g., Ellis and McGuire, 1996, 1986). Early and recent papers have analyzed provider incentives by payment amount and degree of retrospectiveness through a theoretical lens (e.g., Acquatella, 2021; Ma and Mak, 2019, 2015), but so far empirical evidence estimating causal impacts of the use of retrospective or prospective payments has been scant. Some exceptions have modeled optimal contracts in the context of dialysis medications (Gaynor et al., 2023), utilized system-wide changes in payments such as transitions from FFS to managed care systems (Kuziemko et al., 2018; Aizer et al., 2007), leveraged changes in payments within FFS systems (Clemens and Gottlieb, 2014) or within capitation systems (Ho and Pakes, 2014b,a), and exploited changes in how the government reimburses hospitals under bundled payments (Einav et al., 2022). We complement this work by estimating the plausibly causal effect of the use

[2] The APGAR score is a measure of infant health during the birthing process. This score is on a scale from 1 to 10, with a higher score reflecting better infant health.

of FFS exploiting variation in contracting decisions within insurer-provider pairs in a managed care system.

We also contribute to the literature on determinants of c-section rates. With c-sections becoming more prevalent in the last two decades across the globe (Betrán et al., 2016), contributing to rising health care costs (Sakala et al., 2013), and being the top reason for hospitalization among women (AHRQ, 2018), understanding how delivery decisions are made has become central to the health policy debate. Some work has studied place-based drivers of c-section rates (Robinson et al., 2024; Card et al., 2023) and the role of physician moral hazard (Foo et al., 2017; Godager and Wiesen, 2013; Johnson and Rehavi, 2016). We complement this work by showing that contracts between insurers and providers have significant effects on c-section rates. Moreover, we show that incentives depend on the contract under which each delivery procedure is covered, and that regulation of payments for a single delivery service may generate very different effects depending on how the alternative service is reimbursed.

The remainder of this paper is structured as follows: Section 2 describes our institutional setting and raw data, Section 3 presents our empirical design, Section 4 shows our main results, and Section 5 concludes.

2 Setting and Data

We study the contracts established between insurers and providers participating in the contributory health care system in Colombia, which covers the half of the population in the country who pay payroll taxes. In this system, private insurers provide one national health insurance plan to their beneficiaries through a network of providers. Other elements of the insurance plan such as premiums, cost-sharing, and service coverage are strictly regulated. Providers in our empirical setting include hospitals, clinics, and physician practices, but not stand-alone doctors. Providers operate in a single market or municipality.

Insurers negotiate with providers over contracts for each individual health service covered by the national insurance plan. This means that different services such as c-sections and vaginal deliveries may be covered under different contracts even for the same insurer-provider pair. Insurers and providers can choose from a set of three contract types: FFS, bundled payments, and capitation.[3] We detail how payments are made under each of these contract types in Section 2.2.

To describe these contracts, we use health claims data for all pregnant women in the contributory system who had a childbirth between 2013 and 2015. These data report the type of service, service date, provider, insurer, negotiated service price, and contract type for each claim. There are 521,408 deliveries in the raw data. We categorize women as having a high-risk pregnancy if they received a diagnosis indicating that their pregnancy was high risk any time in the 9 months of pregnancy.[4] We also categorize women as having

[3]Bundled payments involve payments from the insurer to the provider for an episode of care or for a particular diagnosis.

[4]These ICD-10 codes include: supervision of high-risk pregnancies (O09), infections of genitourinary tract in pregnancy (O23), pregnancy-induced hypertension or pre-eclampsia (O10-O16), hemorrhage due to threatened abortion (O20), excessive vomiting during pregnancy (O21), venous complications and hemorrhoids in pregnancy (O22), diabetes mellitus in pregnancy, childbirth, and the puerperium (O24), malnutrition in pregnancy, childbirth and the puerperium (O25), abnormal findings on antenatal screening of mother (O28), and

severe maternal morbidity (SMM) based on the services and diagnosis codes rendered between 1 week and 1 month after childbirth following the Centers for Disease Control and Prevention's definition.[5] We exclude women with multiple or breech pregnancies for whom c-sections are medically recommended and keep deliveries associated with a woman's first childbirth. Our final analysis sample contains 343,967 deliveries. Whenever a delivery is performed at a provider organization such as a hospital, a clinic, or a physician practice, we observe the organization's ID number, but we do not have information on the physician who performed the delivery procedure at this organization.

Table 1: Summary Statistics of Analysis Sample

	C-sections	Vaginal deliveries
Delivery characteristics, mean (s.d.)		
Fee-for-service (FFS)	0.754 (0.430)	0.671 (0.470)
Age$\leq$27	0.669 (0.471)	0.788 (0.408)
Age$>$ 27	0.331 (0.471)	0.212 (0.408)
High pregnancy risk	0.522 (0.500)	0.449 (0.497)
Severe maternal morbidity (SMM)	0.049 (0.216)	0.032 (0.176)
Birth weight (grams)	3040 (568.7)	3071 (414.4)
1-minute APGAR score	8.236 (0.887)	8.273 (0.859)
Spending on delivery day†	545.3 (557.2)	442.2 (336.9)
Spending up to delivery day†	625.2 (654.7)	531.9 (490.3)
Number of observations, count (%)		
Deliveries	166,665 (48.5)	177,302 (51.5)
Insurer-provider	1986 (58.3)	2919 (85.8)
Provider	760 (52.6)	1258 (87.1)
Insurer	14 (100)	14 (100)

Note: Table shows the mean and standard deviation (in parenthesis) of the main variables in our analysis sample stratified by delivery procedure. An observation is a delivery. The 1-minute APGAR score reflects how well the infant tolerated the birthing process and is measured on a scale from 1 to 10, a higher score reflecting better infant health. †Measures of spending are reported in USD.

complications of anesthesia during pregnancy (O29)

[5]See https://www.cdc.gov/maternal-infant-health/php/severe-maternal-morbidity/icd.html

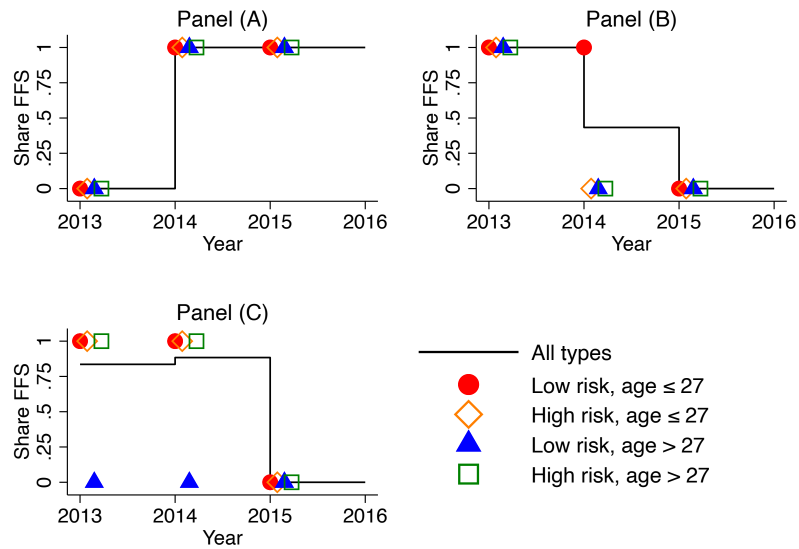
Table 1 presents summary statistics of our analysis sample. An observation is a delivery claim. The c-section rate equals 49%. Most deliveries are reimbursed under a FFS contract, although the use of FFS is more prevalent for c-sections than for vaginal deliveries. More than half of deliveries performed correspond to women of age 27 or younger. Around half of pregnancies are classified as high risk, but the rate of high risk pregnancies is greater among c-sections. Likewise, spending up to and on the day of the delivery is higher for c-sections, as are rates of SMM. Infant health outcomes including the 1-minute APGAR score and birth weight are marginally better for vaginal deliveries relative to c-sections.

2.1 How are Contracts Chosen?

Insurers and providers typically negotiate contracts at the beginning of every calendar year, but the end of these negotiations and the timing of when contracts are enacted can vary. Insurer-provider pairs negotiate contracts for each delivery procedure and may choose to index contract types to patient characteristics such as the woman's age and pregnancy risk. In the analysis to follow, we define a woman type or contracting unit for an insurer-provider pair in a given year as a combination of age-group ($\leq$27, $>$27) and pregnancy risk (high vs. low). We choose the age of 27 as the age group cutoff since this corresponds to the average age in the analysis sample. We assign to each insurer-provider-woman type-year its modal contract type; this modal contract equals the observed contract for 96% of deliveries in our data.[6]

[6] For the 4% of observations for which the modal contract does not equal the observed contract, we impose this modal contract. Our results are robust to excluding these obser-

Figure 1: Examples of Contract Variation within Insurer-Provider Pairs



Note: Each panel corresponds to a different insurer-provider pair. The solid black line shows the share of c-sections in each year that are reimbursed under FFS. The markers show the contract type for each woman type in each year, which will be either fully FFS or fully capitated in a given year.

The primary source of variation that we use in our empirical analyses is in the contract type used by an insurer-provider pair for a given woman type over time. Figure 1 illustrates this variation in the c-section contract for three of the 3,404 insurer-hospital pairs in our analysis sample. For each insurer-provider pair there is variation in contracts within a woman type. For example, Panel A shows an insurer-provider pair where no c-sections were reimbursed under a FFS contract in 2013, but by 2014 all c-sections were covered under FFS. Panel B shows an insurer-provider pair where, in 2013, 100% of c-sections were reimbursed under FFS. In 2014, only c-sections for low-risk women less than 28 years old were reimbursed under FFS, and by 2015, no c-sections were

vations as seen in section 4.4.

reimbursed under FFS. Across the three panels, there is a trend toward either a fully CAP system or a fully FFS system within insurer-provider pair, but the direction of the trend varies across pairs.

TABLE 2: Summary of Changes in Contracts over Time

Service	2013	2014	2015
C-section			
Aggregate	*0.79*	*0.76*	*0.70*
Market	0.73	0.74	0.69
	(0.13)	(0.13)	(0.16)
Insurer	0.63	0.60	0.53
	(0.07)	(0.11)	(0.27)
Provider	0.77	0.75	0.68
	(0.28)	(0.31)	(0.28)
Vaginal delivery			
Aggregate	*0.71*	*0.68*	*0.61*
Market	0.68	0.72	0.62
	(0.12)	(0.11)	(0.10)
Insurer	0.71	0.73	0.61
	(0.12)	(0.13)	(0.33)
Provider	0.67	0.67	0.58
	(0.34)	(0.33)	(0.28)

Note: Table shows trends in the use of FFS for c-sections in the top panel and for vaginal deliveries in the bottom panel. "Aggregate" reports the fraction of procedures that are covered under FFS every year. "Market" reports the average fraction of insurer-provider pairs in a municipality that cover the procedure under FFS weighted by the number of deliveries. "Insurer" reports the average fraction of providers within an insurer's network that are covered under FFS weighted by the number of deliveries. "Provider" reports the average fraction of insurers that use a FFS contract with a given provider weighted by the number of deliveries. Standard deviations are reported in parenthesis.

Table 2 shows that, in aggregate, use of FFS fell over the sample period by approximately 10 p.p. from a baseline of 79% for c-sections and 71% for vaginal deliveries. This pattern holds across insurer-provider pairs within a market (row labelled "Market"): we see reductions in the average fraction of pairs that use FFS in a market equal to 4 p.p for c-sections and 6 p.p. for vaginal deliveries. These reductions are mainly coming from some insurers switching most of the providers in their network to CAP. Note that for both procedures the table shows a decreasing trend in the fraction of providers within an insurer's network that are covered under FFS (row labelled "Insurer") as well as in the fraction of insurers that use a FFS contract with a given provider (row labelled "Provider"). However, there is greater heterogeneity in contracts across providers than across insurers. For instance, in 2014, the standard deviation of the fraction of insurers that cover a given provider under FFS was three times higher than the standard deviation of the fraction of providers under FFS in a given insurer's network (both reported in parenthesis).

The greater heterogeneity across providers than across insurers in the use of FFS suggests that insurers have higher bargaining power, potentially because they are on the short-side of the market.[7] In any given year, more than 50% of insurers contract with at least 50 providers, while more than 50% of providers contract with only one insurer. Endogenous selection of insurers into FFS based on unobserved changes in their relative bargaining power (as demonstrated in Serna (2024)) threatens the validity of our empirical design, since we exploit the temporal variation in contracts within insurer-provider-woman type.

For this temporal variation to be plausibly exogenous, it must be that changes in contract type are uncorrelated with time-varying factors within an insurer-provider-woman type that might also affect our outcomes of interest.

[7] In our conversations with agents who work for the contracting areas of insurers and providers we also learned that insurers typically make take-it-or-leave-it offers to providers for delivery services, corroborating our intuition on the relative bargaining power.

To show whether this is the case, we collapse the data to the insurer-provider-woman type-year level, calculating average per capita measures of health care spending and health outcomes. We estimate the following equation:

$$Switch^{s}_{jhmt} = \mathbf{x'}_{jhmt}\beta + \gamma_{jhm} + \delta_j t + \tau_t + \varepsilon_{jhmt} \tag{1}$$

where $Switch^{s}_{jhmt}$ is a binary indicator for whether insurer j and provider h switch their contract type for delivery service $s \in \{\text{c-section, vaginal delivery}\}$ for women of type m relative to year $t-1$. $\mathbf{x}_{jhmt}$ is a set of lagged average per capita delivery characteristics. We include insurer-provider-woman type fixed effects γ_{jhm} and year fixed effects τ_t, as well as insurer-specific linear time trends $\delta_j t$. These time trends account for the evolution of unobserved insurer characteristics such as their relative bargaining power, which likely determine the choice of contracts. Because providers practice only in one market, γ_{jhm} also implicitly controls for market fixed effects. Standard errors are clustered at the insurer-provider-woman type level. We estimate equation (1) separately for switches in the contract for c-sections and in the contract for vaginal deliveries. A lack of statistically significant associations between these lagged characteristics and contract switching is evidence in favor of contracts driving changes in delivery practices and against causality in the other direction.

The estimation results are provided in Table 3. For both c-sections and vaginal deliveries, we find zero correlation between contract switches and lagged average per capita delivery characteristics. We include in our specification one contemporaneous measure given by the share of deliveries attributable to each woman-type within an insurer-provider-year. Correlation between contract switches and these contemporaneous delivery shares could reflect changes in diagnostic practices, such as upcoding, that are meant to select women into particular contract types. We find no statistically significant effects for any measures across either specification. Appendix Table 1 shows similar results when we include insurer-by-year fixed effects rather than insurer-specific linear time trends. These results suggest that changes in contract type are quasi-random and are perhaps determined by idiosyncratic managerial preferences which have been shown in other settings to influence hospital performance (see e.g., Otero and Munoz, 2022).

2.2 What are the Incentives under Each Contract?

The three contracts available to insurer-provider pairs in our setting are FFS, bundled payments, and capitation. If a service is covered under FFS, then the insurer negotiates a price to be paid every time the service is provided to one of its beneficiaries only after it has been provided. Thus, FFS is fully retrospective. If a service is covered under capitation, then the insurer negotiates a price per person, a service frequency per person, and a target population, and pays the provider the product of these three elements before services are provided to their beneficiaries (typically at the beginning of the calendar year). Thus, capitation is fully prospective. Low-complexity services such as primary care consultations and blood tests tend to be capitated.

Under bundled payments, the insurer negotiates a price for a bundle of services used during an episode of care. This price is paid every time an episode of care occurs. Episodes of care typically covered under bundled payments

Table 3: Correlates of Contract Switching

	Contract switch	
	(1) C-section	(2) Vaginal delivery
Woman-type share of deliveries	-0.139	0.0410
	(0.0834)	(0.0850)
Lagged delivery characteristics		
SMM	-0.000341	0.0199
	(0.0774)	(0.126)
Birth weight (kilograms)	-0.0218	0.0140
	(0.0296)	(0.0395)
1-minute APGAR score	0.00298	0.0166
	(0.0161)	(0.0172)
Log spending on delivery day	-0.0286	0.0450
	(0.0375)	(0.0479)
Log spending up to delivery day	-0.0252	-0.0189
	(0.0189)	(0.0231)
Woman-type share of deliveries	-0.0472	-0.0636
	(0.0898)	(0.0943)
Constant	0.538	-0.156
	(0.360)	(0.416)
N	2106	1668
R^2	0.14	0.22

Note: Table shows the correlates of contract switches for c-sections and for vaginal deliveries in columns 1 and 2, respectively. An observation is an insurer-provider-woman type-year. A woman type is a combination of age group (age $\leq$ 27, age $>$ 27) and pregnancy risk (high vs. low). Lagged delivery characteristics correspond to the lagged average per capita measure. Specifications include insurer-provider-woman type and year fixed effects as well as insurer-specific time trends. Standard errors in parenthesis are clustered at the insurer-provider-woman type level. Measures of spending are averages per capita.

tend to be highly standardized such as hysterectomies or appendectomies. Bundled payments are less than fully retrospective because the insurer does not pay for additional services provided during the episode of care that are not included in the bundle, and because payments will often not vary with the type of service provided. For example, if childbirth is covered under a bundled payment contract, then the provider will be reimbursed the same amount regardless of whether the infant is delivered by vaginal delivery or c-section. Given that capitation and bundled payments function similarly with respect to retrospectiveness in this setting, we will for the remainder of this study distinguish only between FFS and non-FFS, where non-FFS refers to both capitation and bundled payments and will be referred to as simply CAP.

In the case of childbirth, and under fully retrospective contracts like FFS, the provider has an incentive to provide the service with the highest profit margin. Under prospective contracts like capitation and bundled payments, the provider has an incentive to provide the lowest cost treatment as payments do not vary with the services actually rendered. These incentives have been well documented in other settings where contracts may vary over time, but do not vary across services (e.g., Kuziemko et al., 2018; Adida et al., 2017; Ho and Pakes, 2014a). However, in our setting where contracts may differ across delivery procedures, incentives will depend on the combination of contracts.

To see these incentives more clearly, we propose a simple model of procedure choice. Let the price of a c-section be p_C and its marginal cost be m_C and let the corresponding values for vaginal deliveries be p_V and m_V. We assume that $m_C > m_V$. The unit of payment depends on which contract type the procedure is covered under. When a procedure is covered under CAP, we summarize the product of frequency of use and target population as the "expected rate of use," normalized to one. Thus, the expected rate of use is π_C for c-sections and $1 - \pi_C$ for vaginal deliveries when they are covered under CAP.

We assume providers make a procedure choice to maximize their expected profits per woman type. The profit from choosing procedure $s \in \{C, V\}$

conditional on the contracts is

$$E[\Pi_s|\mathbf{FFS},\mathbf{x}] = (E[R_s|\mathbf{FFS}] - m_s)D(\mathbf{x})$$

where

$$\begin{aligned} E[R_s|\mathbf{FFS}] =& (\mathrm{FFS}_C \mathbf{1}[s=C] + (1-\mathrm{FFS}_C)\pi_C)p_C \\ &+ (\mathrm{FFS}_V(1-\mathbf{1}[s=C]) + (1-\mathrm{FFS}_V)(1-\pi_C))p_V \end{aligned}$$

R_s is the provider's revenue per woman, FFS_s is an indicator for whether service s is covered under FFS, $\mathbf{1}[s=C]$ is an indicator for choosing c-sections, and $\mathbf{x}$ is a vector of observed insurer, provider, and woman characteristics. We assume provider demand $D(\mathbf{x})$ is independent of FFS_s because contracts are not directly observable to women conditional on $\mathbf{x}$.[8] Demand is also independent of procedure prices for simplicity. The provider's problem is

$$\max_{s\in\{C,V\}} \{E[\Pi_C|\mathbf{FFS}], E[\Pi_V|\mathbf{FFS}]\}$$

In this model, payment retrospectiveness affects only expected revenues but not costs, since costs are incurred only when a procedure is actually rendered. Payment retrospectiveness impacts expected revenues in two ways. First, the expected rate of use summarizes the ex-ante number of procedures that insurers and providers expect to render to their beneficiaries and thus

[8] Patients in Colombia do not observe the contracts between insurers and providers, but may select into providers based on characteristics that correlate with the use of FFS such as provider size or quality, which we capture in the model with $\mathbf{x}$. Thus, our assumption is that conditional on $\mathbf{x}$, provider demand does not vary with contract types.

applies only when procedures are covered under CAP. For example, if both procedures are covered under CAP then the expected revenue per woman is $\pi_C p_C + (1 - \pi_C)p_V$, but when both procedures are covered under FFS the expected revenue per woman is independent of π_C. Second, even if a procedure is covered under CAP but is not chosen, providers still get paid for it ex-ante. For example, if conditional on c-sections being covered under FFS and vaginal deliveries being covered under CAP, the provider chooses c-sections, then its expected revenue per woman equals $p_C + (1 - \pi_C)p_V$.

TABLE 4: Example of Contracting Scenarios and Incentives

Scenario (1)	Contract (2)	Service (3)	Expected rate of use (4)	Expected provider profit (5)	Choice (6)
1	CAP	C-section	π_C	$\pi_C p_C + (1 - \pi_C)p_V - m_C$	0
	CAP	Vaginal delivery	$1 - \pi_C$	$\pi_C p_C + (1 - \pi_C)p_V - m_V$	1
2	FFS	C-section	—	$p_C - m_C$	?
	FFS	Vaginal delivery	—	$p_V - m_V$	
3	FFS	C-section	—	$p_C + (1 - \pi_C)p_V - m_C$	1
	CAP	Vaginal delivery	$1 - \pi_C$	$(1 - \pi_C)p_V - m_V$	0
4	CAP	C-section	π_C	$\pi_C p_C - m_C$	0
	FFS	Vaginal delivery	—	$\pi_C p_C + p_V - m_V$	1

Note: Table shows examples of the profit maximizing choice of procedure for the provider under four possible contracting scenarios in which c-sections and vaginal deliveries are either covered under the same contract or under different contracts. The price and marginal cost of a c-section are p_C and m_C, respectively, the corresponding values for vaginal delivery are p_V and m_V. The expected rate of use π_C summarizes the product of frequency of use and target population in a CAP contract, normalized to 1. The procedure choice is the one that maximizes profits for the provider conditional on each contracting scenario.

Table 4 summarizes providers' expected profits per woman under the four possible contract type combinations. This table highlights that the choice of procedure will depend on prices, marginal costs, and payment retrospectiveness. For instance, in the case where both procedures are capitated in scenario 1, providers will choose to render vaginal deliveries because of their lower marginal cost. Covering vaginal delivery under FFS as in scenario 4 only heightens incentives to provide vaginal deliveries as now revenues increase in the number of vaginal deliveries provided. In scenario 2, we see that covering both c-section and vaginal delivery under FFS generates ambiguous predictions on the choice of procedure as the decision depends on whether the profit margin for c-section exceeds that for vaginal delivery. In scenario 3, where c-sections are covered under FFS and vaginal deliveries under capitation, the provider would choose to perform c-sections because of the degree of payment retrospectiveness. Note that in this scenario even if c-sections and vaginal deliveries had the same price and marginal cost, the profit maximizing choice of procedure would still be c-sections. Therefore, this choice of procedure can only be explained by the degree of payment retrospectiveness.

Table 4 also provides insight into the overall effect of using retrospective contracts like FFS and how those effects can be expected to differ for c-section and vaginal delivery. To estimate the effect of covering c-section under FFS we can:

- Compare scenario 3 to scenario 1, where the contract for vaginal deliveries is fixed to CAP. In the latter scenario, incentives are strongly in favor of providing vaginal delivery; therefore, the effect of covering c-section under FFS is very likely to increase the c-section rate.

- Compare scenario 2 to scenario 4, where the contract for vaginal deliveries is fixed to FFS. As before, incentives in scenario 4 favor vaginal delivery; hence, covering c-section under FFS likely increases the use of c-section in this exercise.

Similarly, to estimate the effect of covering vaginal delivery under FFS we can:

- Compare scenario 4 to scenario 1, where the contract for c-sections is fixed to CAP. Because in both scenarios incentives are aligned with the provision of vaginal delivery, we should expect to find no impact of FFS on the c-section rate.

- Compare scenario 2 to scenario 3, where the contract for c-section is fixed to FFS. Given that incentives under scenario 3 are aligned with the provision of c-sections, the effect of covering vaginal deliveries under FFS on the c-section rate is ambiguous.

In our empirical application, we take advantage of variation in contracts over time and across procedures to estimate whether insurers and providers are responsive to these asymmetric incentives. Failure to account for differential responses to FFS depending on the procedure can lead to underestimation of the desired treatment effect. Consider for example a comparison of scenarios 2 and 3 to scenarios 1 and 4. In the former group, c-sections are covered under FFS while in the latter they are covered under CAP. In scenarios 1 and 4 the provider would always choose vaginal delivery, thus covering c-section under FFS in scenarios 2 and 3 would likely lead to higher c-section rates. While the direction of the effect is consistent with our intuition, this naïve estimation confounds the impact of changes in the contract for vaginal deliveries. If vaginal deliveries are more likely to be provided in scenario 4, then the magnitude of the impact of the c-section contract on c-section rates will likely be attenuated.

We note that our model is a simplified representation of the world and may not capture several other ways in which profits may vary and in which agents may respond to contracts. On the one hand, providers may be altruistic, considering both profits and patient utility when making procedure choices. However, to the extent that the degree of altruism does not vary across contracting scenarios, it will not affect the predictions of our model. On the other hand, providers under CAP may have incentives to select low-risk patients who are less likely to need health care, while insurers may have incentives to steer high-risk patients towards providers covered under CAP. These risk selection incentives can make provider demand depend on the contracts, which would lead us to underestimate the impact of FFS on the use of c-sections. For example, if insurers send all the high-risk women who potentially need c-sections to providers where c-sections are covered under CAP, then the impact of FFS on the c-section rate will be biased toward zero.

There may be other unobserved determinants of delivery procedure choice such as provider convenience or capacity constraints, which we do not consider. Nonetheless, if provider moral hazard is the predominant way in which contracts affect delivery procedure choice and outcomes, we can expect to estimate results in line with the incentives outlined in Table 4.

3 Empirical Design

To study the effect of contracts on medical treatment decisions, we take advantage of the quasi-random variation in contracts within a procedure, insurer, provider, and type of woman over time. Note that because contracts are established at the procedure level, it may the case that c-sections are more likely to be covered under FFS than vaginal deliveries across all insurer-provider pairs. It may also be the case that women for whom c-sections are medically recommended tend to be covered under FFS. This suggests that a comparison of c-section rates across insurer-provider pairs and services would yield a mechanically higher c-section rate under FFS than CAP because of selection into contracts. Thus, the appropriate research design should instead exploit the variation in contracts within procedure, insurer, provider, and type of woman, which the Colombian setting uniquely provides.

We estimate models as follows:

$$y_{jhmt} = \alpha \mathrm{FFS}^{s}_{jhmt} + \gamma_{jhm} + \delta_j t + \tau_t + \varepsilon_{jhmt} \tag{2}$$

where y_{jhmt} is an outcome for insurer j, provider h, and type of woman m in year t, FFS^{s}_{jhmt} is an indicator for whether procedure $s \in \{\text{c-section, vaginal delivery}\}$ is covered under FFS, γ_{jhm} is an insurer-provider-woman type fixed effect, $\delta_j t$ is an insurer-specific linear time trend, and τ_t is a year fixed effect. The coefficient of interest is α, which captures the plausibly causal effect of FFS contracts relative to CAP. To estimate these models we collapse the claim-level data to the $jhmt$-level, calculating average per capita outcome measures. Appendix A details our data cleaning process. We cluster our standard errors at the level of insurer-provider-woman type, which defines a contract.

We include fixed effects at the level of insurer-provider-woman type to control for time-invariant factors that may be correlated with the use of FFS. For example, if insurer or provider quality are correlated with contract type, then the estimated effect of FFS would be biased. As with our model in equation (1), these fixed effects also capture market-level variation in outcomes, since providers practice only in one market. τ_t accounts for any aggregate trend or seasonality in contracting decisions. Lastly, the inclusion of an insurer-specific time trend accounts for changes in their relative bargaining power, quality, or enrollee composition over time that may be correlated with our outcomes of interest.

Appendix Figures 1 and 2 plot the residuals from a regression of contract types on insurer-provider-woman type fixed effects, year fixed effects, and insurer-specific time trends for c-section and vaginal delivery, respectively. Both figures show that substantial variation in contract type remains to estimate our effects of interest.

4 Results

Table 5 presents results of the impact of FFS on procedure choice comparing scenarios 1 and 3 from Table 4 in the case of c-sections and scenarios 1 and 4 in the case of vaginal deliveries.[9] The dependent variable in both specifications is the c-section rate. In column (1) we find that covering c-sections under FFS generates a 27% increase in the c-section rate. Column (2) shows that there is no statistically or economically significant effect of covering vaginal delivery under FFS. These findings are in line with our model's predictions: conditional on vaginal deliveries being covered under CAP, covering c-sections under FFS increases incentives to provide c-sections, while the same is not true for vaginal deliveries.

Table 5: Impact of FFS Use on C-section Rates for Different Contract Combinations

	C-section rate	
	(1)	(2)
C-section FFS	0.131	
	(0.0376)	
Vaginal delivery FFS		-0.0228
		(0.0457)
Constant	0.421	0.395
	(0.0296)	(0.0392)
Sample	Scenarios 1 & 3	Scenarios 1 & 4
% change in outcome	27.0	-4.7
N	469	463
R^2	0.83	0.80

Note: Table shows the impact of the use of FFS for c-sections and for vaginal deliveries on the c-section rate in columns 1 and 2, respectively. Definition of samples is provided in Table 4. An observation is an insurer-provider-woman type-year. A woman type is a combination of age group (age ≤ 27, age > 27) and pregnancy risk (high vs. low). Specifications include insurer-provider-woman type and year fixed effects as well as insurer-specific time trends. Standard errors in parenthesis are clustered at the insurer-provider-woman type level. Table reports the percentage change in the outcome due to covering the procedure under FFS, calculated as the coefficient on FFS divided by the average c-section rate in the sample.

[9] The exercises in Table 5 might not include all the observations for a given contracting unit. For example, if for a given jhm the contract type for both c-sections and vaginal deliveries is CAP in year t and is FFS for both procedures in year $t+1$, then the $t+1$ observation is not included in the estimation in column (1).

Appendix Table 2 provides additional estimates of the effect of contract types across the scenarios summarized in Table 4. For example, comparing scenarios 2 and 3, findings show that the c-section rate is invariant to the c-section contract, while conditional on scenarios 2 and 4 we find that covering vaginal deliveries under CAP raises the c-section rate. In general, our findings indicate both that service profitability determines treatment decisions over and beyond patients' underlying health status, and that whether and how contracts impact delivery procedure choice is not symmetric across services.

Table 6: Impact of FFS Use on C-section Rates

	C-section rate	
	(1)	(2)
C-section FFS	0.0316	
	(0.0124)	
Vaginal delivery FFS		-0.00002
		(0.0106)
Constant	0.519	0.375
	(0.0130)	(0.0119)
Sample	Full	Full
% change in outcome	6.5	-0.004
N	6354	6557
R^2	0.76	0.81

Note: Table shows the impact of the use of FFS for c-sections and for vaginal deliveries on the c-section rate in columns 1 and 2, respectively. An observation is an insurer-provider-woman type-year. A woman type is a combination of age group (age ≤ 27, age > 27) and pregnancy risk (high vs. low). Specifications include insurer-provider-woman type and year fixed effects as well as insurer-specific time trends. Standard errors in parenthesis are clustered at the insurer-provider-woman type level. Table reports the percentage change in the outcome due to covering the procedure under FFS, calculated as the coefficient on FFS divided by the average c-section rate in the sample.

Table 6 further corroborates the intuition developed from Table 4. In column (1), we test that across all scenarios, the effect of covering c-section under FFS will be biased toward zero because of the confounding effect of changes in the contract for vaginal deliveries. We find indeed that the effect of a FFS contract for c-section on the c-section rate is positive but equals only 6.5% compared to the 27% reported in Table 5. We find similar results in column (2) where the coefficient on the use of FFS for vaginal deliveries is negative but substantially smaller in magnitude compared to the corresponding estimate in Table 5.

If payments to providers are higher under FFS than under CAP, then the responsiveness to contracts that we estimate above may be attributable to payment amount rather than to each contract's degree of payment retrospectiveness. Put differently, the contracts considered here are two-dimensional, varying in both *when* payments are made – ex ante vs ex post relative to the time of treatment – and the payment *amount*. We are interested in evaluating providers' responsiveness to the first dimension, conditional on payment amount. To do so, we estimate equation (2) controlling for the log of average per capita health care spending incurred during the delivery day. This variable is the sum of ex-ante and ex-post payments for services provided in the delivery day, so it captures payments for both services covered under CAP and under FFS. We use this variable as a proxy for payment amount for two reasons. First, a direct comparison of prices across contract types is not possible because prices have different units across contract types. Second, using prices as independent variable would constitute a "bad control" since prices depend on the delivery procedure contract. Our choice of using spending on the delivery day alleviates these two concerns.[10]

[10]There may still be reasons to believe that this more aggregated measure of spending constitutes a "bad control" if providing a c-section leads to more services being provided during the delivery day. Therefore, we take our results in Table 7 only as suggestive evidence of the effect of payment retrospectiveness.

Results are presented in Table 7. Column (1) shows that the impact of covering c-section under FFS is largely unaffected by the inclusion of this variable relative to Table 5. In this case, the use of FFS for c-section is related to a 26% increase in the c-section rate. Column (2) further indicates that the contract for vaginal deliveries has no significant effect on the outcome. Taken together, these findings suggest that medical treatment decisions are mainly influenced by the retrospective nature of payments.

TABLE 7: Impact of FFS Use on C-section Rates Controlling for Spending

	C-section rate	
	(1)	(2)
C-section FFS	0.123	
	(0.0359)	
Vaginal delivery FFS		-0.0348
		(0.0451)
Log spending on delivery day	0.0946	0.0693
	(0.0366)	(0.0385)
Constant	-0.149	-0.0161
	(0.221)	(0.222)
Sample	Scenarios 1 & 3	Scenarios 1 & 4
% change in outcome	25.5	-7.2
N	469	463
R^2	0.84	0.80

Note: Table shows the impact of the use of FFS for c-sections and for vaginal deliveries on the c-section rate controlling for the log of average per capita health care spending incurred during the delivery day. Definition of samples is provided in Table 4. An observation is an insurer-provider-woman type-year. A woman type is a combination of age group (age $\leq$ 27, age $>$ 27) and pregnancy risk (high vs. low). Specifications include insurer-provider-woman type and year fixed effects as well as insurer-specific time trends. Standard errors in parenthesis are clustered at the insurer-provider-woman type level. Table reports the percentage change in the outcome due to covering the procedure under FFS, calculated as the coefficient on FFS divided by the average c-section rate in the sample.

4.1 Marginal Women and Marginal Providers

If FFS causes a significant change in c-section rates, who are the women affected by the contracting decision and which providers drive the effect? In this subsection, we characterize marginal women and providers by exploring whether there are heterogeneous treatment effects. In Table 8 we explore heterogeneity in treatment effects by pregnancy risk, estimating equation (2) with the inclusion of an interaction between FFS^{s}_{jhmt} and an indicator for high-risk pregnancy. In both columns results indicate that there are no significant differences in responsiveness to contract type across high- and low-risk pregnancies. This finding is troubling given that for high-risk pregnancies, changes in delivery procedure choice due to financial incentives may lead to worse health outcomes compared to low-risk pregnancies.

We now move to evaluating which providers are more likely to respond to the financial incentives generated by FFS contracts. We do so by estimating heterogeneous treatment effects by provider size, as measured by the total number of deliveries each provider renders over the sample period. Our hypothesis is that relatively small provider organizations where administrators or managers who negotiate contracts are more likely to influence physicians' treatment decisions drive our main results.

We estimate equation (2) with the inclusion of an interaction between FFS^{s}_{jhmt} and an indicator for whether the provider renders less than 250 deliveries annually ("small provider"). In Table 9 we find that small providers are much more responsive to contracts: they are 52% more likely to use c-section

Table 8: Impact of FFS Use on C-section Rates by Pregnancy Risk

	C-section rate	
	(1)	(2)
C-section FFS × High risk	-0.0673	
	(0.0759)	
C-section FFS	0.169	
	(0.0495)	
Vaginal delivery FFS × High risk		-0.0604
		(0.0843)
Vaginal delivery FFS		-0.00131
		(0.0393)
Constant	0.424	0.396
	(0.0293)	(0.0393)
Sample	Scenarios 1 & 3	Scenarios 1 & 4
% change in outcome for high risk pregnancies	21.0	-12.7
% change in outcome for low risk pregnancies	34.9	-0.3
N	469	463
R^2	0.83	0.80

Note: Table shows the impact of the use of FFS for c-sections and for vaginal deliveries on the c-section rate by pregnancy risk in columns 1 and 2, respectively. Definition of samples is provided in Table 4. An observation is an insurer-provider-woman type-year. A woman type is a combination of age group (age ≤ 27, age > 27) and pregnancy risk (high vs. low). Specifications include insurer-provider-woman type and year fixed effects as well as insurer-specific time trends. Standard errors in parenthesis are clustered at the insurer-provider-woman type level. Table reports the percentage change in the outcome due to covering the procedure under FFS for high- and low-risk pregnancies, calculated as the coefficient on the interaction divided by the average c-section rate in the sample.

when it is reimbursed under FFS compared to 26% for large providers. Small providers are also 32% less likely to use c-section when vaginal delivery is reimbursed under FFS, while large providers are not more likely to use c-section in this case.

Our result that contracts substantially influence the c-section rate at a provider organization suggests that contracts ultimately influence physicians' decisions within these organizations. This can happen if provider organizations give incentives to physicians to perform procedures when they are covered

TABLE 9: Impact of FFS Use on C-section Rates by Provider Size

	C-section rate	
	(1)	(2)
C-section FFS × Small provider	0.126	
	(0.0396)	
C-section FFS	0.125	
	(0.0387)	
Vaginal delivery FFS × Small provider		-0.160
		(0.0513)
Vaginal delivery FFS		0.000679
		(0.0497)
Constant	0.422	0.396
	(0.0297)	(0.0390)
Sample	Scenarios 1 & 3	Scenarios 1 & 4
% change in outcome for small providers	52.0	-32.9
% change in outcome for large providers	25.9	0.1
N	469	463
R^2	0.83	0.80

Note: Table shows the impact of the use of FFS for c-sections by provider size. Definition of samples is provided in Table 4. A small provider is defined as one that provides fewer than 250 deliveries over the sample period. An observation is an insurer-provider-woman type-year. A woman type is a combination of age group (age $\leq$ 27, age $>$ 27) and pregnancy risk (high vs. low). Specifications include insurer-provider-type of woman and year fixed effects as well as insurer-specific time trends. Standard errors in parenthesis are clustered at the insurer-provider-woman type level. Table reports the percentage change in the outcome due to covering the procedure under FFS for small and large providers, calculated as the coefficient on the interaction divided by the average c-section rate in the sample.

under FFS relative to CAP. Some examples of these incentives include better control over schedules and salaries and more flexibility in medical practice style.[11] In fact, the finding that treatment effects are significant only among relatively small providers where the distinction between who owns the provider organization and who performs medical procedures is less marked substantiates this claim.

[11] See https://www.ama-assn.org/medical-residents/transition-resident-attending/it-matters-how-your-practice-paid-primer-young

4.2 Risk Selection

As discussed in section 2.2, contracts may influence incentives to engage in both risk selection and moral hazard. Under moral hazard, conditional on the pool of enrollees, providers will choose to perform the most profitable procedure. Under risk selection, insurers will choose to enroll women with low risk scores or low ex-ante health care costs if in-network providers are disproportionately covered under FFS. Our exercises so far show evidence of substantial provider moral hazard, as c-section rates vary significantly with payment retrospectiveness conditional on enrollment and women's characteristics.

In this subsection, we explore whether contracts generate outcomes consistent with risk selection. First, we show whether enrollment responds to contracts by regressing the switch-in rate for insurer j on our FFS indicators for this insurer. The switch-in rate is the fraction of women enrolled with j in year t who were enrolled with $-j$ in $t-1$. Second, we show whether insurers respond to risk selection incentives by regressing the log of lagged average per capita health care spending on our FFS indicators. This specification captures risk selection on the intensive margin as in Brown et al. (2014), as a negative correlation with the FFS indicators would suggest that insurers tend to send lower-cost patients to providers covered under FFS.

Table 10 presents the results. We find no relation between women's switching decisions and contracting decisions, and therefore no significant evidence that insurers engage in risk selection on the intensive margin. These results are both a test of the plausible exogeneity of contract variation within insurer,

Table 10: Impact of FFS Use on Risk Selection Incentives

	Switch-in rate		Log lagged cost	
	(1)	(2)	(3)	(4)
C-section FFS	0.00266		-0.0414	
	(0.0157)		(0.0778)	
Vaginal delivery FFS		0.0101		-0.486
		(0.0642)		(0.340)
Constant	0.0830	0.0735	13.84	13.88
	(0.0243)	(0.0327)	(0.152)	(0.267)
Sample	Scenarios 1 & 3	Scenarios 1 & 4	Scenarios 1 & 3	Scenarios 1 & 4
N	469	463	338	336
R^2	0.59	0.52	0.80	0.77

Note: Table shows the impact of the use of FFS for c-sections and for vaginal deliveries on the fraction of women who switch into an insurer (columns 1 and 2) and the log of lagged average health care cost per capita (columns 3 and 4). Definition of samples is provided in Table 4. For year t the fraction of women who switch-in is calculated as the number of women who change their insurer in t relative to $t-1$ divided by the total number of women enrolled with the insurer in t. To calculate the log of lagged average health care cost per capita we first compute for every woman her total annual health care cost and then we average the lagged measure across all women within a contracting unit. An observation is an insurer-provider-woman type-year. A woman type is a combination of age group (age $\leq$ 27, age $>$ 27) and pregnancy risk (high vs. low). Specifications include insurer-provider-woman type and year fixed effects as well as insurer-specific time trends. Standard errors in parenthesis are clustered at the insurer-provider-woman type level.

provider, and woman type and an evidence that contracts affect treatment choices and intensity of care mainly through provider moral hazard.

4.3 Health Outcomes

The public policy sphere has had a longstanding concern regarding the impacts on maternal and infant health outcomes from using c-sections among women who have a low suitability for the procedure (California Health Care Foundation, 2022). In fact, evidence on whether c-sections affect maternal health outcomes is mixed (Card et al., 2023; Fischer et al., 2023). In light of our results showing that contracts affect treatment decisions above and beyond patient health status, we investigate whether quasi-random variation in contracts for c-sections and for vaginal deliveries impact health.

In Table 11 we estimate equation (2) using as outcomes the average per capita birth weight and 1-minute APGAR score, and the fraction of women who experience SMM within 1 week and 1 month after giving birth. We find no impacts on any of these measures of maternal or infant health. Thus, the effects of contracts on delivery procedure choice, while economically significant, do not necessarily affect health outcomes.

Table 11: Impact of FFS Use on Maternal and Infant Health Outcomes

	Birth weight		1-min APGAR score		Severe maternal morbidity	
	(1)	(2)	(3)	(4)	(5)	(6)
C-section FFS	-0.0286		-0.0126		-0.00773	
	(0.0519)		(0.0543)		(0.0114)	
Vaginal delivery FFS		-0.00531		0.0197		-0.0424
		(0.0905)		(0.0790)		(0.0415)
Constant	3.034	3.047	8.334	8.316	0.0494	0.0284
	(0.0479)	(0.0625)	(0.0641)	(0.0922)	(0.00807)	(0.0140)
Sample	Scenarios 1 & 3	Scenarios 1 & 4	Scenarios 1 & 3	Scenarios 1 & 4	Scenarios 1 & 3	Scenarios 1 & 4
% change in outcome	-1.0	-0.2	-0.2	0.2	-20.9	-114.4
N	461	455	461	455	469	463
R^2	0.63	0.63	0.87	0.76	0.57	0.56

Note: Table shows the impact of the use of FFS for c-sections and for vaginal deliveries on maternal and infant health outcomes. Definition of samples is provided in Table 4. We consider as outcomes the average per capita infant's birth weight and 1-minute APGAR score, as well as the fraction of women who experience severe maternal morbidity (SMM) between 1 week and 1 month after giving birth. We construct the SMM indicator following the CDC definition. An observation is an insurer-provider-woman type-year. A woman type is a combination of age group (age ≤ 27, age > 27) and pregnancy risk (high vs. low). Specifications include insurer-provider-woman type and year fixed effects as well as insurer-specific time trends. Standard errors in parenthesis are clustered at the insurer-provider-woman type level. Table reports the percentage change in the outcome due to covering the procedure under FFS, calculated as the coefficient on FFS divided by the average c-section rate in the sample. Variation in the number of observations across specifications is due to missing values in some of our outcome variables.

4.4 Robustness Checks

We perform several robustness checks of our sample selection criteria and estimation procedure to provide suggestive evidence that our results capture causal effects. Appendix Figure 3 shows that our estimates are robust to using randomization inference on either the coefficient on c-section FFS or vaginal delivery FFS (see MacKinnon and Webb, 2020; Carpenter and Churchill, 2023). Therefore, our estimated coefficients are statistically larger than if they were generated by chance. Appendix Table 3 shows that our main estimates are robust to including woman type-specific time trends, which may capture changes in delivery practice styles for women with certain conditions across insurers and providers. Appendix Table 4 also shows that our main estimates are robust to replacing the insurer-specific time trends with insurer-by-year fixed effects. Insurer-by-year fixed effects will better capture non-linear variation over time in relative bargaining power which may confound the choice of contracts and the choice of procedure. In Appendix Table 5 we replace the insurer-specific time trends with insurer-woman type-specific linear time trends, finding that our results are robust.

To account for potential endogenous changes in the composition of women over time – such as changes generated by women switching across insurers based on the contracts – we construct our analysis sample focusing only on women who never switched their insurer over the sample period. Results in Appendix Table 6 show that our main estimates are robust to this sample selection criteria. We also provide results excluding the 4% of observations for which the modal contract does not equal the observed contract. Appendix Table 7 shows that our main estimates are unaffected by dropping these observations.

5 Conclusions

There is an impressive amount of theoretical literature establishing the incentives that health insurers and health care providers face under different contracts such as fee-for-service, capitation, and bundled payments. Yet, empirical evidence on whether contracts causally influence medical treatment decisions as the theory predicts is scant. Moreover, most empirical studies do not observe the contracting unit, the timing of contracts, and in many cases the contracts themselves. With our data and setting, we are able to overcome each of these limitations.

In this paper, we estimate the plausibly causal effect of FFS contracts relative to CAP contracts in the context of delivery procedures. We exploit within insurer-provider pair variation in contracts for c-sections and for vaginal deliveries using data from Colombia. In the Colombian health care system, insurers and providers may cover c-sections and vaginal deliveries under different contracts. We describe how contracts are determined in this setting, what are the resulting incentives for insurers and providers, and how these incentives vary across delivery procedures.

We find that covering c-sections under FFS increases the c-section rate by 27% while covering vaginal deliveries under FFS does not impact delivery procedure choice. We show that these effects are explained mainly by the degree of payment retrospectiveness of each contract rather than by the payment amount negotiated between insurers and providers. Our results indicate that bilateral contracting decisions between insurers and providers causally impact the type, cost, and intensity of care that women receive. While we find no effect of contracts on short-run measures of maternal and infant health, we are unable to test whether there are long-run impacts on health outcomes. This causal relation between contracts and medical treatment decisions matters for health policy as improving patient and population health outcomes may require regulating insurer-provider bilateral contracting decisions.

References

ACQUATELLA, A. (2021): "Evaluating Prospective Payment Contracts," Working paper.

ADIDA, E., H. MAMANI, AND S. NASSIRI (2017): "Bundled Payment vs. Fee-for-Service: Impact of Payment Scheme on Performance," *Management Science*, 63, 1606–1624.

AGARWAL, R., J. M. LIAO, A. GUPTA, AND A. S. NAVATHE (2020): "The Impact Of Bundled Payment On Health Care Spending, Utilization, And Quality: A Systematic Review," *Health Affairs*, 39, 50–57.

AHRQ (2018): "Healthcare Cost and Utilization Project Statistical Brief 281," https://hcup-us.ahrq.gov/reports/statbriefs/sb281-Operating-Room-Procedures-During-Hospitalization-2018.jsp.

Aizer, A., J. Currie, and E. Moretti (2007): "Does Managed Care Hurt Health? Evidence from Medicaid Mothers," *The Review of Economics and Statistics*, 89, 385–399.

Baker, L. C. (1997): "The Effect of HMOs on Fee-for-Service Health Care Expenditures: Evidence from Medicare," *Journal of Health Economics*, 16, 453–481.

——— (1999): "Association of Managed Care Market Share and Health Expenditures for Fee-for-Service Medicare Patients," *Journal of the American Medical Association*, 281, 432–437.

Betrán, A. P., J. Ye, A.-B. Moller, J. Zhang, A. M. Gülmezoglu, and M. R. Torloni (2016): "The Increasing Trend in Caesarean Section Rates: Global, Regional and National Estimates: 1990-2014," *PLOS One*, 11, e0148343.

Brown, J., M. Duggan, I. Kuziemko, and W. Woolston (2014): "How Does Risk Selection Respond to Risk Adjustment? New Evidence from the Medicare Advantage Program," *American Economic Review*, 104, 3335–3364.

California Health Care Foundation (2022): "Reducing unnecessary cesarean-section deliveries in California," https://www.chcf.org/wp-content/uploads/2017/12/PDF-ReducingCSectionsFlier.pdf.

Card, D., A. Fenizia, and D. Silver (2023): "The Health Impacts of Hospital Delivery Practices," *American Economic Journal: Economic Policy*, 15, 42–81.

Carpenter, C. S. and B. F. Churchill (2023): "'There She Is, Your Ideal': Negative Social Comparisons and Health Behaviors," NBER Working Paper no. 31156.

Clemens, J. and J. D. Gottlieb (2014): "Do Physicians' Financial Incentives Affect Medical Treatment and Patient Health?" *American Economic Review*, 104, 1320–1349.

Einav, L., A. Finkelstein, Y. Ji, and N. Mahoney (2022): "Voluntary Regulation: Evidence from Medicare Payment Reform," *The Quarterly Journal of Economics*, 137, 565–618.

Ellis, R. P. and T. G. McGuire (1986): "Provider Behavior under Prospective Reimbursement: Cost Sharing and Supply," *Journal of Health Economics*, 5, 129–151.

——— (1996): "Hospital Response to Prospective Payment: Moral Hazard, Selection, and Practice-Style Effects," *Journal of Health Economics*, 15, 257–277.

Fischer, S. J., H. Royer, and C. D. White (2023): "Health Care Centralization: The Health Impacts of Obstetric Unit Closures in the US," Forthcoming, *American Economic Journal: Applied Economics*.

Foo, P. K., R. S. Lee, and K. Fong (2017): "Physician Prices, Hospital Prices, and Treatment Choice in Labor and Delivery," *American Journal of Health Economics*, 3, 422–453.

Gaynor, M., N. Mehta, and S. Richards-Shubik (2023): "Optimal Contracting with Altruistic Agents: Medicare Payments for Dialysis Drugs," *American Economic Review*, 113, 1530–1571.

Godager, G. and D. Wiesen (2013): "Profit or Patients' Health Benefit? Exploring the Heterogeneity in Physician Altruism," *Journal of Health Economics*, 32, 1105–1116.

Ho, K. and A. Pakes (2014a): "Hospital Choice, Hospital Prices, and Financial Incentives to Physicians," *American Economic Review*, 104, 3841–3884.

——— (2014b): "Physician Payment Reform and Hospital Referrals," *American Economic Review: Papers and Proceeedings*, 104, 200â205.

Johnson, E. M. and M. M. Rehavi (2016): "Physicians Treating Physicians: Information and Incentives in Childbirth," *American Economic Journal: Economic Policy*, 8, 115–141.

Kuziemko, I., K. Meckel, and M. Rossin-Slater (2018): "Does Managed Care Widen Infant Health Disparities? Evidence from Texas Medicaid," *American Economic Journal: Economic Policy*, 10, 255–83.

Ma, C.-t. A. and H. Y. Mak (2015): "Information Disclosure and the Equivalence of Prospective Payment and Cost Reimbursement," *Journal of Economic Behavior & Organization*, 117, 439–452.

——— (2019): "Incentives in Healthcare Payment Systems," *Oxford Research Encyclopedia, Economics and Finance*.

MacKinnon, J. G. and M. D. Webb (2020): "Randomization Inference for Difference-in-Differences with Few Treated Clusters," *Journal of Econometrics*, 218, 435–450.

Otero, C. and P. Munoz (2022): "Managers and Public Hospital Performance," Working paper.

Press, M. J., R. Rajkumar, and P. H. Conway (2016): "Medicareâs New Bundled Payments: Design, Strategy, and Evolution," *Journal of the American Medical Association*, 315, 131–132.

Robinson, S., H. Royer, and D. Silver (2024): "Geographic Variation in Cesarean Sections in the United States: Trends, Correlates, and Other Interesting Facts," *Journal of Labor Economics*, 42, S219 – S259.

Sakala, C., S. Delbanco, and H. Miller (2013): "The Cost of Having a Baby in the United States," Truven Health Analytics Marketscan Study, https://www.nationalpartnership.org/our-work/resources/health-care/maternity/archive/the-cost-of-having-a-baby-in-the-us.pdf.

Serna, N. (2024): "Exogenous Exits, Market Structure, and Equilibrium Contracts in Health Care," Working paper.

Appendix A Data Cleaning

The raw data contains all the delivery procedures for women enrolled with insurers participating in the contributory health care system during 2013 to 2015. The total number of observations in this data is N=521,408. To obtain our final analysis sample we proceed in the following steps:

1. Keep singleton, non-breech, first-time births (N=437,191).

2. Keep deliveries that are reimbursed under fee-for-service, capitation, or bundled payment contracts (N=343,967).

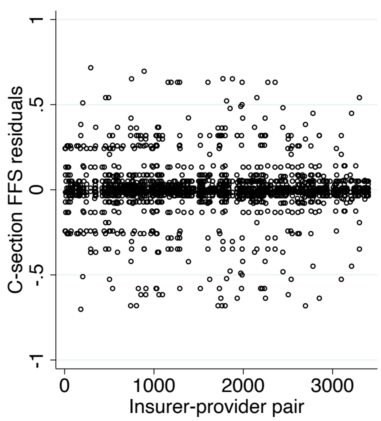
3. Obtain the modal contract per service (c-sections and vaginal deliveries), insurer, provider, type of woman, and month. A type of woman is defined by a combination of age group ($<=27$, >27) and pregnancy risk. The age group cutoff represents the average age in the data.

4. Collapse the data to the insurer, provider, type of woman, and year, retaining the modal contract and computing c-section rates and average per-capita measures of health care spending and of maternal and infant health outcomes.

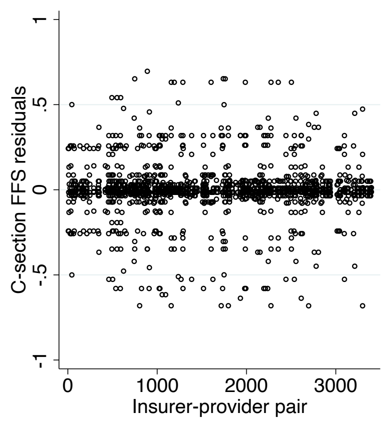
Appendix B Residual Variation in Contracts

Appendix Figure 1: Residual Variation in Use of FFS for C-sections within Insurer-Provider

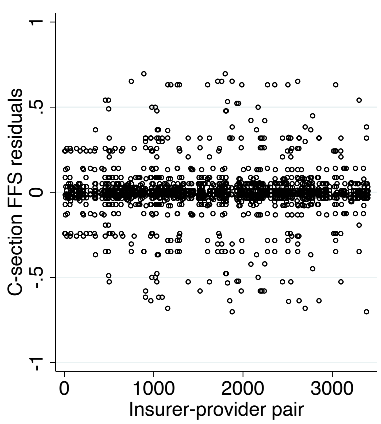
(a) Age≤27, Low risk



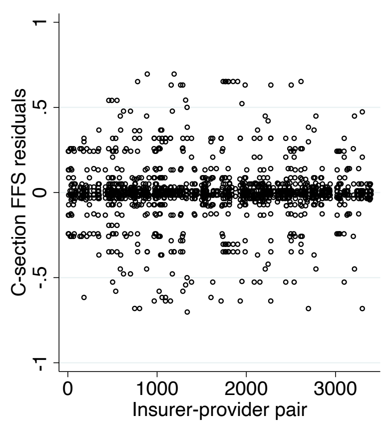
(b) Age≤27, High risk



(c) Age >27, Low risk



(d) Age >27, High risk



Note: Figure shows the distribution of residuals from a regression of our predicted contract for c-sections on insurer-provider-woman type fixed effects, year fixed effects, and insurer-specific time trends. We report the distribution conditional on the woman's age group and pregnancy risk.

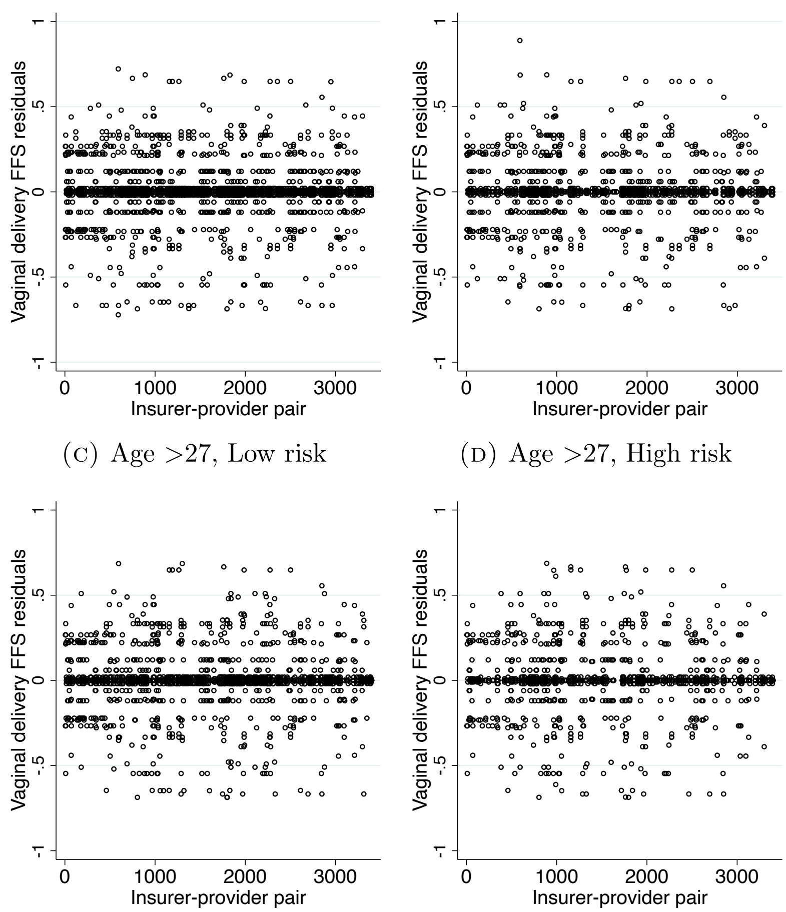
Appendix Figure 2: Residual Variation in Use of FFS for Vaginal Deliveries within Insurer-Provider

(a) Age≤27, Low risk

(b) Age≤27, High risk

(c) Age >27, Low risk

(d) Age >27, High risk



*Note:*Figure shows the distribution of residuals from a regression of our predicted contract for vaginal deliveries on insurer-provider-woman type fixed effects, year fixed effects, and insurer-specific time trends. We report the distribution conditional on the woman's age group and pregnancy risk.

Appendix C Robustness Checks

Appendix Table 1: Correlates of Contract Switching Controlling for Insurer-Year Fixed Effects

	Contract switch	
	(1)	(2)
	C-section	Vaginal delivery
Woman-type share of deliveries	-0.139	0.0410
	(0.0835)	(0.0852)
Lagged delivery characteristics		
SMM	-0.000341	0.0199
	(0.0774)	(0.127)
Birth weight (kilograms)	-0.0218	0.0140
	(0.0296)	(0.0395)
1-minute APGAR score	0.00298	0.0166
	(0.0162)	(0.0172)
Log spending on delivery day	-0.0286	0.0450
	(0.0375)	(0.0479)
Log spending up to delivery day	-0.0252	-0.0189
	(0.0189)	(0.0231)
Woman-type share of deliveries	-0.0472	-0.0636
	(0.0898)	(0.0945)
Constant	0.519	-0.221
	(0.346)	(0.400)
N	2106	1668
R^2	0.58	0.62

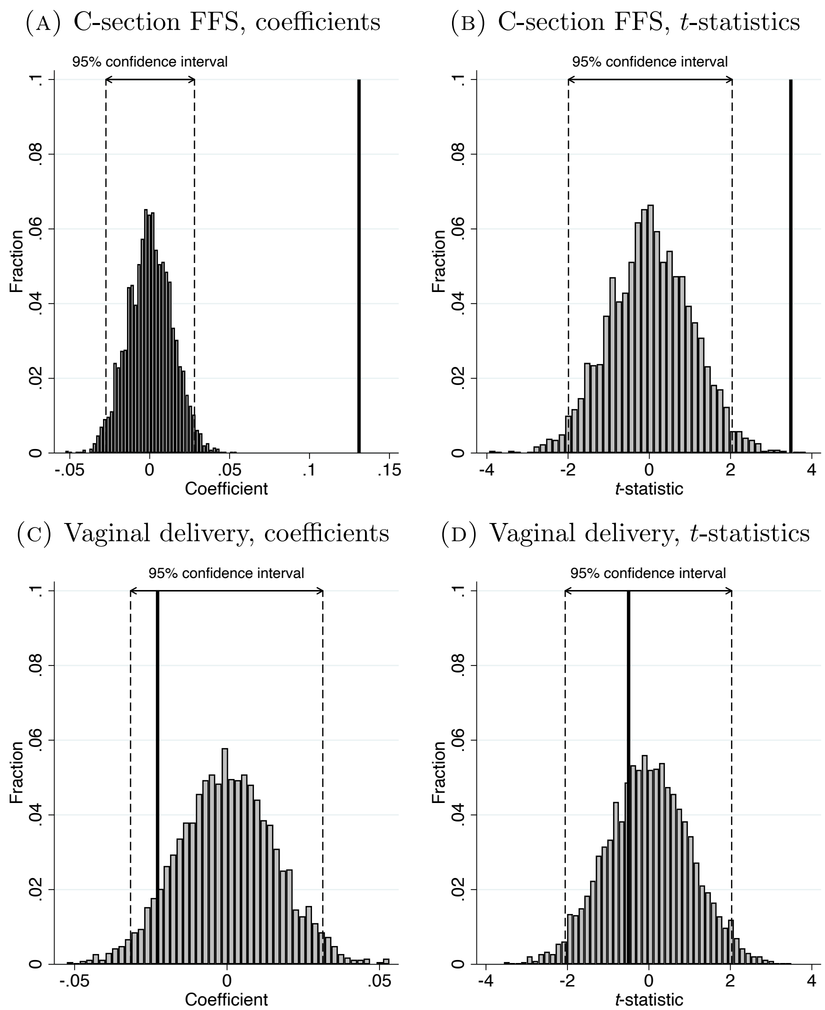
Note: Table shows the correlates of contract switches for c-sections and for vaginal deliveries in columns 1 and 2, respectively. An observation is an insurer-provider-woman type-year. A woman type is a combination of age group (age ≤ 27, age > 27) and pregnancy risk (high vs. low). Lagged delivery characteristics correspond to the lagged average per capita measure. Specifications include insurer-provider-woman type and year fixed effects as well as insurer-year fixed effects. Standard errors in parenthesis are clustered at the insurer-provider-woman type level. Measures of spending are averages per capita.

Appendix Table 2: Impact of Capitation Use on C-section Rates for Different Contract Combinations

	C-section rate	
	(1)	(2)
C-section Capitated	-0.0238	
	(0.0248)	
Vaginal delivery Capitated		0.266
		(0.0962)
Constant	0.488	0.218
	(0.00945)	(0.0646)
Sample	Scenarios 2 & 3	Scenarios 2 & 4
% change in outcome	-4.9	54.8
N	3482	92
R^2	0.78	0.79

Note: Table shows the impact of the use of capitation for c-sections and for vaginal deliveries on the c-section rate in columns 1 and 2, respectively. Definition of samples is provided in Table 4. An observation is an insurer-provider-woman type-year. A woman type is a combination of age group (age $\leq$ 27, age $>$ 27) and pregnancy risk (high vs. low). Specifications include insurer-provider-woman type and year fixed effects as well as insurer-woman type-specific time trends. Standard errors in parenthesis are clustered at the insurer-provider-woman type level. Table reports the percentage change in the outcome due to covering the procedure under FFS, calculated as the coefficient on FFS divided by the average c-section rate in the sample.

Appendix Figure 3: Randomization Inference on Main Model Coefficients



Note: Figures show the distribution of placebo coefficients and cluster-robust t-statistics obtained from randomly assigning contract types 5,000 times and estimating Equation 2 using the full analysis sample and using the c-section rate as the dependent variable. The solid black line in each panel shows the estimated coefficient or t-statistic from using the actual contract type. These coefficients can be found in Table 6.

Appendix Table 3: Impact of FFS Use on C-section Rates Controlling for Woman Type-Specific Time Trends

	C-section rate	
	(1)	(2)
C-section FFS	0.131	
	(0.0379)	
Vaginal delivery FFS		-0.0209
		(0.0428)
Constant	0.387	0.341
	(0.0439)	(0.0495)
Sample	Scenarios 1 & 3	Scenarios 1 & 4
% change in outcome	27.1	-4.3
N	469	463
R^2	0.84	0.80

Note: Table shows the impact of the use of FFS for c-sections and for vaginal deliveries on the c-section rate in columns 1 and 2, respectively. Definition of samples is provided in Table 4. An observation is an insurer-provider-woman type-year. A woman type is a combination of age group (age ≤ 27, age > 27) and pregnancy risk (high vs. low). Specifications include insurer-provider-woman type and year fixed effects as well as insurer- and woman type-specific time trends. Standard errors in parenthesis are clustered at the insurer-provider-woman type level. Table reports the percentage change in the outcome due to covering the procedure under FFS, calculated as the coefficient on FFS divided by the average c-section rate in the sample.

Appendix Table 4: Impact of FFS Use on C-section Rates Controlling for Insurer-Year Fixed Effects

	C-section rate	
	(1)	(2)
C-section FFS	0.138	
	(0.0394)	
Vaginal delivery FFS		-0.0153
		(0.0462)
Constant	0.431	0.440
	(0.00539)	(0.00528)
Sample	Scenarios 1 & 3	Scenarios 1 & 4
% change in outcome	28.5	-3.2
N	467	463
R^2	0.83	0.80

Note: Table shows the impact of the use of FFS for c-sections and for vaginal deliveries on the c-section rate in columns 1 and 2, respectively. Definition of samples is provided in Table 4. An observation is an insurer-provider-woman type-year. A woman type is a combination of age group (age ≤ 27, age > 27) and pregnancy risk (high vs. low). Specifications include insurer-provider-woman type fixed effects, year fixed effects, and insurer-year fixed effects. Standard errors in parenthesis are clustered at the insurer-provider-woman type level. Table reports the percentage change in the outcome due to covering the procedure under FFS, calculated as the coefficient on FFS divided by the average c-section rate in the sample.

Appendix Table 5: Impact of FFS Use on C-section Rates Controlling for Insurer-Woman Type Specific Time Trends

	C-section rate	
	(1)	(2)
C-section FFS	0.129	
	(0.0381)	
Vaginal delivery FFS		-0.0114
		(0.0384)
Constant	0.376	0.308
	(0.0764)	(0.0982)
Sample	Scenarios 1 & 3	Scenarios 1 & 4
% change in outcome	26.7	-2.3
N	469	463
R^2	0.84	0.81

Note: Table shows the impact of the use of FFS for c-sections and for vaginal deliveries on the c-section rate in columns 1 and 2, respectively. Definition of samples is provided in Table 4. An observation is an insurer-provider-woman type-year. A woman type is a combination of age group (age ≤ 27, age > 27) and pregnancy risk (high vs. low). Specifications include insurer-provider-woman type and year fixed effects as well as insurer-woman type-specific time trends. Standard errors in parenthesis are clustered at the insurer-provider-woman type level. Table reports the percentage change in the outcome due to covering the procedure under FFS, calculated as the coefficient on FFS divided by the average c-section rate in the sample.

Appendix Table 6: Impact of FFS Use on C-section Rates for Subsample of Inertial Women

	C-section rate	
	(1)	(2)
C-section FFS	0.124	
	(0.0373)	
Vaginal delivery FFS		-0.0109
		(0.0781)
Constant	0.424	0.388
	(0.0327)	(0.0428)
Sample	Scenarios 1 & 3	Scenarios 1 & 4
% change in outcome	25.3	-2.2
N	469	463
R^2	0.82	0.75

Note: Table shows the impact of the use of FFS for c-sections and for vaginal deliveries on the c-section rate in columns 1 and 2, respectively. Definition of samples is provided in Table 4. Regressions are estimated using the subsample of inertial women who do not change their insurer during the sample period. An observation is an insurer-provider-woman type-year. A woman type is a combination of age group (age ≤ 27, age > 27) and pregnancy risk (high vs. low). Specifications include insurer-provider-woman type and year fixed effects as well as insurer-specific time trends. Standard errors in parenthesis are clustered at the insurer-provider-woman type level. Table reports the percentage change in the outcome due to covering the procedure under FFS, calculated as the coefficient on FFS divided by the average c-section rate in the sample.

Appendix Table 7: Impact of FFS Use on C-section Rates Excluding Contracts that do not Equal Mode

	C-section rate	
	(1)	(2)
C-section FFS	0.144	
	(0.0425)	
Vaginal delivery FFS		-0.0407
		(0.0469)
Constant	0.404	0.387
	(0.0317)	(0.0387)
Sample	Scenarios 1 & 3	Scenarios 1 & 4
% change in outcome	30.0	-8.5
N	469	463
R^2	0.82	0.80

Note: Table shows the impact of the use of FFS for c-sections and for vaginal deliveries on the c-section rate in columns 1 and 2, respectively. Definition of samples is provided in Table 4. Deliveries for which the observed contract type does not equal the modal contract type for the insurer-provider-woman type-year group to which that delivery belongs are dropped. An observation is an insurer-provider-woman type-year. A woman type is a combination of age group (age ≤ 27, age > 27) and pregnancy risk (high vs. low). Specifications include insurer-provider-woman type and year fixed effects as well as insurer-specific time trends. Standard errors in parenthesis are clustered at the insurer-provider-woman type level. Table reports the percentage change in the outcome due to covering the procedure under FFS, calculated as the coefficient on FFS divided by the average c-section rate in the sample.